Large-Scale Applications of

Multi-Reference Methods in Chemistry

and

Development of a Multi-Reference Møller-Plesset

Perturbation Theory Program

Dissertation zur Erlangung des

naturwissenschaftlichen Doktorgrades

der Bayerischen Julius-Maximilians-Universität Würzburg

vorgelegt von

Patrick Musch

aus Obernburg a. Main

Würzburg 2003

Eingereicht am: _____

bei der Fakultät für Chemie und Pharmazie

1. Gutachter: _____

2. Gutachter: _____

der Dissertation

1. Prüfer: _____

2. Prüfer: _____

3. Prüfer: _____

des Öffentlichen Promotionskolloquiums

Tag des Öffentlichen Promotionskolloquiums: _____

Doktorurkunde ausgehändigt am: _____

meiner Familie

"A man of genius makes no mistakes; his errors are volitional
and are the portals of discovery."

James Joyce (1882 - 1941)

"Experience is the name everyone gives to their mistakes."

Oscar Wilde, Lady Windermere's Fan, 1892, Act III

Parts of the present work are published in journals[108,109,110] and poster contributions.

Acknowledgement

Contents

List of Figures

List of Tables

List of Listings

Chapter 1

Introduction

1.1 The Contribution of Multi-Reference Methods to the Future in Computational Chemistry

The field of computational chemistry is tremendously influenced by trends in chemistry towards material and biological sciences. Both fields of material science and biological chemistry are characterized by large assemblies of molecules with complex compositions in the form of supermolecules or enzymes. The complexity of these systems poses novel challenging questions to the field of computational chemistry. This challenge is taken and the successful development of more efficient computational methods and their application contributes to a more thorough understanding of the factors that control material properties or enzymatic activities with certain substrates or to determine their mode of action.

The methodological approaches to model enzyme structures and binding sites for substrates is traditionally based on force field methods.[1] These approaches are accurate enough to give some insights into equilibrium structures and provide a qualitative understanding of the binding sites, but reliable predictions for reaction profiles or already for the geometrical structure and the electronic states for many transition-metal containing enzymes are very difficult to obtain. Therefore well-founded predictions based on reliable theoretical methods are the challenge of computational chemistry today and in the future.

For enzymes a combination of quantum-mechanical and force field methods can be ap-
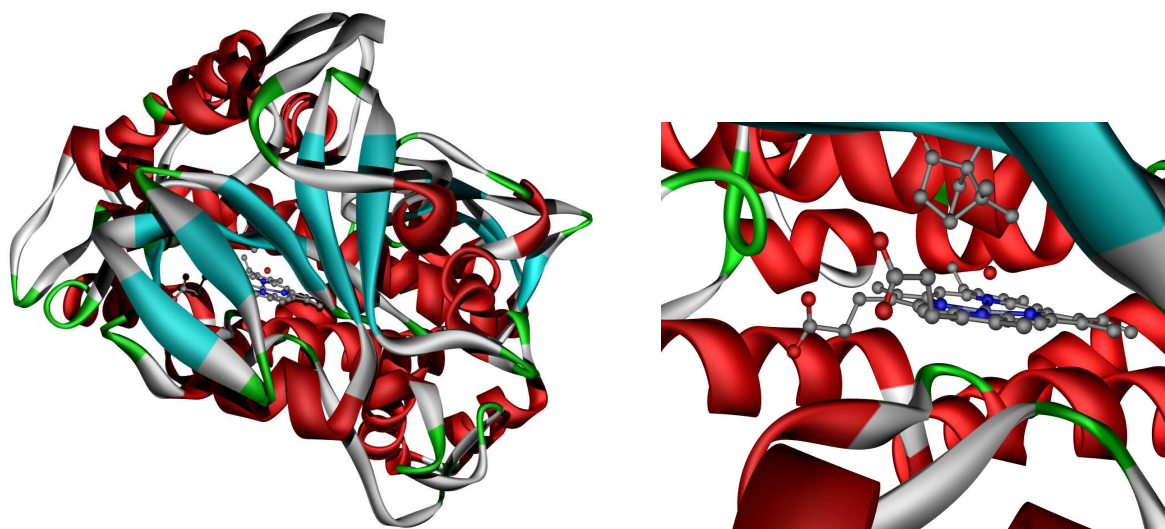
Figure 1.1: X-ray structure of a Cytochrome P450 enzyme with camphor as oxidized compound (left-hand side) and the Fe-porphyrine active site of the enzyme (right-hand side). The structure is available from the Brookhaven Protein Database, PDB code 1DZ9.[24]

plied[2] where the reaction at the active site and the closest surrounding is described by quantum-mechanical methods (QM part) and the remaining protein is modeled by means of molecular mechanics (MM part). The approach abbreviated as QM/MM helped substantially to understand enzyme reactions like the mode of action of a *Cytochrome P450*.[3, 4] Recent developments for MP2,[5–9] Coupled Cluster[10–13] or DFT[14–17] methods have reduced the computational cost for large-scale systems to a linear-dependency allowing the treatment of systems up to some hundred atoms and up to a few thousand basis functions. The accessibility of methods treating systems of that size is very appealing and lead to fruitful studies in many cases.

The accuracy of these methods, however, is intimately connected with the assumption that a wavefunction, consisting of only one Slater determinant, already represents a very good approximation, *i.e.* encompasses the nature of the electronic state under consideration to 80% or more. For many systems, *e.g.* radicals, diradicals,[18] or transition metal compounds with open shells,[19–23] this assumption does not hold anymore. Here a one determinant wavefunction reflects much less of the character of the electronic state and as a consequence, all methods which base on this single-reference starting point become quite unreliable.

One example of a system that possesses such a difficult electronic structure is the iron-porphyrine active site in Figure 1.1 oxidizing a camphor substrate. The QM part was

described by DFT methods although the substantial difficulties with open-shell transition metals are known.[19–23] For these systems multi-reference methods like multi-reference configuration interaction (MR-CI) or multi-reference perturbation theory (MR-PT) should provide more accurate results because they rely on a more thorough theoretical foundation. At least, such methods are needed to validate approaches that are supposed to be inferior in their theoretical foundation. Applications of multi-reference methods to systems of that size which usually do not possess any symmetry constraints have not been applied routinely[1] but the development of more efficient algorithms to establish these methods is in steady progress.

1.2   A Brief Overview on Multi-Reference Methods

The advantage of multi-reference methods is their great flexibility which enable these approaches to describe reaction paths or excited states very accurately.[28, 29] In the beginning of the field of computational chemistry multi-reference methods like MR-CI focused on the properties of atomic or small molecular systems. The choice of the systems was originally limited by the available computer power and restricted to the access to a few "supercomputers". The nowadays application of multi-reference methods to problems which are of interest to the experimental chemist or physicist, would have been impossible without the rapidly growing computational performance at a reasonable cost of ownership and without the development of new concepts and algorithms to reduce the scaling of the computational cost without a (significant) sacrifice of accuracy.

One of these algorithms is the selecting MR-CI of Buenker and Peyerimhoff[30, 31] which was later improved by Hanrath and Engels[32] in form of the DIESEL program package. In individual selection schemes an estimate of the energy contribution of each configuration to the total energy is used to select those configurations which are important and enter the CI routines. The contributions of the neglected space are extrapolated by the computed energy estimates employing perturbational approaches. This ansatz surely requires a certain experience but reproduces as accurate results as full MR-CI procedures. This

---

[1]In this context computations of porphyrine systems are commonly used for "large-scale" benchmark computations[25–27] but are usually performed in D_{2h} symmetry, which reduces the computational cost (integrals and time) by about a factor roughly corresponding to the order of the point group.

individual selection concept was transferred later by Grimme to a MR-MP2 program.[33] The experience which is required to use the individual selection methods in a reliable way originates from the nearly arbitrary composition of the selected interaction space. In the selected space basically every combination of interacting configurations has to be evaluated explicitly unless sophisticated data structures to represent the selected configuration spaces are applied as in the approach of Hanrath and Engels. The full evaluation requires a large amount of computational time once the selected spaces grow rapidly as for large molecules but the required effort is reduced by the latter approach.

Another approach to reduce the computational cost of a MR-CI was proposed by Meyer[5, 34] who internally contracted the reference wavefunction. The interaction space was generated by annihilation and creation operators on the whole reference wavefunction and not on each configuration. The internal contraction scheme reduces the computational cost to that of a single-reference CI treatment, but increases the complexity of the evaluation of interactions between the spin eigenfunctions. The internally contracted methods are not exact but represent approximations to the solution of the full problem with negligible errors.[5] Werner and Knowles also applied a partitioning scheme of the configuration space which is necessary to achieve an efficient implementation of an integral-direct method. Modified algorithms of these concepts were also applied to an internally contracted MR-PT2[27] in analogy to the CASPT2 ansatz of Andersson and Roos.[35–37]

The ansatz of local correlation treatments,[5, 38–41] which was successfully applied for single-reference post Hartree-Fock methods like MP2,[5–9, 42–44] CCSD,[5, 10, 11, 45, 46] or higher Coupled Cluster schemes,[5, 12, 13, 47] also entered the multi-reference field. Local correlation treatments have their basis in the locality of correlation effects. This means that only a rather small assembly of orbitals in close neighbourhood is supposed to contribute significantly to the correlation energy. This approach therefore requires the use of local instead of canonical orbitals and those orbitals assigned to one assembly is called an orbital domain. The localization procedure is rather crucial for the performace of the method, because the computational cost for the treatment of one domain possesses the same scaling as the canonical method. Energy contributions of interactions between different domains are extrapolated by less expensive methods (MP2 terms in the local Coupled Cluster approach or multipole approximations for the local MP2). The choice of the domains is therefore substantial for the computational cost of the local correlation treatment and the reliability of the results.

Based on these concepts, local MR-PT2[48] and local MR-CI[9, 49] approaches have been developed. The computational cost was reduced to a certain degree, but a reliable estimate on the errors coming along with this approximation has not yet been assessed for reaction paths or for systems with a difficult electronic or a bulky geometrical structure.[50]

1.3   Intensions and Objects of this Work

Beside their significance for the answer on chemical questions, the application of multi-reference methods on systems with typical multi-reference character is important to evaluate the scope of single-reference approaches with respect to their capability to describe such complex systems. The reactivity and electronic structures of three systems representing typical multi-reference cases were characterized in this project and in each case the reliability of the computational approaches was assessed. The computations concerning the first system explore the reactivity of the biologically active Neocarzinostatin chromophore, which is known as an anti-tumor antibiotics. Compounds like the enyne-ketenes are investigated to analyse the similarites and differences of related classes of compounds. The role of the allenic, diradical, and zwitterionic intermediates, which are found in the cyclization of the enyne-allenes and enyne-ketenes is further investigated for the example of the 1,2,4-cyclohexatriene compound and its hetero-substituted analogues. The computations were very successful since the results could answer some open questions and delivered some very surprising new knowledge about the studied systems.

The experience with various packages like the MOLCAS,[51] MOLPRO[52] or DIESEL[53] suite of programs, however, showed the limitations for the application of multi-reference perturbation theory and multi-reference configuration interaction approaches for large molecules, which forced the use of simpler approaches. These limitations originate from the integral handling, the implementations of the interacting space, and memory considerations and are inherently founded in the design of the algorithms. The computational limits in the DIESEL package concerning the MR-MP2 and MR-CI parts are mainly attibuted to the integral handling. The program package was originally designed for small to medium-sized molecules possessing up to 150 basis functions and requires the accessibility to all MO integrals in memory. This limits the maximum system size to about 250 basis functions on

commodity 32 bit hardware. Further on, parallelization was only taken into account for the selecting MR-CI procedure and implemented for shared-memory computers. The memory requirements during the iterations to store the coefficient vectors set the limit to systems which possess up to 100 million configuration state functions.

For large-scale multi-reference computations and in particular for the more efficient MR-MP2 approach, an improved code was of interest. The major object of this project is to provide a modern MR-MP2 code applicable to large molecules at a reasonable computational cost. The idea of local correlation treatments was not considered, because the definition of orbital domains for complex electronic systems is supposed to be of the same difficulty to the user as applying some selection criteria for a truncation of the interaction space.

The requirements for a code capable to perform large-scale MR-MP2 computations define the scope of this project and are listed below:

1. The program should be applicable to molecules with up to 1000 basis functions and configuration spaces of up to a few billion configuration state functions.

2. The structure of the program should be designed to easily incorporate extensions to the code with respect to flexible selection schemes or other methods like MR-CI. The selection schemes could either employ an individual selection of configurations or a scheme based on the contributions of particular external orbitals.

3. The integrals should be only used if needed and the specific integral properties of the MR-MP2 method should be applied to reduce integral storage as much as possible.

4. The large configuration spaces and coefficient vectors should be treated in parallel to distribute the computational cost.

5. The storage requirements for the vectors in a parallel computation should also be reduced as much as possible.

6. For large-scale computations the construction of the Hamilton matrix is the most time-consuming step and should be optimized as much as possible.

An appropriate modification of the DIESEL package with respect to a parallelization routine for distributed memory systems[54] and improved integral handling was intended

first, but it turned out that the implementation into the existing package and porting the C++ sources to recent C++ standard conform compilers would cost as much effort as an implementation from scratch using some parts of the DIESEL package. Because of this reason the decision to program a new code from scratch was taken.

Chapter 2

Multi-Reference Perturbation Theory

The computation of reaction profiles and molecular properties constitutes the major part of quantum-chemical applications with the intension to gain more insight and a more thorough understanding about certain classes of molecules or reactions. Before the advent of appropriate functionals in Density Functional Theory (DFT), the most common approaches to perform these calculations were *ab initio* methods like Hartree-Fock (HF), second-order Møller-Plesset perturbation theory (MP2), Configuration Interaction (CI) and the singles and doubles Coupled Cluster ansatz (CCSD) including perturbative triples (CCSD(T)). The HF approach is situated at the low end of the methodical hierarchy of *ab initio* methods and provides the basis for all post Hartree-Fock theories (*e.g.* MP2, CI or CCSD(T)). The latter ones and DFT are preferably used, since the inclusion of electron correlation is necessary for a reliable description of the desired quantities. To give an impression about the amount of the electron correlation, its relative value is rather small compared to the total energy. The Hartree-Fock approach already accounts for 99% of the exact energy, but the energetic profiles of chemical processes or molecular properties are determined by differences in the remaining fraction of 1% of the correlation energy. In cases with near-degeneracies single-determinant based correlation methods and DFT cannot describe the systems correctly and multi-reference methods have to be employed. Like the MP2 ansatz, the multi-reference perturbation theory (MR-PT) is the most cost-effective multi-reference approach compared to multi-reference CI (MR-CI) and multi-reference Coupled Cluster (MR-CC).

Prior to the formulation of the Møller-Plesset perturbation theory and its generalization

to multi-reference wavefunctions, a brief outline[1] of Hartree-Fock theory is given to point out some issues that are relevant for perturbation theories. An overview on other existing multi-reference perturbation theory approaches finishes the chapter.

2.1   The Hartree-Fock Approach

The Hartree-Fock approach represents the most simple *ab initio* method to approximate the solution of the time-independent Schrödinger equation (eq. 2.1) for many-electron systems.

$$\hat{H}|\psi\rangle = E|\psi\rangle \tag{2.1}$$

Eq. 2.1 is an eigenvalue equation of the Hamilton operator $\hat{H}$ acting on the electronic wavefunction $|\psi\rangle$ and the energy E as the eigenvalue for a specific $|\psi\rangle$. For a normalized wavefunction $|\psi\rangle$ the energy expectation value (eq. 2.2) is obtained by projection of eq. 2.1 onto the complex conjugate $\langle\psi|$.

$$E = \langle\psi|\hat{H}|\psi\rangle \tag{2.2}$$

The most simple approximation for a many-electron wavefunction $|\psi\rangle$ that obeys the Pauli principle is a single Slater determinant Φ_{Slater} (eq. 2.3). In the given example, the spin-orbitals are ordered row-wise, while the electronic coordinates are ordered column-wise.

$$\Phi_{Slater} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \cdots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \cdots & \phi_N(2) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \phi_1(N) & \phi_2(N) & \cdots & \phi_N(N) \end{vmatrix} \tag{2.3}$$

The evaluation of the energy expression from eq. 2.2 using the single determinant wavefunction in eq. 2.3 reduces to sums over a vector $\boldsymbol{h}$ and the Coulomb and Exchange matrices $\boldsymbol{J}$ and $\boldsymbol{K}$, respectively[2].

---

[1]Details about the HF approach can be found in standard quantum-chemistry textbooks.[55, 56]

[2]Notation: Many-particle operators are denoted in capital letters and single-particle operators in minor letters. Matrices are given in bold faced letters and their corresponding matrix elements are subscripted

$$E \;=\; \sum_i^N h_i + \sum_i^N \sum_{j>i}^N \left(J_{ij} - K_{ij} \right) + V_{nn} \tag{2.4}$$

$$=\; \sum_i^N h_i + \frac{1}{2} \sum_i^N \sum_j^N \left(J_{ij} - K_{ij} \right) + V_{nn} \tag{2.5}$$

The indices i and j run over the occupied spin-orbitals. Equation 2.5 can also be expressed in terms of Coulomb and Exchange operators, from which the Fock operator and the Hartree-Fock equation are derived by a variational procedure with boundary conditions[55, 56] that minimizes the energy expectation value with respect to variational parameters. The variational parameters are the molecular orbitals or in case of an LCAO ansatz the coefficients of the expansion in the atomic orbital basis.[57]

$$E \;=\; \sum_i^N \langle \phi_i | \hat{h}_i | \phi_i \rangle + \frac{1}{2} \sum_{ij}^N \left(\langle \phi_j | \hat{j}_i | \phi_j \rangle - \langle \phi_j | \hat{k}_i | \phi_j \rangle \right) + V_{nn} \tag{2.6}$$

$$\hat{j}_i | \phi_j(2) \rangle = \langle \phi_i(1) | \hat{r}_{12}^{-1} | \phi_i(1) \rangle | \phi_j(2) \rangle \tag{2.7}$$

$$\hat{k}_i | \phi_j(2) \rangle = \langle \phi_i(1) | \hat{r}_{12}^{-1} | \phi_j(1) \rangle | \phi_i(2) \rangle \tag{2.8}$$

$$\hat{f}_i = \hat{h}_i + \sum_j^N \left(\hat{j}_{ij} - \hat{k}_{ij} \right) \tag{2.9}$$

The Fock operator $\hat{f}_i$ is an effective one-particle operator and is associated with the variation of the energy but not with the energy itself and only depends on the occupied molecular orbitals (MO). The variational procedure finally leads to the Hartree-Fock equation, which represents a pseudo-eigenvalue equation (eq. 2.10) in the basis of canonical molecular orbitals with the orbital energies ϵ as eigenvalues.

$$\hat{f}_i | \phi_i \rangle = \epsilon_i | \phi_i \rangle \tag{2.10}$$

$$\epsilon_i = \langle \phi_i | \hat{f}_i | \phi_i \rangle \tag{2.11}$$

For the discussion of Møller-Plesset perturbation theory (see Section 2.2) it is helpful to

---

with the respective indices.

formulate the HF energy in terms of the orbital energies, which are given as the expectation values of the Fock operator $\hat{f}_i$ (eq. 2.11). From eq. 2.5

$$
\begin{aligned}
E &= \sum_i^N \left(h_i + \frac{1}{2} \sum_j^N (J_{ij} - K_{ij}) \right) \\
&= \sum_i^N \underbrace{\left(h_i + \sum_j^N (J_{ij} - K_{ij}) \right)}_{\epsilon_i} - \sum_i^N \left(\frac{1}{2} \sum_j^N (J_{ij} - K_{ij}) \right) \\
&= \sum_i^N \epsilon_i - \left(\frac{1}{2} \sum_i^N \sum_j^N (J_{ij} - K_{ij}) \right)
\end{aligned}
\tag{2.12}
$$

the total energy E is given as a sum over orbital energies of the occupied MO and the corrections to the electron-electron interaction which is counted twice because of the Fock operator.

The spin functions can be integrated out from the spin-orbitals and spatial functions are obtained, which allow an easier numerical treatment.[56] In the closed-shell case, eq. 2.5 as well as the Fock operator simplify to the spin-averaged quantities given in eq. 2.13 and 2.14.

$$
\hat{f}_i = \hat{h}_i + 2 \sum_j^{N/2} \left(\hat{j}_{ij} - \frac{1}{2} \hat{k}_{ij} \right)
\tag{2.13}
$$

$$
E = 2 \sum_i^N h_i + 2 \sum_i^N \sum_j^N \left(J_{ij} - \frac{1}{2} K_{ij} \right) + V_{nn}
\tag{2.14}
$$

The computational treatment directly based on MO is quite cumbersome and was simplified by the introduction of a basis set expansion of the MO into atomic orbitals (AO) (eq. 2.15). This concept finally leads to the Roothaan-Hall equation and can be found again in the formulation of the multi-reference perturbation theory (see Section 2.3). Because of the basis set expansion, the MO integrals are substituted by AO integrals and the MO coefficients are absorbed into the density matrix $D_{\gamma\delta}$. Eq. 2.16 and eq. 2.17 are eqivalent, but given in two different notations. The former one is in the Dirac (physicist's) notation (indexing of electrons is $\langle 1, 2 | 1, 2 \rangle$) and the latter given in Mulliken (chemist's) notation $((1, 1 | 2, 2))$.

$$
\phi_i = \sum_\alpha^N c_{\alpha i} \chi_i
\tag{2.15}
$$

$$
\begin{aligned}
\langle \chi_\alpha | \hat{f}_i | \chi_\beta \rangle &= \langle \chi_\alpha | \hat{h}_i | \chi_\beta \rangle + 2 \sum_j^{N/2} \langle \chi_\alpha | \hat{j}_{ij} - \frac{1}{2} \hat{k}_{ij} | \chi_\beta \rangle \\
&= \langle \chi_\alpha | \hat{h}_i | \chi_\beta \rangle + 2 \sum_j^{N/2} \left(\langle \chi_\alpha \phi_j | \chi_\beta \phi_j \rangle - \frac{1}{2} \langle \chi_\alpha \phi_j | \phi_j \chi_\beta \rangle \right) \\
&= \langle \chi_\alpha | \hat{h}_i | \chi_\beta \rangle + 2 \sum_j^{N/2} \sum_\gamma^{AO} \sum_\delta^{AO} c_{\gamma j} c_{\delta j} \left(\langle \chi_\alpha \chi_\gamma | \chi_\beta \chi_\delta \rangle - \frac{1}{2} \langle \chi_\alpha \chi_\gamma | \chi_\delta \chi_\beta \rangle \right) \\
&= \langle \chi_\alpha | \hat{h}_i | \chi_\beta \rangle + \sum_\gamma^{AO} \sum_\delta^{AO} D_{\gamma\delta} \left(\langle \chi_\alpha \chi_\gamma | \chi_\beta \chi_\delta \rangle - \frac{1}{2} \langle \chi_\alpha \chi_\gamma | \chi_\delta \chi_\beta \rangle \right) & (2.16) \\
&= (\chi_\alpha | \hat{h}_i | \chi_\beta) + \sum_\gamma^{AO} \sum_\delta^{AO} D_{\gamma\delta} \left((\chi_\alpha \chi_\beta | \chi_\gamma \chi_\delta) - \frac{1}{2} (\chi_\alpha \chi_\delta | \chi_\gamma \chi_\beta) \right) & (2.17) \\
D_{\gamma\delta} &= 2 \sum_j^{N/2} c_{\gamma j} c_{\delta j} & (2.18)
\end{aligned}
$$

2.2 Rayleigh-Schrödinger and Møller-Plesset Perturbation Theory

2.2.1 Rayleigh-Schrödinger Perturbation Theory

The Hartree-Fock ansatz accounts for about 99% of the exact electronic energy of an atom or a molecule. The remaining part though is crucial for the computation of reaction profiles as the energies involved in breaking and building bonds are lower than this fraction that is associated with electron correlation. When comparing relative energies, the errors in HF partially cancel out, but for many cases methods that encounter electron correlation are necessary. Starting from the HF ansatz with a single determinant wavefunction, the consensus in all post Hartree-Fock methods is to augment the wavefunction with additional Slater determinants[3]. These determinants are generated by applying a certain excitation scheme to the HF determinant. With this extended wavefunction one is now able to recover electron correlation. The degree of the accounted electron correlation and the method how the multi-determinant problem is solved depends on the excitation scheme and distinguishes

---

[3]Spin correctness has to be ensured, so if a spin-free formulation of the method is not possible, the spin-coupling schemes have to be evaluated explicitly to obtain the matrix elements. This can be done on the level of determinants but also on the level of Configuration State Functions (CSF), which possess the correct spin properties.[55, 56]

the post HF methods from each other.

The formulation of many-body perturbation theory starts from the Fock operator and not the Hartree-Fock energy, but the treatment of the correlation energy as a perturbation is rationalized by the previously given assumption, that the reference energy (*e.g.* HF accounts for 99% of the exact energy) is a close approximation to the exact energy and the remaining fraction is small compared to the reference energy. In this approximation, the Hamilton operator $\hat{H}$ consequently can be split up into two parts. The first part represents the unperturbed Hamitonian $\hat{H}_0$ while the second part is the perturbation operator $\hat{V}$ (eq. 2.19). The degree of the perturbation is encountered by the perturbation parameter λ, with $\lambda = 0$ corresponding to the unperturbed system.

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{2.19}$$

$$\hat{H} = \hat{H}_0 + \lambda\hat{V} \tag{2.20}$$

$$\hat{H}_0|\Phi_i\rangle = E_i|\Phi_i\rangle \tag{2.21}$$

The solution of the Schrödinger equation in this approximation is facilitated by the requirement that the unperturbed system is represented by an eigenvalue equation (eq. 2.21).

$$\left(\hat{H}_0 + \hat{V}\right)|\psi\rangle = E|\psi\rangle \tag{2.22}$$

As the degree of perturbation λ can be increased continuously from 0 to a specific value, the energy E and the wavefunction $|\psi\rangle$ of the system also must change accordingly, and therefore can be written in a Taylor expansion as in eqn. 2.23 and 2.24 with λ having the order of the perturbation in the exponent.

$$E = \lambda^0 E_0 + \lambda^1 E_1 + \lambda^2 E_2 + \ldots \tag{2.23}$$

$$\psi = \lambda^0 \psi_0 + \lambda^1 \psi_1 + \lambda^2 \psi_2 + \ldots \tag{2.24}$$

After the Taylor expansion, the wavefunction and also its norm depend on the power of the parameter λ. It is normalized by *intermediate normalization* that defines the overlap of the unperturbed (reference) wavefunction $|\Phi_0\rangle$ with the total wavefunction as 1. The overlap of the unperturbed wavefunction and the reference wavefunction is 1 and thus the overlap of the corrections to $|\psi\rangle$ with $|\Phi_0\rangle$ is zero. The intermediate normalization facilites the evaluation of the energy and wavefunctions expressions.

$$\langle \psi | \Phi_0 \rangle \;=\; 1 \tag{2.25}$$

$$\langle \lambda^0 \psi_0 + \lambda^1 \psi_1 + \lambda^2 \psi_2 + \ldots | \Phi_0 \rangle \;=\; 1 \tag{2.26}$$

$$\lambda^0 \langle \psi_0 | \Phi_0 \rangle + \lambda^1 \langle \psi_1 | \Phi_0 \rangle + \lambda^2 \langle \psi_2 | \Phi_0 \rangle + \ldots \;=\; 1 \tag{2.27}$$

$$\langle \psi_{i \neq 0} | \Phi_0 \rangle \;=\; 0 \tag{2.28}$$

With the expansion of the Taylor series in eq. 2.23 and 2.24, the Schrödinger equation 2.22 equals the expression in eq. 2.29. Those terms including all perturbations up to a given order are grouped together in eq. 2.30 to 2.33.

$$\left(\hat{H}_0 + V \right) \left(\lambda^0 |\psi_0\rangle + \lambda^1 |\psi_1\rangle + \lambda^2 |\psi_2\rangle + \ldots \right) \;=\;$$
$$\left(\lambda^0 E_0 + \lambda^1 E_1 + \lambda^2 E_2 + \ldots \right) \left(\lambda^0 |\psi_0\rangle + \lambda^1 |\psi_1\rangle + \lambda^2 |\psi_2\rangle + \ldots \right) \tag{2.29}$$

$$\lambda^0 \;:\; \hat{H}_0 |\psi_0\rangle = E_0 |\psi_0\rangle \tag{2.30}$$

$$\lambda^1 \;:\; \hat{H}_0 |\psi_1\rangle + \hat{V} |\psi_0\rangle = E_0 |\psi_1\rangle + E_1 |\psi_0\rangle \tag{2.31}$$

$$\lambda^2 \;:\; \hat{H}_0 |\psi_2\rangle + \hat{V} |\psi_1\rangle = E_0 |\psi_2\rangle + E_1 |\psi_1\rangle + E_2 |\psi_0\rangle \tag{2.32}$$

$$\lambda^n \;:\; \hat{H}_0 |\psi_n\rangle + \hat{V} |\psi_{n-1}\rangle = \sum_{i=0}^{n} E_i |\psi_{n-i}\rangle \tag{2.33}$$

The solution for the nth order correction to the energy is obtained by projecting eq. 2.33 onto $\langle \Phi_0 |$ and using the intermediate normalization and the hermicity of $\hat{H}_0$ to simplify the expression for the evaluation of the energy correction.

$$\langle \Phi_0 | \hat{H}_0 | \psi_n \rangle + \langle \Phi_0 | \hat{V} | \psi_{n-i} \rangle \;=\; E_n \langle \Phi_0 | \psi_0 \rangle + \sum_{i=0}^{n-1} E_i \langle \Phi_0 | \psi_{n-1} \rangle \tag{2.34}$$

$$E_0\langle\Phi_0|\psi_n\rangle + \langle\Phi_0|\hat{V}|\psi_{n-1}\rangle \;=\; E_n\langle\Phi_0|\psi_n\rangle \tag{2.35}$$

$$E_n \;=\; \langle\Phi_0|\hat{V}|\psi_{n-1}\rangle \tag{2.36}$$

The operators $\hat{H}_0$, $\hat{V}$ and the wavefunction ψ are unspecified so far. Employing the hermicity of $\hat{H}_0$ and the property of the solutions of the unperturbed Schrödinger equation (eq. 2.21) to form a complete set (eq. 2.37), the first-order correction to the wavefunction can be expanded into this complete set. The set of equations derived from this procedure is known as *Rayleigh-Schrödinger* perturbation theory.

$$\psi_1 = \sum_i c_i \Phi_i \tag{2.37}$$

The equation for the first-order correction (eq. 2.31) is expanded to eq. 2.39.

$$\left(\hat{H}_0 - E_0\right)|\psi_1\rangle + \left(\hat{V} - E_1\right)|\Phi_0\rangle \;=\; 0 \tag{2.38}$$

$$\left(\hat{H}_0 - E_0\right)\left(\sum_i c_i|\Phi_i\rangle\right) + \left(\hat{V} - E_1\right)|\Phi_0\rangle \;=\; 0 \tag{2.39}$$

The projection of eq. 2.39 onto the reference wavefunction $\langle\Phi_0|$ yields the first order correction to the energy (eq. 2.44), which is an average of the perturbation operator $\hat{V}$ acting on the unperturbed wavefunction.

$$\sum_i c_i \left(\langle\Phi_0|\hat{H}_0 - E_0|\Phi_i\rangle\right) + \left(\langle\Phi_0|\hat{V} - E_1|\Phi_0\rangle\right) \;=\; 0 \tag{2.40}$$

$$\sum_i c_i\langle\Phi_0|\hat{H}_0|\Phi_i\rangle - E_0\sum_i c_i\langle\Phi_0|\Phi_i\rangle + \langle\Phi_0|\hat{V}|\Phi_0\rangle - E_1\langle\Phi_0|\Phi_0\rangle \;=\; 0 \tag{2.41}$$

$$\sum_i c_i E_i\langle\Phi_0|\Phi_i\rangle - c_0 E_0 + \langle\Phi_0|\hat{V}|\Phi_0\rangle - E_1 \;=\; 0 \tag{2.42}$$

$$c_0 E_0 - c_0 E_0 + \langle\Phi_0|\hat{V}|\Phi_0\rangle - E_1 \;=\; 0 \tag{2.43}$$

$$E_1 \;=\; \langle\Phi_0|\hat{V}|\Phi_0\rangle \tag{2.44}$$

The term $\langle\Phi_0|\hat{H}_0|\Phi_i\rangle$ on the left hand side of eq. 2.41 utilizes the eigenfunction property of $\hat{H}_0$ acting on $|\Phi_0\rangle$ and this equation simplifies to eq. 2.42. The projection onto the excited states with $\Phi_{j\neq0}$ (eq. 2.47) is used to compute the first-order correction $|\psi_1\rangle$ to the total wavefunction. Projecting out the reference wavefunction $|\Phi_0\rangle$ ensures the two spaces

(reference space and excited space) to be orthogonal to each other.

$$\sum_i c_i \left(\langle \Phi_j | \hat{H}_0 - E_0 | \Phi_i \rangle \right) + \left(\langle \Phi_j | \hat{V} - E_1 | \Phi_0 \rangle \right) = 0 \tag{2.45}$$

$$\sum_i c_i \left(\langle \Phi_j | \hat{H}_0 - E_0 | \Phi_i \rangle \right) = - \left(\langle \Phi_j | \hat{V} - E_1 | \Phi_0 \rangle \right) \tag{2.46}$$

$$\sum_i c_i \left(\langle \Phi_j | \hat{H}_0 - E_0 | \Phi_i \rangle \right) = - \left(\langle \Phi_j | \hat{V} | \Phi_0 \rangle \right) \tag{2.47}$$

A similar procedure can be applied to derive the equation for the second-order correction to the energy (eq. 2.48) that is given as

$$E_2 = \sum_i c_i \langle \Phi_0 | \hat{V} | \Phi_i \rangle \tag{2.48}$$

2.2.2 Møller-Plesset Perturbation Theory

The equations for the wavefunction and energy corrections have been provided by the Rayleigh-Schrödinger perturbation theory, but a suitable choice of $\hat{H}_0$ was not specified so far. In the Møller-Plesset (MP) ansatz[58,59] that uses the single determinant HF wavefunction as the reference wavefunction $|\Phi_0\rangle$, $\hat{H}_0$ is chosen to be the sum of Fock operators over the occupied molecular orbitals. This partitioning counts the average electron-electron interaction twice (see eq. 2.12) which has to be compensated by an appropriate choice of the perturbation operator $\hat{V}$ (eq. 2.51).

$$\hat{H}_0 = \sum_{i=1}^{N} \hat{f}_i \tag{2.49}$$

$$= \sum_{i=1}^{N} \left(\hat{h}_i + \sum_{j=1}^{N} \left(\hat{j}_{ij} - \hat{j}_{ij} \right) \right)$$

$$= \sum_{i=1}^{N} \hat{h}_i + 2 \langle V_{ee} \rangle \tag{2.50}$$

This choice of $\hat{H}_0$ gives E_{MP0} as the sum over orbital energies and by using eq. 2.50 the energy up to the first-order correction E_{MP1} is the HF energy. The final MP2 energy is the sum of the HF energy and the second-order correction (eq. 2.48).

$$\hat{V} = \hat{V}_{ee} - 2\langle \hat{V}_{ee} \rangle \tag{2.51}$$

$$E_1 = \langle \Phi_0 | \hat{V} | \Phi_0 \rangle$$

$$= \langle \hat{V}_{ee} \rangle - 2\langle \hat{V}_{ee} \rangle$$

$$= -\langle \hat{V}_{ee} \rangle \tag{2.52}$$

$$E_{\text{MP0}} = E_0$$

$$= \langle \Phi_0 | \hat{H}_0 | \Phi_0 \rangle$$

$$= \sum_i^N \langle \Phi_0 | \hat{f}_i | \Phi_0 \rangle$$

$$= \sum_i^N \epsilon_i \tag{2.53}$$

$$E_{\text{MP1}} = E_{\text{MP0}} + E_1$$

$$= \sum_{i=1}^N \left(\epsilon_i - \frac{1}{2} \sum_{j=1}^N \left(\hat{j}_{ij} - \hat{k}_{ij} \right) \right)$$

$$= E_{\text{HF}} \tag{2.54}$$

$$E_{\text{MP2}} = E_0 + E_1 + E_2$$

$$= E_{\text{HF}} + E_2$$

$$= E_{\text{HF}} + \sum_i c_i \langle \Phi_0 | \hat{V} | \Phi_i \rangle \tag{2.55}$$

The evaluation of the second-order correction can be simplified by employing the diagonal structure of the matrix associated with the Fock operator in the basis of canonical MO. The coefficients for the first-order wavefunction in eq. 2.47 are determined from eq. 2.56. The left term in eq. 2.57 yields only non-vanishing matrix elements for identical configurations i and j and reduces the matrix element to E_j. The coefficient c_j is evaluated by a matrix element of the perturbation operator and an energy difference between the expectation values of $\hat{H}_0$ of the reference and the excited space wavefunctions.

$$\sum_i c_i \left(\langle \Phi_j | \hat{H}_0 - E_0 | \Phi_i \rangle \right) = - \left(\langle \Phi_j | \hat{V} | \Phi_0 \rangle \right) \tag{2.56}$$

$$\underbrace{\sum_i c_i \langle \Phi_j | \hat{H}_0 | \Phi_i \rangle}_{E_j \delta_{ij}} - E_0 \underbrace{\sum_i c_i \langle \Phi_j | \Phi_i \rangle}_{=c_j} = -\langle \Phi_j | \hat{V} | \Phi_0 \rangle \tag{2.57}$$

$$c_j E_j - c_j E_0 = -\langle \Phi_j | \hat{V} | \Phi_0 \rangle \tag{2.58}$$

$$c_j = \frac{\langle \Phi_j | \hat{V} | \Phi_0 \rangle}{E_0 - E_j} \qquad (2.59)$$

The second-order energy correction to the energy (eq. 2.60) is obtained by inserting the coefficients c_j of the first-order correction (eq. 2.59) into the expression for E_2 (eq. 2.48).

$$E_2 = \sum_{i \neq 0} \frac{\langle \Phi_0 | \hat{V} | \Phi_i \rangle \langle \Phi_i | \hat{V} | \Phi_0 \rangle}{E_0 - E_i} \qquad (2.60)$$

Eq. 2.60 runs over the excited determinants generated from the HF wavefunction. For MP2, the reference wavefunction and the excited determinants interact directly via the $\hat{r}_{12}^{-1}$ operator. The maximum excitation level therefore is limited to single and double excitations. As shown in eqn. 2.61-2.63 single excitations do not contribute due to zero overlap and the Brioullin theorem.[55,56] For the final energy expression only the interactions between the HF determinant and the doubly excited determinants have to be considered to compute the MP2 energy.

$$\langle \Phi_0 | \hat{V} | \Phi_i^a \rangle = \langle \Phi_0 | \hat{H} - \sum_{j=1}^{N} \hat{f}_j | \Phi_i^a \rangle \qquad (2.61)$$

$$= \langle \Phi_0 | \hat{H} | \Phi_i^a \rangle - \sum_{j=1}^{N} \langle \Phi_0 | \hat{f}_j | \Phi_i^a \rangle \qquad (2.62)$$

$$= \underbrace{\langle \Phi_0 | \hat{H} | \Phi_i^a \rangle}_{=0, Brioullin} - \underbrace{\epsilon_a \langle \Phi_0 | \Phi_i^a \rangle}_{=0} \qquad (2.63)$$

The energy denominator in eq. 2.60 can be expressed in terms of expectation values of the Fock operator and reduces to differences of MO energies. Note that the matrix elements notation changes from physicists notation to chemists notation from eq. 2.65 to 2.66. The latter notation is preferred here to simplify the comparison with multi-reference MP2.

$$E_2 = \sum_{i<j}^{occ.MO} \sum_{a<b}^{virt.MO} \frac{\langle \Phi_0 | \hat{V} | \Phi_{ij}^{ab} \rangle \langle \Phi_0 | \hat{V} | \Phi_{ij}^{ab} \rangle}{E_0 - E_{ij}^{ab}} \qquad (2.64)$$

$$= \sum_{i<j}^{occ.MO} \sum_{a<b}^{virt.MO} \frac{|\langle \phi_i \phi_j | \phi_a \phi_b \rangle - \langle \phi_i \phi_j | \phi_b \phi_a \rangle|^2}{E_0 - E_{ij}^{ab}} \qquad (2.65)$$

$$= \sum_{i<j}^{occ.MO} \sum_{a<b}^{virt.MO} \frac{|(\phi_i \phi_a | \phi_j \phi_b) - (\phi_i \phi_b | \phi_j \phi_a)|^2}{E_0 - E_{ij}^{ab}} \qquad (2.66)$$

$$
= \sum_{i<j}^{occ.MO} \sum_{a<b}^{virt.MO} \frac{|(\phi_i\phi_a|\phi_j\phi_b) - (\phi_i\phi_b|\phi_j\phi_a)|^2}{\sum_{k\neq a,b}^{occ.MO} \langle\phi_k|\hat{f}_k|\phi_k\rangle - \sum_{k'\neq i,j}^{occ.MO} \langle\phi_{k'}|\hat{f}_{k'}|\phi_{k'}\rangle} \tag{2.67}
$$

$$
= \sum_{i<j}^{occ.MO} \sum_{a<b}^{virt.MO} \frac{|(\phi_i\phi_a|\phi_j\phi_b) - (\phi_i\phi_b|\phi_j\phi_a)|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \tag{2.68}
$$

The expression with MO energies simplifies the computation of the energy but is disadvantageous, if some MO become nearly degenerate. Consequently the denominator gets close to zero and the energy contribution from the few nearly degenerate MO dominates the sum over all excited determinants and spoils the MP2 energy leading to unphysical results. For multi-reference cases where near-degeneracies are quite ubiquitous, the MP theory has to be generalized to remedy this severe problem.

2.3 Generalization to Multi-Reference Møller-Plesset Perturbation Theory (MR-MP)

The generalization of single-reference Møller-Plesset perturbation theory to multi-reference wavefunctions employs the same partitioning scheme of the Hamiltonian (eq. 2.19) as for the single-reference case and also assumes the reference wavefunction to be an eigenfunction of the unperturbed system (eq. 2.21). The concept to generalize the single-reference to the multi-reference case is illustrated for the example of the Pulay and Wolinski[60,61] perturbation-theory and its closely related variant from Murphy and Messmer.[62] The implementation used in this work is based on the latter formulation.

Since the choice of $\hat{H}_0$ is crucial for the convergence of the MP calculation, $\hat{H}_0$ should be chosen to be close to $\hat{H}$ but also to be close to diagonal form to facilitate an efficient computation of the correction terms. In section 2.2 the $\hat{H}_0$ operator for the single-reference case was defined as the sum over one-electron Fock operators $\hat{f}$ (cf. eq. 2.9) and the Fock matrix is diagonal in the basis of canonical MO. The Fock operator was partitioned into the one-electron operator $\hat{h}$ and the Coulomb and Exchange operators $\hat{j}$ and $\hat{k}$, respectively with the definitions given in eq. 2.7 and 2.8.

$$
\hat{H}_0 = \hat{F} = \sum_i \hat{f}_i \tag{2.69}
$$

$$\hat{f} = \hat{h} + \hat{j} - \hat{k} \tag{2.70}$$

For a given general multi-reference wavefunction (eq. 2.71) the expectation value of the Fock operator $\hat{f}$ acting on the multi-reference wavefunction $|\Phi_0\rangle$ is given in eq. 2.73. The evaluation of the expectation value is determined by the sums over K and L of the Configuration State Functions (CSF) in the reference wavefunction. In contrast to the single-reference approach, different $|\Phi_K\rangle$ and $|\Phi_L\rangle$ do interact through the effective one-electron Fock operator and the matrix elements do not vanish for the off-diagonal part of the Fock matrix.

$$|\Phi_0\rangle = \sum_K c_K |\Phi_K\rangle \tag{2.71}$$

$$
\begin{aligned}
\langle\Phi_0|\hat{f}|\Phi_0\rangle &= \langle\sum_K c_K\Phi_K|\hat{f}|\sum_L c_L\Phi_L\rangle \tag{2.72}\\
&= \sum_K\sum_L c_K c_L \langle\Phi_K|\hat{h}+\hat{j}+\hat{k}|\Phi_L\rangle \\
&= \sum_K\sum_L c_K c_L \langle\Phi_K|\hat{h}|\Phi_L\rangle + \\
&\quad\ \sum_K\sum_L c_K c_L \langle\Phi_K|\hat{j}|\Phi_L\rangle - \\
&\quad\ \sum_K\sum_L c_K c_L \langle\Phi_K|\hat{k}|\Phi_L\rangle \tag{2.73}
\end{aligned}
$$

As a consequence the expectation value for a multi-reference wavefunction is not only computed from the diagonal element, but also has contributions from off-diagonal elements. The diagonal structure of the Fock matrix in the single reference case has been completely lost and the full matrix has to be evaluated. Therefore the simple choice for $\hat{H}_0$ as the sum of Fock operators violates the condition in eq. 2.21 and is subject of the modifications described later.

In analogy to eq. 2.14, the reduction from the expression on the configurational level to the orbital level yields two generalized Coulomb and Exchange operators, which comprise the projection of the reference coefficients c_K and c_L to the MOs (k and l) occupied in the reference configurations $|\Phi_K\rangle$ and $|\Phi_L\rangle$.

$$\hat{j}_{MR} = \sum_K \sum_L c_K c_L \hat{j}_{k \in K, l \in L} \tag{2.74}$$

$$\hat{k}_{MR} = \sum_K \sum_L c_K c_L \hat{k}_{k \in K, l \in L} \tag{2.75}$$

Without proof the evaluation of Fock matrix elements for the multi-reference case from eq. 2.72 is equivalent to the formulation with generalized Coulomb and Exchange operators and projects the coefficients of the reference wavefunction into a density matrix D_{kl} of the MO that are occupied in the reference wavefunction.

$$f_{ab} = h_{ab} + \sum_{kl} D_{kl} \left[(ab|kl) - \frac{1}{2} (al|kb) \right] \tag{2.76}$$

D_{kl} contains the compressed information of the reference wavefunction and is equivalent to the HF density matrix, if a HF wavefunction is applied (see eq. 2.17). The Fock matrix elements for the multi-reference case depend on the reference density matrix and the Coulomb and Exchange integrals with the interacting MO a and b and the MO i and j in the reference space. The evaluation of the Fock matrix elements for the multi-reference case requires a sum over integrals multiplied with the corresponding elements of the density matrix in contrast to a single determinant MP possessing a single value. Interestingly, the integrals needed by single-reference MP2 are solely of the type $(ak|bl)$ with a and b denoting virtual MO and k and l MO occupied in the HF wavefunction, while for the generalized Fock matrix the type of the Coulomb $(ab|kl)$ and the Exchange integrals $(al|kb)$ differ. The latter type corresponds to single-reference MP2, but the existence of the former necessitates a somewhat different integral transformation.

In analogy to single-reference MP theory the Fock operator for the many-electron system can be formulated as the sum over the single-electron Fock operators.

$$\hat{F} = \sum_i \hat{f}_i \tag{2.77}$$

The eigenvalue equation condition and the required orthogonality of the excited to the reference space cannot be fulfilled, if $\hat{H}_0$ is simply set to $\hat{F}$. The introduction of projection operators to the simple choice $\hat{H}_0 = \hat{F}$ ensures the orthogonality of the interacting

spaces. The projection operators chosen by Pulay and Wolinski[60, 61] (eq. 2.78) ensure the orthogonality of the reference space and the first-order interacting space. The reference wavefunction for this MR-MP ansatz was a MCSCF wavefunction from which the excited space was generated by application of one- and two-electron operators on the whole MC-SCF reference wavefunction. The partitioning of the interacting spaces is comparable to the internal contraction scheme applied by Werner and Knowles.[5] This choice of the wavefunction lets the single excitation space vanish because of the generalized Brioullin theorem but requires the explicit orthogonalization of the interacting space including double excitations out of the reference space.

$$\hat{H}_0 = \hat{P}_0\hat{F}\hat{P}_0 + \hat{P}_S\hat{F}\hat{P}_S + \hat{P}_D\hat{F}\hat{P}_D + \dots \tag{2.78}$$

$$\hat{P}_0 = |\Phi_0\rangle\langle\Phi_0| \tag{2.79}$$

$$\hat{P}_S = |\sum_K \Phi_K\rangle\langle\sum_L \Phi_L| \tag{2.80}$$

$$\hat{P}_D = |\sum_O \Phi_O\rangle\langle\sum_P \Phi_P| \tag{2.81}$$

Murphy and Messmer[62] used a similar partitioning as Pulay and Wolinski but generated the excited space by single and double excitations out of the configurations of the reference wavefunction but not the whole wavefunction itself. The introduction of modified projection operators for the reference and the single/double excitation space again ensure the orthogonality condition.

$$\hat{H}_0 = \hat{P}_0\hat{F}\hat{P}_0 + \hat{P}_{SD}\hat{F}\hat{P}_{SD} + \dots \tag{2.82}$$

$$\hat{P}_0 = |\Phi_0\rangle\langle\Phi_0| \tag{2.83}$$

$$\hat{P}_{SD} = |\sum_K \Phi_K\rangle\langle\sum_L \Phi_L| \tag{2.84}$$

$$\tag{2.85}$$

The action of the projection operators on the reference wavefunction $|\Phi_0\rangle$ and an excited

space configuration $|\Phi_i\rangle$ is illustrated in the following equations, that give an impression on how the orthogonality of the reference and the first-order interaction space is preserved.

$$
\begin{aligned}
\langle\Phi_0|\hat{H}_0|\Phi_0\rangle &= \langle\Phi_0|\hat{P}_0\hat{F}\hat{P}_0|\Phi_0\rangle + \langle\Phi_0|\hat{P}_{SD}\hat{F}\hat{P}_{SD}|\Phi_0\rangle + \ldots & (2.86)\\
&= \langle\Phi_0|\Phi_0\rangle\langle\Phi_0|\hat{F}|\Phi_0\rangle\langle\Phi_0|\Phi_0\rangle + \sum_K\sum_L\langle\Phi_0|\Phi_K\rangle\langle\Phi_K|\hat{F}|\Phi_L\rangle\langle\Phi_L|\Phi_0\rangle + \ldots & (2.87)\\
&= \langle\Phi_0|\hat{F}|\Phi_0\rangle & (2.88)\\
\langle\Phi_i|\hat{H}_0|\Phi_j\rangle &= \langle\Phi_i|\hat{P}_0\hat{F}\hat{P}_0|\Phi_j\rangle + \langle\Phi_i|\hat{P}_{SD}\hat{F}\hat{P}_{SD}|\Phi_j\rangle + \ldots & (2.89)\\
&= \langle\Phi_i|\Phi_0\rangle\langle\Phi_0|\hat{F}|\Phi_0\rangle\langle\Phi_0|\Phi_j\rangle + \sum_K\sum_L\langle\Phi_i|\Phi_K\rangle\langle\Phi_K|\hat{F}|\Phi_L\rangle\langle\Phi_L|\Phi_j\rangle + \ldots & (2.90)\\
&= \langle\Phi_i|\hat{F}|\Phi_j\rangle & (2.91)
\end{aligned}
$$

This partitioning scheme ensures the generalized perturbation theory to be equivalent to the single-reference case once a single-reference wavefunction is applied. The evaluated energy expressions therefore are identical (eq. 2.95). The only difference is the computation of the first-order correction $|\psi_1\rangle$ to the wavefunction that is obtained from the solution of eq. 2.96.

$$
\begin{aligned}
E_0 &= \langle\Phi_0|\hat{H}_0|\Phi_0\rangle & (2.92)\\
E_1 &= \langle\Phi_0|\hat{V}|\Phi_0\rangle & (2.93)\\
E_2 &= E_1 + \langle\Phi_0|\hat{V}|\psi^1\rangle & (2.94)\\
&= E_1 + \sum_j c_j^1\langle\Phi_j^1|\hat{V}|\Phi_0\rangle & (2.95)
\end{aligned}
$$

With the use of projection operators $|\Phi_0\rangle$ can be considered as an eigenfunction of $\hat{H}_0$, but for the CSF generated in the excited space $|\Phi_{j\neq0}\rangle$ and eq. 2.47 has to be solved explicitely for c_j^1 (see eq. 2.96) and cannot be evaluated directly as for the single-reference case. Solving of the large but sparse linear equation system in eq. 2.96 cannot be done with direct methods anymore but is done with an iterative scheme. The problem can be formulated as the projection of the matrix $\mathbf{A}$ onto the solution vector $\vec{c}$ that is solved if the condition $-\vec{I}$ is fulfilled. A more detailed introduction into the topic of iterative solvers and how the equation $\mathbf{A}\cdot\vec{c} = \vec{I}$ is solved can be found in Chapter 3.

$$\underbrace{\sum_j c_j^1 \langle \Phi_i^1 | \hat{H}_0 - \hat{E}_0 | \Phi_j^1 \rangle}_{\mathbf{A} \cdot \vec{c}} = - \underbrace{\langle \Phi_i^1 | \hat{V} | \Phi_0 \rangle}_{=Inhomogenity, \vec{I}} \tag{2.96}$$

The perturbation operator $\hat{V}$ generally is evaluated by projecting $\hat{V} | \Phi_0 \rangle$ onto $\langle \Phi_j^1 |$ (eq. 2.100) with $\hat{V}$ being defined as $\hat{H} - \hat{H}_0$.

$$
\begin{aligned}
\langle \Phi_j^1 | \hat{V} | \Phi_0 \rangle &= \langle \Phi_j^1 | \hat{H} - \hat{H}_0 | \Phi_0 \rangle & (2.97) \\
&= \langle \Phi_j^1 | \hat{H} | \Phi_0 \rangle - \langle \Phi_j^1 | \hat{H}_0 | \Phi_0 \rangle & (2.98) \\
&= \langle \Phi_j^1 | \hat{H} | \Phi_0 \rangle \\
&\quad - \langle \Phi_j^1 | \Phi_0 \rangle \langle \Phi_0 | \hat{F} | \Phi_0 \rangle \langle \Phi_0 | \Phi_0 \rangle & (2.99) \\
&\quad + \sum_K \sum_L \langle \Phi_j^1 | \Phi_K \rangle \langle \Phi_K | \hat{F} | \Phi_L \rangle \langle \Phi_L | \Phi_0 \rangle \\
&= \langle \Phi_j^1 | \hat{H} | \Phi_0 \rangle & (2.100)
\end{aligned}
$$

Employing the orthogonality of the reference and the first-order interacting space introduced by the projection operators, the perturbation operator is the Hamilton operator for the full system and the expressions for the perturbation energies are given in eq. 2.101 to 2.104.

$$
\begin{aligned}
E_0 &= \langle \Phi_0 | \hat{H}_0 | \Phi_0 \rangle & (2.101) \\
E_1 &= \langle \Phi_0 | \hat{V} | \Phi_0 \rangle & (2.102) \\
&= \langle \Phi_0 | \hat{H} | \Phi_0 \rangle & (2.103) \\
E_2 &= E_1 + \sum_j c_j^1 \langle \Phi_j^1 | \hat{V} | \Phi_0 \rangle & (2.104) \\
&= E_1 + \sum_j c_j^1 \langle \Phi_j^1 | \hat{H} | \Phi_0 \rangle & (2.105)
\end{aligned}
$$

The second-order correction to the energy is evaluated by multiplying the coefficient vector of the iteratively solved first-order correction to the wavefunction with the inhomogenity $\vec{I}$ (see eq. 2.96). In contrast to single-reference MP2 the coefficient vector is not computed directly but is obtained after explicit solution of a sparse linear equation system.

2.4   Overview on MR-PT Approaches

To put the two multi-reference perturbation theory (MR-PT) approaches that have been discussed in the last section into a wider context, a brief conceptional outline on various MR-PT approaches is given. All of these approaches were formulated with the intension to find the best possible partitioning scheme for MR-PT as the choice of the unperturbed Hamilton operator in MR-PT is not unique and a balance between accuracy and efficiency had to be found.[63]

2.4.1   MR-PT Methods Based on MCSCF/CASSCF Wavefunctions

The major part of MR-PT approaches published in literature is based on MCSCF or CASSCF wavefunctions and belongs to the class "diagonalize-then-perturb". The use of CASSCF wavefunctions is believed to be more advantegous than MCSCF wavefunctions as they show a better convergence behavior but have the disadvantage to create large perturbation spaces.[64] The most popular ansatz is the CASPT2 method proposed by Kerstin Andersson[35,36] and briefly described in the following section.

The CASPT2 ansatz describes a method for a single-reference state given by a multi-configuration CASSCF wavefunction that is assumed to account for all near-degeneracies. Starting from the CASSCF wavefunction a first-order interacting space is generated that is block diagonal with respect to the zeroth-order Hamiltonian and non-orthogonal.

The choice of $\hat{H}_0$ is made to assure the orthogonality relation of the reference space to the excited spaces by projection operators (eq. 2.85). This technique was also used for the approaches of Pulay-Wolinski and Murphy-Messmer.

$$\hat{H}_0 = \hat{P}_0\hat{F}\hat{P}_0 + \hat{P}_K\hat{F}\hat{P}_K + \hat{P}_{SD}\hat{F}\hat{P}_{SD} + \hat{P}_{TQ...}\hat{F}\hat{P}_{TQ...} \tag{2.106}$$

The zeroth-order Hamiltonian is partioned into the occupied reference space of the CASSCF wavefunction denoted as 0, the space K is the orthogonal complement to the occupied CASSCF space that generates the CASSCF wavefunction, the interacting space of the single and double excitations to the reference wavefunction, and the space with the higher

excitations which are neglected. The single and double excited space is divided into eight groups of two internal, three seminternal and three external excitation spaces, which blocks the Hamilton matrix into 8×8 blocks.

The resulting matrix of the singles and doubles space

$$\langle i|\hat{H}_0 - \hat{E}_0|j\rangle \tag{2.107}$$

is further simplified by the definition that the Fock matrix only includes diagonal terms of a one-particle operator while the off-diagonal terms are treated in the perturbation. The interactions between $\langle i|$ and $|j\rangle$ thus vanish if they belong to different blocks and a block-diagonal structure of the Hamilton matrix is obtained. The resulting blocks then can be easily diagonalized with standard methods. In CASPT2 the choice of the Fock operator can be modified defining different model Hamiltonians.

The appropriate choice of $\hat{F}$ is given by

$$\hat{F}_{pq\,\sigma} = \hat{a}_{p\,\sigma}\left[\hat{H}, \hat{a}_{q\,\sigma}\right] - \hat{a}_{p\,\sigma}\left[\hat{H}, \hat{a}_{q\,\sigma}\right] \tag{2.108}$$

with a spin-averaged expectation value denoted by

$$f_{pq} = \frac{1}{2}\sum_{\sigma}\langle 0|\hat{F}_{pq\,\sigma}|0\rangle \tag{2.109}$$

The evaluation of this term yields the expression for the matrix elements

$$f_{pq} = h_{pq} + \sum_{rs} D_{rs}\left[(pq|rs) - \frac{1}{2}(ps|rq)\right] \tag{2.110}$$

The standard CASPT2 procedure can also be extended to work with state-averaged CASSCF wavefunctions called multi-state CASPT2.[37]

A similar approach that also employs diagonal Fock matrices was proposed by Hirao,[65] while also methods with more general Hamiltonians for MCSCF wavefunctions and such employing the properties of CAS-CI wavefunctions were investigated.[66–68]

A method also employing the Møller-Plesset partitioning scheme was developed by Pulay and Wolinski[60,61] already described in the previous section. They employed projection operators on a general MCSCF wavefunction generating the perturbation space by excitations acting on the internally contracted wavefunction. Applications of larger molecules

that employ this ansatz can be found in a recent work of van Lenthe *et al.*[69] The variant of Murphy and Messmer[62] replaced the internally contracted MCSCF reference wavefunction by a Generalized Valence Bond (GVB) wavefunction. The interacting space was generated by excitations from each of the reference configurations and does not need the orthogonalization steps as it is necessary in the method of Pulay and Wolinski. This approach is also applicable for any general MCSCF wavefunction without substantial loss in accuracy.

Partitioning schemes different to the previous ones were also employed. Nakano[70, 71] defines a partitioning of the Hamilton matrix into two orthogonal spaces. The zeroth-order wavefunctions are state-averaged CASSCF wavefunctions for target states and define the P space. The complementary eigenfunctions of the CAS-CI Hamiltonian and the CSF generated from excitations to virtual orbitals span the R space. The Hamilton matrix is diagonal in the R space and $\mathbf{H}$ can be block-diagonalized into the form

$$\begin{pmatrix} \mathbf{H}_{PP} & \mathbf{H}_{PR} \\ \mathbf{H}_{PR} & \mathbf{H}_{RR} \end{pmatrix} \tag{2.111}$$

Davidson[64] defined a similar partitioning

$$\mathbf{Hc} = \begin{pmatrix} \mathbf{H}_{PP} & \mathbf{H}_{PR} \\ \mathbf{H}_{PR} & \mathbf{H}_{RR} \end{pmatrix} \begin{pmatrix} \mathbf{c}_P \\ \mathbf{c}_R \end{pmatrix} = E \begin{pmatrix} \mathbf{c}_P \\ \mathbf{c}_R \end{pmatrix} \tag{2.112}$$

in which $\mathbf{H}_{PP}$ is further partitioned into $\mathbf{H}_{P'P'}$, $\mathbf{H}_{P'P''}$, $\mathbf{H}_{P'P''}$, and $\mathbf{H}_{P''P''}$. $\mathbf{H}_{P'P'}$ contains matrix elements between configurations dominating the reference wavefunction. The space P' is called *reduced model space* and is perturbed by interacting configurations which facilitates the blocking of the matrix.

Other partitioning schemes were proposed by Witek, Nakano and Hirao[72] who tried to improve the zeroth-order energy by constraining the correction of $|\psi^{(n+1)}\rangle = 0$ in the order n series. The partitioning of the Hamilton operator is not only possible within the Møller-Plesset or the Ebstein-Nesbet framework, but also partitioning schemes were developed that mix both approaches.[73, 74]

2.4.2 Large-Scale Techniques to Improve Performance

The computational efficiency of MR-PT methods compared to MR-CI approaches[4] is advantegous to treat large systems. The bottleneck of the size of the perturbation space remains. Comparing to multi-reference CI methods, similar algorithms to reduce the problem size have been developed for perturbation theories. The formulation of a MR-PT ansatz that consequently used the ansatz of internal contractions reduced the size of the multi-reference problem to that of the single-reference case.[5,27] A different approach was done by Grimme[28,33,75] who adapted the individually selecting approach from the MR-CI[32] to MR-MP and reduced the computational effort substantially. As multi-reference approaches cannot be considered as "black-box" methods, the thresholds used in the indivdually selecting ansatz and the necessary extrapolation add an additional degree of freedom for the user and require an additional amount of experience.

The use of localized orbitals in MR-PT[48] instead of canonical ones reduces the interactions in the Hamilton matrix. This method is considered to be more "black-box" than individually selecting procedures, but the choice of the localized orbitals still remains crucial and retaining a balanced descrpition for the potential energy surface is considered to be similarly or even more difficult.

2.4.3 Methodological Properties

A major issue concerning perturbation theory is the question about size-consistency or size-extensivity. It was formerly believed, that MR-PT is size-consistent if the reference wavefunction is also size-consistent. It was pointed out by Davidson,[76] that no multi-reference perturbation theory then can be size-consistent, because no multi-determinant wavefunction can be an eigenfunction of a Hamiltonian containing only one-particle operators. A study for the CASPT2 and the MR-MP2 of Pulay and Wolinski showed that the errors related to size-inconsistency can be safely neglected for practical applications. MR-MP2 can be made size-consistent by imposing a modified projection scheme and a strict orthog-

---

[4]The efficiency or MR-CI approaches improved substantially by indidually selecting algorithms truncating the CI space to the most important configurations and extrapolation of the neglected space[32] or the development of an MR-CI based on localized orbitals in combination with pseudo-spectral integral evaluation.[9,49] Note that both approaches base on truncated CI spaces but on a different selection scheme.

onalization of the interacting spaces onto each other employing a pivoted Householder QR orthogonalization.[77, 78]

The convergence of MR-PT is strongly related to the quality of the reference wavefunction and the intruder state problem.[79] Approaches to avoid the intruder state problem introduce shift parameters into the equations of the perturbation series and force the orbital energies to diverge. These larger orbital energy differences then improve the convergence behavior.[80, 81]

Compared to MR-CI, MR-MP possesses a larger sparsity of the matrix which improves the efficiency, but MR-MP in contrast is not variational and does not give an upper limit for the total energy. For practical considerations the variational condition is not considered to be of major importance anymore.

Many of the various MR-PT approaches were "proofs of principles" to improve the existing partitioning schemes and the applications mostly did not exceed test cases and were far away from real chemical applications. Some other approaches like CASPT2, Pulay-Wolinski, and Murphy-Messmer probably do not possess the optimum partitioning scheme but can be applied to chemically relevant problems.

Chapter 3

Iterative Methods for Large Linear Equation Systems

3.1 Introduction

In many areas of scientific and industrial applications, most of the problems related with many variable systems are connected to the solution of large matrices. The matrices constructed from the equations of these systems can possess eigenpairs of eigenvectors and eigenvalues or the optimum solution is sought for a certain condition. Especially in quantum chemical applications the computation of eigenvectors and eigenvalues plays a very important role. Eigenvalue problems have to be solved in the HF procedure diagonalizing the Fock matrix to yield the MO as eigenvectors and the MO energies as eigenvalues, in the computation of frequencies or in the computation of excited states in a (MR-)CI procedure. For the latter application the full diagonalization to yield all of the eigenvectors is only possible for the smallest systems possessing a few electrons. The paradigm of direct solvers that compute all eigenpairs changed once larger systems were treated and only the desired eigenvalues and eigenvectors were necessary to be computed. For this purpose Davidson developed an iterative method related to the Lanczos algorithm (see Section 3.3.3) that computed desired eigenpairs in a (MR-)CI problem. The basic algorithm[82;83] was extended and refined to increase its efficiency and is now widely used in quantum chemical applications.[84–86]

A second area of numerical algorithms is related with the solution of large linear equation systems like geometry optimizations, force field algorithms, wavefunction optimization, non-canonical perturbation theory (*e.g.* local MP2 or MR-MP2) and many more. Linear equation systems and eigenvalue problems are different in their conception, but are commonly based on finding a vector, that fulfills a certain condition to a matrix, which can be built for both problems. The basic methods to solve the numerical tasks resemble reach other very much and originate from the same basic techniques, which are briefly described in the next sections. Prior to that, the equations for the multi-reference perturbation theory are given in a matrix form in the following section.

3.2   The Linear Equation System in MRMP

The present work mainly deals with improved algorithms for a multi-reference perturbation theory and focuses on the algorithmic details to obtain the solution vector.

$$\underbrace{\sum_j c_j^1 \langle \Phi_i^1 | \hat{H}_0 - \hat{E}_0 | \Phi_j^1 \rangle}_{\mathbf{A} \cdot \vec{x}} = \underbrace{-\langle \Phi_i^1 | \hat{V} | \psi_0 \rangle}_{\vec{b}} \tag{3.1}$$

The main equation (eq. 3.1, see also eq. 2.96) to solve the first-order correction to the wavefunction represents a set of linear equations. The general structure of such a system is best described as

$$\mathbf{A} \cdot \vec{\tilde{x}} = \vec{b} \tag{3.2}$$

with the Hamilton matrix $\mathbf{A}$, the solution vector $\vec{\tilde{x}}$ and the inhomogenity $\vec{b}$. The solution vector $\vec{x}$ is formally obtained by inverting the matrix $\mathbf{A}$.

$$\vec{\tilde{x}} = \mathbf{A}^{-1} \cdot \mathbf{b} \tag{3.3}$$

The matrix $\mathbf{A}$ is sparse because $\hat{H}_0$ represents a one-electron operator but usually is very large for chemically interesting molecules (dimension of about 10^6 to 10^9). The direct inversion of $\mathbf{A}$ using Gauss elimination[78,87] or single value decomposition (SVD)[78,87] is feasible for medium-sized matrices, but impractical for large and sparse matrices, since these techniques manipulate the rows and columns of the matrix and spoil the sparsity. The direct access to the needed matrix elements during the manipulation requires the matrix to

be stored in memory, a condition that can be only fulfilled for small systems. The loss of the sparsity in the course of the manipulation reinforces this problem. Therefore iterative techniques have to be applied which do not change the matrix, but iteratively improve an approximated trial vector $\vec{x}$ to minimize the difference to the optimum solution vector $\vec{\tilde{x}}$.[87]

3.3   General Approximation Techniques

The iterative methods can be divided into classic techniques which employ matrix partitioning and subspace methods that are based on projection approaches.

3.3.1   Classic Iteration Schemes

By definition[87] the iterative scheme to solve $\mathbf{A}\vec{x} = \vec{b}$ is given as

$$\mathbf{M}\vec{x}_n = \vec{b} + \mathbf{S}\vec{x}_{n-1} \tag{3.4}$$

which employs the partition of $\mathbf{A}$ into

$$\mathbf{A} = \mathbf{M} - \mathbf{S} \tag{3.5}$$

and yields

$$\mathbf{M}\vec{x}_n = \vec{b} - \mathbf{A}\vec{x}_{n-1} - \mathbf{M}\vec{x}_{n-1} \tag{3.6}$$

If $\mathbf{M}$ is non-singular and using the definition of the residual vector $\vec{r}$ (eq,. 3.7) one gets eq. 3.8.

$$\vec{r}_n = \vec{b} - \mathbf{A}\vec{x}_n \tag{3.7}$$

$$\vec{x}_n = \mathbf{M}^{-1}\vec{r}_{n-1} + \vec{x}_{n-1} \tag{3.8}$$

After the evaluation of eq. 3.8, which requires the matrix $\mathbf{M}$ to be easily inverted, the residual $\vec{r}_n$ is computed for the next iteration step.

The partition scheme of $\mathbf{A}$ now defines the different iterative schemes of Gauss-Seidel[78,87] and Jacobi.[78,87]

$$\text{Jacobi} \quad \mathbf{M}_{Jacobi} = \mathbf{D} \quad \mathbf{S}_{Jacobi} = -(\mathbf{L} + \mathbf{U}) \tag{3.9}$$

$$\text{Gauss} - \text{Seidel} \quad \mathbf{M}_{GS} = \mathbf{D} + \mathbf{L} \quad \mathbf{S}_{GS} = -\mathbf{U} \tag{3.10}$$

with the matrix diagonal denoted as $\mathbf{D}$, the strict lower triangle $\mathbf{L}$ and the strict upper triangle $\mathbf{U}$.

3.3.2 General Projection Methods

More elaborate techniques than these classic iteration schemes utilizing matrix partitioning are projection methods. Projection techniques onto $\mathcal{K}$ and $\mathcal{L}$, which are defined as subspaces of the space spanned by the matrix $\mathbf{A}$, find an approximate solution $\vec{\tilde{x}}$ that belongs to $\mathcal{K}$ by imposing the condition that the new residual vector is orthogonal to $\mathcal{L}$ (eq. 3.11 and 3.12).[87, 88]

$$\vec{\tilde{x}} \in \mathcal{K} \tag{3.11}$$

$$\vec{b} - \mathbf{A}\vec{\tilde{x}} \perp \mathcal{L} \tag{3.12}$$

The vector $\vec{\tilde{x}}$ can be written as $\vec{\tilde{x}} = \vec{x} + \delta$ with $\delta \in \mathcal{K}$ and the equation of the residual becomes

$$\begin{aligned} 0 &= \vec{b} - \mathbf{A}\vec{\tilde{x}} \tag{3.13} \\ &= \vec{b} - \mathbf{A}\left(\vec{x} + \delta\right) \perp \mathcal{L} \\ &= \vec{r}_0 - \mathbf{A}\delta \perp \mathcal{L} \end{aligned}$$

$$\vec{r}_0 - \mathbf{A}\delta \perp \mathcal{L} \tag{3.14}$$

which defines the approximate solution. The matrix is projected onto the residual in each step and a new orthogonal residual that leads to a new trial vector is created for the next iteration step. A new enlarged subspace is also built. Note that for the two subspaces $\mathcal{K}$ and $\mathcal{L}$ being identical, the projection method is equivalent to the Gauss-Seidel procedure.[88]

3.3.3 Krylov Subspace Methods

The most commonly used subspace methods are derived from the Krylov subspace method[87, 88] in which the subspace of the matrix $\mathbf{A}$ and the vector $\vec{r}$ is defined to be

$$\mathcal{K}_m\left(\mathbf{A}, \vec{r}_0\right) = span\left\{\vec{r}_0, \mathbf{A}\vec{r}_0, \mathbf{A}^2\vec{r}_0, \ldots, \mathbf{A}^{m-1}\vec{r}_0\right\} \tag{3.15}$$

The approximation to the solution $\vec{\tilde{x}}$ obtained from the Krylov subspace is of the form

$$\mathbf{A}^{-1}\vec{b} \approx \vec{x}_m = \vec{r}_0 + q_{m-1}\left(\mathbf{A}\right)\vec{r}_0 \tag{3.16}$$

and gives an approximation for the inversion of $\mathbf{A}$ with a certain polynomial q of degree $m-1$. For the simplest case of $\vec{x}_0 = 0$ and $\vec{r}_0 = \vec{b}_0 - \mathbf{A}\vec{x}_0$, the exact solution is approximated by this polynomial.

$$\mathbf{A}^{-1}\vec{b} \approx q_{m-1}\left(\mathbf{A}\right)\vec{b} \tag{3.17}$$

The Krylov subspace method is applicable to any matrix. More specialized derivatives are the Arnoldi method that is optimal for non-symmetric matrices and the Lanczos method that is formulated for symmetric matrices.[88]

3.4   The Method of Conjugate Gradients

3.4.1   Methodic Background and Algorithm

The method of Conjugate Gradients[78, 88, 89] is a well known iterative technique for positive-definite sparse linear equation systems involving Hermitian matrices. Basically, it is an orthogonal projection technique onto the Krylov subspace $\mathcal{K}_m\left(\mathbf{A}, \vec{r}_0\right)$, which can be considered as a derivative of the Lanczos method.

For the functional $\mathcal{Q}$ given in eq. 3.18 the solution $\vec{\tilde{x}}$ (eq. 3.19) of the linear equation system represents its minimum.[89]

$$\mathcal{Q} = \frac{1}{2}\vec{x}^T\mathbf{A}\vec{x} - \vec{b}^T\vec{x} \tag{3.18}$$

$$\vec{\tilde{x}} = \mathbf{A}^{-1}\vec{b} \tag{3.19}$$

The idea to determine the minimum of $\mathcal{Q}$ starts from an approximate solution vector $\vec{x}^k$ and a vector $\vec{p}^k$ defining a search direction. For the iterative procedure

$$\vec{x}^{k+1} = \vec{x}^k + \alpha_k \vec{p}^k \tag{3.20}$$

with the real value α_k is determined to fulfill the condition

$$\mathcal{Q}\left(\vec{x}^{k+1}\right) = \mathcal{Q}\left(\vec{x}^k + \alpha_k \vec{p}^k\right) = min\mathcal{Q}\left(\vec{x}^k + \alpha_k \vec{p}^k\right) \tag{3.21}$$

The vectors $\vec{p}^k$ are conjugate to $\mathbf{A}$ and fulfill the orthogonality relation

$$(\vec{x}, \mathbf{A}\vec{y}) = \vec{x}^T \mathbf{A}\vec{y} = \delta_{xy} \tag{3.22}$$

if $\vec{x} \neq \vec{y}$. The vectors $\vec{p}$ are linearly independent.[89] Furthermore it can be shown, that the residual vectors $\vec{r}$ are orthogonal to each other (eq. 3.23).

$$\left(\vec{r}^k, \vec{r}^j\right) = \left(\vec{r}^k\right)^T \cdot \vec{r}^j = 0 \tag{3.23}$$

For the calculation of α_k it is necessary to derive a formula for the residual $\vec{r}^{k+1}$ in the next iteration step.

$$\begin{aligned}
\vec{r}^{k+1} &= \mathbf{A}\vec{x}^{k+1} - \vec{b} \tag{3.24}\\
&= \mathbf{A}\left(\vec{x}^k + \alpha \vec{p}^k\right) - \vec{b}\\
&= \left(\mathbf{A}\vec{x}^k - \vec{b}\right) + \alpha_k \mathbf{A}\vec{p}^k\\
&= \vec{r}^k + \alpha \mathbf{A}\vec{p}^k \tag{3.25}
\end{aligned}$$

With eq. 3.25 α_k can now be calculated. The projection of $\vec{r}^k$ onto eq. 3.25 from the right-hand side and solving for α_k becomes

$$\underbrace{\left(\vec{r}^{k+1}, \vec{r}^k\right)}_{=0} = \left(\vec{r}^k, \vec{r}^k\right) + \alpha_k \left(\mathbf{A}\vec{p}^k, \vec{r}^k\right) \tag{3.26}$$

$$\alpha_k = -\frac{\left(\vec{r}^k, \vec{r}^k\right)}{\left(\mathbf{A}\vec{p}^k, \vec{r}^k\right)} \tag{3.27}$$

The new search direction is defined by the projection method to

$$\vec{p}^{k+1} = \vec{r}^{k+1} + \beta_k \vec{p}^k \tag{3.28}$$

The denominator in eq. 3.27 can be rewritten using eq. 3.28 and employing the conjugacy relation eq. 3.22 to

$$\left(\mathbf{A}\vec{p}^{k}, \vec{r}^{k}\right) = \left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k} - \beta_{k-1}\vec{p}^{k-1}\right) \tag{3.29}$$

$$= \left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k}\right) - \beta_{k-1}\underbrace{\left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k-1}\right)}_{=0} \tag{3.30}$$

$$= \left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k}\right) \tag{3.31}$$

The optimization parameter α_k finally is expressed as

$$\alpha_k = -\frac{\left(\vec{r}^{k}, \vec{r}^{k}\right)}{\left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k}\right)} \tag{3.32}$$

The scaling parameter for the new search direction is derived from eq. 3.28 by projection with $\mathbf{A}\vec{p}^{k}$ from the right.

$$\underbrace{\left(\vec{p}^{k+1}, \mathbf{A}\vec{p}^{k}\right)}_{=0} = \left(\vec{r}^{k+1}, \mathbf{A}\vec{p}^{k}\right) + \beta_k\left(\vec{p}^{k}, \mathbf{A}\vec{p}^{k}\right) \tag{3.33}$$

$$\beta_k = -\frac{\left(\vec{r}^{k+1}, \mathbf{A}\vec{p}^{k}\right)}{\left(\vec{p}^{k}, \mathbf{A}\vec{p}^{k}\right)} \tag{3.34}$$

β_k is further simplified by solving eq. 3.25 to $\mathbf{A}\vec{p}^{k}$, inserting this equation into the nominator of the preliminary form of β_k (eq. 3.34) and finally substituting α_k by eq. 3.32.

$$\beta_k = -\frac{1}{\alpha_k}\frac{\left(\vec{r}^{k+1}, \vec{r}^{k+1} - \vec{r}^{k}\right)}{\left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k}\right)} \tag{3.35}$$

$$= -\frac{1}{\alpha_k}\frac{\left(\vec{r}^{k+1}, \vec{r}^{k+1}\right)}{\left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k}\right)}$$

$$= \frac{\left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k}\right)}{\left(\vec{r}^{k}, \vec{r}^{k}\right)}\frac{\left(\vec{r}^{k+1}, \vec{r}^{k+1}\right)}{\left(\mathbf{A}\vec{p}^{k}, \vec{p}^{k}\right)}$$

$$= \frac{\left(\vec{r}^{k+1}, \vec{r}^{k+1}\right)}{\left(\vec{r}^{k}, \vec{r}^{k}\right)} \tag{3.36}$$

The algorithm for the Conjugate Gradient method can now be formulated

- Choose start vector $\vec{x}_0$

- $\vec{p}_0 = \vec{r}_0 = \mathbf{A}\vec{x}_0 - \vec{b}$

- $(\gamma_0)^2 = (\vec{r}_0, \vec{r}_0)$

- For $k = 0, 1, \ldots$

  1. $\vec{z}^k = \mathbf{A}\vec{p}^k$

  2. $\alpha_k = -\dfrac{\gamma_k^2}{\left(\vec{p}^k, \vec{z}^k\right)}$

  3. $\vec{x}^{k+1} = \vec{x}^k + \alpha_k \vec{p}^k$

  4. $\vec{r}^{k+1} = \vec{r}^k + \alpha_k \vec{z}^k$

  5. $\left(\gamma_{k+1}\right)^2 = \left(\vec{r}^{k+1}, \vec{r}^{k+1}\right)$

  6. $\beta_k = \dfrac{(\gamma_{k+1})^2}{(\gamma_k)^2}$; if $\gamma_{k+1} \leq \epsilon$ stop

  7. $\vec{p}^{k+1} = \vec{r}^{k+1} + \beta_k \vec{p}^k$

3.4.2　Preconditioning Techniques

The application of the Conjugate Gradient method given in its basic formulation is possible for sparse linear equation systems characterized by symmetric positive-definite matrices, but it is not recommended since the convergence behavior changes substantially with the matrix structure and the method looses its robustness. The best matrix condition to achieve fast convergence is a matrix that is close to the unit matrix. This disadvantage can be overcome if preconditioners are introduced, that manipulate the matrix $\mathbf{A}$ before the first step in the Conjugate Gradient algorithm and therefore enhance the efficiency of the algorithm by reducing the necessary iterations to achieve convergence.

The concept of the manipulation of the linear equation system by preconditioners is illustrated in the following set of equations leading to a modified algorithm. The original matrix equation is expanded by the unit matrix, which is then replaced by the resolved identity $\mathbf{S}^T\mathbf{S}^{-T}$. In the next step, the matrix $\mathbf{S}$ is multiplied from the left and a similarity transformed set of equations is built.

$$\mathbf{A}\vec{x} = \vec{b}$$
$$\mathbf{A}\mathbf{S}^T\mathbf{S}^{-T}\vec{x} = \vec{b}$$
$$\underbrace{\mathbf{S}\mathbf{A}\mathbf{S}^T}_{=\hat{\mathbf{A}}}\underbrace{\mathbf{S}^{-T}\vec{x}}_{=\hat{\vec{x}}} = \underbrace{\mathbf{S}\vec{b}}_{=\hat{\vec{b}}}$$

The transformed matrix $\hat{\mathbf{A}}$ and the vectors $\hat{\vec{x}}$ and $\hat{\vec{b}}$ can be applied to the Conjugate Gradient method, but a more efficient algorithm can be deduced from the original one that avoids an

explicit computation of $\hat{\mathbf{A}} = \mathbf{S}^T \mathbf{A} \mathbf{S}^{-T}$ and $\vec{\hat{b}} = \mathbf{S}\vec{b}$.[89]

The algorithm for the Preconditioned Conjugate Gradient method can be formulated as

- Choose start vector $\vec{x}_0$

- $\vec{r}_0 = \mathbf{A}\vec{x}_0 - \vec{b}$

- $\mathbf{M}\vec{\hat{r}}_0 = \vec{r}_0$; *i.e.* $\vec{\hat{r}}_0 = \mathbf{M}^{-1}\vec{r}_0$

- $\vec{p}_0 = \vec{\hat{r}}_0$

- $(\gamma_0)^2 = \left(\vec{\hat{r}}_0, \vec{r}_0\right)$

- For $k = 0, 1, \ldots$

  1. $\vec{z}^k = \mathbf{A}\vec{p}^k$

  2. $\alpha_k = -\dfrac{\gamma_k^2}{\left(\vec{p}^k, \vec{z}^k\right)}$

  3. $\vec{x}^{k+1} = \vec{x}^k + \alpha_k \vec{p}^k$

  4. $\vec{r}^{k+1} = \vec{r}^k + \alpha_k \vec{z}^k$

  5. Solve $\mathbf{M}\vec{\hat{r}}^{k+1} = \vec{r}^{k+1}$ for $\vec{r}^{k+1} \Rightarrow \vec{\hat{r}}^{l+1} = \mathbf{M}^{-1}\vec{\hat{r}}^{k+1}$

  6. $(\gamma_{k+1})^2 = \left(\vec{\hat{r}}^{k+1}, \vec{r}^{k+1}\right)$

  7. $\beta_k = \dfrac{(\gamma_{k+1})^2}{(\gamma_k)^2}$; If $\gamma_{k+1} \leq \epsilon$ stop

  8. $\vec{p}^{k+1} = \vec{\hat{r}}^{k+1} + \beta_k \vec{p}^k$

For $\mathbf{M} = \mathbf{I}$, the algorithm equals the unpreconditioned form. The main problem that still exists is solving the equation $\mathbf{M}\vec{\hat{r}}^{k+1} = \vec{r}^{k+1}$. The matrix $\mathbf{M}$ also should transform the matrix $\mathbf{A}$ close to the identity matrix for an optimal convergence behavior. The simplest choice for a preconditioning matrix $\mathbf{M}$ is the matrix containing the main diagonal of $\mathbf{A}$. This matrix can be easily inverted and is readily accessible since the explicit construction of an additional preconditioner matrix is unnecessary and improves the convergence behavior. Other techniques to construct preconditioner matrices are the incomplete Cholesky decompostion[88, 89] and domain decomposition methods.[89]

3.4.3 Parallelization

The Conjugate Gradient method can be easily parallelized. The degree of parallelization mainly depends on the efficiency of a parallel matrix vector product to form $\vec{z}^k = \mathbf{A} \cdot \vec{p}^k$. The vector updates of $\vec{x}^{k+1}$, $\vec{r}^{k+1}$ and $\vec{p}^{k+1}$ can be done parallelly without any interprocess communication. The scalar products to calculate α_k and β_k are partially computed on each node and then summed up to the final value.

- Choose start vector $\vec{x}_0$

- $\vec{p}_0 = \vec{r}_0 = \mathbf{A}\vec{x}_0 - \vec{b}$

- $(\gamma_0)^2 = (\vec{r}_0, \vec{r}_0)$

- For $k = 0, 1, \ldots$

   1. Compute $\vec{z}^k = \mathbf{A}\vec{p}^k$ in parallel. Redistribute the solution to all nodes.

   2. Compute partial scalar product $\left(\vec{p}^k, \vec{z}^k\right)$ in parallel and collect to $\alpha_k = -\frac{\gamma_k^2}{\left(\vec{p}^k, \vec{z}^k\right)}$. Distribute α_k to all nodes.

   3. Compute $\vec{x}^{k+1} = \vec{x}^k + \alpha_k \vec{p}^k$ in parallel.

   4. Compute $\vec{r}^{k+1} = \vec{r}^k + \alpha_k \vec{z}^k$ in parallel.

   5. Compute parts of $(\gamma_{k+1})^2 = \left(\vec{r}^{k+1}, \vec{r}^{k+1}\right)$ parallel and collect; If $\gamma_{k+1} \leq \epsilon$ stop

   6. $\beta_k = \frac{(\gamma_{k+1})^2}{(\gamma_k)^2}$ is also computed in parallel mode and send to all nodes.

   7. Compute $\vec{p}^{k+1} = \vec{r}^{k+1} + \beta_k \vec{p}^k$ in parallel.

   8. Exchange components of $\vec{p}^{k+1}$ between all nodes.

Chapter 4

Concepts and Algorithm for Large-Scale MR-MP2 Computations

4.1 Basic Concepts

The iterative scheme in a MR-MP2 procedure to obtain the first-order correction to the wavefunction and finally the second-order energy require the repeated use of the Hamilton matrix, which is identical in each iteration step. The reference wavefunction in combination with the externals orbitals governs the matrix dimension and the size of the solution vector. The representation of the reference wavefunction and the total wave function is either possible with Slater determinants or with Configuration State Functions (CSF). A configuration, from which a set of CSF is generated, is a scheme of occupied orbitals and can be also regarded as a linear space spanned by Slater determinants with all possible combinations of α and β spin in the open-shell orbitals.[90] The usage of CSF instead of Slater determinants reduces the size of the Hamilton matrix (see Section 4.2.1). Without any further simplifications the number of matrix elements in the Hamilton matrix scales quadratically with the number of CSF or Slater determinants.

For large systems the dimension of the hermitian Hamilton matrix exceeds the main memory capacities of the largest computers and the fastest scheme in which the matrix is kept in the main memory and used in each iteration for the multiplication with converging trial vectors has to be abandoned. Comparing the CPU performance of nowadays computers

with disc-based storage devices, disc storage is disadvanteous and it is more efficient to construct the Hamilton matrix incrementally in each iteration and to multiply the generated matrix elements directly with the required part of the trial vector.

The two main concepts that did come up with direct methods were the integral-driven approach for which the evaluation of the matrix elements is controlled by the present integrals[5] and a configuration-driven approach.[90] The integral-driven scheme starts from a set of integrals and evaluates all matrix elements connected to these integrals independent from their position in the matrix. In the configuration-driven approach the start addresses of the coefficients of the trial vector are easier to determine from the position of the interacting configurations. Both approaches take advantage of sorted interacting spaces discussed in Section 4.3 but the configuration-driven approach was preferred for this MR-MP2 project to facilitate parallelization and the efficient construction of the Hamilton matrix. The efficient construction of the Hamilton matrix is crucial for the performance of the method since the most time-consuming step in the scheme is the combined generation of $\mathbf{H_0}$ and the multiplication of the matrix elements with the trial vector. The amount of memory needed to store the necessary integrals for the iterative part of the MR-MP2 procedure is negligible as will be pointed out later in this chapter.

An efficient computation of the Hamilton matrix requires elaborate schemes, which reduce the overhead of computing matrix elements that vanish by symmetry or by virtue of the operator properties and also an efficient computation of the matrix elements themselves. For the computation of the matrix elements the Symmetric Group Approach was chosen, which was already successfully applied in the DIESEL program package.[53, 54] An analysis of the structure of the Hamilton matrix on various levels and the classification of interactions between pairs of configurations illustrate the concepts to reduce the computational cost for the built of the Hamilton matrix in a MR-MP2 procedure.

4.2   The Computation of Matrix Elements in the Symmetric Group Approach (SGA)

4.2.1   The Symmetric Group Approach and Characterization of Interactions

4.2.1.1   The SGA Framework

The framework used in this work to represent matrix elements of the Hamilton matrix is the same as in the MR-CI approaches implemented by Volker Pless and Michael Hanrath[54, 91] and bases on the SGA.[92, 93] The Symmetric Group is an abelian group of permutations, that can be used to construct the antisymmetric spin eigenfunctions for a given multiplicity S, that are eigenfunctions to the $\hat{S}^2$ and the $\hat{S}_z$ operators. These eigenfunctions are used to build up Configuration State Functions (CSF) from configurations. Configurations are schemes of orbital occupations and do not represent a mathematical object in the sense that it can be used to compute matrix elements. The CSF for a given configuration are computed from the spatial part of the wavefunction multiplied by the spin eigenfunctions. The degree of degeneracy $f_{S,\nu}^{\mathrm{SGA}}$ of a certain configuration for a given multiplicity $2S+1$ and number of unpaired electrons ν is given as

$$f_{S,\nu}^{\mathrm{SGA}} = \begin{pmatrix} \nu \\ \frac{1}{2}\nu - S \end{pmatrix} - \begin{pmatrix} \nu \\ \frac{1}{2}\nu - S - 1 \end{pmatrix} \tag{4.1}$$

In the formalism of the SGA a CSF is expressed as

$$\Phi_n^{\mathrm{SGA}} = \hat{A} \prod_{j=1}^{n} \phi_{n(j)}^{\mathrm{MO}}(\vec{r}_j) \cdot \Theta_k^{S,M}(\omega_1, \ldots, \omega_n) \tag{4.2}$$

with the antisymmetrizing operator $\hat{A}$, the occupied MO $\phi_{n(j)}^{\mathrm{MO}}$ and the spin eigenfunctions $\Theta_k^{S,M}$. The spin eigenfunctions $\Theta_k^{S,M}$ depend on the total spin S with the $\langle \hat{S}^2 \rangle$ eigenvalue $S(S+1)$, the $\langle \hat{S}_z \rangle$ eigenvalue M, and the degeneracy index k. The spin eigenfunctions $\Theta_k^{S,M}$ can be generated as linear combinations of products of primitive spin functions Ω. They need not to be explicitly generated but can be evaluated by more efficient schemes like Young-Tableaux.[92]

$$\Theta_k^{S,M} = \sum_i \gamma_i \prod_j \Omega(\omega_j) \tag{4.3}$$

The computation of matrix elements[93] between two CSF Φ_n^{SGA} and Φ_m^{SGA} can be computed as

$$
\begin{aligned}
\langle \Phi_n^{\mathrm{SGA}} | \hat{H} | \Phi_m^{\mathrm{SGA}} \rangle &= \frac{1}{n!} \sum_{P \in S_n} (-1)^P \\
&\cdot \underbrace{\left\langle \prod_{j=1}^n \phi_{n(j)}^{\mathrm{MO}}(\vec{r}_j) | \hat{H}\hat{P} | \prod_{j=1}^m \phi_{m(j)}^{\mathrm{MO}}(\vec{r}_j) \right\rangle}_{\text{spatial part}} \\
&\cdot \underbrace{\left\langle \Theta_l^{S,M}(\omega_1, \ldots, \omega_n) | \hat{P} | \Theta_k^{S,M}(\omega_1, \ldots, \omega_n) \right\rangle}_{\text{spin part } U_{kl}}
\end{aligned}
\tag{4.4}
$$

with $P \in S_n$ denoting all n possible permutations of the permutation operator $\hat{P}$ amongst the open-shell MO S_n. As seen from eq. 4.4, the computation of a matrix element in the SGA is completely separated into a spatial part and a spin part. Matrix elements between the spin parts of two configurations with a certain degree of degeneracy constitute a matrix of coupling coefficients U_{kl} between the corresponding CSF called representation matrix $\mathbf{U}$.

The matrix elements could also be evaluated by linear combinations of Slater determinants that represent eigenfunctions to the $\hat{S}^2$ and the $\hat{S}_z$ operator (Table CI). However, the use of CSF in the SGA framework instead of Slater determinants is advantageous as the number spin eigenfunctions in the SGA and therefore the size of the Hamilton matrix is smaller than the number of Slater determinants that have to linearly combined (compare eq. 4.1 and eq. 4.5).

$$f_{S,\nu}^{\mathrm{Slater}} = \begin{pmatrix} \nu \\ \frac{1}{2}\nu - S \end{pmatrix} \tag{4.5}$$

In the SGA the spin functions are generated first and then antisymmetrized afterwards, while linear combinations of Slater determinants are proceeded the other way round.

4.2.1.2   Interaction Schemes

The computation of matrix elements in the form of eq. 4.4 depends on a sum over $n!$ permutations described by a permutator that brings two configurations into maximum alignment (line-up permutation). Each interaction between two configurations can be described by a permutation operator, but the assignment of the permutation operators to the interaction classes is not unique. A simpler classification scheme like the Slater rules is more desirable and it was shown[91] that the representation matrices computed in the framework of the SGA can be assigned to the labels of the Table CI procedure.[94–96]

The Table CI classifies the interactions between two configurations into various cases characterized by the labels ΔK, P, Q, R, Q_l and Q_r. This classification originates from the integration of the spin part and the antisymmetrization of the spatial and the spin part.

1. $\Delta K \in \{-2, -1, 0, 1, 2\}$ gives the difference in the number of open shells between two configurations $|\phi_R\rangle$ and $|\phi_L\rangle$.

2. $P \in \{1, \ldots, 5\}$ characterizes the type of the required integrals.

3. $R \in \{1, \ldots, 6\}$ codes the number and the position of closed-shell orbitals with respect to the interacting open-shell orbitals.

4. Q_r and Q_l with any integer number represent the positions (p_i) of the singly occupied interacting orbitals n_s in a configuration and are computed according to

$$Q_r, Q_l \;\; = \;\; 1 + \sum_{i=1}^{n_s} \begin{pmatrix} p_i - 1 \\ n_s - i + 1 \end{pmatrix}$$

These labels generate 15 major interaction classes for interactions between two configurations. In some of these classes finer interaction classes related to the interacting orbitals and required integrals can be distinguished. The unique classification of two interacting configurations is denoted in the following as a Table CI case.[91] The configuration with the same or a larger number of open shells is denoted as $|\phi_\mu\rangle$ and the other one as $|\phi_\lambda\rangle$.

Table 4.1: Table CI cases taken from the Ph.D. thesis of Volker Pless,[91] Table 3.2. The first configuration is $|\mu\rangle$, the second one $|\lambda\rangle$. Note that in each configuration the order of the orbitals is $a < b < c < d$.

Table Case	Configurations	Integrals	ΔK	P	R
1	c^2 a^2	$(ca\|ca)$	0	4	1
2	b^1c^2 a^1b^2	$(cb\|ca)$	0	2	1
3	c^2 a^1b^1	$(cb\|ca)$	1	2	2
4	a^2b^2 $a^1b^1c^2$	$(cb\|ca)$	1	2	1
5	c^1d^1 a^1b^1	$(db\|ca), (da\|cb)$	1	2	1
6	$c^1d^1a^2b^2$ a^1b^1	$(db\|ca), (da\|cb)$	0	1	6
7a	$c^1d^1a^2$ $a^1b^1c^2$	$(dc\|ba), (db\|ca)$	0	1	2
7b	$c^1d^1b^2$ $a^1b^1c^2$	$(dc\|ba), (da\|cb)$	0	1	3
7c	$c^1d^1a^2$ $a^1b^1d^2$	$(dc\|ba), (da\|cb)$	0	1	5
7d	$c^1d^1b^2$ $a^1b^1c^2$	$(dc\|ba), (db\|ca)$	0	1	5
8a	d^1a^2 $a^1b^1c^1$	$(dc\|ba), (db\|ca)$	1	1	4
8b	d^1b^2 $a^1b^1c^1$	$(dc\|ba), (da\|cb)$	1	1	5
8c	d^1c^2 $a^1b^1c^1$	$(dc\|ba), (da\|cb)$	1	1	6
9a	$d^1b^2c^2$ $a^1b^1c^1d^2$	$(dc\|ba), (db\|ca)$	1	1	1

9b	$d^1a^2c^2$ $a^1b^1c^1d^2$	$(dc|ba), (da|cb)$	1	1	2
9c	$d^1a^2b^2$ $a^1b^1c^1d^2$	$(db|ca), (da|cb)$	1	1	3
10a	a^2b^2 $a^1b^1c^1d^1$	$(db|ca), (da|cb)$	2	1	1
10b	a^2c^2 $a^1b^1c^1d^1$	$(dc|ba), (da|cb)$	2	1	2
10c	a^2d^2 $a^1b^1c^1d^1$	$(dc|ba), (db|ca)$	2	1	3
10d	b^2c^2 $a^1b^1c^1d^1$	$(dc|ba), (db|ca)$	2	1	4
10e	b^2d^2 $a^1b^1c^1d^1$	$(dc|ba), (da|cb)$	2	1	5
10f	c^2d^2 $a^1b^1c^1d^1$	$(db|ca), (da|cb)$	2	1	6
11	c^1 a^1		0	3	1
12	a^1c^2 c^1a^2		0	3	2
13a	a^2 a^1b^1		1	3	1
13b	b^2 a^1b^1		1	3	2
14			0	5	1

The labels of a Table CI Case can be determined by means of difference configurations that are represented by a common part of two interacting configurations, a second part with missing orbitals in $|\phi_L\rangle$ and a third part with missing orbitals in $|\phi_R\rangle$.

The given example illustrates the concept of a difference configuration created from the two interacting configurations $|\phi_R\rangle$ and $|\phi_L\rangle$.

$$
\begin{aligned}
|\phi_L\rangle &= |a^2 b^1 c^2 d^1\rangle \\
|\phi_R\rangle &= |a^2 b^1 c^1 d^2\rangle
\end{aligned}
$$

$$
\begin{aligned}
\text{common part} &\ :\ |a^2 b^1\rangle \\
\text{missing in } |\phi_L\rangle &\ :\ |c^1 d^2\rangle \\
\text{missing in } |\phi_R\rangle &\ :\ |c^2 d^1\rangle
\end{aligned}
$$

The position of the interacting open shells p_i are

$$
\begin{aligned}
\text{missing in } |\phi_L\rangle &\ :\ (1) \\
\text{missing in } |\phi_R\rangle &\ :\ (1)
\end{aligned}
$$

and the excitation order is computed to be a single excitation.

$$
|c^1 d^2\rangle \rightarrow |c^2 d^1\rangle
$$

The determination of the Table CI Case can be done very efficiently from the scheme in Figure 4.1[54] and requires only a maximum of three orbital comparisons for the $\hat{H}$ operator. For the $\hat{H}_0$ operator in the case of a MR-MP2 treatment this scheme further simplifies to a maximum of only one orbital comparison (Figure 4.2) and increases the efficiency to create the Table CI labels that characterize the representation matrices.

4.2.2   Explicit Computation of Matrix Elements

4.2.2.1   Matrix Elements for a Two-Particle Operator ($\hat{H}$)

The computation of the spatial part of the matrix elements for a two-particle Hamilton operator $H_{|\lambda\rangle,|\mu\rangle}$ (with the two interacting configurations $|\phi_\lambda\rangle$ and $|\phi_\mu\rangle$) is needed in the computation of the inhomogenity (see eq 2.96). Open-shell orbitals that are common in both configurations are denoted as S_i, while doubly occupied orbitals common in $|\lambda\rangle$ and $|\mu\rangle$ are denoted as D_i and their corresponding numbers as n_s and n_d. Interacting orbitals are given in minor letters a, b, c and d. $\mathbf{U}_S^{n_s}$ are the representation matrices of the permutation
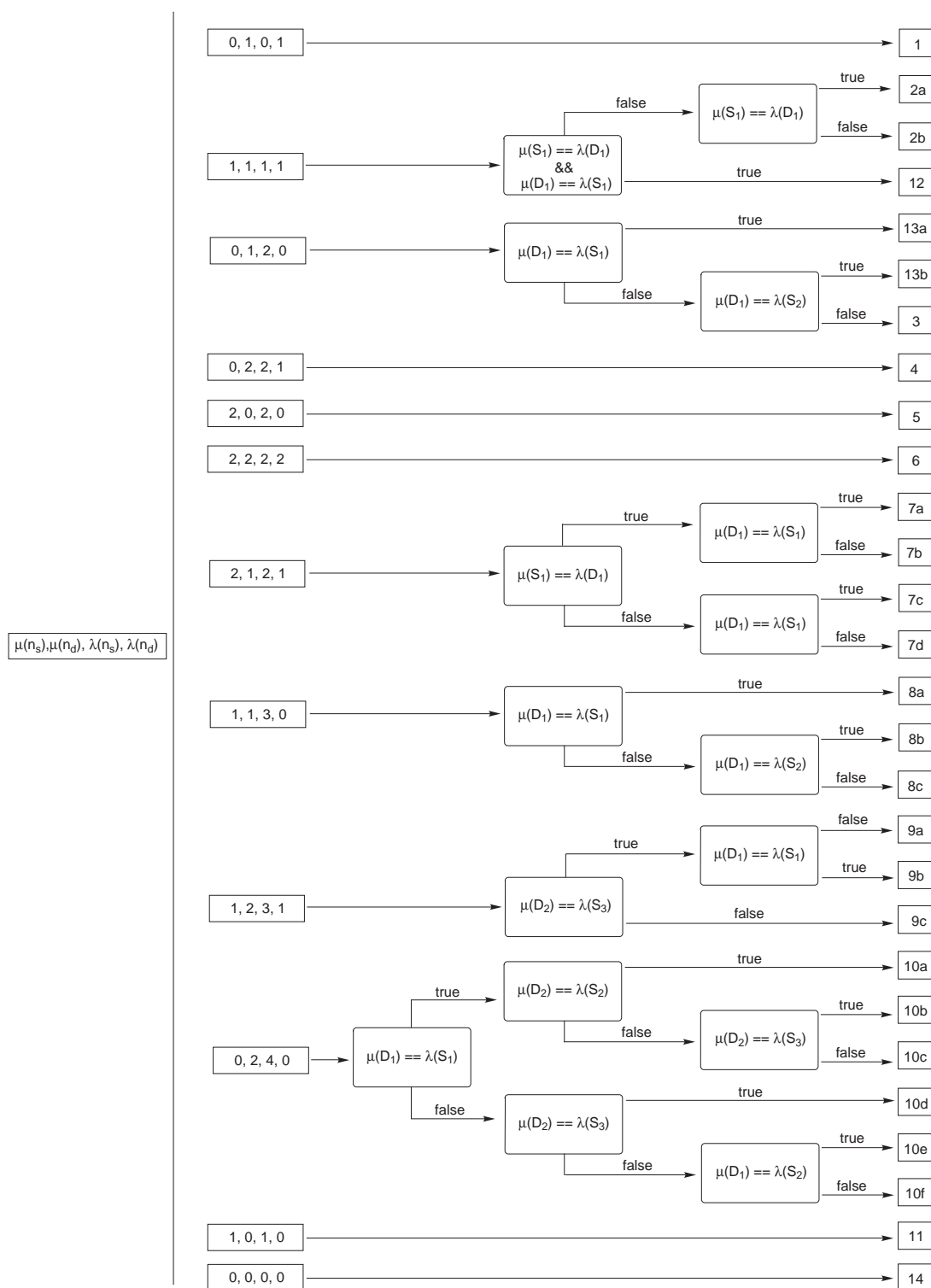
Figure 4.1: Scheme to determine the Table CI cases according to Michael Hanrath.[54]

operator concerning (ai) (transposition) and $\hat{P}_q^s$ of the line-up permutation in the open-shell orbitals.
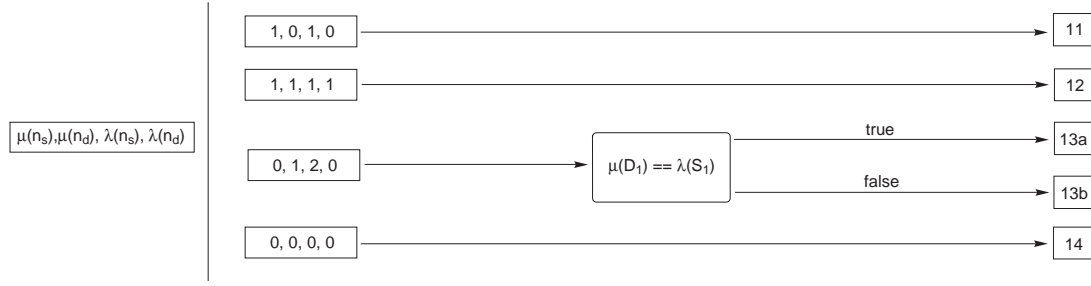
Figure 4.2: Scheme to determine the Table CI cases for an effective one-particle operator.

1. Identical configurations (diagonal elements, case 14):

$$
\begin{aligned}
H_{|\lambda\rangle,|\mu\rangle} = \ & \Big[\sum_{i=1}^{n_s}(S_i|S_i) \\
& + \sum_{i=1}^{n_d}\left(2\cdot(D_i|D_i) + (D_iD_i|D_iD_i)\right) \\
& + 2\sum_{i=1}^{n_s}\sum_{j=1}^{n_d}(S_iS_i|D_jD_j) + 4\sum_{i=1}^{n_d}\sum_{j=i+1}^{n_d}(D_iD_i|D_jD_j) \\
& - \sum_{i=1}^{n_s}\sum_{j=1}^{n_d}(S_iD_j|D_jS_i) - 2\sum_{i=1}^{n_d}\sum_{j=i+1}^{n_d}(D_iD_j|D_jD_i) \\
& + \sum_{i=1}^{n_s}\sum_{j=i+1}^{n_s}(S_iS_i|S_jS_j)\Big]\cdot\mathbf{1} \\
& + \sum_{j=1}^{n_s-1}\sum_{i=j+1}^{n_s}(S_iS_j|S_jS_i)\cdot\mathbf{U_S^{n_s}}\left((ij)\right)
\end{aligned}
\tag{4.6}
$$

2. Three cases can be distinguished for single excitations with the interacting orbitals a and b. The common part in all cases is given as I:

$$
\begin{aligned}
I = \ & \Big[(a|b) \\
& + \sum_{i=1}^{n_s}(ab|S_iS_i) \\
& + \sum_{i=1}^{n_d}\left(2\cdot(ab|D_iD_i) - (aD_i|D_ib)\right)\Big]
\end{aligned}
\tag{4.7}
$$

(a) Case 11 $[b \leftrightarrow a]$

$$
H_{|\lambda\rangle,|\mu\rangle} = \Big[I + \sum_{i=1}^{n_s}(aS_i|S_ib)\cdot\mathbf{U_S^{n_s}}\left((S_1^{|\lambda\rangle)}i)\right)\Big]\cdot\mathbf{U_S^{n_s}}\left(\hat{P}_q^s\right)
\tag{4.8}
$$

(b) Case 12 $[ab^2 \leftrightarrow a^2b]$

$$
H_{|\lambda\rangle,|\mu\rangle} = \Big[I - (ab|aa) - (ab|bb) + \sum_{i=1}^{n_s}(aS_i|S_ib)\cdot\left[\mathbf{1} + \mathbf{U_S^{n_s}}\left((ai)\right)\right]\Big]\cdot\mathbf{U_S^{n_s}}\left(\hat{P}_q^s\right)
\tag{4.9}
$$

(c) Case 13 $[a^2 \leftrightarrow ab]$ or $[b^2 \leftrightarrow ab]$

$$
H_{|\lambda\rangle,|\mu\rangle} = \Big[I + (ab|aa) + \sum_{i=1}^{n_s}(aS_i|S_ib)\Big]\cdot\mathbf{U_S^{n_s}}\left((ai)\right)\Big]\cdot\mathbf{U_S^{n_s}}\left(\hat{P}_q^s\right)
\tag{4.10}
$$

3. Double excitation (cases 1-10):

$$H_{|\lambda\rangle,|\mu\rangle} = \left[a_{|\lambda\rangle,|\lambda\rangle}(ab|cd) + A_{|\lambda\rangle,|\lambda\rangle}(ad|cb)\right] \cdot \mathbf{U}_{\mathbf{S}}^{\mathbf{n_s}}\left(\hat{P}_q^s\right) \tag{4.11}$$

4.2.2.2   Matrix Elements for an Effective One-Particle Operator ($\hat{H}_0$)

The rules of the SGA are also applied to compute the matrix elements for effective one-particle operators like the $\hat{H}_0$ operator in MR-MP2 but the computation of the matrix elements simplifies substantially compared to a two-particle operator. For the computation of the matrix elements of the $\mathbf{H_0}$ matrix the form of the $\hat{H}_0$ operator in terms of the many-electron Fock operator in eq. 4.12 is recalled. The matrix elements for $\hat{f}_l$ are given in eq. 4.13 (see also (eq. 2.77 and 2.76 in Section 2.3). The projection operators in $\hat{H}_0$ are already applied by the restriction that the interacting configurations belong uniquely to the first-order interacting space from which the reference space is excluded.

$$\hat{F} = \sum_l \hat{f}_l \tag{4.12}$$

$$f_{ab} = h_{ab} + \sum_{ij} D_{ij} \left[(ab|ij) - \frac{1}{2}(aj|ib)\right] \tag{4.13}$$

The effective one-particle operator restricts the maximum excitation level between two configurations to one (Table CI cases 11-14) and only the effective one-electron integrals in eq. 4.6 - 4.10 are involved. The matrix elements generally can be computed as a sum over the one-electron Fock operators $\hat{f}_l$ and the corresponding matrix elements that connect the two configurations $|\lambda\rangle$ and $|\mu\rangle$ via the interacting orbitals $\langle\phi_{|\lambda\rangle}^{\mathrm{a}}|$ denoted as a and $|\phi_{|\mu\rangle}^{\mathrm{b}}\rangle$ which is denoted as b.

$$\langle\lambda|\hat{H}_0|\mu\rangle = \langle\lambda|\hat{F}|\mu\rangle \tag{4.14}$$

$$= \langle\lambda|\sum_l^{\#\mathrm{elec}}\hat{f}_l|\mu\rangle \tag{4.15}$$

The application of eq. 4.15 on eq. 4.13 allows to derive explicit formulas for the matrix elements.

1. Identical configurations (diagonal elements, Table CI case no. 14)

$$H_{0,|\lambda\rangle,|\mu\rangle} = \left[\sum_a^{MO\in|\lambda\rangle}\left(h_{aa} + \sum_{ij}D_{ij}\left[(aa|ij) - \tfrac{1}{2}(aj|ia)\right]\right)\right]\cdot\mathbf{1} \qquad (4.16)$$

2. Single excitations (Table CI cases 11 to 13)

$$H_{0,|\lambda\rangle,|\mu\rangle} = \left[\left(h_{ab} + \sum_{ij}D_{ij}\left[(ab|ij) - \tfrac{1}{2}(aj|ib)\right]\right)\right]\cdot\mathbf{U_S^{n_s}}\left(\hat{P}_q^s\right) \qquad (4.17)$$

Matrix elements between identical configurations are evaluated as a sum of the Fock matrix elements over all occupied orbitals multiplied with the identity matrix. The major part of matrix elements between two configurations that differ in one orbital are expressed by a single term connecting these two orbitals multiplied with the representation matrix of the Coulomb part.

4.3 Levels of Configuration Treatment

Summarizing the preceding section, the evaluation of individual matrix elements of a Hamilton matrix in this work bases on Configuration State Functions (CSF). The CSF are generated from configurations, which simply denote orbital occupation schemes and represent the spatial part of the CSF. They are obtained by multiplication of the integrals connected to the configuration with appropriate spin eigenfunctions of the $\hat{S}^2$ and the $\hat{S}_z$ operators. The construction of all CSF for a given multiplicity and configuration leads to a set of CSF. The interaction between two sets of CSF representing a block of matrix elements could be regarded as the interaction between the configurations the CSF are generated from and a corresponding block of coupling coefficients betweeen the spin parts. The configurations are created by excitation operators acting onto the configurations of the reference wavefunctions. The classification of two interacting configurations and the evaluation of the spin-coupling matrices are crucial for an efficient implementation.

4.3.1 Internal-External Separation

A full evaluation of the Hamilton matrix requires the comparison of all possible configurations generated from the reference space. The number of comparisons then scales quadrat-

ically with the number of configurations. Only configurations that are connected to each other by less than a double excitation in the case of $\hat{H}_0$ and less than a triple excitation for $\hat{H}$ contribute non-zero matrix elements to the matrix.

For reasonable basis sets the number of external orbitals usually exceeds the number of occupied orbitals in the reference space by about an order of magnitude. Bunches of configurations can be sorted into groups that only differ in the external orbitals but possess a common rest of orbitals occupied in the reference space. Once sets of configurations are sorted in this way the common rest with internal orbitals of a configuration denoted as *internal* rest can be separated from the *external* part. A set of configurations can then be constructed by appending a series of external configuration rests to an internal part. The internal-external separation of configurations allows a prelimimary classification of the excitation scheme on the much smaller internal level and decides about whether a block of matrix elements contributes non-zero elements. If pairs of internal configurations differ by a double excitation for $\hat{H}_0$ or a higher excitation for $\hat{H}$, the complete block of all external configurations belonging to an internal rests yields matrix elements of zero and the external interactions need not to be evaluated explicitely. This separation reduces the quadratic dependency of the number of configurations.

The internal configurations are created by application of creation and annihilation operators on the reference configurations and the external configurations by creation operators on the vacuum.

Excitation level	Configuration	Space
0	$|\lambda_{Ref}\rangle$	[internal]
1	$a^*|\lambda_{Ref}\rangle$	[internal-1]
2	$a^*b^*|\lambda_{Ref}\rangle$	[internal-2]
$\vdots$	$\vdots$	

$a,b \in \mathrm{MO_{int}}$ and $k,l \in \mathrm{MO_{int}} + \mathrm{MO_{ext}}$

The internal-external separation was used in the DIESEL[53,54] package and was shown that the quadratic scaling reduces to approximately 1.5 to 1.6 depending on the reference space. For large-size molecules or large reference spaces with many occupied orbitals in the reference configurations the number of [internal-x] configuration grows quite large and the cost to compare the internal rests gets significant. This separation scheme can be

improved by an additional separation of the internal space that further reduces the number of compared configurations.

4.3.2 Inactive-Active-External Separation

A partitioning scheme which is superior to the internal-external separation is obtained if all orbitals doubly occupied in all reference configurations are separated from those with possess varying occupation numbers. The part of the configuration with all common orbitals doubly occupied will be denoted as *inactive*, while the remaining rest will be denoted as *active*.[5] For a set of arbitrary reference configurations this is illustrated in the next scheme.

Reference Configuration	Inactive-Part	Active-Part
$\lvert a^2b^2c^2d^2e^2f^2i^2j^2\rangle$		$\lvert i^2j^2\rangle$
$\lvert a^2b^2c^2d^2e^2f^2j^2k^2\rangle$		$\lvert j^2k^2\rangle$
$\lvert a^2b^2c^2d^2e^2f^2i^2j^1k^1\rangle$	$\lvert a^2-f^2\rangle$	$\lvert i^2j^1k^1\rangle$
$\lvert a^2b^2c^2d^2e^2f^2i^1j^2k^1\rangle$		$\lvert i^1j^2k^1\rangle$
$\lvert a^2b^2c^2d^2e^2f^2i^1j^1k^2\rangle$		$\lvert i^1j^1k^2\rangle$

The set of inactive configurations needed to construct the interacting space ([inactive-x], x=0,1,2) are created by application of annihilation operators to an inactive configuration $\lvert\lambda_{inactive}\rangle$ which is common for all reference configurations.

Excitation level	Configuration	Space
0	$\lvert\lambda_{inactive}\rangle$	[inactive]
1	$a^*\lvert\lambda_{inactive}\rangle$	[inactive-1]
2	$a^*b^*\lvert\lambda_{inactive}\rangle$	[inactive-2]
$\vdots$	$\vdots$	

$a,b \in \mathrm{MO}_{inactive}$

The generated [inactive] subsets can either be represented in the natural way of occupied orbitals for which the size of the configuration depends on the number of occupied orbitals or in terms of holes limited to a number up to the maximum allowed excitation order. The cost to compare two configurations scales linearly with number of occupied orbitals and the representation in terms of holes decreases the computational cost substantially. This

is proven by some example computations given in Table 4.2 which show that the effort for computing the inactive interactions scales much more favorable with the number of inactive orbitals in the two compared inactive configurations if they are implemented in terms of holes. The computational time still scales quadratically with the number of [inactive-x] configurations, if all [inactive-x] interactions are evaluated. Further improvements that reduce the number of configuration comparisons are presented later.

Table 4.2: Efficiency of computing [inactive-x]/[inactive-x'] interactions in particle or hole representation for a set of four molecules A to B taken from a simple comparison of the interactions in a $\mathbf{H_0}$ matrix. Timings were taken on a Dual Athlon MP 1200 MHz compiled with Intel C++ V 7.1 and MPICH 1.2.5.

Molecule	#inactive MO	#[inactive-1]	#[inactive-2]	t(Particle)[s]	t(Hole)[s]	Factor
A	22	22	231	38	7	5.4
B	35	35	595	340	41	8.3
C	51	51	1275	2100	180	11.7
D	83	83	3403	22000	1300	16.9

The configuration rests belonging to the active space are created in the same way as the internal configuration rests in the previous section. In the following scheme a and x are creation and annihilation operators acting on the orbitals in the active space. The active part of the reference configurations will be denoted as *Ref*.

[active±x']	Configuration
⋮	⋮
+2	$\prod_{i=1}^{n} \prod_{k=1}^{n+2} a_i^* x_k^\dagger |\lambda_{Ref}\rangle$
+1	$\prod_{i=1}^{n} \prod_{k=1}^{n+1} a_i^* x_k^\dagger |\lambda_{Ref}\rangle$
0	$\prod_{i=1}^{n} \prod_{k=1}^{n} a_i^* x_k^\dagger |\lambda_{Ref}\rangle$
-1	$\prod_{i=1}^{n} \prod_{k=1}^{n-1} a_i^* x_k^\dagger |\lambda_{Ref}\rangle$
-2	$\prod_{i=1}^{n} \prod_{k=1}^{n} -2 a_i^* x_k^\dagger |\lambda_{Ref}\rangle$
⋮	⋮

The created active and inactive configurations are not stored in an arbitrary manner but sorted into sets that obey the rules to create adjacent entries of similarly structured configurations.

1. The configuration rests are divided into subsets according to the number of electrons and the resulting irreducible representation.

2. Inside such a subset, configurations with less open shells precede the ones with more open shells.

3. Within this subset, the next sorting criterion is the position of the open-shell orbitals.

4. Finally, configurations with identical open-shell orbitals are sorted according to the position of the closed-shell orbitals.

The configurations spanning the internal space of a MR-CI-SD and MR-MP2 are obtained by appropriate combinations of inactive ([inactive-x']) and active ([active±y']) configuration rests given in Table 4.3. The construction of the internal rests imposes the restriction that all generated internal configurations do not exceed an excitation order n ($n = 2$ for MR-MP2 and MR-CI-SD) with respect to the reference configurations.

Table 4.3: Possible combinations of [inactive-x'] and [active±y'] restricted to the maximum of a double excitation between the created configurations and the reference space. [ref] denotes the set of active parts of the reference configurations.

[internal-x]	[inactive-x']/[active±y']
[internal]	[inactive] / [active]^a
	[inactive − 1] / [active + 1]^b
	[inactive − 2] / [ref + 2]
[internal-1]	[inactive] / [active − 1]
	[inactive − 1] / [active]^b
	[inactive − 2] / [ref + 1]
[internal-2]	[inactive] / [ref − 2]
	[inactive − 1] / [ref − 1]
	[inactive − 2] / [ref]

^{a} [active] $\not\subset$ [ref] in MR-MP2; including single
     and double excitations from [ref].
^{b} [active] includes [ref] and single excitations
     inside the active reference space.

The splitting of the internal rest into an inactive and an active part allows an efficient evaluation of the interactions for the internal rests. On the level of the sets of [inactive-x] configurations, interactions can be classified quite simply. If the interactions on this level already exceed the maximum excitation order then the complete block of active rests and the corresponding external configurations for each of the active parts yields only zero matrix-elements. The selection of surviving interactions already at the inactive-level reduces the computation cost and improves the scaling significantly.

In the *inactive-active* separation scheme, the usually larger inactive configuration rest is split off and the efficiency concerning the comparison between two active configurations is improved compared to internal rests.

In the same way as the internal-external separation generates sets of external configurations that are appended to the internal rests, the *inactive-active* separation introduces a similar but even finer structure of the configuration space similarly. Each [inactive-x] configuration is appended by a particular set of active rests to yield a certain category of [internal-y] configurations (Table 4.3), with y denoting the number of missing electrons. The recurrence of the active rests for all existing inactive configurations in a certain inactive set can be utilized to avoid the computation of each interaction in the active space as the sets of active configurations with other active parts form interaction patterns that repeat for each inactive interaction. These patterns have to be computed once for a particular set and can be used throughout the whole calculation reducing the effort of configuration comparisons to simple input-output procedures during the major part of computation in the iterative steps.

4.3.3   Patterns

The recurring sequences of core and active interactions are treated with patterns for which the interactions at a certain level (inactive, active, or external) are precomputed and the non-vanishing entries are stored. The storage scheme includes the difference configuration as it enters the determination of the table case and therefore the representation matrices in the SGA. It basically represents the excitation order between the two interacting configurations as well as the interacting orbitals that determine the integrals which enter the computation of matrix elements. The excitation order of a configuration rest is a criterion on the existence

and non-existence of non-zero matrix elements and also influences the interacting orbitals. In this scheme the excitation order and the interacting orbitals are computed with each difference configuration and stored in the pattern.

The patterns are used directly to incrementally build up the full interaction between two configurations from the core, active, and external interactions. The implementation is based on difference configurations (see Section 4.2.1) which are the basic objects for the computation of matrix elements.

One difficulty in the use of patterns comes along with addressing of the coefficient vector needed for the matrix-vector multiplication. In a full evaluation of *e.g.* active interactions, the interaction is classified and once it vanishes, the new offset for the coefficient address is computed from the active interaction and the corresponding external rests. The number of open shells determines the manifold of the spin eigenfunctions and is the quantity needed to address the coefficient vector correctly. In the patterns discussed above the information about the incremental offsets connected with the vanishing interactions cannot be stored, as the offsets do not solely depend on the entries in the pattern (*e.g.* interactions between active configurations) but also on the external rests that are missing in the pattern. The recovery of this information is implemented with configuration indices for all configurations occuring in the pattern and their number of open shells. The position entry is necessary to compute the "distance" between two pattern elements and the number of open shells for each of the neglected entries in combination with the number of corresponding external configurations is used to compute the address of the coefficient vector. The computational effort to obtain the new offsets is negligible compared to the computation of a difference configuration and increases the efficiency of the algorithm compared to a full treatment.

4.4 Matrix Structure of H_0 and Inhomogenity

4.4.1 Inhomogenity

The inhomogenity vector $\vec{I}$ represents the projection of the diagonalized reference wavefunction $|\Phi_0\rangle$ onto the interaction space (left-hand side of Figure 4.3). The columns in Figure 4.3 denote the configurations of the multi-configurational reference wavefunction ($|\Phi_0^i\rangle$) while

the rows represent the configurations of the interacting space. The matrix elements between two individual interacting configurations are evaluated by means of the SGA. The numbers of comparisons between two configurations grow rapidly for increasing interacting spaces and reference wavefunctions. The reference configurations determine the size of the interacting space in conjunction with the number of external MO and the number of comparisons is $N_{|\Phi_k^1\rangle} \times N_{|\Phi_0^i\rangle}^{ref}$ if no advantage of the structure is taken[1].
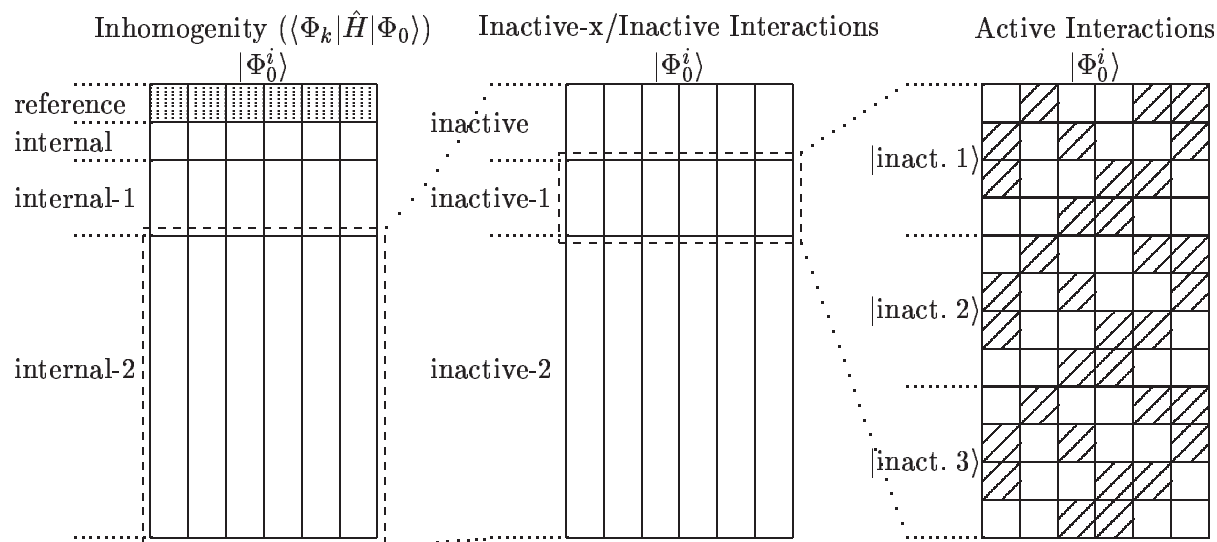
Figure 4.3: Structure of the MR-MP inhomogenity vector. The left-hand side represents an internal-external separation of the inhomogenity. The space of the reference configurations is excluded from the interacting space. The inactive-active partitioning of the internal regions (internal-2 depicted in the center) further refines the interaction classes. The right-hand side depicts the interactions of the active rests for three arbitrary inactive parts. Vanishing blocks of matrix elements are denoted in black.

An efficient scheme to reduce the computational cost of the evaluation of $\vec{I}$ first utilizes the internal-external separation already introduced at the beginning of this chapter. This first level is classified according to the number of external MO in the configurations of the interacting space and groups configurations together that possess a similar structure. Each internal region is further divided into a inactive regime and the corresponding active part of the respective internal rests (center of Figure 4.3). Each inactive configuration in a particular inactive regime (here: [inactive-1]) is appended with a set of active rests (right-

---

[1]For a MR-MP2 corrected selecting MR-CI the computation of the inhomogenity is expensive. The computational cost increases with the size of the selected CI wavefunction as it represents the reference wavefunction for the inhomogenity.[54]

hand of Figure 4.3) that finally corresponds to a set of internal configurations. To exploit the splitting into a inactive and an active part, the internal rests are sorted according to identical inactive parts which generates a repeating sequence of active configuration rests. In the cases of [internal-1] and [internal-2] spaces, a set of external configurations is appended to the internal parts.

The use of matrix patterns presented in Section 4.3.3 for the classification of interactions on the active level allows to determine those entries with vanishing contributions at a very early level. The time-consuming comparison of configurations is thus avoided. This aspect, that plays a minor role in case of [internal] configurations, becomes significant once whole columns of external configurations are appended to each [internal-x] rest and vanishing matrix elements are recognized already on the active level of configuration treatment.

4.4.2   The MR-MP-H_0 matrix

4.4.2.1   General Structure

The hermitian $\mathbf{H_0}$ matrix (Figure 4.4) encounters the interacting configurations generated from the reference space. It can be divided by means of the internal-external separation into three regions (Figure 4.4, I) characterized by the number of annihilated electrons x denoted as [internal-x] and the corresponding external configurations. The reference space is excluded from the interacting space. The regions inside the $\mathbf{H_0}$ matrix are listed in Table 4.4. In the $\mathbf{H_0}$ matrix interactions between two configurations connected to each other by more than a single excitation vanish, since in contrast to $\hat{H}$ of the inhomogenity, $\hat{H}_0$ represents an effective one-electron operator. Consequently, the matrix block [internal]/[internal-2] (C in Figure 4.4, I and in Table 4.4) is zero by virtue of this property. In the following discussion only the upper triangular part of the hermitian submatrices are regarded.[2].

Each internal region is further split up (Figure 4.4, II) into 3×3 inactive regions ([inactive], [inactive-1] and [inactive-2] interacting with each other). The submatrices of the inactive configuration sets are symmetric for internal regions with the same number of electrons (A, D, and F) but non-symmetric if the numbers of electrons differ in the internal rests. The

---

[2]The multiplication of the matrix elements of the lower and the upper triangle with the corresponding parts of the coefficient vector are executed simultaneously.
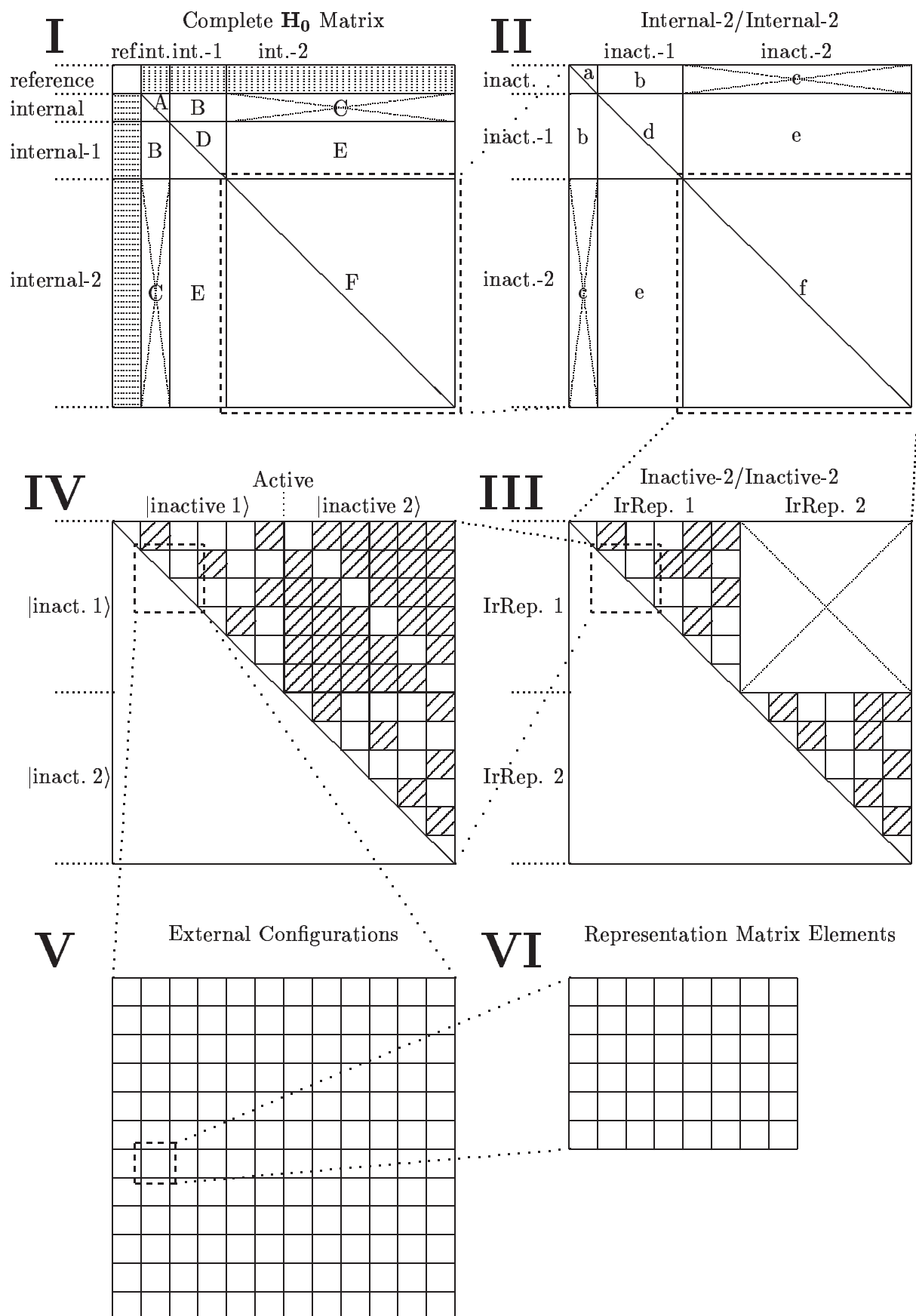
Figure 4.4: Structure of the MR-MP-$\mathbf{H_0}$ matrix.

Table 4.4: Regions in the MR-MP-$\mathbf{H_0}$ matrix according to Figure 4.4 (I).

Region	[Internal-x]	[Internal-x']	External	External'	minimum excitation level
A	Internal	Internal	-	-	0
B	Internal	Internal-1	-	$\lvert a^1 \rangle$	1
C	Internal	Internal-2	-	$\lvert a^2 \rangle$, $\lvert a^1 b^1 \rangle$	2
D	Internal-1	Internal-1	$\lvert a^1 \rangle$	$\lvert a^1 \rangle$	0
E	Internal-1	Internal-2	$\lvert a^1 \rangle$	$\lvert a^2 \rangle$, $\lvert a^1 b^1 \rangle$	1
F	Internal-2	Internal-2	$\lvert a^2 \rangle$, $\lvert a^1 b^1 \rangle$	$\lvert a^2 \rangle$, $\lvert a^1 b^1 \rangle$	0

analyis of interactions of large parts of the matrix on the coarse-grain inactive level is the basis for an efficient evaluation of the finer interaction structures of the active and external configurations in the sparse $\mathbf{H_0}$ matrix.

Table 4.5: Structure of the matrix blocks inside the MR-MP $\mathbf{H_0}$ matrix. Identical internal configurations are excluded.
[a] Not treated in stored patterns but implemented directly with orbitals.

[Internal-y]	[Internal-x]	[Inact.-y']	[Inact.-x']	Inact.-Struct.	Active-Struct.	Ext. Struct.
Internal	Internal	Inact.	Inact.	Single	Pattern	-
Internal	Internal	Inact.	Inact.-1	Row	Pattern	-
Internal	Internal	Inact.	Inact.-2	-	-	-
Internal	Internal	Inact.-1	Inact.-1	All	Pattern/Diag.	-
Internal	Internal	Inact.-1	Inact.-2	Pattern	Pattern	-
Internal	Internal	Inact.-2	Inact.-2	Pattern	Pattern/Diag.	-
Internal	Internal-1	Inact.	Inact.	Single	Pattern	Row
Internal	Internal-1	Inact.	Inact.-1	Row	Pattern	Row
Internal	Internal-1	Inact.	Inact.-2	-	-	-
Internal	Internal-1	Inact.-1	Inact.	-	-	-
Internal	Internal-1	Inact.-1	Inact.-1	Diag.	Pattern	Row
Internal	Internal-1	Inact.-1	Inact.-2	Pattern	Diag.	Row
Internal	Internal-1	Inact.-2	Inact.	-	-	-
Internal	Internal-1	Inact.-2	Inact.-1	-	-	-
Internal	Internal-1	Inact.-2	Inact.-2	Pattern	Diag.	Row
Internal-1	Internal-1	Inact.	Inact.	Single	Pattern	Diag.

				Row	Pattern	Diag.
Internal-1	Internal-1	Inact.	Inact.-1	Row	Pattern	Diag.
Internal-1	Internal-1	Inact.	Inact.-2	-	-	-
Internal-1	Internal-1	Inact.-1	Inact.-1	All	Pattern/Diag.	Diag.
Internal-1	Internal-1	Inact.-1	Inact.-2	Pattern	Pattern	Diag.
Internal-1	Internal-1	Inact.-2	Inact.-2	Pattern	Pattern/Diag.	Diag.
Internal-1	Internal-2	Inact.	Inact.	Single	Pattern	Pattern[a]
Internal-1	Internal-2	Inact.	Inact.-1	Row	Pattern	Pattern[a]
Internal-1	Internal-2	Inact.	Inact.-2	-	-	-
Internal-1	Internal-2	Inact.-1	Inact.	-	-	-
Internal-1	Internal-2	Inact.-1	Inact.-1	Diag.	Pattern	Pattern[a]
Internal-1	Internal-2	Inact.-1	Inact.-2	Pattern	Pattern	Pattern[a]
Internal-1	Internal-2	Inact.-2	Inact.	-	-	-
Internal-1	Internal-2	Inact.-2	Inact.-1	-	-	-
Internal-1	Internal-2	Inact.-2	Inact.-2	Diag.	Pattern	Pattern[a]
Internal-2	Internal-2	Inact.	Inact.	Single	Pattern	Diag.
Internal-2	Internal-2	Inact.	Inact.-1	Row	Pattern	Diag.
Internal-2	Internal-2	Inact.	Inact.-2	-	-	-
Internal-2	Internal-2	Inact.-1	Inact.-1	All	Pattern/Diag.	Diag.
Internal-2	Internal-2	Inact.-1	Inact.-2	Pattern	Pattern	Diag.
Internal-2	Internal-2	Inact.-2	Inact.-2	Pattern	Pattern/Diag.	Diag.

Generally, all interactions between two inactive configurations that correspond to a higher excitation than a single excitation give zero matrix elements ([inactive]/[inactive-2], block c in Figure 4.4, II). The inactive-active separation possesses an additional useful feature that allows a simple decision which sets of inactive configurations interact to contribute non-zero matrix elements. Do both sets of inactive configurations are used to build up configurations of same internal spaces (internal, internal-1, and internal-2), all interactions have a possibility to exist. If the inactive sets are associated with different internal regions (B and E in Figure 4.4, I), the number of electrons in the inactive configurations decides about the existence of whole matrix blocks. Two inactive sets constituting the rows ([inactive-y]) and the columns ([inactive-x]) of a certain submatrix can only contribute to non-zero matrix elements if the number of electrons in [inactive-y] is larger or equal the number of electrons in [inactive-x] (y≤x). These blocks in regions B an E (Figure 4.4, I) correspond to the labels a, b, d, e, and f in the upper triangle of Figure 4.4, II. If y>x (less electrons in [inactive-

y] than in [inactive-x]) the complete block including active and external configuration rest vanishes (b and c in the lower triangle of Figure 4.4, II; compare also with Table 4.5).

A simple example to illustrate this feature is given in equation 4.18 for the interaction of two configurations of the type [internal]/[internal-1]. The configuration of [internal] is constructed from [inactive-2][active+2] and the one of [internal-1] of [inactive-1][active][external]. Orbitals up to c are inactive orbitals, m, n and o denote active orbitals and x denotes an external orbital appended to the [internal-1] configurations.

$$\langle \underbrace{a^2 b^1 c^1}_{inactive} \underbrace{m^2 n^2 o^2}_{active} |\hat{H}_0| \underbrace{a^2 b^2 c^1}_{inactive} \underbrace{m^2 n^1 o^1}_{active} \underbrace{x^1}_{ext.} \rangle \;\; = \;\; \langle b^1 n^2 o^2 |\hat{H}_0| b^2 n^1 o^1 x^1 \rangle \tag{4.18}$$

The active rests that are necessary to build up a configuration of certain internal character differ by at least a double excitation and therefore the whole inactive-blocks vanishes.

The structure of the matrix was so far created by virtue of the properties of the inactive-active partitioning. If the molecule of interest possesses symmetry, the symmetry properties of the configuration rests have to be considered (Figure 4.4, III). Each inactive region of a certain internal submatrix (Figure 4.4, II) is further blocked according to the irreducible representations of the point group.

Two interacting inactive configurations belonging to different irreducible representations are connected by at least a single excitation because the interacting orbitals possess a different symmetry. The condition for the full configurations (inactive+active+external part) to be of identical symmetry forces the active or external parts to belong to different irreducible representations to compensate for the different inactive symmetries. The excitation order between the two active or external rests again has to at least one. Therefore, blocks of inactive configuration sets that belong to different irreducible representations (IrRep. 1/IrRep. 2 in Figure 4.4, III) vanish.

Up to the present considerations, the $\mathbf{H_0}$ matrix is blocked by the internal-external separation, the inactive-active partitioning and the symmetry considerations at the inactive-level. The next level (Figure 4.4, IV) which introduces an even finer structure considers sets of inactive configurations which belong to the same irreducible representation of the point group. The interaction between configuration rests without an excitation out of the inactive orbitals ([inactive]/[inactive]) is trivial as the [inactive] set only consists of a single configuration and belongs to the totally symmetric irreducible representation of the point group.

Does an [inactive-1] set interacts with itself (block d in Figure 4.4, II), all interactions between configurations of this set contribute to non-vanishing contributions since they can only differ by a single excitation and therefore the interaction depends on the active and/or external configuration rests. Interactions of two sets of the type [inactive-2]/[inactive-2] (block d in Figure 4.4, II) are depicted schematically in Figure 4.4, III. For each of the interactions between two inactive configurations an appropriate interaction scheme exists. If two identical inactive configurations interact, the final contribution is determined by the active and external rests. Only active configurations have to be considered with less than a double excitation between each other. Active rests connected by a single excitation have to possess identical external rests. Only if the active rests are identical the external rests can differ in one MO which then about the interaction.

If two inactive configurations are connected by a single excitation, non-zero matrix elements are only found if the active and the external rests that belong to the respective configurations are identical. This classification of interactions already reduces the effort to evaluate the H_0 matrix but the efficiency still can be improved especially if large molecules are considered. The number of [inactive-1] configurations scales linearly and the number of [inactive-2] scales quadratically with the number of occupied MO. A substantial saving is achieved if the patterns of [inactive-2]/[inactive-2] interactions are precomputed prior to the evaluation of the H_0 matrix and used to access the inactive matrix blocks directly, so that it is possible to avoid any cost arising from non-interacting inactive configurations. The same procedure is also done for the active interactions in analogy to the computation of the inhomogenity. The previously mentioned possible diagonal structure of blocks of active configurations is a special case of a pattern structure and can be easily coded in a general way.

The interactions discussed so far take place at the internal level. The external interactions for each pair of interacting internal configuration rests (Figure 4.4, V) represent the last level before the computation of the matrix elements (Figure 4.4, VI).

The interactions on the external configuration level depend on the excitation order between the internal configuration rests. If the number of electrons in the internal rests (inactive and active) are the same and the internal excitation order equals one, the non-vanishing contribution of the external configurations is reduced to a diagonal in the external matrix

block. This kind of external interaction contributes the major part of the sparsity of the
Hamilton matrix and the spatial part of the matrix elements is determined by the internal rest alone. Two other kinds of external submatrices are less trivial to evaluate. The
external interactions for identical internal configurations allow single excitations in the external configurations and cause a more dense interaction pattern. Matrix blocks belonging
to [internal-1]/[internal-2] are also more difficult to evaluate as a more complicated structure
exists (see Section 4.4.2.2). The use of patterns to evaluate these external interactions efficiently is disadvantageous as the number of externals rapidly grows with the basis set size.
For this reason a direct implementation of the external interactions was preferred. For each
internal interaction a block of representation matrix elements exists that finally constitutes
the matrix elements of the $\mathbf{H_0}$ matrix.

4.4.2.2   Direct Implementation of External Configuration Interactions

The interactions between external configurations contribute the major part of the computational work in the construction of the $\mathbf{H_0}$ matrix. The occuring interactions are more
restricted than for the active level and of the same complexity as the inactive interactions.
If two internal configuration parts are connected to each other by a single excitation, both
external configuration rests have to be the same to yield non-vanishing matrix elements.
In a matrix representation these configurations represent diagonal entries in the external
submatrix. Two identical internal configuration rests allow the external configurations to
differ by a single excitation to yield non-vanishing matrix elements. Basically, this would
require the full evaluation of all interactions between each pair of external configurations.
This situation holds for matrix blocks that consist of configurations with only one external orbital, but not for interacting external configurations rests with two occupied external
orbitals. For these external rests the number of double excitations exceeds the number of
single excitations and a full evaluation reduces the computational efficiency. An analogous
situation is also found for interactions between the [internal-1] and [internal-2] configuration
spaces. The two external configuration rests, one with a single external orbital and the other
with two occupied external orbitals, can only contribute to non-vanishing matrix elements
if one orbital is identical in both rests.

   The patterns as they are used for the inactive and active levels to locate the positions of the

interacting configurations are not preferrably applied since the required prototypes of the external submatrices commonly extend the storage limits in main memory. If the computed molecule belongs to a higher-order point group than C_1, patterns for each combination of the irreducible representation had to be evaluated and efficiently accessed. An approach to avoid the explicit computation of all interactions or to apply the memory-consuming alternative of patterns is the direct implementation of the non-vanishing interactions in the external submatrix. This approach certainly utilizes the limited possibilities of interactions (one or two occupied external orbitals) present in a MR-MP2 ansatz.

	a^2	b^2	c^2	d^2	e^2	f^2	ab	ac	ad	ae	af	bc	bd	be	bf	cd	ce	cf	de	df	ef
a^2							$\binom{a}{b}$	$\binom{a}{c}$	$\binom{a}{d}$	$\binom{a}{e}$	$\binom{a}{f}$										
b^2							$\binom{b}{a}$					$\binom{b}{c}$	$\binom{b}{d}$	$\binom{b}{e}$	$\binom{b}{f}$						
c^2								$\binom{c}{a}$				$\binom{c}{b}$				$\binom{c}{d}$	$\binom{c}{e}$	$\binom{c}{f}$			
d^2									$\binom{d}{a}$				$\binom{d}{b}$			$\binom{d}{c}$			$\binom{d}{e}$	$\binom{d}{f}$	
e^2										$\binom{e}{a}$				$\binom{e}{b}$			$\binom{e}{c}$		$\binom{e}{d}$		$\binom{e}{f}$
f^2											$\binom{f}{a}$				$\binom{f}{b}$			$\binom{f}{c}$		$\binom{f}{d}$	$\binom{f}{e}$
ab								$\binom{b}{c}$	$\binom{b}{d}$	$\binom{b}{e}$	$\binom{b}{f}$	$\binom{a}{c}$	$\binom{a}{d}$	$\binom{a}{e}$	$\binom{a}{f}$						
ac									$\binom{c}{d}$	$\binom{c}{e}$	$\binom{c}{f}$	$\binom{a}{b}$				$\binom{a}{d}$	$\binom{a}{e}$	$\binom{a}{f}$			
ad										$\binom{d}{e}$	$\binom{d}{f}$		$\binom{a}{b}$			$\binom{a}{c}$			$\binom{a}{e}$	$\binom{a}{f}$	
ae											$\binom{e}{f}$			$\binom{a}{b}$			$\binom{a}{c}$		$\binom{a}{d}$		$\binom{a}{f}$
af															$\binom{a}{b}$			$\binom{a}{c}$		$\binom{a}{d}$	$\binom{a}{e}$
bc													$\binom{c}{d}$	$\binom{c}{e}$	$\binom{c}{f}$	$\binom{b}{d}$	$\binom{b}{e}$	$\binom{b}{f}$			
bd														$\binom{d}{e}$	$\binom{d}{f}$	$\binom{b}{c}$			$\binom{b}{e}$	$\binom{b}{f}$	
be															$\binom{e}{f}$		$\binom{b}{c}$		$\binom{b}{d}$		$\binom{b}{f}$
bf																		$\binom{b}{c}$		$\binom{b}{d}$	$\binom{b}{e}$
cd																	$\binom{d}{e}$	$\binom{d}{f}$	$\binom{c}{e}$	$\binom{c}{f}$	
ce																		$\binom{e}{f}$	$\binom{c}{d}$		$\binom{c}{f}$
cf																			$\binom{c}{d}$	$\binom{c}{e}$	
de																				$\binom{e}{f}$	$\binom{d}{f}$
df																					$\binom{d}{e}$
ef																					

Figure 4.5: Interactions between two-electron external configurations in the totally symmetric irreducible representation occuring for identical internal configurations. a-f denote external orbitals.

The direct implementation of the external contributions is supposed to come along with a minimal computational effort that only addresses those configurations that actually interact but requires the explicit classification of the occurring interactions. This classification also has to consider the various irreducible representations and can be done once the configu-

rations are sorted by their external orbitals in a particular but non-unique scheme. In this work, configurations that consist of only one external orbital are sorted with respect to the irreducible representation they belong to and the orbital number while configurations with two occupied external orbitals are sorted in a manner presented in the first line given in Figure 4.5. The ordered set starts with closed-shell configurations followed by the ones with two singly occupied external orbitals. The submatrix in Figure 4.5 represents the external interactions in the totally symmetric irreducible representation for two identical internal rests. These blocks are located along the main diagonal of the hermitian $\mathbf{H_0}$ matrix and thus only the upper triangle part is considered.

This submatrix starts with the block of the external configurations a^2 to f^2 that do not interact with each other because they represent a double excitation. However, they interact with the open-shell configurations a^1b^1 (denoted ab) to ef. The open-shell configurations are ordered in the following scheme. First the external orbital at the beginning of an irreducible representation (a) is set, then the remaining higher orbitals (b to f) in this representation are added to build the configurations and the sequence is continued with the next orbital b.

Based on this ordering the interactions of the closed-shell configuration on the left-hand side ($|\Phi_L\rangle$) and the open-shell configurations on the right-hand side ($|\Phi_R\rangle$) are reduced to two simple rules. If the orbital of the closed-shell configuration $|\Phi_L\rangle$ and the first orbital in the open-shell configuration $|\Phi_R\rangle$ is the same then a set of contiguous matrix elements in one row is present ($b^2/(bc$, bd, be, and bf in the second line of Figure 4.5). If the second orbital of the open-shell configuration $|\Phi_R\rangle$ and the MO in $|\Phi_L\rangle$ are identical, then a single entry results (b^2/ab). The entries represent the interacting orbitals that enter the matrix element computation with the orbital on the left-hand side superscripted to the orbital on the right-hand side. The number of elements in this part of the submatrix is computed to #ext.×(#ext.-1) An algorithm that determines the position of the respective elements is given in Listing 4.1.

The positions of the single entries depend on the position of the second orbital in the open-shell configuration. The first loop iterates over all irreducible representations of the point group (not shown in Figure 4.5 for the sake of simplicity) to encompass all combinations of orbitals that contribute to the totally symmetric irreducible representation. For each MO on the left-hand side (moL_1) the corresponding row is now treated. First the matrix

elements are computed for the current position (starts at position 0 for the first entry) and then the number of single entries is determined from the position of moL_1 alone.

Listing 4.1: Part I of the algorithm to compute the interacting between two-electron external configurations in the totally symmetric irreducible representation occuring for identical internal configurations. Closed-shell configurations in $|\Phi_L\rangle$.

```
// 1. closed shell left/ 2 open shells right
for (IrRep = group.begin(); IrRep < group.end(); IrRep++)
{  mos = get_external_mos(IrRep);
  // loop over closed-shells on left side;
  for (moL1 = mos.begin(); moL1 = mos.end(); moL1++)
  {  // Gaps preceed contiguous part in one row!
    number_of_gaps = pos(moL1) - 1;
    moR1 := mos.begin();
    moR2 = moL1;
    for (count = 0; count < number_of_gaps; count++)
    {  // calculate matrix elements from interacting MO L1, R1 and R2!
      MRMPHOMatElements(tablecase,interactingMO(L1,L2=L1,R1,R2));
      offset = tmp_offset + ( pos(moR2) - pos(moR1) - 1 ) * (spineigen);
      // calculate new offset
      tmp_offset += (mos.size() - pos(moR1)) * (spineigen);
      moR1++;
    }
    // create (e.g. |b2> -> |bc>)
    moR1 = moL1;
    moR2 = moR1 + 1;
    for (moR2; moR2 = mos.end(); moR2++)
    {  // Calculate matrix elements from interacting MO L1, L2, R1 and R2!
      MRMPHOMatElements(tablecase,interactingMO(L1,L2,R1,R2));
      // calculate new offset
      offset += (spineigen);
    }
  }
}
```

The offset for the next entry is now computed by the position of the MO moR_1 and moR_2 and the number of spin eigenfunctions (*spineigen*). In this loop also the starting position for the contiguous part of the row is computed. The interacting MO are set again and a loop from $moR_2 = moL_1 + 1$ to the highest MO treats the remaining interacting configurations. The number of non-vanishing entries per row in this submatrix equals the number or orbitals in the irreducible representation.

The structure of the interactions between two external open-shell configurations is also

shown in Figure 4.5. This part of the submatrix consists of small triangular blocks if the first orbitals in both configurations $|\Phi_L\rangle$ and $|\Phi_R\rangle$ are the same and a pattern in analogy to the closed-shell/open-shell case exists if these differ. Listing 4.2 shows the corresponding part of the algorithm.

Listing 4.2: Part II of the algorithm to compute the interacting between two-electron external configurations in the totally symmetric irreducible representation occuring for identical internal configurations. Open-shell configurations in $|\Phi_L\rangle$.

```
// 2. 2 open shell left/ 2 open shells right
// Get initial offset from previous part
for (IrRep = group.begin(); Irrep = group.end(); IrRep++)
{   mos = get_external_mos(IrRep);
  // set initial MO (e.g. |ab>)
  moL1 = mos.begin();
  moL2 = moL1;
  moL2++;
  // Build prototype difference configuration of 2 open shells / 2 open shells
  // Set configuration on the right: (e.g. |ac>)
  // Loop over open shells on left side;
  for (moL1 = mos.begin(); moL1 = mos.end(); moL1++)
  {   for (moL2 = moL1+1; moL2 = mos.end(); moL2++)
    {   // increment |ab> to |ac>
      offset += (spineigen);
      // Do contigous part of row before gaps->contary to closed/2 open
      for (moR2 = moR1+1) -> mos.end())
      {   // Calculate matrix elements from interacting MO L1, L2, R1 and R2!
        MRMPHOMatElements(tablecase,interactingMO(L1,L2,R1,R2));
        // calc new offset
        offset += (spineigen);
      }
      int number_of_gaps = pos(moL2) - pos(moL1) - 1;
      moR1 = moL1 + 1;
      moR2 = moL2 ;
      for (count = 0; count < number_of_gaps; count++)
      {
        // see previous procedure
      }
      // do contiguous part
      if (moL2 < lastMO) // if L2 is last mo -> no gapless block exists
      {   moR1 = moL2;
        moR2 = moR1;
        moR2++; // -> (build e.g. |bc> in row of |b2>
        for (moR2; moR2 = mos.end(); moR2++)
        {
          // as done previously
        }
```

```
        }
      }
    }
}
```

A pattern for external configurations in the other irreducible representation than the totally symmetric one can also be derived. Now all combinations between orbitals of the corresponding irreducible representation that result in the desired representation have to be treated. In the example in Figure 4.6 only orbitals of two irreducible representations are given for simplicity (orbitals $a - f$ and $i - l$).

	ai	aj	ak	al	bi	bj	bk	bl	ci	cj	ck	cl	di	dj	dk	dl	ei	ej	ek	el
ai		$\binom{i}{j}$	$\binom{i}{k}$	$\binom{i}{l}$	$\binom{a}{b}$				$\binom{a}{c}$				$\binom{a}{d}$				$\binom{a}{e}$			
aj			$\binom{j}{k}$	$\binom{j}{l}$		$\binom{a}{b}$				$\binom{a}{c}$				$\binom{a}{d}$				$\binom{a}{e}$		
ak				$\binom{k}{l}$			$\binom{a}{b}$				$\binom{a}{c}$				$\binom{a}{d}$				$\binom{a}{e}$	
al								$\binom{a}{b}$				$\binom{a}{c}$				$\binom{a}{d}$				$\binom{a}{e}$
bi						$\binom{i}{j}$	$\binom{i}{k}$	$\binom{i}{l}$	$\binom{b}{c}$				$\binom{b}{d}$				$\binom{b}{e}$			
bj							$\binom{j}{k}$	$\binom{j}{l}$		$\binom{b}{c}$				$\binom{b}{d}$				$\binom{b}{e}$		
bk								$\binom{k}{l}$			$\binom{b}{c}$				$\binom{b}{d}$				$\binom{b}{e}$	
bl												$\binom{b}{c}$				$\binom{b}{d}$				$\binom{b}{e}$
ci										$\binom{i}{j}$	$\binom{i}{k}$	$\binom{i}{l}$	$\binom{c}{d}$				$\binom{c}{e}$			
cj											$\binom{j}{k}$	$\binom{j}{l}$		$\binom{c}{d}$				$\binom{c}{e}$		
ck												$\binom{k}{l}$			$\binom{c}{d}$				$\binom{c}{e}$	
cl																$\binom{c}{d}$				$\binom{c}{e}$
di														$\binom{i}{j}$	$\binom{i}{k}$	$\binom{i}{l}$	$\binom{d}{e}$			
dj															$\binom{j}{k}$	$\binom{j}{l}$		$\binom{d}{e}$		
dk																$\binom{k}{l}$			$\binom{d}{e}$	
dl																				$\binom{d}{e}$
ei																		$\binom{i}{j}$	$\binom{i}{k}$	$\binom{i}{l}$
ej																			$\binom{j}{k}$	$\binom{j}{l}$
ek																				$\binom{k}{l}$
el																				

Figure 4.6: Matrix entries between two-electron external configurations in a non-totally symmetric irreducible representation occuring for identical internal configurations.

Triangular matrix blocks are present for identical first orbitals but once the first orbital differs, matrix entries only survive if they possess an identical second MO and a diagonal structure of the submatrix with identical interacting MO is found. The algorithm to compute the positions of the interacting external configurations is given in Listing 4.3.

Listing 4.3: Algorithm to compute the interaction between two-electron external configurations in a non-totally symmetric irreducible representation occuring for identical internal configurations.

```
for (IrRep1 = group.begin(); IrRep1 = group.end(); IrRep1++)
{  mosL = get_external_mos(IrRep1);
  IrRep2 = group.getProduct(externalsymm,IrRep1);
  if (IrRep2 > Irep11)
  {  mosR = get_external_mos(IrRep2);
    // do loop over 1st mo left // all in irrep1
    for (moL1 = mosL.begin(); moL1 = mosL.end(); moL1++)
    {  for (moL2 = mosr.begin(); moL2 = mosR.end(); moL2++)
        moR = mosL.begin();
        moR2 = moL2;
        // do line in triangle
        for (moR2; moR2 = mosR.end(); moR2++)
          // do sparse part!
          // distance between single elements is number of entries(ir2);
          // number is of entries(IrRep1) - pos(moL1)
          // Calculate matrix elements from interacting MO L1, L2, R1 and R2!
          MRMPHOMatElements(tablecase,interactingMO(L1,L2,R1,R2));
          // calc new offset
          offset += (spineigen);
        }
        for(gaps = pos(moL1)+1; gaps = mosL.size(); gaps++)
        {  moR1++;
          // calc new offset
          offset = tmp_offset + (pos(moL2)-1) *
                  (*spineigen)(internal_openshellsL + 2);
          tmp_offset += mosR.size() * (*spineigen)(internal_openshellsL + 2);
          // Calculate matrix elements from interacting MO L1, L2, R1 and R2!
          MRMPHOMatElements(tablecase,interactingMO(L1,L2,R1,R2));
        }
      }
    }
  }
}
```

If the external parts of two configurations belong to two different irreducible representations of the point group, the full configurations must differ by a double excitation. The first excitation arises from the different irreducible representation of two interacting external orbitals. The second excitation must occur in the internal part (inactive or active) since the symmetry of both full configurations must be the same. These external blocks of configurations vanish.

A slightly different situation arises if the external rests of the interacting configuration sets

[internal-1]/[internal-2] are regarded. For these external submatrices rectangular blocks exist depending on the possible combinations of irreducible representations. For the [internal-1]a^1/[internal-2]a^1b^1 part with a and b denoting external orbitals, three cases can be distinguished. The first case (Figure 4.7) summarizes the arising interactions, if the external configuration rests belong to the totally symmetric irreducible representation. The algorithm is identical to the one in Figure 4.5.

	a^2	b^2	c^2	d^2	e^2	ab	ac	ad	ae	bc	bd	be	cd	ce	de
a	(a)					(b)	(c)	(d)	(e)						
b		(b)				(a)				(c)	(d)	(e)			
c			(c)				(a)			(b)			(d)	(e)	
d				(d)				(a)			(b)		(c)		(e)
e					(e)				(a)			(b)		(c)	(d)

Figure 4.7: Matrix entries of external configurations in [internal-1]/[internal-2] matrix blocks. The external part of the [internal-2] set belongs to the totally symmetric irreducible representation.

	ai	aj	ak	al	bi	bj	bk	bl	ci	cj	ck	cl	di	dj	dk	dl	ei	ej	ek	el
a	(i)	(j)	(k)	(l)																
b					(i)	(j)	(k)	(l)												
c									(i)	(j)	(k)	(l)								
d													(i)	(j)	(k)	(l)				
e																	(i)	(j)	(k)	(l)

Figure 4.8: Matrix entries of external configurations in [internal-1]/[internal-2] matrix blocks. The external part of the [internal-2] configuration set belongs to the non-totally symmetric irreducible representation. The singly occupied MO in $|\Phi_L\rangle$ belong to the same irreducible representation as the first MO in $|\Phi_R\rangle$.

	ai	aj	ak	al	bi	bj	bk	bl	ci	cj	ck	cl	di	dj	dk	dl	ei	ej	ek	el
i	(a)				(b)				(c)				(d)				(e)			
j		(a)				(b)				(c)				(d)				(e)		
k			(a)				(b)				(c)				(d)				(e)	
l				(a)				(b)				(c)				(d)				(e)

Figure 4.9: Matrix entries of external configurations in [internal-1]/[internal-2] matrix blocks. The external part of the [internal-2] configuration set belongs to the non-totally symmetric irreducible representation. The singly occupied MO in $|\Phi_L\rangle$ belong to the same irreducible representation as the second MO in $|\Phi_R\rangle$.

The second case and the third case are depicted in Figures 4.8 and 4.9. The configurations $|\Phi_R\rangle$ belong to the non-totally symmetric irreducible representation and if the occupied orbital in the configuration on the left-hand side is identical to the first one in the configuration on the right-hand side ($|\Phi_R\rangle$ in Figure 4.8) then a band structure of the sparse submatrix exists. If the second orbital in $|\Phi_R\rangle$ is the same as the orbital in $|\Phi_L\rangle$, then the structure changes to a diagonal for each orbital block of the submatrix. The number of non-vanishing matrix entries is computed to $n^{MO(IrRep1)} \cdot n^{MO(IrRep2)}$. The algorithm for both cases is given in Listing 4.4

Listing 4.4: Algorithm to address non-vanishing matrix entries of external configurations in [internal-1]/[internal-2] matrix blocks. The [internal-2] set belongs to the non-totally symmetric irreducible representation.

```
// calc start of ext1/ext2
mosL = get_external_mos(extSymmL);
moL1 = mosL.begin();

for (IrRep1 = group.begin(); IrRep1 = group.end(); IrRep1++)
{  IrRep2 = group.getProduct(extSymmL,IrRep1);

  // one set of mos on right-hand side equals mosL; otherwise excitation order is
> 1!
  mosR1 = get_external_mos(IrRep1);
  mosR2 = get_external_mos(IrRep2);
  if (IrRep2 > Irep11)
  {  moR1 = mosR1.begin();
    moR2 = mosR2.begin();
    if (extSymmL == symm(moR1)) // band structure
    {  // do loop over 1st mo left // all in irrep1
      for (moL = mosL.begin(); moL = mosL.end(); moL++)
      {  // do bands
        tmp_offset += (pos(moL) - 1)* mosR1.size() * (spineigen);
        for (moR1; moR1 = mosR.end(); moR1++);
        {  // calc offset
          offset += (spineigen);
          // Calculate matrix elements from interacting MO L1, L2, R1 and R2!
          MRMPHOMatElements(tablecase,interactingMO(L1,L2,R1,R2));
        }
      }
    }
    else if (extSymmL == symm(moR2)) // diagonal structure
    {  for (moL = mosL.begin(); moL = mosL.end(); moL++)
      {  for (moR1 = mosR1.begin(); moR1 = mosR1.end(), moR1++)
        {  // do single entries
```

```
        tmp_offset += (pos(moR1) - 1)* (spineigen);
        offset = tmp_offset + mosR1.size()* (spineigen);
        for (moR2 = mosR2.begin(); moR2 = mosR2.end(); moR2++)
        {  // use multiple offset from temp_offset;
          // Calculate matrix elements from interacting MO L1, L2, R1 and R2!
          MRMPHOMatElements(tablecase,interactingMO(L1,L2,R1,R2));
        }
      }
      // calc new offset of next row
    }
  }
}
}
```

The given structures of the external submatrices were used to derive algorithms that determine the interacting configurations and their positions to compute the matrix elements with non-zero contributions. The algorithms described here in pseudo-code represent a model in which only the non-vanishing external interactions are evaluated. The implemented model builds up the interaction directly from the external orbitals, without going the detour of creating the interacting external configurations and computing the interaction.

4.5   Integral Handling

The integrals needed for a MR-MP2 computation can be gathered from eqn.4.19 (see also eq. 2.96) and 4.22 (eq. 2.76).

$$\sum_j c_j^1 \langle \Phi_i^1 | \hat{H}_0 - \hat{E}_0 | \Phi_j^1 \rangle = - \underbrace{\langle \Phi_i^1 | \hat{H} | \psi_0 \rangle}_{Inhomogenity} \tag{4.19}$$

The integrals needed to compute the inhomogenity arise from interactions of the type

$$\langle \Phi_i^1 | \hat{H} | \Phi_0 \rangle = \langle ab | kl \rangle - \langle ab | lk \rangle \tag{4.20}$$

$$= (ak | bl) - (al | bk) \tag{4.21}$$

with a and b denoting internal and external orbitals and k and l representing internal orbitals only. Eq. 4.20 is given in physicist's notation and eq. 4.21 given in chemist's notation. The Coulomb and Exchange integrals are of the same type $(all, internal | all, internal)$ and correspond to the single-reference MP2 integrals, except for an increased number of internal and a decreased number of external orbitals.

The integrals needed on the left-hand side of eq. 4.19 are given in eq. 4.22 and are of the type $(all, all | internal, internal)$ for the Coulomb and $(all, internal | internal, all)$ for the Exchange part.

$$f_{ab} = h_{ab} + \sum_{kl} D_{kl} \left[(ab|kl) - \frac{1}{2} (al|kb) \right] \qquad (4.22)$$

The appearing Coulomb part in eq. 4.22 rules out the use of the programs which perform a standard MP2 integral transformation but requires a 4-index transformation of the type $(all, all | all, internal)$. The storage requirement of the four-index integrals can be reduced if only integrals of the type $(int, int | int, int)$, $(ext, int | int, int)$, $(ext, ext | int, int)$, and $(ext, int | ext, int)$ are stored. The permutational symmetry of the integrals has already been considered. The four-index integrals are required for the evaluation of the inhomogenity, which is computed once, stored and used to compute the second-order correction to the energy. The four-index integrals are therefore only needed once and can be deleted after the computation of the inhomogenity. The matrix elements needed in the iterative steps concerning $\hat{H}_0$ (left-hand side of eq. 4.19) are computed from effective one-electron integrals which are reduced by contracting of the four-index integrals in conjunction with the density matrix D_{kl}.

A more compact way of integral handling is based on the use of three-index (RI) integrals instead of four-index integrals. RI integrals of the type $(ab|aux)$ recover the full four-index quantity by a scalar product of the $(ab|aux)$ vector with the dimension of the auxiliary basis functions aux and reduce the storage requirements. The orbitals a and b denote all internal and external orbitals and a separated storage like for the four-index integrals scheme does not have significant advantages.

Two different schemes to evaluate the four-index integrals from the RI integrals are possible. The four-index integrals can be constructed and stored prior to the computation of the inhomogenity in the first step of the iterative procedure and then used throughout the computation or the integrals can be generated from the RI integrals on the fly each time a four-index integral is needed. Storage of the necessary four-index integrals is only advantageous for medium-size systems (about 300 basis functions) and requires a substantial amount of disc-space for large molecules and the storage in main memory is not feasible. The generation of the four-index integrals from RI integrals on the fly seems to be time-consuming

but with nowadays CPU speeds the reconstruction of the integrals does not cost as much as reading the integrals from disc-based IO-devices. The second approach is more advantegous since not all four-index integrals are needed and the iterative process requires only effective one-electron integrals. However, both of the two methods are implemented in the program ot ensure compatibility with existing integral codes (MOLCAS and TURBOMOLE).

The expression of the Fock matrix elements in MR-MP2 (eq. 4.22) shows f_{ab} to depend only on the orbitals a and b but not directly on the full four-index integrals. The four-index integrals together with the two-index integrals can be absorbed into effective one-electron integrals of type $(a|b)$. The contraction requires the reference space dependent density matrix and the four-index integrals connected with the orbitals a and b and all internal orbitals k and l. The spatial parts of the matrix elements are computed from the reduced one-electron integrals as needed in case of non-diagonal interactions and sums over all occupied orbitals in case of diagonal interactions. The scheme was already implemented in the DIESEL package but not yet published. In the iterative procedure the four-index integrals or the RI integrals are not necessary anymore and can be removed.

4.6 Parallelization Scheme

4.6.1 Parallelization Models

Parallelization distributes the cost of a computation on more than one CPU and reduces the over-all time from starting a calculation to receiving the result. The parallelization models are determined by the system architecture. Multi-CPU computers that access the same main memory address space (*Symmetric Multi Processing, SMP*) and those having a unique address space for each CPU (*Distributed Memory*) differ substantially in their architecture and different parallelization concepts have to be applied. Parallelization on multi-processor computers with distributed memory, which are much more cost-efficient than those with shared-memory, is only effective if the time a part of the computation needs on one CPU of the parallel environment is at least a magnitude larger than the exchange of the data needed on each CPU. For SMP systems the parallelization is done with threads. Threads are child processes started by a mother process, but do not have the administration overhead of

stand-alone processes[3]. Threads and processes have either access to an indivual part of the shared-memory or access the same (shared) data. The implementation can be done manually with system calls (*e.g.* `fork`) or bases on more elaborate extensions of some compilers like *OpenMP*.[97–99] A third architecture that is becoming quite popular at the moment (AMD Opteron) is called *Non Uniform Memory Architecture, NUMA*. Each processor accesses an assigned part of the memory but also has some slower access to the memory of the other processors (*e.g. Hypertransport*). NUMA and the Distributed Memory architectures require parallelization models which minimize communication. The communication times of NUMA systems are much shorter than for cluster networks but need a manually optimized parallelization scheme for optimum performance too.

The model for SMP architectures cannot be applied to distributed memory systems (one computer or a cluster of individual computers connected via a fast network interface) because the data transfer from a memory part of one node to another is not possible without any further levels of communication. The transportation of data can either be done with emulated shared-memory environments like *Global Arrays* or *PVM (Parallel Virtual Machine)* or directly bases on interprocess communication on different nodes like the *MPI (Message-Passing Interface)*.[100–102] MPI requires an explicit implementation of the data structures but allows a low-level parallelization. The distributed memory models can also be applied for SMP or NUMA systems but might loose some efficiency caused by communication overhead.

For SMP systems parallelization models like OpenMP (see Listing 4.5) the compiler gets information on how the data in the memory has to be treated and how a block has to be parallelized. Shared data in the main memory is accessible for all threads while private data is only accessible for the threads that owns that particular data. The structure of a sequential program is kept, added with single compiler directives for the parallelization.

Listing 4.5: OpenMP program for a parallel matrix-vector multiplication. The memory occupied by the matrix and the vectors x and y, respectively are shared in the memory.

```
#include <omp.h>
// set number of threads
void omp_set_num_threads(int nthreads);
```

---

[3]In Linux kernels up to release 2.4 threads are implemented as POSIX threads which actually are stand-alone processes. From kernel 2.6 on, the threading libraries do support real threads.

```
void do_matrix_vector_multiplication()
{
  vector<double> x, y;
  // begin parallel region
  #pragma omp parallel private(j, k, nthreads, myid) default(shared)
  {  nthreads = omp_get_num_threads();
    myid = omp_get_thread_num();
    // do parallel loop
    for (k = 0; j < N; k++)
    {   double d = 0.0;
      for(j = 0; k < N; j++)
        d = matrix[k][j] * y[j];
      x[k] = d;
    }
  }     // end of parallel region
}
```

In Message-Passing environments the data has to be exchanged explicitely at the cost of a more complicated program logic. The advantage of MPI still is the portability to other platforms and a high control on the data structures to reduce communication overhead. An example program that illustrates the basics of MPI is given in Listing 4.6.

Listing 4.6: Basic MPI Program to illustrate the main concept of interprocess communication. Every sent data necessarily requires a corresponding receiving procedure. The given example for a matrix-vector multiplication is not the optimal code but illustrates the basic ideas.

```
#include "mpi.h"
#include <mpi++.h>
  MPI::Init(argc, argv); // initialize MPI
  MPI::Status status; // get MPI status
  int numprocs = MPI::COMM_WORLD.Get_size(); // number of processors
  int myid = MPI::COMM_WORLD.Get_rank(); // id: 0=master; 1-... = slaves;
  vector<double> tmp, x, y;
    if (myid == master) // master part
    {   // BroadCast to all slaves (1)
      MPI::COMM_WORLD.Bcast(&N, 1, MPI::INT, master);
      MPI::COMM_WORLD.Bcast(&(*matrix),N*N,MPI::DOUBLE,master);
      // send part of y to specific slave (2)
      MPI::COMM_WORLD.Send(&(*y), length, MPI::DOUBLE, destination, tag);
      //receive part of x result from slave(3)
      MPI::COMM_WORLD.Recv(&(*tmp), length,
                  MPI::DOUBLE, MPI::ANY_SOURCE,
                  MPI::ANY_TAG, status);
      // ... reduce to final x[k]
      x[k] += tmp[k];
```

```
    }
    else // slave part
    {  // Receive BroadCast (1')
      MPI::COMM_WORLD.Bcast(&N,1,MPI::INT,master);
      MPI::COMM_WORLD.Bcast(&(*matrix),N*N,MPI::DOUBLE,master);
      // receive from a certain node (here: master) (2')
      MPI::COMM_WORLD.Recv(&(*y), length,
                   MPI::DOUBLE, MPI::ANY_SOURCE,
                   MPI::ANY_TAG, status);
      // ... do something with data, e.g. partial matrix-vector multiplication
      for (k=offset; k < partN; k++)
        for (j=offset; j < partN; j++)
          tmp[k] += matrix[k][j] * y[j];
      //send result back to master (3')
      MPI::COMM_WORLD.Send(&(*tmp), length, MPI::DOUBLE, master, tag);
    }
  MPI::COMM_WORLD.Barrier();
  MPI::Finalize();
```

The program flow expresses the so-called master-worker scheme (see Figure 4.10). The master distributes the main work among the workers and collects the returned results. Control of data exchange is done via a central instance of the master, which basically is a process to distribute the data and to control the computational flow with insignificant computational load.

4.6.2   Message-Passing Based Parallelization of MR-MP2

The computational bottlenecks in a MR-MP2 computation are the computational cost to compute the $\mathbf{H_0}$ matrix and the matrix-vector product with the trial vector during the iterations. For large-scale computations the dimensions of the matrix and the trial vector are of about one billion or more. Parallelization in a MR-MP2 procedure is advantegous for the computation of the inhomogenity and the matrix-vector multiplication of the $\mathbf{H_0}$ matrix elements with the coefficient vector.

The parallelization scheme of the MR-MP2 program encounters only those blocks at the inactive level that are non-zero and distributes the parts of the coefficient vector that is needed for the matrix-vector multiplication. The inhomogenity and the $\mathbf{H_0}$ are divided into blocks that correspond to the maximum dimension of a vector to be stored in memory. Only those blocks of inactive configurations are distributed to the workers that contribute non-zero entries on this level. The division employs the inactive patterns described earlier to exclude
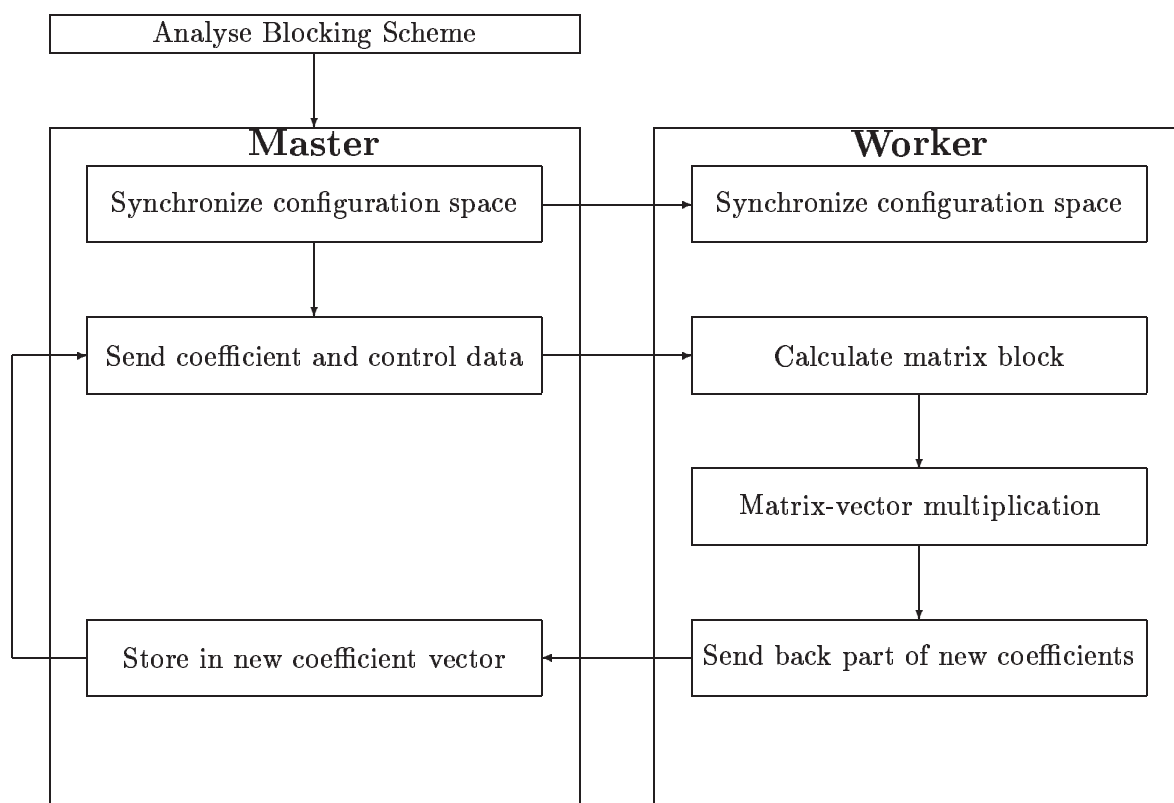
Figure 4.10: Master-Worker scheme for a parallelized matrix-vector multiplication in a MR-MP2 procedure.

the treatment of empty blocks. The treatment of empty blocks increases the communication overhead (evaluation of the subblocks with only vanishing entries; eventually requesting the appropriate coefficients, ...) with no actual computational effort. The blocking scheme treats same regions in one block (*e.g.* [inactive], [inactive-1], [inactive-2]) and also groups inactive difference configurations in the same irreducible representation together. The distribution scheme for the inhomogenity is given in Listing 4.7. The (much more complicated) scheme for $\mathbf{H_0}$ is not given.

With the dynamical distribution of the coefficient vectors only the required parts of the vector are sent to the workers where the appropriate part of the matrix elements are constructed and the matrix-vector multiplication is handled. The applied scheme utilizes fast sequential routines that consume a large share of the computational time but blocks the problem into smaller submatrices whose corresponding coefficients can be handled in memory. Disc storage of the coefficient vectors on each worker thus can be avoided.

Listing 4.7: Parallelization scheme for the computation of the inhomogenity.

```cpp
#include "Inhomogenity.h"
#include "Pattern.h"
InhomogenityMPIBlocking()
{
  vector<int> blocked:
  // 0: empty
  // 1: internal-x1
  // 2: core-x1 // -> -1 -> intern-extern separation
  // 3: irrep1
  // 4: core1 start
  // 5: core1 end
  // 6: CSFStart1
  // 7: Delta CSF1
  for (internal1 = 0; internal1 <= mrspace->getExcitationOrder(); internal1++)
  {
    code[1] = internal1;
    for (core1 = 0; core1 <= mrspace->getExcitationOrder(); core1++)
    {
      code[2] = core1;
      for (irrep1 = 1; irrep1 <= mrspace->getGroupOrder(); irrep1++)
      {
        code[3] = irrep1;
        counter1 = 1;
        if (mrspace->getCoreIndex(core1,irrep1,0) > 0)
        {
          finish = 0;
          firstindex1 = mrspace->getCoreIndex(core1,irrep1,counter1);
          lastindex1 = firstindex1;
          probeindex1 = 0;
          while (!finish)
          {  // get previous CSF = start of new one
             start1 = mrspace->getCSFStart(internal1,firstindex1-1);
            lastindex1 = mrspace->getCoreIndex(core1,irrep1,counter1);
            deltaCSF1 = mrspace->getCSFStart(internal1,lastindex1) - start1;
if (lastindex1 < mrspace->getCoreIndex(core1,irrep1,0))
              probeindex1 = mrspace->getCSFStart(internal1,
                      mrspace->getCoreIndex(core1,irrep1,counter1+1));
          // CSF number too large in the next pass -> send and start with next
block
            if (probeindex1 - start1  > vectormem)
            {  code[4]  = firstindex1;
              code[5]  = lastindex1;
              code[6]  = start1;
              code[7]   = deltaCSF1;
              firstindex1 = mrspace->getCoreIndex(core1,irrep1,counter1+1);
              for (int i = 0; i < infolength; i++)
                blocked.push_back(code[i]);
            }
```

```cpp
          // fits into 1 block
          if (counter1 == mrspace->getCoreIndex(core1,irrep1,0))
        {   code[4] = firstindex1;
            code[5] = lastindex1;
            code[6] = start1;
            code[7] = deltaCSF1;
            for (int i = 0; i < infolength; i++)
              blocked.push_back(code[i]);
            finish = 1;
        }
        counter1++;
      }
    }
   }
  }
 }
 return blocked;
}
```

Chapter 5

The DIESEL-MP Program

5.1 Object-Oriented Design and Implementation in C++

Object-oriented programming languages like the commonly used C++ or Java emphasize the data structure in form of objects in contrast to procedural or modular programming languages (Fortran, C and many more). For each kind of objects, which are defined as classes, methods that act on instances of the objects[1] are attributed. The main idea of this concept arises from the rather natural consideration that each object (*e.g.* a human being) possesses certain natural methods like eating, sleeping, walking, or swimming while other classes possess different ones (*e.g.* birds possess the "method" flying). The advantage of object-oriented languages basically comes from the opportunity to abstract the common features of two classes (*e.g.* birds also eat, sleep or might swim) to a base class from which the data structure and the methods of all other more specialized classes can be inherited. The class design allows a great flexibility of the code and provides a program architecture that can easily be extended to other methods. The MR-MP2 program could be easily converted to a MRCI program as the basic classes for both programs are comparable.

From possible object-oriented programming languages, C++ was chosen for the implementation of the MR-MP2 code, as the object-oriented design eases well-maintainable and extentable codes. C++ further produces faster binaries compared to Java but also allows the mixing of different languages (C, Fortran) for special routines. The need of a parallel

---

[1]The allocated object of a class is called an instance of an object.

implementation also restricted the choice of language to Fortran, C or C++ particularly in the context of the parallel *Message-Passing Interface* (MPI) environment. The basic classes needed for a MR-MP2 are now discussed in detail and the program flow is presented.
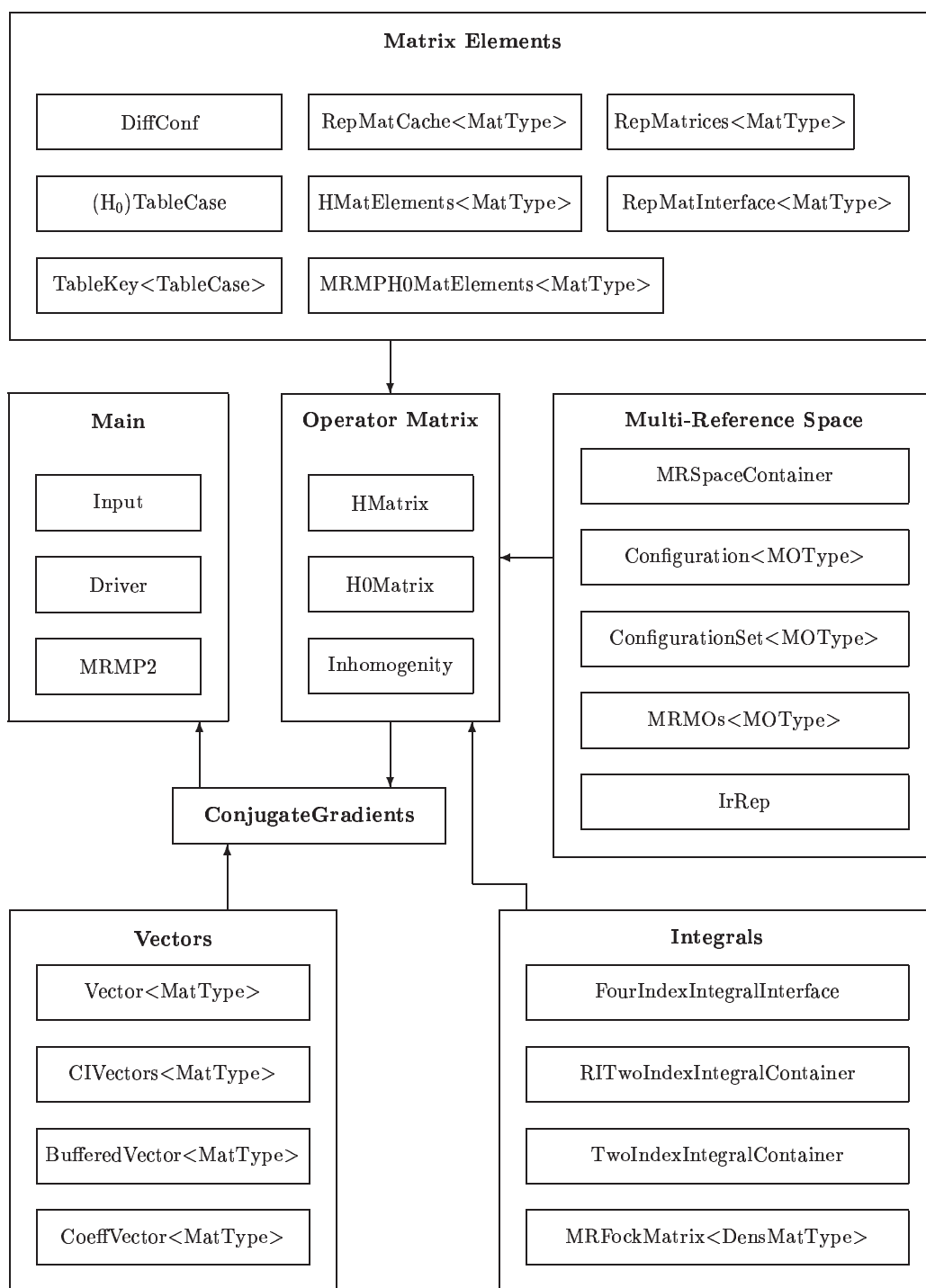
5.1.1 Class Diagram for a MR-MP2 Procedure

Figure 5.1: Class Diagram for a MR-MP2 procedure.

The basic classes for a MR-MP2 procedure which can also be generally applied to any multi-reference method are illustrated in Figure 5.1. The central part of the program are the classes that are used to build up the multi-reference space. The multi-reference space consists of configurations handled in the class `Configuration` and is stored in an object of the class `MRSpaceContainer`. Both classes depend on the representation of the molecular orbitals and therefore template implementations were chosen to ensure that the class not only works for MO represented by integer numbers but also for different types (*e.g.* complex orbitals). Symmetry operations for abelian groups are considered in the class `IrRep`, which also belongs to this complex of classes.

A second complex of classes handles the storage of the required integrals. Both of the mentioned complexes are needed for the implementation of the Hamilton matrices which can be generally regarded as matrices with elements depending on the interaction of operators (a two-electron operator in `HMatrix` or `Inhomogenity` or an effective one-electron operator in the case of `HOMatrix`). These complexes are not features of object-oriented languages but are grouped together to assist a logical structure of the code.

The elements of these operator matrices are provided from the complex of classes used to compute the Hamilton matrix elements. The implementation of the matrix element classes were partially taken from the DIESEL-MR-CI and MR-MP2 codes of Michael Hanrath[53, 54] but were thoroughly ported from the *GNU 2.9x* compiler series to the ANSI conform *GNU 3.x* and the *Intel* compiler. The Fortran routines[91] for the computation of the representation matrices in the SGA framework were left unchanged and used for this program. The classes `DiffConf`, `TableCase`, `HOTableCase` and `TableKey` were completely modified to work with recent C++ compilers. The classes that explicitly compute the matrix elements (`HMatElements` and `MRMPHOMatElements`) were orginally designed for an individually selecting MR-CI and not optimal for a treatment of the full multi-reference space. The concerned methods were overloaded (several methods with same names but different signatures) to better account for the changing requirements in the course of matrix element computation (identical configurations, single excitations with differing internal MO or with differing external MO, etc.). These requirements are also reflected in the class `DiffConf` that computes the difference configurations (see Section 4.2.1).

The Hamilton matrices are finally instantiated from the class `Conjugate Gradients` whose

methods perform the interative process to obtain the first-order correction of the MR-MP2 wavefunction. In an MR-CI this class would be replaced by a class `Davidson`[82,84,85] but the structure of the basic class design is kept. The `ConjugateGradients` class is used by the `Driver` and `MRMP2` classes that control the program flow (input, restarts, final output, etc.). Before the program flow of the implemented MR-MP2 procedure is discussed, the basic classes for the configuration handling and setting up the multi-reference space are presented in some more detail in the following section.

5.1.2   C++ Implementation of Basic Classes in a MR-MP2

5.1.2.1   Class `Configuration`

Listing 5.8: Interface of the class `Configuration`. The implementation is neglected for the sake of simplicity.

```
template <class MOType>
class Configuration
{
  public:
    Configuration();
    Configuration(const Configuration&); // copy-constructor
    Configuration(const string&); // create conf from string input
    ~Configuration(); //destructor
// --------------------------- base functions ----------------------- //
    Configuration& create(MOType mo); // create one electron
    Configuration& annihilate(MOType mo); // annihilate one electron

    int getNumberOfOpenShells() const;
    int getNumberOfClosedShells() const;
    MOType getOpenShell(int index) const;
    MOType getClosedShell(int index) const;

#ifdef MPIPARALLEL
    void MPIsend(int destination);
    void MPIreceive(int source);
#endif
// --------------------------- operators ---------------------------- //
    Configuration<MOType> operator+(const Configuration<MOType>&);
    Configuration<MOType> operator-(const Configuration<MOType>&);

    Configuration& operator++(int); // iterator
    ostream& operator<<(ostream&);
```

```
  private:
    vector<MOType> open;
    vector<MOType> closed;

    void sort(); // from Sedgewick
    void quicksort(vector<MOType>& orbs, int l, int r); // from Sedgewick
};
// ---------------------- comparing operators  ---------------------- //
// operator ==
inline bool operator==(const Configuration<MOType>& _a, const
Configuration<MOType>& _b);
// operator <
inline bool operator<(const Configuration<MOType>& _a, const
Configuration<MOType>& _b);
```

The basic object for the computation of matrix elements is the configuration. The class exists of two dynamically allocated arrays for the open-shell and the closed-shell orbitals, respectively. In the DIESEL package configurations are implemented as fixed arrays and the dimensions were hard-coded during the compilation. Arrays with fixed dimensions are processed on the *stack* and not on the *heap*. Both represent a different processing of data in memory with the *stack* being better performant concerning memory access than the *heap*. In the DIESEL-MP program configurations with up to several hundred occupied orbitals should be treated and the slightly slower, but more flexible dynamic arrays were chosen. The memory overhead of fixed arrays with too large dimensions is thus avoided.

An efficient treatment of configurations (adding and substracting configurations, creating and annihilating orbitals) is facilitated by sorted entries of orbitals. For this purpose a quicksort routine[103] was implemented. The performance of dynamic arrays is about 20% increased over the more natural representation employing sets. Operators to add and to substract configurations are implemented with regard to the calculation of difference configurations. Comparing operators are used to sort the configurations once inserted into sets of configurations. These sets are used to ensure that a configuration exists only once in the multi-reference space and provides an ordered scheme of all inserted configurations (see Section 4.3).

5.1.2.2   Class `DiffConf`

The difference configuration (compare Section 4.2.1) is employed to determine the Table CI labels (see class `TableCase`) and is set up on the configuration class. The data structure

is represented by three configurations and the relative position of the open-shell orbitals (see Section 4.2.1). The maximum number of open shells is hard-coded to 20, because the size of a representation matrix for 20 open shells has a dimension of about 10000. Each representation matrix would thus require about 400 MB of memory which is far beyond a reasonable size. The construction of a difference configuration can be done with the basic method `calc` or can be built directly (`build`). The latter procedure is used for the efficient implementation of the external configuration parts (see Section 4.4.2.2).

Listing 5.9: Class `DiffConf` to handle difference configurations that are the basis for the computation of the matrix elements in the SGA framework and are used to determine the Table CI case (class `TableCase`).

```cpp
class DiffConf
{
  public:
    DiffConf(); // default constructor
    DiffConf(const Configuration<MOType> &_core);
    DiffConf(Configuration<MOType>&, Configuration<MOType>&);
    ~DiffConf(); // default destructor
// --------------------------- functions  --------------------------- //
    void calc(const Configuration<MOType>&, const Configuration<MOType>&);
   // set DiffConf directly (needed for direct evaluation of externals)
     void build(MOType mo_same1, MOType mo_same2,
           MOType mo_left1, MOType mo_left2,
             MOType mo_right1, MOType mo_right2);

    const Configuration<MOType>& getSame() const ;
    const Configuration<MOType>& getFrom() const ; // DiffConf L
    const Configuration<MOType>& getTo() const; // DiffConf M

    UINT32 getKey(bool&); // return key for TableCase
    int calcExcitationOrder();
    void calcInteractingMO(MOType *interacting_mo);
// --------------------------  operators --------------------------- //
    ostream& operator<<(ostream&);
  // for istream for Patterns
    friend istream& operator>>(istream& strm, DiffConf& diff);

    // add two DiffConfs; result is *this! faster than operator+
    void add(DiffConf&, DiffConf&);
    DiffConf operator+(DiffConf&); // add two DiffConfs

  private:
    Configuration<MOType> same;
    Configuration<MOType> diffL;
```

```
        Configuration<MOType> diffM;
        int posL[MAXOPEN];
        int posM[MAXOPEN];

        bool switchSides;
        int excitationLevel;
        Configuration<MOType> core; // default core conf

        void calcPosOpenShells();
};
```

Get methods for the individual parts of the difference configurations are used for the matrix element computation, the correct assignment of the required integrals and for the determination of the Table CI cases. The correct assignment of integrals for the computation of the inhomogenity and the $\mathbf{H_0}$ matrix can also be directly determined by an analysis of interacting MO in the difference configuration and the excitation level. These features are widely used in conjunction with the patterns to build up the matrices.

5.1.2.3 Table CI Cases

The difference configurations are the basis for the classification of Table CI Cases. The class `TableCase` analyzes the Table CI case(see Table 4.1) for a difference configuration and sets the appropriate indices of the needed integrals. The determination follows the scheme given in Figure 4.1.

Listing 5.10: Class `TableCase` to determine the Table CI cases based on difference configurations. For the $\mathbf{H_0}$ matrix a more efficient version exists (class `H0TableCase`).

```
class TableCase
{
  public:
    TableCase();
    TableCase(BinomialCoefficient *);
    TableCase(int _openShells, int _dK, int __P, int _R, int _qR, int _qL);
    ~TableCase();
// --------------------------- functions ------------------------- //
    void calc(DiffConf& dc);

    const char* getName() const;
    int  getNumberOfMoreOpenShells() const;
    int  getdK() const;
    int  getP() const;
```

```
    int  getR() const;
    int  getqR() const;
    int  getqL() const;

    const CbExIntegralIndex& getCbExIndex() const;
    TwoElectronIntegralIndex getTwoElectronIntegralIndex() const;

  private:
    BinomialCoefficient  *binom; // used to calc qR and qL
    int  openShells;  // #(open shells) in conf. of higher super category
    int  dK;        // difference between supercategories ("delta K")
    int  P;          // type of integrals
    int  R;           // classification of interaction
    int  qR;        // class. of interacting open shells in "right" conf.
    int  qL;        // class. of interacting open shells in "left" conf.
    CbExIntegralIndex CbExIndex; // index for integrals

    void calcQ(const DiffConf & dc);
    int calcQFromPos(const int *ipos, int npos);
};
```

The Table CI Cases for the $\mathbf{H_0}$ matrix concern only single excitations. This much simpler scheme (Figure 4.2) is also implemented in the class `H0TableCase` that improves the efficiency building up the $\mathbf{H_0}$ matrix.

5.1.2.4   Class `MRSpaceContainer`

The main container used to store the configurations, indices and start values of CSF for the matrix-vector multiplication is implemented in the class `MRSpaceContainer`. This container generates configuration spaces for full configurations, internal-external separated and core-active-external configurations. The different spaces of MR-CI (including the reference wavefunction) and of MR-MP2 (excluding the reference wavefunction) were also considered in the implementation.

Listing 5.11: Storage class for the multi-reference space depending on the configuration level and the employed method.

```
class MRSpaceContainer
{
  public:
    MRSpaceContainer();
    MRSpaceContainer(const string _mode, const string _method);
    MRSpaceContainer(  const string _mode, const string _method,
```

```
                MOContainer<MOType,IrrepType> _mos,
                int _electrons, int _symmetry, int _multiplicity,
                int _excitationOrder, ConfigurationSet _refs);
    ~MRSpaceContainer();
//-----------------------------------------------------------------------//
    // internal: internal-x; core: core-x;
    // irrep: IrRep x; pos: position in array ([0] = #entries)
    Configuration<MOType>& getCore(int internal, int irrep, int pos) const;
    Configuration<MOType>& getRef(int pos) const;
    Configuration<MOType>& getActive(int internal,int core,
                        int irrep, int pos) const;
    Configuration<MOType>& getExternal(int internal, int irrep, int pos) const;

    int getCSFStart(int internal, int index) const;
    int getNumberOfCSF() const; // total CSF;

    int getSymmetry() const;
    int getMultiplicity() const;
    int getConfigurationLevel() const;
//-----------------------------------------------------------------------//
    enum ConfigurationLevel { Full,Intern,CoreActive,CoreActiveHole };
    enum Method { MRMP2,MRCI };
//-----------------------------------------------------------------------//
#ifdef MPIPARALLEL
    void MPIsend(int destination);
    void MPIreceive(int source);
#endif
//-----------------------------------------------------------------------//
  private:
    ConfigurationLevel treatment;
    Method method;
    ConfigurationSet refs;
    Irrep irrep;
    int groupOrder, multiplicity, symmetry, index, totalCSFs, externalMOs;

// indexed containers for Configuration storage
    Configuration<MOType> *core;
    Configuration<MOType> *active; // also for full and internal configurations
    Configuration<MOType> *refactive;
    Configuration<MOType> *external;

    int ***indexCore, ****indexActive, ***indexExternal, *indexRef;
    int **CSFCoreIndex, *CSFActiveIndex;
//-----------------------------------------------------------------------//
    void calcCSFStart();

    void createExternalConfs(MRTreeInternExtern* externTree);
    void insertCoreConfs(MRTreeCoreActive* coreTree);
    void insertActiveConfs(MRTreeConf* tree);
};
```

5.1.3 MR-MP2 Program Flow

The flow of the MR-MP2 program (Figure 5.2) consists of seven major steps implemented in the classes `Driver` and `MRMP2`. The first step reads the input data used to control the program (reference configurations, multiplicity, desired roots, irreducible representations, provided memory, etc.). In the next step the one-electron and two-electron integrals are read and stored either in memory or on disc.
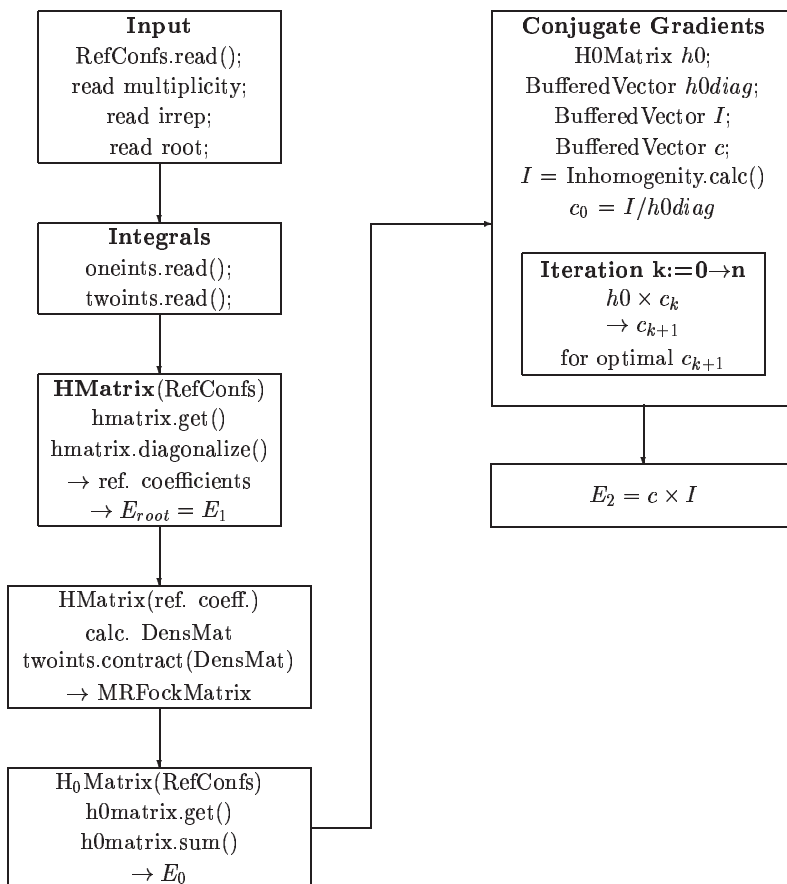

Figure 5.2: Flow for a MR-MP2 program.

The coefficients of the reference wavefunction are the elements of the eigenvector obtained by a diagonalization of the reference matrix built from the reference configurations. The eigenvalue from the diagonalization is the first-order correction to the energy E_1 (eq. 2.102). The density matrix obtained from the reference matrix is used to contract the two-electron integrals into the effective one-electron integrals (eq 2.76), object of class `MRFockMatrix`).

The zero-order energy E_0 (eq. 2.101) is computed by summing up the matrix elements of the $\mathbf{H_0}$ matrix spanned by the reference configurations.

In the Conjugate Gradient class, the $\mathbf{H_0}$ matrix is initialized and the inhomogenity is evaluated. After the evaluation of the inhomgenity, the two-electron integrals are not needed anymore and deallocated. The iterative steps to obtain the final solution vector for the first-order correction of the wavefunction are done with the Conjugate Gradient method. This solution vector is then multiplied with the inhomogenity to yield the second-order correction to the energy E_2 (eq. 2.104).

5.1.4   Provision for Possible Extensions

The object-oriented implementation facilitates an easy extension of modified algorithms like the incorporation of selection schemes to reduce the dimension of the solution vectors and the Hamilton matrix. The selection schemes may be based on an individual selection of configurations sor an orbital selection of the external MO.

In the case of an individual selection the energy contributions of the $\mathbf{H_0}$ matrix elements along the main diagonal are evaluated and configurations below an individually chosen threshold do not enter the interaction space. The reduced interaction space is solved employing the standard Conjugate Gradient procedure and the energy contributions of the neglected configurations are used to extrapolate the energy of the reduced space to that of the full interaction space. The use of an individual interaction scheme requires lists of external configurations for each internal configuration rest. The class `MRSpaceContainer` is already prepared for this feature.

A different scheme, that leaves the logic structure of the $\mathbf{H_0}$ matrix intact, applies the selection of only the external orbitals discarding certain external orbitals completely from the interaction space. In analogy to the previous scheme, the energy contributions of the configurations along the main diagonal of the $\mathbf{H_0}$ matrix are used and are attributed to the external orbitals. If the sum of energy contributions for a particular external orbital exceeds a user-specified threshold then the orbital is considered in the interaction space. The energy contributions of the neglected configurations are again used to extrapolate the energy from the iteratively solved reduced MR-MP2 computation to the full MR-MP2 space. This scheme does not spoil the structure of the external interactions (see Section 4.4.2.2) as these only depend on the relative position of the external orbitals.

The design of the basic classes also allows an efficient implementation of a MR-CI procedure with or without the selection schemes described above. The external interactions could also be implemented directly, but the interactions get more complicated as not only single excitations contribute to non-zero matrix elements but also double excitations. The number of distinguished external matrix blocks exceeds the number present in a MR-MP2, with some being characterized by the same pattern of non-vanishing matrix elements.

5.2   Interfaces to TURBOMOLE and MOLCAS

The DIESEL-MP program possesses interfaces to two commercial quantum-chemical codes. The first interface utilizes the integrals from the MOLCAS suite of programs[51] and originates from Michael Hanrath written for the DIESEL[53] package. The integrals are stored as four-index integrals in single precision and stored on disc limiting the maximum number of basis functions to a dimension of about 200.

Larger systems can be treated by the RI[104–106] interface developed by Stefan Grimme and provided with the TURBOMOLE source code.[107] This interface was adapted in conjunction with Svetlana Stepanenko to the latest version of the TURBOMOLE package and saves the one-electron and the two-electron three-index RI integrals on disc. The large file support enables the use of integral files up to 1000 basis functions. This limit comes intrinsically from restrictions in the TURBOMOLE since all orbitals have to be set occupied in the course of the RI transformation using the RI-MP2 module.

The current limit of the RI integral transformation can be overcome if an orbital dependent selection scheme is combined with a modified integral transformation. For a given reference wavefunction the integrals needed for the evaluation of the main diagonal elements are of the type $(all, all|internal, internal))$, with the first two integral indices restricted to be the same. The second type of integrals, which are needed to compute the inhomogenity vector, are identical to those necessary in the RI-MP2 method with the orbitals of the multi-reference wavefunction entering the RI-MP2 integral transformation. In the next step the external orbitals can be selected by their energy contributions of the diagonal elements and the selected external orbitals enter a modified integral transformation, in which only those selected orbitals are considered.

5.3 Hardware Considerations

5.3.1 System Architecture

The DIESEL-MP program is designed for large-scale MR-MP2 parallel computations on commodity hardware clusters like the HOMER PC cluster in the Institute of Organic Chemistry at the University of Würzburg. The crucial issues concerning good performance are interprocess communication and storage requirements as well in memory as on hard-disc. The parallel algorithm is designed such that the interprocess communication is kept at a reasonable level limiting the storage requirements on the computing nodes (workers). The master has to compensate for that but in modern cluster architectures with master nodes that are equipped with large memory and hard-discs this distribution policy does not become a problem.

The nowadays developments in hardware technology with increasing CPU performance for low-cost 32 bit architectures and the slow but steady establishment of medium-price 64 bit systems (*AMD Opteron*) are reflected in the implementation of the program. For 32 bit and 64 bit architectures the program can be compiled with the *GNU 3.2* or later compiler suites (32 bit on Intel and AMD platforms, 64 bit on HP ALPHA, AMD Opteron and many more) or the Intel compilers (32 bit on Intel or AMD platforms or 64 bit on Itanium systems). Both compiler suites show comparable performance for the *DIESEL-MP* program.

5.3.2 Parallel Requirements

The parallel implementation allows the simultaneous distribution of the matrix parts and execcution of the matrix-vector multiplications. For this step the distribution of the co-efficient vectors is crucial for the performance. One possibility is the distribution of the complete vector among all workers which then have access to all vector elements. This requires large disc spaces on all nodes as not only the original vector, but also the parts of the newly formed vectors after the matrix-vector multiplication have to be stored. In one iteration step the vector (in a dimension of 10^8 to 10^9 entries or about 800 MB to 8 GB) has to be sent to all workers, stored and the new vector with the same storage requirement allocated. On state-of-the-art 32 bit systems the memory is limited to 2 GB per process

and the vectors are hardly kept in memory, but have to be stored on disc. The new vector for the next iteration step has to be contracted from all the partial vectors on the master node and distributed again amongst all workers. The send and receive processes of the vectors cause system load via the network interface and for a matrix-vector multiplication no computational work can be done during the transmission processes. In the time span the master collects the result vectors and contracts them to a new trial vector the workers are idle and computational efficiency is lost. Therefore this approach is not perfect.

The network transfer performance of current Gigabit networks is of the same order as the IO performance of mass market hard-disc and the difference between writing data to or reading data from disc is comparable to network transfers of smaller parts of the vectors needed at a certain worker node. Therefore this scheme was implemented in the current code as it does not require any vector storage on the worker nodes. The parallel distribution scheme presented in Section 4.6 is governed by the size of the vector dimension which depends on the memory on the workers. The maximum vector size is controlled by the user's input. The idle time of the workers is reduced by a an efficient storage handling of the needed vector parts. The master reads a bunch of required coefficients, sends them to a worker and and immediately prefetches the set for the next parallel block. Once the master receives the result vector from the worker, the next vector is sent and the result vector is stored afterwards. The time for reading and writing the coefficient vectors is done while the workers perform the matrix-vector multiplication.

A test computation with one worker node and a medium-sized molecule of about 42 correlated electrons, 150 basis functions and 16 inactive and 9 active orbitals forming a multi-reference space dimension of 67 million CSF (about 512 MB disc storage needed for each vector) requires only about 800 MB of memory storage alltogether in contrast to the 1.6 GB needed by the MR-MP2 program in the DIESEL package. In the DIESEL MR-MP2 program the solution vectors are stored in memory. This distribution scheme thus can be applied for large-scale computations. The performance of the program is comparable to the DIESEL program package. Significantly larger tests systems, for which the new program is designed for, cannot be handled with old DIESEL package.

Chapter 6

Applications of Multi-Reference Methods for Complex Electronic Systems

Multi-reference methods like MR-CI and MR-PT do not belong to the standard methods commonly used in computational chemistry to study reactivities and reaction mechanisms of compounds interesting to the experimentally working chemist. Their most severe disadvantage from most user's point of view is the missing *black box* character, which requires a distinct experience with the methods and a certain caution to interpret the results. Despite this apparent disadvantage they are necessary for diradical species and offer a deep insight into problems that would be never achieved with methods like DFT or MP2. The obtained results helped to get experience about the capabilities and limitations of the latter methods for difficult cases.

In this chapter three projects are presented, which, if they were independently regarded, do not seem to be related to each other, except for diradical intermediates occuring in all of the three. The first project studies the Neocarzinostatin chromophore (NCS) which is a representative for an entire class of biologically active anti-tumor antibiotics. The structural feautures that are essential for its efficacy now can be understood.[108] Its efficacy is governed by the structural element of a monocyclic enyne[3]-cumulene that is unusual for natural products. The occurence of this and similar compounds, however, initiated the

search for model systems that allowed the basic understanding of the reaction mechanism and contributed to the discovery of new classes of compounds with potential efficacy. One example that shows a broad spectrum in its reactivity and regioselectivity is the presented class of the non-natural and short-lived enyne-ketenes.[109] These compounds indeed show a large variety of intermediates with different electronic structures that are difficult to predict. The same electronic structures as they are present for these intermediates can also be found for cyclic allenes derived from the isobenzene 1,2,4-cyclohexatriene. A model that first establishes systematic predictions on the occurence and the stability of the possible electronic states is proposed in the third project.[110]

6.1   The Chemistry of Neocarzinostatin

6.1.1   A Brief Story about Neocarzinostatin and Related Model Compounds

Neocarzinostatin (NCS)[111,112] is an antibiotic 1:1 complex of a protein component called the NCS apoprotein and a chromophore first isolated in 1965 by Ishida[113] from *Streptomyces carzinostaticus* Var. F-41 (Figure 6.1). After it was possible to separate both parts, it was quickly recognized that the major biological activity was caused by the NCS chromophore, while the apoprotein has an important role for the stabilization and transport of the chromophore. The structure of the NCS chromophore having a bicyclo[7.3.0]dodecadiyne framework was elucidated in 1985 by Edo and coworkers.[114] This work gained attention as in the year 1987 two other novel natural product families, Calicheamicin and Esperamicin,[115] with enediyne-moieties were discovered. Two years later Dynemicin A followed. This section briefly reviews the basic role of the Neocarzinostatin chromophore to develop new anti-tumor antibiotics with improved efficacy and summarizes the approaches with various model compounds. The structural features than govern the efficacy are presented in Section 6.1.2.

The combination of the fascinating molecular structure of the active chromophore with the high biological efficacy raised the interest of numerous chemists and biologists. The apoprotein serves as a stabilizer and carrier of the NCS chromophore to the DNA. After
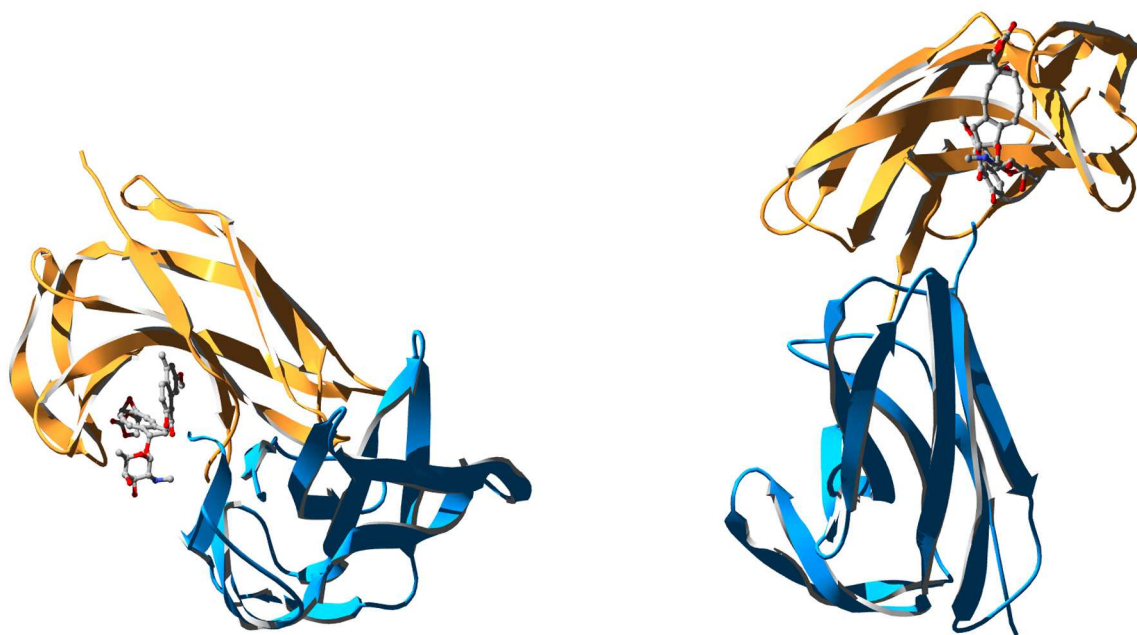
Figure 6.1: X-ray structure of Neocarzinostatin[116] from two different views. The apoprotein is represented as ribbons and the chromophore is shown in ball and stick representation. The Protein Database entry is 1NCO.
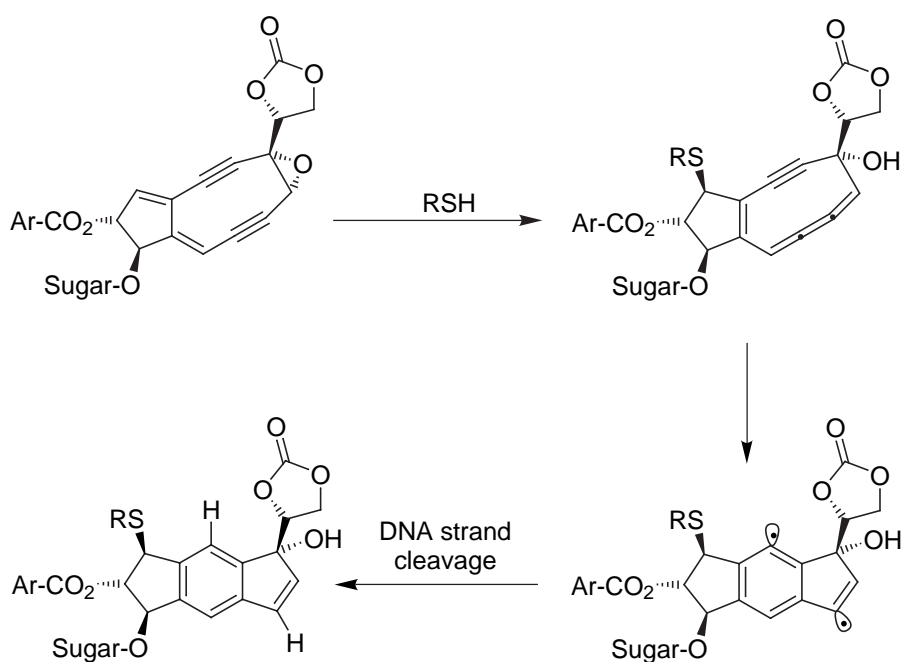

Figure 6.2: Mode of action of the NCS chromophore

the apoprotein is split off, the naphtobenzoate rest (ArCO$_2$ in Figure 6.2) is believed to intercalate in the DNA strand and the bicyclododecadiyne moiety is supposed to be located in the minor groove of the DNA (see Figure 6.3).[111] The mode of action of the NCS chro-

mophore as sketched in Figure 6.2 was first proposed by Myers in 1987.[117] The elucidated structure agrees with the results from a set of cyclization experiments after nucleophilic attack of a thiol[117, 118] causing the enediyne to rearrange to the unstable enyne-cumulene system.[119] The enyne-cumulene system readily reacts in a cyclization process to form a diradical species that abstracts two hydrogens from the $C_{5'}$ centers of adenine and thymine (Figure 6.3). This hydrogen abstraction is responsible for the DNA cleavage and finally leads to cytolysis.[120, 121] The nature of the intermediate was under debate and an alternative mechanism via a zwitterionic intermediate was proposed[122] but ruled out.[119, 121, 123, 124] Comparing NCS with Esperamicin, Calicheamicin and Dynemicin A the modes of action of NCS and these ene-diynes are alike except the preceeding rearrangement reaction to the enyne[3]cumulene[1]. The mode of action is determined by the chromophore serving as a sort of "warhead" that is carried specifically to its target and a "trigger" initiating the formation of the DNA cleaving diradicals.
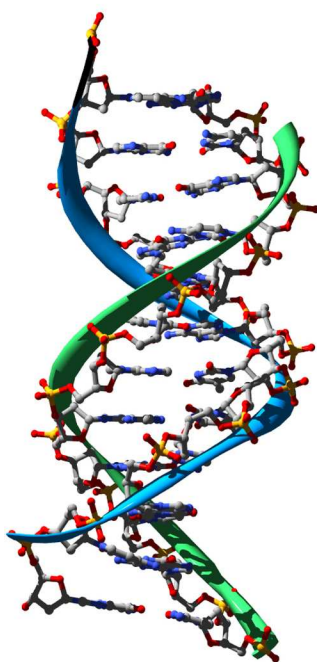
Figure 6.3: b-DNA double strand as target for the NCS chromophore. The NCS chromophore is bound to the minor groove, while the naphtyl moiety is believed to intercalate in the DNA. Hydrogens are abstracted from adenine and thymine.

---

[1]The ene-diynes also readily form cyclic diradicals in the manner of the Bergman reaction[125, 126] discovered years before the isolation and characterization of these natural products. It is interesting to note that before the discovery of this new class of natural products no proposal of diradicals serving as DNA cleaving agents was made.
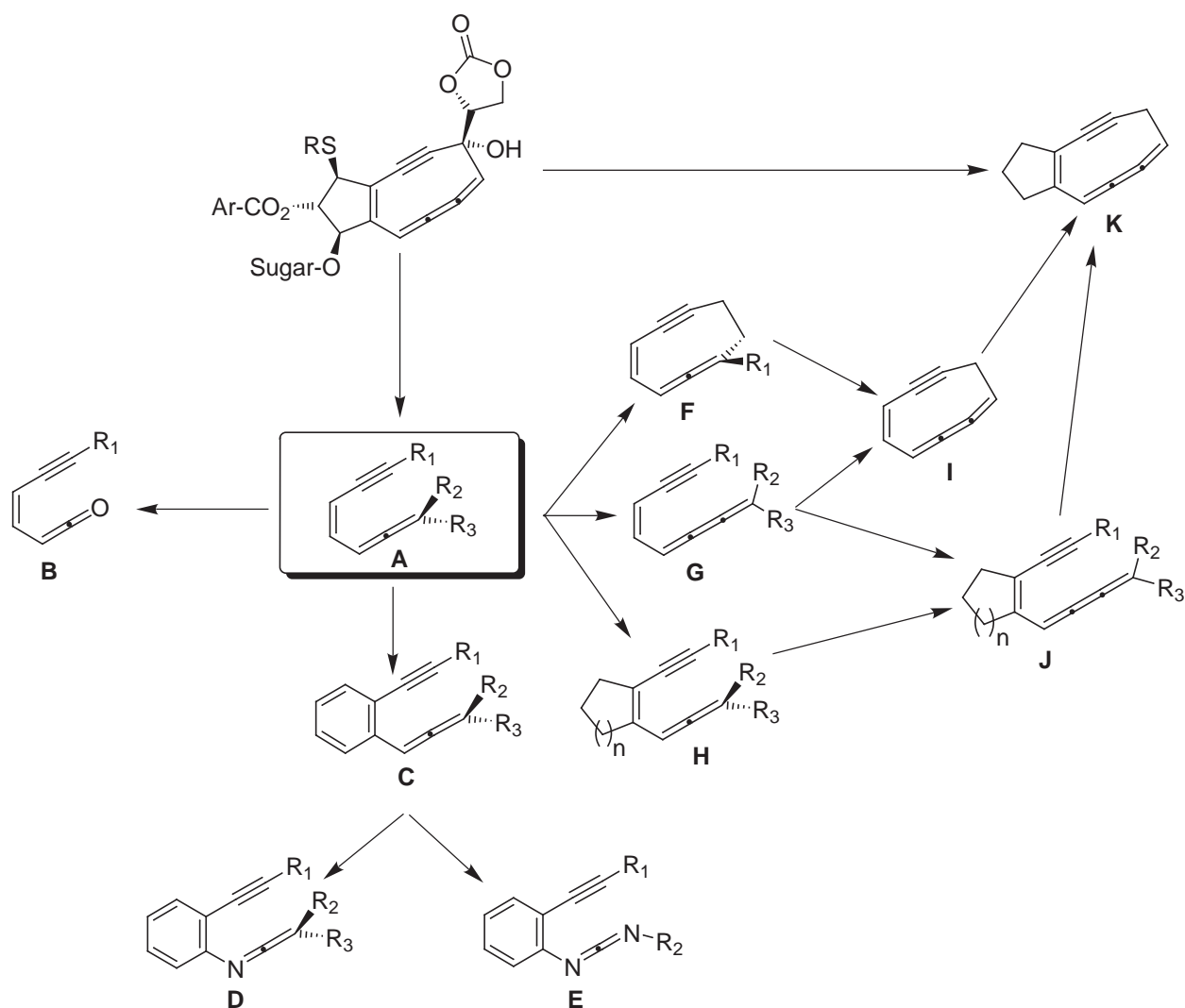
Figure 6.4: Model systems and related compounds derived from the NCS chromophore that were used in experimental and computational studies.

The unusual features and the high efficacy of NCS inspired research on related compounds. To achieve a profound understanding of the mechanism and trying to improve the efficacy of these new antitumor antibiotics, model systems (Figure 6.4) derived from the NCS chromophore were studied. Myers[127] and Saito[128] independently synthesized the first model compound for the NCS chromophore, Z-1,2,4-heptatriene-6-yne (**A** in Figure 6.4), that essentially contains the bond motifs of NCS. The class of enyne-allenes was chosen since it is had an easier synthetic access than the enyne-cumulenes.[129–132] **A** reacts in a thermally induced cyclization reaction yielding toluene as the final product. This novel cyclization showed DNA cleaving activity like the NCS chromophore and had the same regioselectivity as NCS. Trapping experiments and the determination of the thermochemistry established the α,3-didehydrotoluene diradical as the key intermediate formed in the course of the cy-

clization (Figure 6.5).[127, 128, 133–137] This cyclization reaction, mimicking the mode of action of NCS, was finally named after Myers and Saito and is later referred as Myers-Saito reaction. The results obtained from the simple model compound **A** contributed substantially to the understanding of the cyclization mode of the NCS chromophore, which was unkown at the beginning of the studies. A model compound very similar to **A** was introduced by Moore, who investigated the reactivity of an isoelectronic enyne-ketene **B**. The short-lived enyne-ketene also possessed DNA cleaving features, indicating a similar cyclization mechanism. This mechanism will be adressed in Section 6.2.
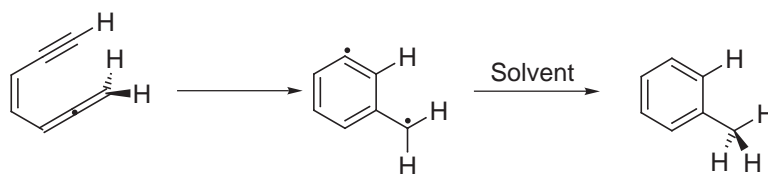
Figure 6.5: Mechanism of the thermal cyclization reaction of Z-1,2,4-heptatriene-6-yne proposed by Myers and Saito that mimics the mode of action of the NCS chromophore

After the proposal of the mode of action of NCS, the experiments of Myers and Saito influenced other groups who tried to utilize the enyne-allene moiety to get synthetic access to biologically active compounds being as effective or superior to NCS. In the group of Schmittel a series of experiments was made that investigated the thermal cyclizations of substituted enyne-allenes (**C** in Figure 6.4).[138] For synthetic reasons benzannulated compounds were employed. Surprisingly, the cyclization of enyne-allenes having sterically demanding ($R^1=^t$Bu) or aromatic ($R^1=p$-Toluene) substituents yielded not the Myers-Saito product but a five-membered ring instead of a six-membered ring[2]. Computations showed that the Myers-Saito pathway is thermodynamically preferred, but the bulky substituents favor a more compact transition state that yields the Schmittel products in the kinetically controlled reaction (Figure 6.6).[139] The same regioselectivity is found for enyne-allenes embedded in a nine- or ten-membered ring (see **F**).[140, 141] The effect of benzannulation on the thermodynamics and the reaction barriers with a more distinct preference for the Schmittel and not the Myers-Saito was also explored.[142–145]

Ring size effects generally influence the cyclization modes of the enyne-allenes as was

---

[2]It is interesting to note, that in the review of Wang[137] one example of an enyne-ketene is given that shows a product distribution of six- and five-membered rings in the case of a phenyl substituent at the alkyne terminus.
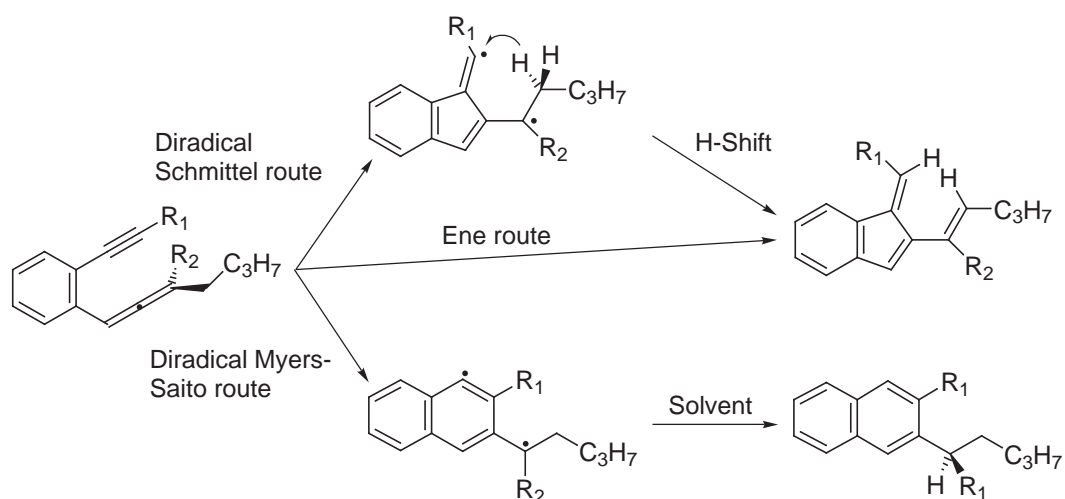
Figure 6.6: Alternative regioselectivities in the thermal cyclization reaction of substituted Z-1,2,4-heptatriene-6-ynes.

shown in a combined experimental and theoretical study for compounds of type **H**.[146] With a cyclohexene or cyclopentene unit attached to the enyne-allene system, the Schmittel route yielding the fulvene was found but if a cyclopentene ring was attached to the enyne-allene the opposite regioselectivity (Myers-Saito mode) was observed (Figure 6.7).
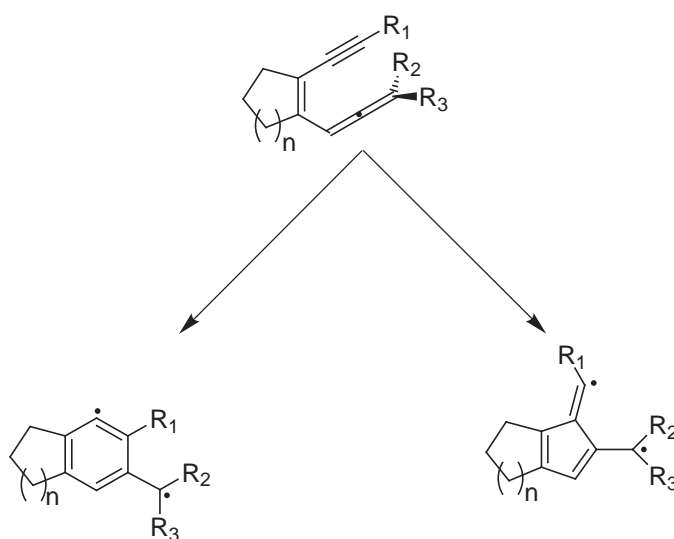

Figure 6.7: Ring strain affects the regioselectivity of the cyclization

The proposed reaction mechanism of the Myers-Saito reaction of **A** *via* the α, 3-didehydro-toluene diradical was supported by computational findings.[147–149] For the Myers-Saito reaction also zwitterionic and allene intermediates were proposed.[150] Alike the Myers-Saito route, the Schmittel reaction was also predicted to proceed *via* a diradical intermediate, if **A** (R^1=H) is regarded,[147, 148] but can change with the substituent R^1. For the benzannulated

system **C** with R^1 =Ph the mechanism *via* the diradical route preferring the Schmittel product is predicted, while for R^1=^{t}Bu both the diradical route as well as a concerted ene-reaction are predicted to be equal.[151] The Schmittel reaction is also shifted away from the diradical reaction mechanism if strong electron-donating substituents as NH_2 are present and a reaction mechanism preferring carbenic intermediates is predicted to occur.[139]

The structural units (enyne-allene moiety, the nine-membered monocycle, and the cyclopentene annulation), which have been described so far for the enyne-allenes, represent the same structural subunits that constitute the NCS chromophore. Bridging the enyne-cumulene **G**[152] by a methylene group yields the cyclic enyne-[3]-cumulene **I**. For **I**, Cramer computed the reaction energy of the cyclization and the stability of the diradical trying to get more insight into the thermodynamic factors, the role of the diradical intermediate of the cyclization, and to compare it with the Myers-Saito reaction.[153] The cyclopentene moiety is found in the open-chain enyne-cumulene **J** that can be thought of being derived from **G** or **H**. **J** is an important step towards the total synthesis of the NCS chromophore for which new methods for an efficient synthesis had to be developed.[154–157] Finally the compound **K** that is related closest to the NCS chromophore was synthesized by Takahashi.[158] It contains all important structural features except for substituents at the bicyclic framework.

Keeping the mode of action of the NCS chromophore in mind, these studies contributed to a basic understanding of the factors that control the regioselectivity of the Myers-Saito and Schmittel cyclization and provided the basis for an understanding of the efficacy of the NCS chromophore which is presented in the next section.

6.1.2   Which Structural Elements Are Relevant for the Efficacy of Neocarzinostatin?[108]

While the general course of the mode of action is clarified, it remains still unclear which structural elements are relevant for the high efficacy of NCS. Based on high-level *ab initio* computations[3], new insight into this topic, which is of great interest for the development of

---

[3]The computational details for obtaining the results are described in the following. The geometric parameters of all stationary points were optimized by employing analytical energy gradients within the density functional theory (DFT) framework as implemented in Gaussian98.[159] For the DFT calculations the B3LYP[160, 161] functional in connection with a 6-31G(d)[162] basis set was employed using a spin and space
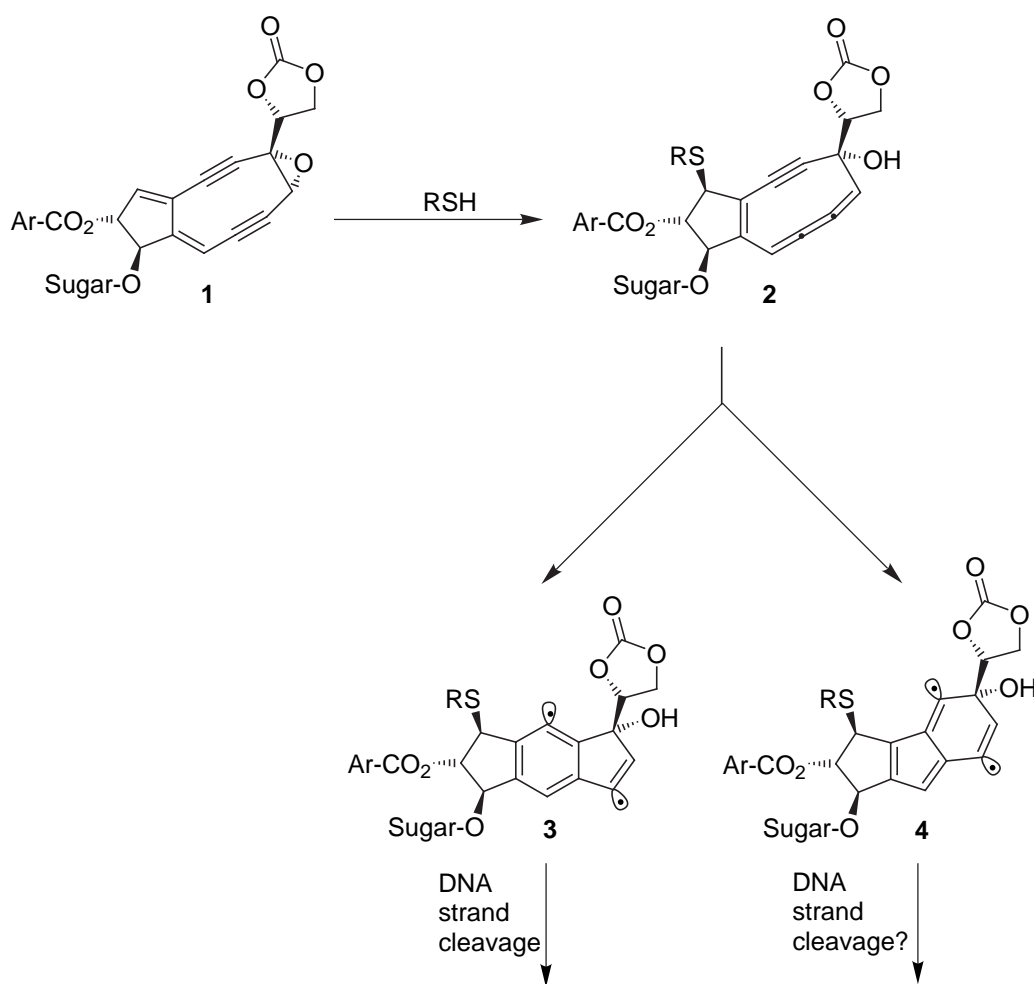
Figure 6.8: Reaction pathway of the NCS chromophore leading to the experimentally confirmed intermediate **3**.

new antitumor antibiotics, is presented. First the six-membered ring cyclization depicted in Figure 6.8 between C^3 and C^7 (see Figure 6.9 for the labelling of the centers), found in

---

unrestricted approach. All stationary points were analyzed by computed harmonic frequencies at the same level of theory. Vibrational, thermal and entropy corrections at 298 K were also computed at the B3LYP/6-31G(d) level of theory. To estimate solvent effects, the COSMO model[163] in combination with the B3LYP level of theory was employed by using the standard values for water. Transition state energies were obtained by single-point computations employing the more reliable closed-shell coupled cluster (CCSD(T)) approach in conjunction with a cc-pVDZ[164] basis set. These computations were performed with the MOLPRO 2000.1 and 2002.1 packages[52, 165] The reliability of the CCSD(T) approach was validated by using the T_1 diagnostics. Reaction energies have been computed by using a multi-reference configuration interaction approach in connection with a Davidson estimate of quadruple excitations (MR-CI+Q) employing a cc-pVDZ basis set as implemented in the DIESEL-MR-CI package.[32] The reference space of the individually selecting MR-CI consisted of up to 27 configuration state functions (CSF) generating a configuration space of up to $223 \cdot 10^6$ CSF. The secular equations to solve were up to the order of $2 \cdot 10^6$.

the mode of action of NCS is compared to the alternative five-membered ring cyclization between the C^2 and C^7 centers. The latter yields the intermediate **4** instead of **3**. Formally both species represent diradicals. For a nine-membered monocyclic enyne-allene, a model compound for the chromophore of NCS (**2**), this alternative route was predicted to be kinetically preferred to the six-membered ring cyclization.[140] From a comparison of both cyclization modes a new model concerning the tasks of the substituents present in **2** arises. To test this model, the influences of the substituents and the solvent on the kinetics of both cyclization modes are discussed. Finally, the capability of the arising diradicals for the subsequent hydrogen abstraction leading to the final DNA cleavage is studied. These investigations provide new insights into the subtle effects that nature uses to obtain powerful antitumor antibiotics. Valuable improvements of these substances can be deduced, since some of the factors that limit the efficacy become clear.

Table 6.1: Transition states of both possible cyclization modes. All values are in kcal·mol^{-1}.

	Method	$C^2 - C^7$ cyclization [a]		$C^3 - C^7$ cyclization [b]	
		$\Delta E^{\ddagger}$	$\Delta G^{\ddagger}$	$\Delta E^{\ddagger}$	$\Delta G^{\ddagger}$
5	UB3LYP/6-31G(d)	+18.6	+17.6	+21.1	+20.8
	CCSD(T)/cc-pVDZ	+13.8		+16.5	
6	UB3LYP/6-31G(d)	+31.9	+31.9	+22.2	+23.0
	CCSD(T)/cc-pVDZ	+29.7		+20.1	
7	UB3LYP/6-31G(d)	+22.8	+21.4	+22.2	+21.9
	CCSD(T)/cc-pVDZ	+22.2		+22.2	
8	UB3LYP/6-31G(d)	+19.5	+18.3	+21.6	+20.7
	CCSD(T)/cc-pVDZ	+16.1		+18.3	

[a] For **6**: $C^3 - C^7$ cyclization. [b] For **6**: $C^3 - C^8$ cyclization.

Initial insight into the topics previously mentioned can be taken from Table 6.1, which summarizes the computed activation barriers of the six-membered $C^3 - C^7$ and of the five-membered $C^2 - C^7$ ring cyclization for various model compounds that are depicted in Figure 6.9. The CCSD(T) computations[4] represent the most reliable approach for the transition

---

[4]Note that in most cases the CCSD(T) approach represents a more reliable method than DFT. While both methods agree nicely in the $\Delta E^{\ddagger}$ values for the open-chain compound **6**, larger differences are found for the monocyclic systems **5** and **8**. However, while both methods give different absolute reaction energies, the correct trends are already obtained with DFT.
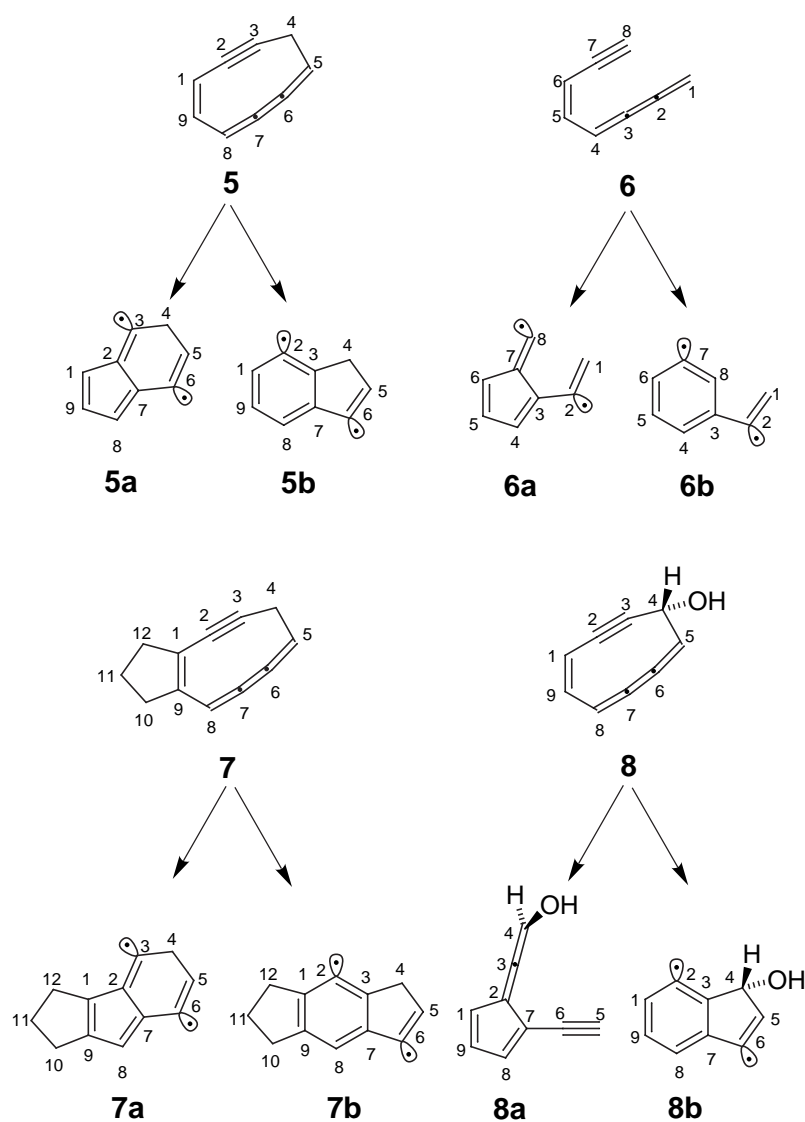
Figure 6.9: Model compounds to investigate the structural effects of the NCS chromophore on the efficacy.

states, since they possess a closed-shell electronic structure as the reactants. At this level of theory the $C^2 - C^7$ cyclization for the parent ring system of the NCS chromophore (**5**, Figure 6.9) is predicted to be kinetically preferred to the $C^3 - C^7$ cyclization ($\Delta E^{\ddagger}(C^2 - C^7)$ = 14 kcal·mol^{-1} vs. $\Delta E^{\ddagger}(C^3 - C^7)$ = 17 kcal·mol^{-1}), which was found to be the only mode of action in NCS.[112] The values computed in this work agree nicely with those calculated by Cramer and Squires employing the BD(T) method.[153] Compared to **5**, the open-chain compound **6** posesses a considerably smaller activation barrier for the $C^3 - C^7$ than the $C^2 - C^7$ cyclization ($\Delta E^{\ddagger}(C^2 - C^7)$ = 30 kcal·mol^{-1} vs. $\Delta E^{\ddagger}(C^3 - C^7)$ = 20 kcal·mol^{-1}). The comparison between **5** and **6** shows that the shift in the regioselectivity is induced by the ring strain present in **5**. It also indicates that this ring strain is important for the

efficacy of **2** since it reduces the activation energy of both cyclization modes considerably. **5b** represents a diradical with a fixed distance of 3.55 Å between the two radical centers. One can imagine that such a fixed distance is also important for the efficacy of the double hydrogen abstraction in the final step of the mode of action of NCS. However, since for NCS only the $C^3 - C^7$ cyclization is found, the substituents present in **2** must shift the regioselectivity from the $C^2 - C^7$ to the $C^3 - C^7$ cyclization.

Information on the importance of substituents for the control of the regioselectivity can also be taken from Table 6.1. The cyclopentene ring in **7** seems to be important for the control of the regioselectivity, since the induced strain increases the activation energy of the $C^2 - C^7$ cyclization more than that of the $C^2 - C^7$ cyclization. A similar trend was found experimentally for enyne-allenes (see Section 6.1.1).[146, 166] In contrast, the OH substituent at C^4 (**8**) increases both activation barriers by about the same amount ($2 \text{ kcal} \cdot \text{mol}^{-1}$) and therefore does not affect the regioselectivity of the cyclization. However, the substituent completely changes the reaction path of the $C^2 - C^7$ cyclization. Along with the ring closure between C^2 and C^7 the ring opens between C^4 and C^5 leading to the final product **8a**. In contrast, the $C^3 - C^7$ cyclization still proceeds *via* a diradical route.

Solvent effects were computed with the COSMO model on the B3LYP/6-31G(d) level of theory.[163] Polar solvents also favor the $C^3 - C^7$ cyclization to some extent, since the activation energy of the $C^3 - C^7$ cyclization for **5** and **7** is lowered by approximately $0.5 \text{ kcal} \cdot \text{mol}^{-1}$, while the activation energy of the $C^2 - C^7$ cyclization is increased by about the same amount. Besides the activation energies, also the reaction energies have to be taken into account for the regioselectivity. The reaction energies given in Table 6.2 show that for the systems considered in this study the $C^3 - C^7$ cyclization is thermodynamically favored. Multi-reference CI computations could only be performed for model system **5** because of the inherent limits in the DIESEL package. The highly endothermic heat of reaction for **5a** is slightly lower than the activation energy and the reverse reaction possesses almost no barrier. The $C^2 - C^7$ cyclization thus is likely to be in an equilibrium preferring the reactant.

These findings suggest that one of the major tasks of the substituents present in the chromophore of NCS is to prevent the five-membered ring cyclization. A possible reason for the control of the regioselectivity can be seen from the computed reaction energies and

Table 6.2: Heats of reaction for both cyclization modes and singlet-triplet gaps E_{S-T} for the corresponding intermediates. All values are in $kcal \cdot mol^{-1}$.

Method	5a	5b	6a	6b	7a	7b
UB3LYP/6-31G(d)	+17.3	+0.8	+16.4	-8.5	+22.2	+2.0
MR-CI+Q/cc-pVDZ	+12.0	-3.7	n.a.	n.a.	n.a.	n.a.
E_{S-T}(UB3LYP/6-31G(d))	+11.9	+4.5	-3.4	+1.3	+13.0	+4.4

the singlet-triplet gap E_{S-T} of both intermediates (Table 6.2)[5]. The intermediate of the $C^2 - C^7$ cyclization **5a** lies energetically so high that its lifetime is considerably limited by the reverse reaction, which possesses nearly no activation barrier. Additionally, the computed E_{S-T} value of about 12 $kcal \cdot mol^{-1}$ indicates that **5a** does not have a distinctive diradical character. In contrast, **5b** lies energetically well below the reactant and represents a true diradical (E_{S-T} around 4 $kcal \cdot mol^{-1}$). The differences concerning lifetime and diradical character indicate that an efficient hydrogen abstraction, which is the important final step in the mode of action of NCS, will only be possible with **5b**. A similar observation is found for **7**, showing that this behavior seems to represent a general motif for bicyclic enyne-cumulene compounds. The small singlet-triplet gap of the acyclic enyne-cumulene **6a** indicates that the ring strain of the cyclic enyne-cumulene system is responsible for the larger values obtained for **5a** and **7a**. Consequently, we expect enyne-cumulene systems in which the $C^2 - C^7$ cyclization is not suppressed to be less efficient antitumor antibiotics.

These investigations show that the chromophore of NCS represents a fascinating example for natural products in which a subtle balance of various influences leads to the desired result. While the monocyclic enyne-cumulene moiety seems to be important to provide a diradical with a fixed distance between both radical centers and a low activation barrier for the cyclization process, the substituents present in **2** are necessary to prevent the competing $C^2 - C^7$ cyclization.

As the effects of the substituents, which have been discussed so far for the NCS chromophore and the family of the enyne-allenes (see Section 6.1.1), govern the reactivities (regioselectivity and the nature of the intermediates), the question is of interest whether the substituent effects are specific for this class of compounds or if general rules are existent.

---

[5]For **5b** the computed values for the singlet-triplet splitting agree nicely with those computed at the BD(T) level of theory.[153]

For this purpose, the reactivity of the class of enyne-ketenes is investigated, which can be considered to be hetero-analogues of the enyne-allenes.

6.2 On the Regioselectivity of the Cyclization of Enyne-Ketenes: A Computational Investigation and Comparison with the Myers-Saito and Schmittel Reaction[109]
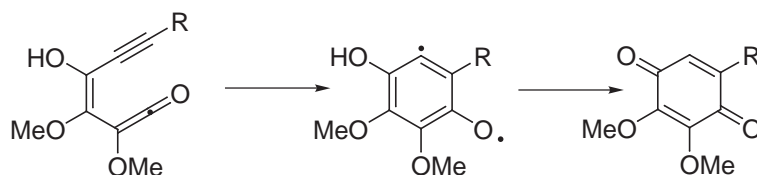
6.2.1 Introduction

Figure 6.10: Moore reaction yielding quinones.

The Moore cyclization of enyne-ketenes[137, 167–170] yielding quinones *via* phenoxy diradicals (see Figure 6.10) belongs to the family of diradical cyclization reactions such as the Bergman reaction of ene-diynes,[125, 126] the cyclization of enyne-cumulenes,[112] and the Myers-Saito reaction of enyne-allenes[127, 128] together with its regioalternative Schmittel cyclization.[138, 171–173]

The Bergman and the Myers-Saito reactions were awarded with substantial interest, because they represent the basis of the antitumor efficacy of ene-diynes and enyne-allenes. The Moore cyclization did not gain as much attention as the formerly mentioned Bergman or Myers-Saito cyclizations, although its DNA strand cleavage ability.

As indicated in Figure 6.11, the cyclization of enyne-ketenes possesses two reaction modes.[137] The first mode is the so-called Moore reaction (Figure 6.11, lower pathway). It forms the new bond between C^2 and C^7. For this reaction pathway, several intermediates have to be considered. In addition to the diradical route leading to the intermediate **9c**, the reaction could also yield the cyclic allene **9d** or a zwitterionic intermediate **9e**. In addition to the
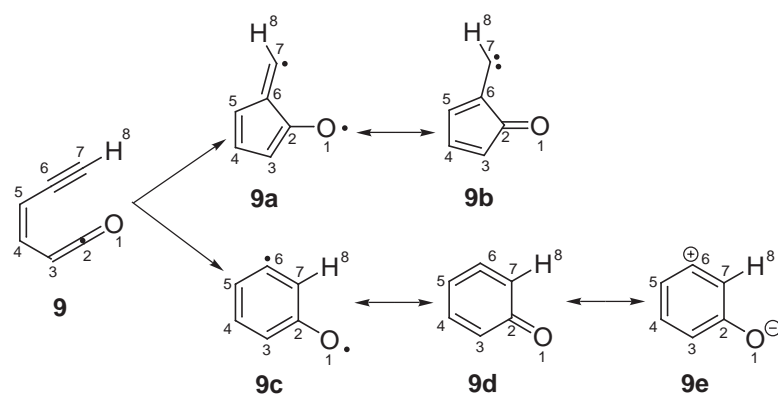
Figure 6.11: Possible reaction routes for **9** and diradical (**9a**, **9c**), carbenelike (**9b**), and allenic (**9e**) intermediates. The zwitterionic species **9e** represents the transition state for the inversion of the allene.

$C^2 - C^7$ cyclization, a five-membered ring between C^2 and C^6 (Figure 6.11, upper path) can be formed *via* the diradical pathway, yielding the σ, π-diradical **9a** or the carbenelike intermediate **9b**.

The mechanisms of both cyclization modes of the enyne-ketenes were unclear, but as the Moore cyclization of enyne-ketenes can be regarded as a heteroanalogue Myers-Saito reaction of enyne-allenes, both the six- and the five-membered ring cyclizations were expected to react *via* a diradical route.[137] However, the differences in the electronic structure of the enyne-allenes and the enyne-ketenes could lead to an energetic preference of the other intermediates. Such a preference was recently seen for the cyclic allenes, where substitution effects reversed the energetic ordering of the diradical and the zwitterionic species.[174, 175]

To answer the open questions, the present work investigates the course of both cyclization modes for the model compounds depicted in Figure 6.12. For this instance, all stationary points of both reaction paths were computed and characterized[6]. To study whether a general

---

[6]Geometric parameters of all stationary points were optimized employing analytic energy gradients within the Density Functional Theory (DFT) framework as implemented in *Gaussian98*.[159] For the DFT calculations, the BLYP and B3LYP[160, 161] functionals in connection with a 6-31G(d), 6-311G(d),[162] and a *cc*-pVDZ[164] basis set was employed using a spin and space unrestricted ansatz. All stationary points were analyzed by computed harmonic frequencies on the same level of theory and tested for wave function instabilities.[176, 177] Vibrational, thermal, and entropy corrections at 298 K were computed on the B3LYP/6-31G(d) level of theory.

Transition state energies of **9** were also obtained by single point computations employing the more reliable closed shell Coupled Cluster (CCSD and CCSD(T)) ansatz[10, 178, 179] in conjunction with a *cc*-pVDZ

motif exists for the family of reactions discussed above, relations between enyne-ketenes and enyne-allene compounds are also investigated.
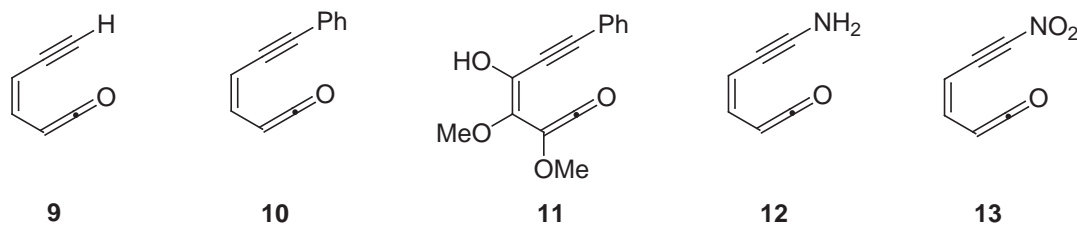

Figure 6.12: Model compounds used in this study.

6.2.2   Investigation of the Parent System

The cyclization of enyne-ketenes in principle possesses two possible regioalternatives. One mode of cyclization is a formation of the five-membered ring between C^2 and C^6 (see Figure 6.11, upper path) on a diradical pathway yielding the σ, π-diradical **9a** or the carbenelike species **9b**. The second mode, known as the Moore reaction, forms the new bond between C^2 and C^7 (see Figure 6.11, lower pathway). Possible intermediates for this reaction mode are the diradical (**9c**), the cyclic allene (**9d**), and the zwitterion (**9e**). Various computational studies prove DFT to be in principle appropriate to treat reactions including diradical intermediates.[140, 147, 181, 182] This finding is helpful since multi-reference methods for open-shell and Coupled Cluster calculations for closed-shell species are only applicable to smaller systems, leaving the computationally less demanding DFT to be the only opportunity to treat systems including all substituents present in the experimentally handled compound. Nevertheless, the results of these computational studies show that a validation is necessary

---

basis set to evaluate the DFT calculations. These computations were peformed with the MOLPRO 2000.1 package.[165] The reliability of the Coupled Cluster ansatz was validated using the T_1 diagnostics.

Reaction energies for **9** have been computed using a multi-reference configuration interaction ansatz in connection with a Davidson estimate of quadruple excitations (MR-CI+Q) as implemented in the DIESEL-MR-CI package,[32] as well as the CASPT2 implementation of the MOLCAS program package.[51] These computations based on CASSCF(10,10) wave functions employing a *cc*-pVDZ basis set and Widmarks *ano-l* basis set[180] in a (8s4p) → [3s1p] contraction for hydrogen and a (14s9p4d) → [4s3p2d] contraction for carbon and oxygen. The contraction schemes were taken from the basis set library of the MOLCAS program package. This basis set will be denoted as VTZP. The reference space of the individually selecting MR-CI consisted of up to 15 configuration state functions (CSF) generating a configuration space of up to 10^8 CSF. The secular equations to solve were up to the order of 10^7.

because various pitfalls are known for systems with complicated electronic structure. Additionally, it cannot be predicted which functional gives the most accurate values for a given class of compounds. For the cyclization of the parent system **9**, such a validation can be taken from Tables 6.3-6.5. Table 6.3 shows the computed activation energies for the reaction routes from **9** to **9a** ($C^2 - C^6$ cyclization) and to **9c-9e** ($C^2 - C^7$ cyclization). The computed geometric parameters can be taken from Table 6.4. Table 6.5 contains the reaction energies.

Table 6.3: Computed barriers of activation for model system **9**. All values are in kcal·mol^{-1}.

Method	$C^2 - C^6$ cyclization			$C^2 - C^7$ cyclization		
	$\Delta E^{\ddagger}$	$\Delta H^{\ddagger}_{298}$	$\Delta G^{\ddagger}_{298}$	$\Delta E^{\ddagger}$	$\Delta H^{\ddagger}_{298}$	$\Delta G^{\ddagger}_{298}$
B3LYP/6-31G(d)	+41.2	+39.4	+41.4	+13.3	+12.1	+13.6
B3LYP/6-311G(d)	+45.6	+43.7	+45.8	+16.7	+15.6	+17.8
BLYP/6-31G(d)	+30.6	+29.1	+31.1	+9.6	+8.7	+10.5
CCSD/cc-pVDZ//B3LYP/6-31G(d)	+48.0			+23.7		
CCSD/cc-pVDZ//BLYP/6-31G(d)	+45.5			+22.4		
CCSD(T)/cc-pVDZ//B3LYP/6-31G(d)	+41.3			+17.6		
CCSD(T)/cc-pVDZ//BLYP/6-31G(d)	+36.2			+16.9		

It is noted, that larger basis sets for the CCSD and CCSD(T) computations were not possible at the time of publication due to numerical limits.

Table 6.4: Selected geometric parameters for **9** and the TS for the $C^2 - C^6$ cyclization [a].

	9		TS $C^2 - C^6$ cyclization	
Parameter	B3LYP	BLYP	B3LYP	BLYP
$d(C^2 - C^6)$	301	304	169	176
$d(O^1 - C^2)$	117	118	120	121
$d(C^6 - C^7)$	121	123	131	129
$\angle(O^1 - C^2 - C^3)$	174	174	121	124
$\angle(C^6 - C^7 - H^8)$	179	179	133	142

[a] Distances are given in pm and angles in degrees.

The transition states (TS) of the $C^2 - C^6$ cyclization to **9a** and of the $C^2 - C^7$ cyclization to **9c** possess almost closed-shell character with only minor spin contamination which allows the use of CCSD and CCSD(T) to validate DFT employing the B3LYP and BLYP functionals. The activation energy for the $C^2 - C^6$ cyclization using B3LYP/6-31G(d) is predicted

to be 41.2 kcal $\cdot$ mol^{-1} with a slight thermodynamic correction to the Gibbs free energy of 41.4 kcal $\cdot$ mol^{-1}. It agrees perfectly with the CCSD(T) energy based on the same geometry of 41.3 kcal $\cdot$ mol^{-1}. With respect to CCSD(T), B3LYP/6-31G(d) underestimates the barrier height of the alternative reaction (C^2−C^7 cyclization) by about 4 kcal $\cdot$ mol^{-1} (13.3 kcal $\cdot$ mol^{-1} vs. 17.6 kcal $\cdot$ mol^{-1}). Corrections to the Gibbs free energy do not change the barrier of activation significantly ($\Delta G_{298}^{\mp} = 13.6$ kcal $\cdot$ mol^{-1} compared to $\Delta E^{\mp} = 13.3$ kcal $\cdot$ mol^{-1}). CCSD yields the same qualitative picture as CCSD(T) but predicts slightly increased barriers of activation for both the C^2−C^6 and C^2−C^7 cyclizations. The BLYP functional seems not to be appropriate to desribe the reaction modes. The computed barriers are much lower that the CCSD(T) counterparts (30.6 kcal $\cdot$ mol^{-1} vs. 41.2 kcal $\cdot$ mol^{-1}). Additionally, the computed geometries point to much more productlike transition states (Table 6.4).

Due to the strong change in the electronic structure along the reaction path, the computation of the reaction energies (**9** → **9a**, **9** → **9c**-**9e**) is quite demanding. Some intermediates represent diradicals and multi-reference methods have to be employed. The quality of the DFT ansatz was also assessed. For the C^2−C^6 cyclization (**9** → **9a**), B3LYP/cc-pVDZ predicts the reaction energy to be 25 kcal $\cdot$ mol^{-1}. This value lies between the values obtained with CASPT2/cc-pVDZ (23 kcal $\cdot$ mol^{-1}) or the MR-CI+Q/cc-pVDZ (27 kcal $\cdot$ mol^{-1}), underlining again the applicability of the B3LYP functional in the present context. Improving the basis set from VDZ to VTZ leads only to small changes in the computed reaction energies if the CASPT2 and the MR-CI+Q approach is employed. Interestingly, the DFT approach possesses a stronger basis set dependency than the more sophisticated approaches.

Similar to the enyne-allenes,[150] the diradical **9c** and the allene structure **9d** of the C^2−C^7 cyclization are found to be very close in energy, whereas the zwitterionic structure **9e** lies much higher in energy. The most reliable values are presumably given by the MR-CI+Q approach in combination with the VTZP basis set. The deviation of the CASPT2 ansatz to the MR-CI results is very small. While the difference between **9c** and **9d** is computed to be 5-6 kcal $\cdot$ mol^{-1}, the energy difference between **9c** and **9e** is about 16 kcal $\cdot$ mol^{-1}. While the B3LYP functional agrees nicely with the more sophisticated approaches, the BLYP functional again deviates by computing **9d** to represent the most stable intermediate of the C^2−C^7 cyclization.

Table 6.5: Computed reaction energies for model system 9[a].

	9a		9c		9d		9e	
Method	$\Delta_R E_{298}$	$\Delta_R G_{298}$	$\Delta_R E_{298}$	$\Delta_R G_{298}$	$\Delta_R E_{298}$	$\Delta_R G_{298}$	$\Delta_R E_{298}$	$\Delta_R G_{298}$
B3LYP/6-31G(d)	+23.4	+24.8	+2.0	+4.8	+6.8	+9.6		
B3LYP/6-311G(d)	+28.8	+30.2	+7.9	+10.7	+11.5	+14.2		
BLYP/6-31G(d)	+27.1	+28.5	+6.3	+8.9	+4.7	+7.4		
B3LYP/*cc*-pVDZ	+25.0	+26.3	+5.4	+8.1	+9.8	+12.4	+13.4	+14.6
B3LYP/*cc*-pVTZ//B3LYP/*cc*-pVDZ	+29.4		+8.8		+11.8		+16.1	
MP2/*cc*-pVTZ//B3LYP/*cc*-pVDZ[b]	+61.1		+33.0		+12.0		+19.5	
(10,10)CASSCF/*cc*-pVDZ//B3LYP/*cc*-pVDZ	+33.2		+11.6		-4.2		+38.4	
CASPT2/*cc*-pVDZ//B3LYP/*cc*-pVDZ	+22.8		+1.0		+7.8		+9.8	
MR-CI+Q/*cc*-pVDZ//B3LYP/*cc*-pVDZ	+26.6		+4.1		+9.8		+18.0	
(10,10)CASSCF/VTZP//B3LYP/*cc*-pVDZ	+24.3		+2.1		-15.0		+27.7	
CASPT2/VTZP//B3LYP/*cc*-pVDZ	+23.7		+1.5		+7.8		+9.0	
MR-CI+Q/VTZP//B3LYP/*cc*-pVDZ	+26.0		+3.0		+7.7		+15.9	

[a] All energies are given in kcal·mol^{-1}. Geometries for **9a** and **9c** were calculated in C_1 symmetry for the 6-31G(d) and 6-311G(d) basis set and in C_s symmetry for the *cc*-pVDZ basis set. **9d** was computed in C_1 symmetry for all basis sets used. The transition state **9e** for the inversion of the cyclic allene **9d** could only be located in C_s symmetry.

[b] Reactant and closed-shell species **9d** and **9e** were calculated using spin restricted MP2 while for the diradicals **9a** and **9c** spin unrestricted UMP2 was applied. Spin-projected MP2 (PMP2) worsens the results dramatically.

Summarizing our evaluation, in the present system, B3LYP is applicable to locate accurate transition structures and reaction energies. As found in other examples, BLYP seems to be inadequate. As expected, the CASSCF and the MP2 approaches lead to completely wrong results showing that both static and dynamical correlation effects have to be accounted for.

Our computations show that for the parent system **9**, the Moore reaction ($C^2 - C^7$ cyclization) is kinetically and thermodynamically strongly favored with respect to the $C^2 - C^6$ cyclization. In comparison to Z-1,2,4-heptatriene-6-yne which represents the parent system of enyne-allenes, **9** possesses a much higher activation energy for the $C^2 - C^6$ cyclization ($41 \, \text{kcal} \cdot \text{mol}^{-1}$ vs. $31 \, \text{kcal} \cdot \text{mol}^{-1}$) while the barrier for the $C^2 - C^7$ cyclization is reduced ($13.3 \, \text{kcal} \cdot \text{mol}^{-1}$ vs. $24.0 \, \text{kcal} \cdot \text{mol}^{-1}$). This behavior may arise from a diminished electron density in the enyne-ketene compared to the enyne-allene which favors the $C^2 - C^7$ cyclization and stipulates the $C^2 - C^6$ cyclization. Both enyne-ketene cyclizations are more endothermic by about $10 \, \text{kcal} \cdot \text{mol}^{-1}$ than their counterparts in the enyne-allenes. While $\Delta_{\text{R}} G_{298}$ of the $C^2 - C^6$ cyclization is predicted to be $27 \, \text{kcal} \cdot \text{mol}^{-1}$ for the enyne-ketene **9a**, a value of about $13 \, \text{kcal} \cdot \text{mol}^{-1}$[139] is found for the enyne-allenes. For the $C^2 - C^7$ cyclization, the corresponding numbers are about $6 \, \text{kcal} \cdot \text{mol}^{-1}$ and $-11 \, \text{kcal} \cdot \text{mol}^{-1}$, respectively[7].

6.2.3   Influence of the substituents

From the structural similarity between enyne-allenes and enyne-ketenes, a comparable reaction scope can be expected. Indeed, although **9** reacts in a $C^2 - C^7$ cyclization leading to 1,4-benzoquinone, for the phenyl-substituted enyne-ketene **11** (see also Figure 6.13), Moore and co-workers found that the product of the competing five-membered $C^2 - C^6$ cyclization represents the major product (46% yield vs. 21%).[170]

This regioselectivity resembles the regioselectivity found for enyne-allene systems[139] but is less pronounced. For an understanding of differences and similarities between these two compound classes, the reaction barriers for various model compounds were computed (Table 6.6). If the hydrogen of the enyne terminus in the parent compound **9** is replaced by a phenyl group (**10**), the activation barrier for the $C^2 - C^6$ cyclization ($\Delta G_{298}^{\neq} = 28.7 \, \text{kcal} \cdot \text{mol}^{-1}$) decreases by about $13 \, \text{kcal} \cdot \text{mol}^{-1}$ compared to the parent system **9** ($\Delta G_{298}^{\neq} = 41.4 \, \text{kcal} \cdot \text{mol}^{-1}$).

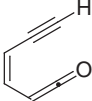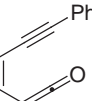[7]MR-CI+Q computations employing a *cc*-pVTZ basis set neglecting the set of f-functions on carbon.[175]
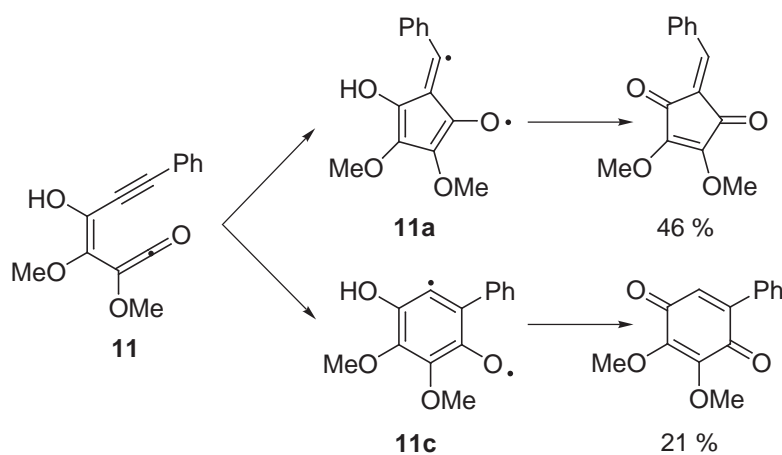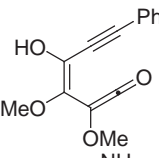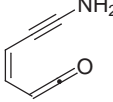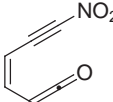
Figure 6.13: Experimental regioselectivity of **11**.

This reduction of the free energy of activation for the $C^2 - C^6$ cyclization of **10** is more pronounced than in the enyne-allene compound for which Ph substitution diminished $\Delta G^{\ne}_{298}$ only from $31.4\,\mathrm{kcal \cdot mol^{-1}}$ to $28.7\,\mathrm{kcal \cdot mol^{-1}}$. The increase of the reaction barrier for the $C^2 - C^7$ cyclization of **10** by about $6\,\mathrm{kcal \cdot mol^{-1}}$ ($13.6\,\mathrm{kcal \cdot mol^{-1}}$ for **9** vs. $19.8\,\mathrm{kcal \cdot mol^{-1}}$ for **10**) is of the same order of magnitude as was computed for the corresponding enyne-allenes ($\Delta G^{\ne}_{298} = 24.0\,\mathrm{kcal \cdot mol^{-1}}$ to $29.8\,\mathrm{kcal \cdot mol^{-1}}$). Although the influence of the phenyl group is considerably larger for the enyne-ketenes than for the enyne-allenes, **10** prefers the Moore reaction by about $10\,\mathrm{kcal \cdot mol^{-1}}$ over the $C^2 - C^6$ cyclization. This prediction also holds, if it is considered that the B3LYP approach seems to underestimate the barrier of the $C^2 - C^7$ cyclization by about $4\,\mathrm{kcal \cdot mol^{-1}}$, while the barrier of the $C^2 - C^6$ cyclization seems to be predicted in a better accord with higher level CCSD(T) computations.

The switch from the $C^2 - C^7$ to the $C^2 - C^6$ cyclization is found, if the OH and OMe substituents being present in the experimentally studied compound, are taken into account (**11**). These oxygen-containing substituents further decrease the free energy of activation for the $C^2 - C^6$ cyclization from $\Delta G^{\ne}_{298} = 28.7\,\mathrm{kcal \cdot mol^{-1}}$ (**10**) to $\Delta G^{\ne}_{298} = +18.7\,\mathrm{kcal \cdot mol^{-1}}$ (**11**) while $\Delta G^{\ne}_{298}$ for the $C^2 - C^7$ cyclization raises slightly from $19.8\,\mathrm{kcal \cdot mol^{-1}}$ (**10**) to $21.9\,\mathrm{kcal \cdot mol^{-1}}$ (**11**), leading to a preference of the $C^2 - C^6$ cyclization and explaining the experimentally found regioselectivity.

The influence of the phenyl group on the reaction barriers of both cyclization modes is much stronger than that found for the enyne-allenes but is not sufficient to induce the shift from the $C^2 - C^7$ to the $C^2 - C^6$ cyclization. As already mentioned above, this difference from the enyne-allene systems is a consequence of the much higher $C^2 - C^6$ cyclization barrier

Table 6.6: Activation barriers for substituted enyne-ketenes **9** - **13** at the B3LYP/6-31G(d) level of theory[a].

System	$C^2 - C^6$ cyclization			$C^2 - C^7$ cyclization		
	$\Delta E^{\ddagger}$	$\Delta H^{\ddagger}_{298}$	$\Delta G^{\ddagger}_{298}$	$\Delta E^{\ddagger}$	$\Delta H^{\ddagger}_{298}$	$\Delta G^{\ddagger}_{298}$
9	+41.2	+39.4	+41.4	+13.3	+12.1	+13.6
10	+28.1	+26.4	+28.7	+18.8	+17.3	+19.8
11	+18.9	+17.4	+18.7	+20.7	+19.2	+21.9
12	+8.8	+7.6	+10.2	+11.4	+10.4	+13.5
13	+27.8	+25.8	+27.4	+16.9	+15.4	+17.5

[a] All energies are given in $kcal \cdot mol^{-1}$.

found for the parent system **9** which in turn may result from the presumably diminished electron density in the enyne-ketene. The shift from the $C^2 - C^7$ to $C^2 - C^6$ cyclization takes place if the lower electron density is compensated by the additional donor groups OH and OMe which reduce the barrier of the $C^2 - C^6$ cyclization and increase the barrier of the $C^2 - C^7$ cyclization. This is in line with the finding of Brunette and Lipton[144] that oxyanion substitution significantly accelerates the $C^2 - C^6$ cyclization in enyne-allenes.

To investigate similarities and differences between enyne-allenes and enyne-ketenes in more detail, we also studied the influence of strong electron-donating substituents (NH_2 in **12**) and of electron-withdrawing substituents (NO_2 in **13**). For enyne-allenes both substituents preferred the $C^2 - C^6$ cyclization but possessed a carbenelike instead of a diradical intermediate.[139, 175]

Compared to the parent compound **9** ($\Delta G^{\ddagger}_{298} = 41.4$ kcal $\cdot$ mol^{-1}), the amino-substituted enyne-ketene **12** possesses a dramatically lower free energy of activation ($\Delta G^{\ddagger}_{298} = 10.2$ kcal $\cdot$ mol^{-1}) for the $C^2 - C^6$ cyclization. In contrast, the activation barrier for the $C^2 - C^7$ cyclization remains almost constant at about 13.5 kcal $\cdot$ mol^{-1}. In agreement with the

Table 6.7: $\Delta_R G_{298}$ for intermediates representing local minima of enyne-ketenes **9** - **13** computed at the B3LYP/6-31G(d) level of theory[a)].

		Intermediate			
	System	a	b	c	d
9		+24.8		+4.8	+9.6
10		+17.7		+9.6	+15.0
11		+0.5		-4.2	+2.7
12			+5.8	-4.5	+4.4
13			+18.4	+3.5	+7.2

[a)] All energies are given in kcal·mol^{-1}. **a** - **d** correspond to the intermediates denoted in Figure 6.11.
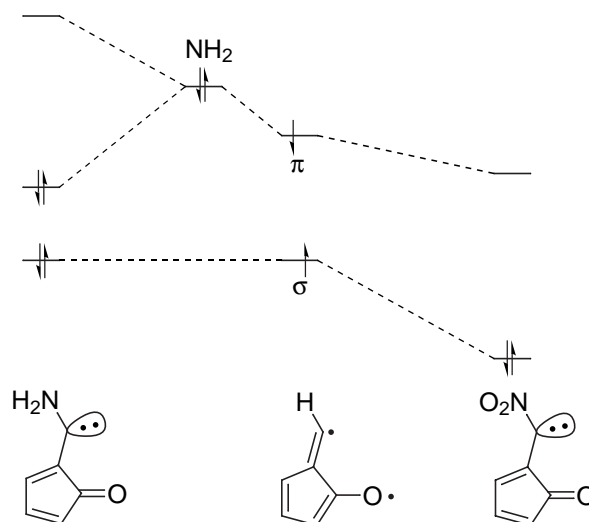
Figure 6.14: Orbital scheme for the effect of electron-donating and electron-withdrawing substituents on diradicals.

enyne-allene analogue, the NH_2 substituent leads to a preference of the $C^2 - C^6$ cyclization, but for the enyne-allene the activation barriers were found to be significantly higher ($\Delta G^{\ne}_{298}$ = 17.8 kcal·mol^{-1} for the $C^2 - C^6$ cyclization and $\Delta G^{\ne}_{298}$ = 22.7 kcal·mol^{-1} for the $C^2 - C^7$

cyclization).[139] If the electron-donating amino group is replaced by the strong electron-withdrawing nitro substituent, the activation barrier for the $C^2 - C^6$ cyclization decreases from 41.4 kcal·mol^{-1} (**9**) to 27.4 kcal·mol^{-1} (**13**). This change is insufficient to switch the regioselectivity of the reaction as the increase in the free energy of activation $\Delta G_{298}^{\neq}$ for the $C^2 - C^7$ cyclization by about 3-4 kcal·mol^{-1} to 17.5 kcal·mol^{-1} (**13**) is too low.

Both the electron-donating NH_2 and the electron-withdrawing NO_2 lower the activation barriers and lead to carbenelike intermediates for the $C^2 - C^6$ cyclization, *i.e.* change the reaction mechanism. The reaction energies are given in Table 6.7. While **9a** was shown to represent a σ, π-diradical, for **12** a carbenelike intermediate is found which possesses a lone-pair σ-orbital centered at the carbon atom in α-position to the NH_2 substituent. Figure 6.14 sketches the results obtained from an NBO analysis.[183] The preference for the carbene structure over the diradical can be explained by π-conjugation due to the planar NH_2 substituent. This interaction lowers the π-orbital and leads to doubly occupied σ and π orbitals (see Figure 6.14). Consequently, in contrast to **9-11**, the $C^2 - C^6$ cyclization of **12** has to be looked upon as a carbenelike reaction mechanism. In contrast, the main intermediate for the Moore reaction still remains a diradical. It is predicted to lie below the cyclic allene by about 9.0 kcal·mol^{-1}.

Also the $C^2 - C^6$ intermediate of **13** possesses a carbenelike electronic structure; the reasons are different however. As indicated by NBO analyses, the NO_2 substituent lowers the σ-orbital by its strong electron-withdrawing effect. As a consequence, the energetic difference between the σ- and π-orbital is enlarged, leading to a double occupancy of the σ-orbital. Similar effects are also found for the enyne-allene.[175]

6.2.4   Summary

In this section, the Moore ($C^2 - C^7$) cyclization and its regioalternative $C^2 - C^6$ cyclization were investigated. While the parent system **9** is studied with DFT and high-level *ab initio* methods (MR-CI+Q, CASPT2, CCSD(T)) to validate the less reliable approaches, the substituent effects are included on the DFT level of theory. To study similarities to the other diradical reactions discussed previously, the present results are compared to diradical cyclizations of enyne-allenes.

For the parent system **9**, the $C^2 - C^7$ cyclization is kinetically and thermodynamically strictly favored with respect to the $C^2 - C^6$ cyclization ($\Delta G^{\ddagger}_{298} = 41$ kcal $\cdot$ mol^{-1} vs. $\Delta G^{\ddagger}_{298}$ $= 18$ kcal $\cdot$ mol^{-1} and $\Delta_R G_{298} = 3$ kcal $\cdot$ mol^{-1} vs. $\Delta_R G_{298} = 26$ kcal $\cdot$ mol^{-1}). The $C^2 - C^7$ cyclization proceeds *via* the diradical intermediate **9c** which is computed to be 5 kcal $\cdot$ mol^{-1} more stable than the cyclic allene **9d** and 13 kcal$\cdot$mol^{-1} lower in energy than the zwitterionic state **9e**.

The main difference between **9** and the parent system of the enyne-allenes (*Z*-1,2,4-heptatriene-6-yne) is the barrier to the $C^2 - C^6$ cyclization (41 kcal$\cdot$mol^{-1} for **9**, 31 kcal$\cdot$mol^{-1} for *Z*-1,2,4-heptatriene-6-yne). Furthermore, both possible cyclization modes for the enyne-ketene are more endothermic than those for the enyne-allene. Both may originate from the diminished electron density in the enyne-ketene compared to the enyne-allene which seems to favor the $C^2 - C^7$ cyclization and disfavors the $C^2 - C^6$ cyclization.

Substituent effects are more pronounced for the enyne-ketenes than for the enyne-allenes. For example, the reaction barrier for the $C^2 - C^6$ cyclization decreases by 13 kcal $\cdot$ mol^{-1} if the hydrogen of the alkyne terminus is substituted by a phenyl group. For the enyne-allene, the barrier decreases by only 3 kcal $\cdot$ mol^{-1}. However, due to the strong difference found for the parent system **9**, the shift from the $C^2 - C^7$ to the $C^2 - C^6$ cyclization only takes place if additional OH and OMe substituents (**11**) are present.

The findings about the influence of strong electron-donating or -withdrawing substituents underline the strong similarities between enyne-ketenes and enyne-allenes. Both NH$_2$ (**12**) and NO$_2$ (**13**) at the alkyne terminus lead to carbenelike intermediates instead of diradicals, but the reasons are different, however.

A model to predict the electronic structure of the intermediates that might occur in the course of the Moore reaction is presented in the next section and derived from a series of cyclic allenes.

6.3   A Comprehensive Model for the Electronic Structure of 1,2,4-Cyclohexatriene and Related Compounds[110]

6.3.1   Introduction

Trends in the periodic table of elements and the concepts related with them have become the foundation of nowadays chemists thinking. Due to their simplicity and flexibility they represent powerful tools to explain variations in chemical structure and reactivity. However, quite often such trends are hidden behind the complexity of the properties and the structures of the various members belonging to a class of molecules. One example is the class of six-membered cyclic allenes represented in Figure 6.15, which recently gathered much interest as intermediates in the cyclization of 1,3-diene-5-ynes,[184] the rearrangement of carbenes[185–188] or as heterobenzene isomers.[189–191] Their complexity is expressed in the chemical properties that, being dependent on the fragment X, range from those of strained allenes (X = CH_2)[174, 189, 192, 193] to those of zwitterionic systems (X = NH).[185, 186, 190] For X = O the chemical behavior varies between both extremes depending on whether the six-membered ring or its benzo-derivative is studied.[174, 194] Systems with heteroatoms from the third row were also investigated to some extent. The role of a cyclic allene with X = S formed by rearrangement of a sulfur-substituted cyclopropylidene was investigated by Shevlin *et al.*[186, 187] Regitz *et al.* isolated a bulky substituted diphosphaisobenzene, which possesses substantial allene character.[191] This finding was supported by quantum chemical computations.[191] Taking into account the rather similar composition of the compounds, the variations in their chemical nature are unexpected.
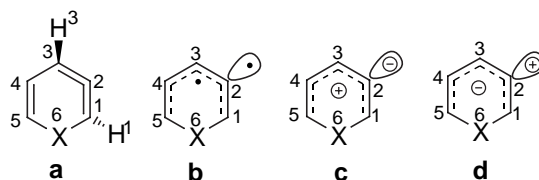

Figure 6.15: Possible electronic structures of cyclic allenes (**a** - **d**). The denotation of the planar zwitterionic species **c** and **d** will be used to reflect the reversed polarity compared to each other.
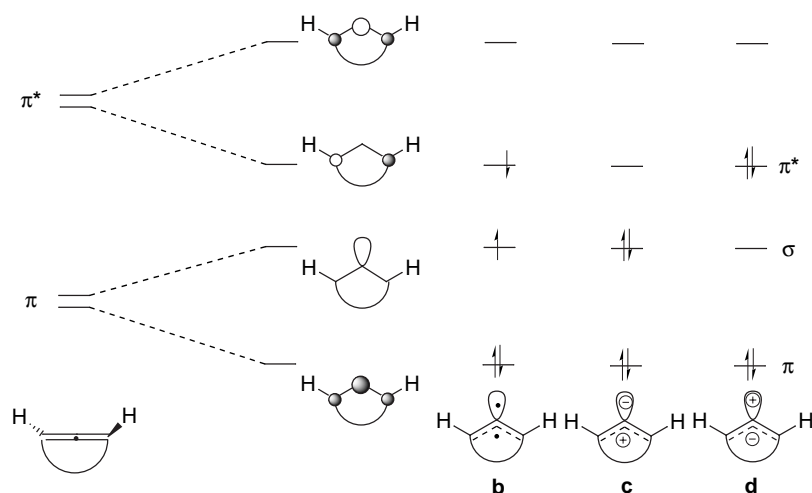
Figure 6.16: Orbital diagram for strain-induced orbital splitting of the allene.[195]

As shown previously,[174, 193] an explanation for the variations in the electronic structure can only be obtained, if, aside from the electronic structure at equilibrium geometry, also the electronic characters of the low lying electronic states at the planar geometry are included into the consideration. The orbital diagram from which these electronic states can be derived (see Figure 6.15) is depicted in Figure 6.16 [8]. The ring strain affects the allene moiety to that extent as it deviates from linearity and the bonding planes of the termini are arranged no longer perpendicular as in the parent allene C_3H_4. For twisted ring geometries the character of an allene is maintained (Figure 6.15, structure **a**) However, as it is shown in Figure 6.16, originally proposed by Johnson[195] for saturated ring systems, geometric constraints lift the degeneracy of the π and π^* orbitals (depicted on the left hand side of Figure 6.16), and a set of 3 π orbitals and 1 σ orbital is obtained. The three lowest lying electronic states at planar geometry are obtained from the occupation pattern of these orbitals. The diradical state **b** originates from double occupation of the π orbital and single occupation of the σ and π^* orbitals. Pairing of the latter two electrons in the σ orbital then leads to the zwitterion **c** possessing a formally negatively charged central carbon atom (C^2 depicted in Figure 6.15) and a formal positive charge in the π system of the allene moiety. In the following all states with this formal charge distribution will be abbreviated as state **c**. If the occupation of the σ orbital and the π^* orbital of **c** is switched, a second zwitterion with reversed polarity (**d**) is obtained.

In the present work the known considerations about the chemical nature of cyclic allenes

---

[8]The discussion of the orbital diagram for a planar allene fragment can be found in [195].

are extended by analyzing variations within the series X = BH (**14**), CH_2 (**15**), NH (**16**), and O (**17**), and the series of the higher homologues with X=AlH (**18**), SiH_2 (**19**), PH (**20**,**21**), and S (**22**). With these data, which were obtained by high-level *ab initio* computations of the low lying electronic states, a comprehensive molecular orbital model is proposed in which the orbitals of the fragments (allenic unit, ethylene moiety and fragment X) are successively combined to the orbitals of the entire system. This model fully explains the variations depending on the fragment X and shows their relationships to trends in the periodic table of elements. The model also reveals an unexpected connection between the class of cyclic allenes and the phenylanion (X = CH^-). Finally, it is analyzed whether a model based on simple orbital energies actually contains all important aspects or if more subtle interactions also have to be considered (*e.g.* geometry relaxation, electron correlation)

6.3.2   Assessment of Theoretical Methods

A reliable computation of energy differences between species that possess a distinctly different electronic structure (closed-shell vs. open-shell) is quite demanding. In the present work, the chemical interpretation of the results crucially depends on a balanced description of the two closed-shell states (allene and zwitterion) and the one open-shell diradical state[9]. To assess the quality of several theoretical approaches, the methods (U)B3LYP,

---

[9]For a reliable description of systems which possess diradical character, a multi-reference treatment is essential in most cases. Since the planar diradical species are significant for the understanding of the chemistry of cyclic allenes,[174, 193] relative electronic energies of all species were computed at the Multi-Reference Configuration Interaction (MR-CI) and the CASPT2 level of theory employing (U)B3LYP geometries. Orbitals for the CASPT2 and MR-CI approaches were obtained by CASSCF calculations using the MOLCAS 5.0 package.[51] A balanced description of the active spaces in the CASSCF procedure chosen for all compounds consisted of the two highest lying occupied π orbitals (a'') with the corresponding number of virtual orbitals and an appropriate number of σ-orbitals (a') including the occupied σ-orbital at C^2. This choice ensures that the allene includes the same σ-orbitals in the active space as the planar species. With this ansatz [6,6] (X = CH^-, NH^{2+}), [8,8] (X = BH, BH^{2-}, CH_2, NH, O, AlH, PH, S) and [10,10] (X = SiH_2, PH with an additional phosphorus in the ring) CASSCF spaces resulted. For the diradical ^1A states of both phosphorus containing compounds the orbitals for the MR-CI computations were taken from the ^3A instead of a ^1A CASSCF calculation. Test computations indicated that the use of these orbitals is appropriate as the error in the MR-CI+Q computations employing optimized triplet orbitals instead of optimized singlet orbitals for the open-shell singlet wavefunction is within the expected error bar for this method of about 1-2 kcal·mol^{-1}. All CASSCF and CASPT2 single-point computations were performed with the MOLCAS

CCSD(T), CASPT2 and MR-CI+Q were employed. (U)B3LYP and CCSD(T) represent single-reference approaches, while CASPT2 and MR-CI+Q are multi-reference methods. The data of the (U)B3LYP, CASPT2, MR-CI+Q and CCSD(T) computations are provided in Table 6.8.

MR-CI+Q, CASPT2 and (U)B3LYP used in this work agree quite well on the energy separation of the states. The energy difference regarding the zwitterion with respect to the allene structure **a** agrees well for MR-CI+Q and CCSD(T) (deviation of maximum $2\,\mathrm{kcal\cdot mol^{-1}}$). As the CCSD(T) ansatz should be the most reliable for well-behaved closed-

---

5.0 package.[51]

The MR-CI approach used in this study is based on a direct individually selecting MR-CI algorithm[32] that computes the MR-CI energy for a reduced dimension of the CI eigenvalue problem while the contributions of the neglected configuration state functions (CSF) were taken into account by the Buenker-Peyerimhoff extrapolation scheme.[30, 31] The influence of quadruple excitations was estimated by the normalized form of the Davidson correction.[196] These calculations abbreviated as MR-CI+Q were performed with the DIESEL-CI program package.[53] The reference spaces of the individually selecting MR-CI computations were determined iteratively and consisted of up to nine reference configuration state functions, from which a configuration space of up to $47 \cdot 10^6$ CSF was generated. The number of secular equations to be solved in the reduced CI problem was up to the order of $12 \cdot 10^6$. The ano-l[180, 197] basis set of triple-ζ quality in a 5s3p1d contraction for the atoms of the second and the third row and a 3s1p contraction scheme for hydrogen was used for all CASSCF, CASPT2 and MR-CI+Q calculations.

CCSD(T) computations for the energy differences between the allene and the zwitterion were done with the MOLPRO package.[52] For these computations Dunnings cc-pVTZ[164] basis set with d-functions omitted for the hydrogen atoms was employed and is denoted as cc-pVTZ(f/p).

The geometric parameters of the stationary points were optimized using analytical gradients of the density functional theory (DFT) utilizing the B3 exchange expression[160, 198, 199] in combination with the correlation functional by Lee, Yang, and Parr (LYP)[161] and Dunning's cc-pVDZ[164] basis set. This approach was found to give more appropriate geometries for a post Hartree-Fock treatment than the CASSCF method.[174, 193] The singlet diradical species were optimized using broken-spin symmetry determinants in the unrestricted ansatz ($\langle S^2 \rangle$ approximately 1), while the planar and the non-planar closed-shell species were computed within a restricted approach. The minima were optimized without symmetry constraints while the geometries of the planar singlet species were determined in C_s symmetry. If C_1 and C_s optimized geometries were identical the latter were used for the subsequent computations. The planar species of the phosphorus systems were optimized constraining the dihedral angle of the allene moiety to $0°$ as the full C_s symmetry because of the sp^3 hybridized phosphorus center could not be retained. The nature of the various stationary points was analyzed by calculation of the DFT harmonic frequencies. Vibrational analyses were also utilized at the (U)B3LYP level to determine the thermal corrections to ΔG_{298}. All DFT calculations were performed with the *Gaussian98* program package.[159]

Table 6.8: Energies of the electronic states of 1,2,4-cyclohexatriene and its hetero analogues as indicated in Figure 6.15[a)].

X	B3LYP[b)]		CASPT2[c)]		MR-CI+Q[d)]		CCSD(T)[e)]
	diradical	zwitterion	diradical	zwitterion	diradical	zwitterion	zwitterion
BH (**14**[f])	+11	+1	+11	+1	+11	+2	+1
CH$_2$ (**15**[g])	+3	+23	+10	+33	+10	+29	+29
NH (**16**[g,h])	+22	0	+22	0	+24	0	n.a.
O (**17**[g])	+16	+1	+15	+4	+18	+3	+2
AlH (**18**[f])	+16	+5	+16	+7	+18	+7	+7
SiH$_2$ (**19**[f,i])	+10	+21	+12	+26	+13	+28	+26
PH (**20**[g])	+8	+14	n.a.	n.a.	+12	+16	+15
PH (**21**[g])	+16	+22	n.a.	n.a.	+21	+21	+22
S (**22**[g])	+9	+7	+11	+8	+14	+10	+10

[a)] All values are given in kcal $\cdot$ mol^{-1} with respect to the allene structure except for

16.

[b)] (U)B3LYP/*cc*-pVDZ//(U)B3LYP/*cc*-pVDZ

[c)] CASPT2/*ano-l*//(U)B3LYP/*cc*-pVDZ

[d)] MR-CI+Q/*cc*-pVDZ//(U)B3LYP/*cc*-pVDZ

[e)] CCSD(T)/*cc*-pVTZ(f/p)//(U)B3LYP/*cc*-pVDZ

[f)] zwitterion **d**, see Figure 6.15

[g)] zwitterion **c**, see Figure 6.15

[h)] zwitterion is ground state

[i)] zwitterion **c** is 49 kcal $\cdot$ mol^{-1} higher than allene **a**

shell systems, this agreement validates the reliability of the MR-CI+Q computations for the allenes and zwitterionic states. CASPT2 shows a slightly inferior performance concerning the energy differences to the zwitterions. (*e.g.* 33 kcal$\cdot$mol^{-1} vs. 29 kcal$\cdot$mol^{-1} (CCSD(T), MR-CI+Q) for **15** or +4 kcal $\cdot$ mol^{-1} vs. +2 kcal $\cdot$ mol^{-1} (CCSD(T)) and +3 kcal $\cdot$ mol^{-1} (MR-CI+Q) for **17**), while it predicts comparable energies for the diradical states compared with MR-CI+Q (*e.g.* 10 kcal $\cdot$ mol^{-1} for CASPT2 and MR-CI+Q for **15**, but a slightly underestimated value of 15 kcal $\cdot$ mol^{-1} compared to 18 kcal $\cdot$ mol^{-1} obtained with MR-CI+Q for X=O). (U)B3LYP does well for the zwitterionic and the diradical states except for

X = CH$_2$, for which both the zwitterion and the diradical are substantially underestimated (3 kcal·mol^{-1} for the zwitterion and 23 kcal·mol^{-1} for the diradical).

From this assessment the MR-CI+Q approach is likely to be the most appropriate to provide a comparative and uniform basis for the discussion of all systems. For this purpose also the computations for the known systems were redone in this work and the MR-CI+Q energies presented here deviate slightly (1 to 2 kcal·mol^{-1}) from previous publications.[174, 193]

6.3.3 Results
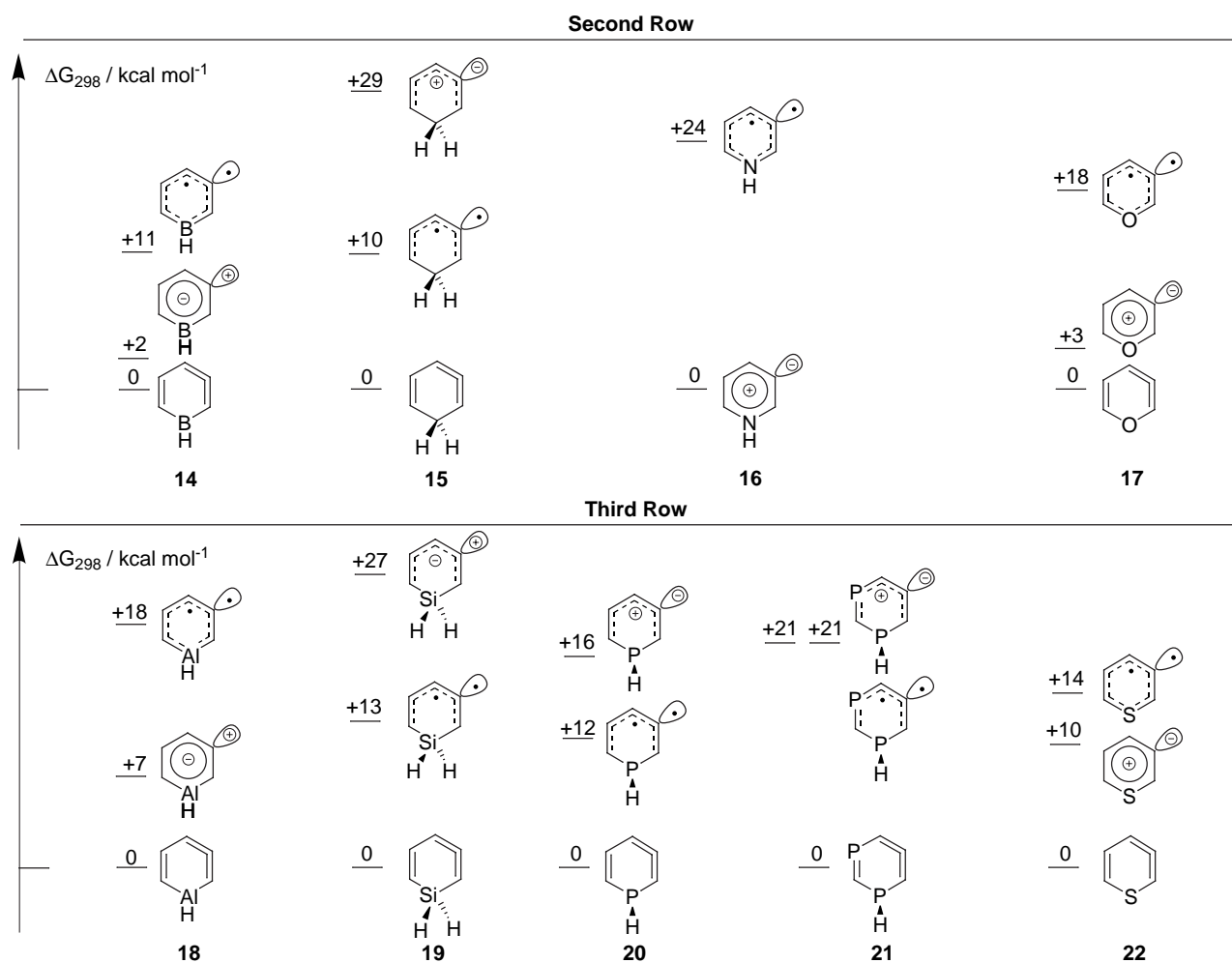

Figure 6.17: Energies and characters of the electronic states of 1,2,4-cyclohexatriene and its hetero analogues.

The influence of X on the electronic structure of the cyclic allenes can be taken from Figure 6.17 which summarizes the computed MR-CI+Q data of the present work. Some geometric parameters are presented in Table 6.9. Figure 6.17 gives the energetic positions

of the low lying electronic states at planar ring geometry with respect to the allene structures which possess puckered ring geometries. In addition to the parent compound 1,2,4-cyclohexatriene and systems with one heteroatom, the parent system of the experimentally synthesized diphosphaisobenzene of Regitz[191] was also studied to evaluate the influence of the second phosphorus center in the ring system.

Figure 6.17 and Table 6.9 show that in the minimum[10] most compounds represent strained allenes with a more or less twisted ring geometry. The second lowest lying species (diradical of zwitterion) represents a transition state for the racemization of the allene. While most systems agree in this respect, the planarisation energies strongly depend on the fragment X. For X = BH (**14**) and X = O (**17**) these energies are very small (2-3 kcal $\cdot$ mol^{-1}) while it vanishes for the isopyridine **16** (X = NH), which hence prefers a zwitterionic ground state (**16c**, see Figures 6.15 and 6.16) as equilibrium structure[185–187,190,193] The planarisation energy amounts 10 kcal $\cdot$ mol^{-1} for X = CH$_2$ (**15**). The planarisation energies correlate with the dihedral angle of the allene moiety H^1 − C^1 − C^3 − H^3, which is a measure for the distortion relative to an unstrained allene possessing a value of 90°. For X = CH$_2$ (**15a**), which possesses the largest planarisation energy for the second row systems, the dihedral angle (see Table 6.9) deviates from the optimum by only 10°. For X = BH (1a). For X = BH (**14a**) and X = O (**4a**), that both possess a smaller barrier, this difference increases to 32° and 49°, respectively.

The planarisation energy rises on going from second to third row fragments X, *e.g.* from X = BH (**14**) to X = AlH (**18**) from 2 kcal $\cdot$ mol^{-1} to 10 kcal $\cdot$ mol^{-1}. This increase in the energy difference is accompanied by a larger diameter of the six-membered monocycle containing third row elements with respect to the second row systems. The diameter rises according to Table 6.9, since the X − C^1 bond length increases *e.g.* from 160 pm for X = BH (**14**) to 201 pm in X = AlH (**18**) and a similar change is obvious for the X − C^5 distance.

Aside from the planarisation energies also the sequences of the electronic states at planar geometry depend on the fragment X. While the diradical state is favored over the zwitterionic one for **14**, **19**, and **20**, the reverse order is found for **14**, **16**, **17**, **18**, and **22**. The energy of

---

[10]The term minimum is used in the context with reference to isomers that have the same connectivity as the structure of the cyclic allenes and represents a local rather than a global minimum on the potential energy surface in most cases considered. The global minima of the C$_6$H$_6$ and C$_5$H$_6$N families, for example, to which **15** (X = CH$_2$) or **16** (X = NH), respectively belong, are surely the benzene and the pyridine.

the diradical and the zwitterionic state are not distinguishable for the phosphorus compound **21**. The additional phosphorus center in the ring (**21** vs. **20**) destabilizes the diradical state by $9 \, \text{kcal·mol}^{-1}$ while the zwitterionic state is destabilized by $5 \, \text{kcal·mol}^{-1}$. The energy gaps between zwitterionic and diradical states range for second row fragments X from $9 \, \text{kcal·mol}^{-1}$ in **14** to $24 \, \text{kcal·mol}^{-1}$ in **17**. No systematic variation in the gaps are found when the second and the third row systems are compared with each other. For instance, X = AlH (**18**) has a slightly larger energy gap ($11 \, \text{kcal} \cdot \text{mol}^{-1}$) than X = BH (**14**, $9 \, \text{kcal} \cdot \text{mol}^{-1}$) but on going from X = O (**17**) to X = S (**22**) the difference between zwitterion and diradical is reduced from $15 \, \text{kcal} \cdot \text{mol}^{-1}$ to $4 \, \text{kcal} \cdot \text{mol}^{-1}$.

The situation is even more complex as the character of the zwitterionic states also depends on X. The electronic structure of the zwitterionic states of the oxygen (**17**) and the nitrogen (**16**) systems is characterized by a formally positively charged ring and a negatively charged central "allene" carbon atom (Figure 6.15, electronic structure **c**).[174, 193] For the zwitterionic state of X = CH_2 the same polarity was found.[174] Concerning X = BH (**14**) the computations predict a reversed polarity, which corresponds to the zwitterionic state **d** in Figure 6.15. The opposite formal polarity of the zwitterions of **14** and **16** is nicely mirrored in the computed dipole moments (B3LYP level of theory). The dipole moment of **14d** is predicted to be 1.9 D with a direction that points through the molecular center towards C^2. This direction of the dipole moment is in line with a formally negatively charged ring and a positively charged central "allene" carbon atom. The dipole moment of **16c** has a value of 5.5 D and points into the opposite direction reflecting the reverse polarity of the zwitterion (see Figure 6.15). The polarity of the zwitterionic states of the second row systems and that of their third row homologues are the same with an interesting exception found for X = SiH_2.

In contrast to the geometric parameters of the allenes **a**, which correlate with the varying planarisation energies, a similar relationship between the trends in the geometric parameters described above and the varying sequences of the planar states is not obvious (see Figure 6.17 and Table 6.9). As found for the allene structure as well as for the states at planar geometries, the $X - C^1$ distances of the various compounds decrease systematically from B to O (*e.g.* 157 pm in **14b** to 140 pm in **17b**) and also from Al to S (*e.g.* 197 pm in **18b** to 178 pm in **22b**). The same sequencees are found for the corresponding $X - C^5$ bond length. While these parameters are expected to depend on X, all other parameters vary less

Table 6.9: Geometric parameters of stationary points [a].

System	d(C^1-C^2)	d(C^2-C^3)	d(X-C^1)	d(X-C^5)	∠(C^1-C^2-C^3)	∠(H^1-C^1-C^3-H^3)
14a	130	135	160	152	152	68
14b	136	140	157	156	128	0
14d	129	136	162	150	155	0
15a	133	134	152	153	131	100
15b	135	140	152	150	128	0
15c	140	143	149	148	110	0
16b	136	139	142	139	128	0
16c	140	142	137	135	111	0
17a	136	139	139	134	116	51
17b	135	139	140	138	127	0
17c	138	142	138	133	111	0
18a	130	134	201	194	158	89
18b	135	141	197	196	137	0
18d	127	136	209	190	162	0
19a	132	133	190	192	150	101
19b	135	140	189	188	134	0
19d	128	135	200	185	160	0
20a	132	133	185	184	144	94
20b	135	139	184	181	133	0
20c	134	140	185	175	130	0
22a	134	135	178	176	132	83
22b	135	139	178	176	132	0
22c	139	142	177	170	115	0

[a] Distances are given in pm and angles in degrees. The denotation of the species and the atomic centers can be taken from Figure 6.15 and Figure 6.17.

with X but depend on the state under consideration. For the diradical states of all systems (second and third row) the $C^1 - C^2$ and the $C^2 - C^3$ distances are computed to be around 135 pm and 140 pm, respectively. In the zwitterionic states **c** these parameters are 138-140 pm ($C^1 - C^2$) and 140-142 pm ($C^2 - C^3$) with the exception of **20c** having a considerably smaller $C^1 - C^2$ distance of 134 pm and a $C^2 - C^3$ bond length of 140 pm. The zwitterionic

states **d**, which show a reversed polarity, possess bond lengths around 128 pm and 136 pm for $C^1 - C^2$ and $C^2 - C^3$, respectively. The most characteristic parameter to describe the structural differences between the zwitterionic states **c** and **d** and the diradical state **b** is the bond angle $C^1 - C^2 - C^3$ of the allene moiety. For all systems with fragments X from the second row, the $C^1 - C^2 - C^3$ angle of the zwitterionic state **c** is nearly constant around 111° while the only zwitterionic state with reversed polarity (X = BH, **14d**) is predicted to be 155°. The diradical states **b** possess bond angles around 128°. A similar trend also exists for the third row systems. The zwitterions **18d** and **19d** (X = AlH and SiH_2) possess similar bond angles $C^1 - C^2 - C^3$ of 160° and 162°, respectively, while for X = PH (**20c**) and X = S (**22c**) this angle is significantly smaller (130° for X=PH and 115° for X = S).
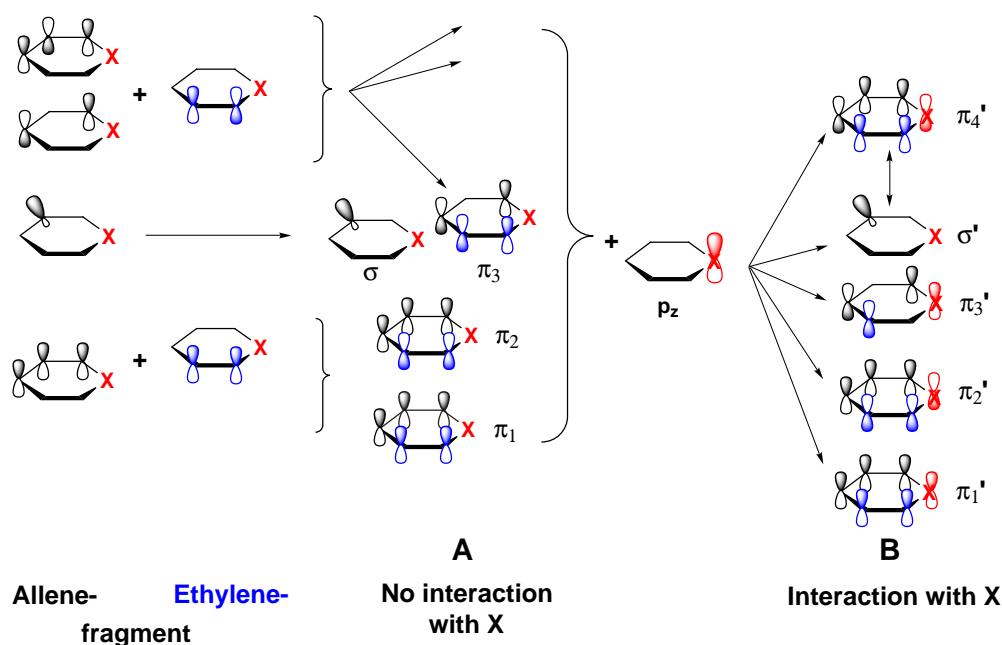
Figure 6.18: Schematic orbital diagram based on the main three bond moieties (allene, ethylene and X). The left-hand side (allene fragment) corresponds to the Figure 6.16.[195] Linear combination of the allene and the ethylene fragment extends the orbital scheme to set **A** in which fragment X has no contribution. Interaction of **A** with fragment X generates the orbital set **B** on the right-hand side.

6.3.4 Discussion

A general model that explains the varying planarisation energies and the sequence of the states at planar geometry can be obtained from an orbital diagram (Figure 6.18) that

builds up the molecular orbitals of the entire system at planar geometry from the orbitals of the fragments (allene moiety, ethylene unit and fragment X). In analogy to Figure 6.16, the orbitals of the allene form **a** can be obtained from the orbitals at planar geometry by relaxation to the lower molecular symmetry of the allene. The diagram starts from the orbitals of the allene moiety given by Johnson[195] (the left hand side of Figure 6.18, see also Figure 6.16 for comparison), which are first combined with the π orbitals of the ethylene fragment leading to orbital set **A**. The positive and negative linear combination of the bonding π orbital of the allene fragment with the bonding π orbital of the ethylene moiety yield two low lying π orbitals denoted as π_1 and π_2. In this model the nonbonding σ orbital remains unchanged. The next π orbital (π_3) arises from the linear combination of the non-bonding orbital of the allene fragment and the anti-bonding orbital of the ethylene moiety. Higher lying orbitals are not relevant for this problem. The combination of the orbitals of the allene and the ethylene moiety in Figure 6.18 (orbital set **A**), is already sufficient to explain the properties of X = CH_2 (**15**), SiH_2 (**19**), and X=PH (**20,21**), as X cannot interact with the π system. One bonding orbital of X = CH_2 or SiH_2 had the correct symmetry (negative linear combination of CH or SiH orbitals) but they represent C-H or Si-H σ bonds which do not interact sufficiently with the π system in orbital set **A** (Figure 6.18). For X=PH no interaction of the lone pair of the phosphorus atom with the π system occurs since the sp^3 hybridization is energetically strongly favored over the sp^2 hybridization.

The relevant states for **15** and **19-21** arise from the occupation pattern of the σ orbital and the π_3 orbital, which are quite close in energy. For the diradical species **b** (cf. Figure 6.15) one electron occupies the π_3 and one the σ orbital at the C^2 center. The two zwitterionic states with reversed polarity result if either the σ (**c**) or π_3 orbital (**d**) is doubly occupied. The polarities correspond to those of the states discussed by Johnson.[195] Within orbital set **A** the energy difference between **b** and **c** is mainly governed by the energy difference between the π_3 and the σ orbital on one hand and the spin pairing energy on the other. Since π_3 and σ are close in energy, the diradical state should be preferred relative to the zwitterion. For X = CH_2 (**15**) and PH (**20,21**) the σ orbital is slightly lower in energy than the π_3 orbital so that the occupation pattern is $\sigma^2\pi_3^0$ and zwitterions of type **c** are found. Surprisingly, the zwitterionic state of X = SiH_2 (**19**), which is about as high in energy as the zwitterionic state of the carbon system **15**, is described by the occupation

pattern $\pi_3^2\sigma^0$ and consequently possesses a reversed polarity. This finding can be related to the silicon β-effect. It comprises hyperconjugation effects which stabilize a positively charged carbocationic center in β position to a silicon center.[200] Due to this effect the zwitterionic state possessing an unoccupied σ orbital is stabilized since only this occupation pattern conserves a positively charged C^2 center in β-position to the silicon center. Indeed, it is found that for the lowest zwitterionic state of the Si system the above mentioned hyperconjugation effects destabilize the σ orbital with respect to the π_3 orbital to that extent that the energetic sequence of both is reversed and the latter gets doubly occupied while the former remains empty. It is interesting to note that this effect only arises for the zwitterionic state. In the diradical state both π and σ orbitals are energetically very close to each other, but possess the sequence (σ below π) found in all other systems. A zwitterionic state with the normal occupation pattern ($\sigma^2\pi_3^0$) also exists for **19** but is much higher in energy (approx. 54 kcal $\cdot$ mol^{-1} with respect to the allene form).

While the explanation of the trends for systems **15** and **19-21** is already possible utilizing orbital set **A**, it is not sufficient for the systems with X = BH, NH, O, AlH, and S, because their fragments X can interact due to their p_z orbitals. The corresponding orbital scheme leading to orbital set **B** (right-hand side of Figure 6.18) is an extension of the previously discussed set **A** and includes its interaction with the p_z orbitals of the fragment X. The new π system now possesses four π orbitals that are relevant for this model. The orbitals π_1', π_2', and π_3', which are constructed from linear combinations of the former π_1, π_2, and π_3 orbitals with the p_z orbital, are stabilized with respect to the latter and lie below the σ orbital. Within a one-electron picture the σ' orbital is not affected. The next π orbital (π_4') is destabilized with respect to the former π_3 so that the energy gap to the σ' orbital is increased in comparison to the energy difference between the former π_3 and the σ orbital. The degree of interaction between the π orbitals π_1, π_2, and π_3 and the p_z orbital of X is relevant for the stabilization of π_1', π_2', and π_3' and the destabilization of π_4'. As a consequence it determines the order of the electronic states at planar geometry. The degree of orbital interaction depends on the radial extension of the p_z orbital as well as on its energy. These properties are mainly governed by the heteroatom in the fragment X and correlate with its electronegativity. The interplay between this interaction and the previously discussed strain in the allene moiety also determines the varying planarisation energies.

The order of the electronic states at planar geometries is first addressed. For all systems

with an appropriate p_z orbital at the fragment X (**14**, **16**, **17**, **18**, **22**) the interaction between this p_z orbital and the π system of orbital set **A** leads to an energy gap between the σ' orbital and the π'_4 orbital, which is larger than the gap between the former σ and π_3. The larger gap favors a double occupation of the σ' orbital, so that the zwitterionic state is preferred with respect to the diradical one. The energy gap between both states is dominated by the degree of interaction. The interaction is strongest for X = NH which also possesses the largest gap ($24 \text{ kcal} \cdot \text{mol}^{-1}$) of all systems in Figure 6.17. The weaker interaction for X = O is expressed by a reduced gap of $15 \text{ kcal} \cdot \text{mol}^{-1}$, which is further reduced to $4 \text{ kcal} \cdot \text{mol}^{-1}$, if X = S is regarded. This finding is in line with the chemical experience about the π donor capability of the heteroatoms. The fragments X = BH and AlH also possess p_z orbitals which interact with the π system of the carbon framework, so that the same orbital diagram as it has been described for X = NH, O, and S is valid. However, the former systems differ from the latter ones as to the number of electrons. Due to their electron deficiency B and Al both cannot contribute any electron to the π system. Therefore only the six electrons contributed from the allene and the ethylene moiety are distributed amongst the three π' orbitals and the σ' orbital. Since the three π' orbitals are lower in energy, they are occupied while the σ' orbital stays empty. As a consequence the resulting zwitterionic state has the same polarity as found for the zwitterionic state **d** defined in Figure 6.15. It has to be kept in mind, that the reasons for the polarity (Figure 6.15: doubly occupied π^*; here: empty π^* and σ orbital) are different.

The orbital occupation patterns also explain the geometric differences between the zwitterionic and diradical states. The bond angle $C^1 - C^2 - C^3$ was found to be the most characteristic parameter to describe the structural differences between the electronic states. The values in Table 6.9 in combination with the occupation pattern of the states show that this angle correlates with the occupation of the σ orbital. For the zwitterionic states with a polarity as **c** (double occupation of σ) the angle is around 111°, while for zwitterionic states with a polarity like **d** (empty σ) the angle varies from 155° to 162°. For the diradical states **b** (singly occupied σ) it was computed to be around 128°. This shows that the angle depends on the spatial extent needed by the empty, the singly or the doubly occupied σ orbital. This is in accordance with the VSEPR model.

To explain the planarisation energy and the related electronic structure of the minimum, the stabilization effects resulting from the orbital interaction between the p_z orbital of X and

the π system in orbital set **A** have to be compared with the opposed strain introduced in the allene moiety. In the planar structure the interaction between the various orbitals but also the strain in the allene moiety can be expected to have a maximum. On going to the allene structure the strain decreases but also the strengths of the interactions are expected to get smaller. So if the p_z orbital and the π system in orbital set **A** interact strongly, a planar structure arises since the strain effects are overcompensated. If the orbital interactions are weaker, the strain effects are dominant and a twisted ring results. The interplay between both effects determines the planarisation energy. For X = NH (**16**) the degree of interaction in the zwitterion overcompensates the strain in the allene moiety and establishes the planar zwitterion as the minimum. The degree of interaction is reduced for X = O, because of the smaller π donor capability of O with respect to N. As a result the stabilization of **c** is decreased with respect to X = NH and consequently the allene is slightly favored over the zwitterion.[174, 193] For X = BH the planarisation energy is similar to the value found for X = O. The question to what extent the different number of electrons of both systems influences the barrier will be addressed later.

The increase in the planarisation energy going from the second to the third row elements results from the larger diameter of the monocycles allowing a better release of the strain effects. A decreased interaction between the $3p_z$ orbitals of the third row heteroatom X and the π system in orbital set **A** could also be responsible for the higher planarisation energy. A reduced interaction, however, would also lead to a smaller energy gap between the zwitterionic and the biradical states. The gap indeed decreases going from X = O to S (15 kcal $\cdot$ mol^{-1} to 4 kcal $\cdot$ mol^{-1}) but stays nearly constant if X = BH (9 kcal $\cdot$ mol^{-1}) is compared to X = AlH (11 kcal $\cdot$ mol^{-1}). Consequently, an expected increase in the planarisation energy for both pairs results from somewhat more subtle reasons.

The orbital diagram also explains the differences between the phosphorus containing systems **20** and **21**. The additional phosphorus center hampers the delocalization in the π system of orbital set **A**, so that both, the diradical and the zwitterionic state, are stabilized with respect to the allene structure. This should also hold for other substitutions of the $C^4 - H$ group since the carbon center enables an optimal delocalization.

The model based on the interplay between the strain in the allene moiety and the stabilization effects introduced by the orbital interaction of the fragment orbitals works very
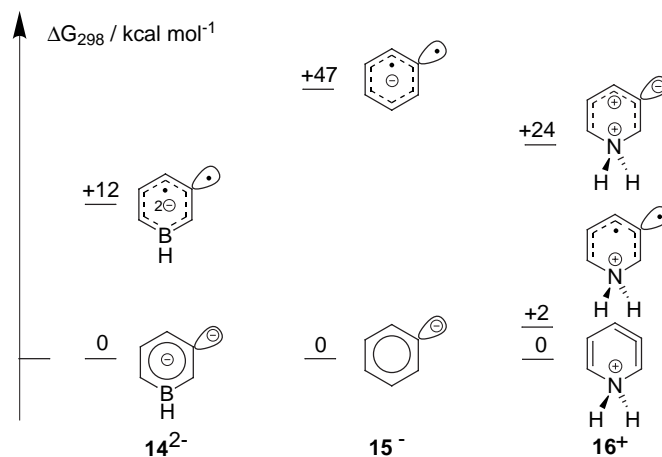
Figure 6.19: Order of the electronic states of the doubly negatively charged system 14^{2-}, the phenlyanion 15^- and the protonated isopyridine 16^+.

well for neutral species so that the question arises whether this model can also be extended to charged species. The computed values for three isoelectronic second row systems ($X = BH^{2-}$, $X = CH^-$, $X = NH_2^+$), obtained with the MR-CI+Q approach, are summarized in Figure 6.19. Within this model the isobenzene ($X = CH_2$) and isopyridine ($X = NH$) systems are the two prototypes, since they are representatives for the different degrees of interactions and the resulting orbital schemes. Deprotonation of the first at the fragment $X = CH_2$ converts the isobenzene to the phenyl anion 15^- ($X = CH^-$).[201, 202] In contrast to 15, 15^- obviously possesses an occupied p_z orbital that can perfectly conjugate with the π system of orbital set $\mathbf{A}$ (see Figure 6.18). Consequently, this orbital model predicts 15^- ($X = CH^-$) to possess a ground state that corresponds to the zwitterionic[11] state $\mathbf{c}$. Its stabilization is expected to be even stronger with respect to the diradical state than that of 16 (X=NH). These predictions are confirmed by the MR-CI+Q computations which predict a planar equilibrium geometry with a closed-shell electronic structure and a diradical state lying $+47$ kcal $\cdot$ mol^{-1} higher in energy than the ground state. This proves the distinctly stronger interaction in the π system in comparison to 16 ($X = NH$) which possesses a gap of $+24$ kcal $\cdot$ mol^{-1}. From the description it is obvious that 15^- behaves like 16 ($X = NH$) and not as his neutral precursor 15 ($X = CH_2$).

Analogous to the change in the electronic structure of the pair $15/15^-$, protonation of

---

[11]In charged species the term zwitterion is not strictly defined. In the present study, this term is used for the states with an electronic structure that resembles the zwitterionic states of the neutral species (see also Figures 6.15 and 6.16).

isopyridine (**16**) to $\mathbf{16}^+(X = NH_2^+)$ also leads to a change in the electronic structure. In this pair, the delocalization found in the π system of **16** is destroyed due to the protonation. The computations indeed predict for $X = NH_2^+$ an allene species having a twisted ring as the equilibrium structure instead of the planar zwitterionic ground state **c** which is found for $X = NH$. The computed planarisation energy is rather small ($2 \text{ kcal} \cdot \text{mol}^{-1}$) but at the planar structure the computations predict the diradical state to be preferred over "zwitterionic" one by $22 \text{ kcal} \cdot \text{mol}^{-1}$, which is in difference to X=NH. This shows that the chemical nature of $\mathbf{16}^+$ is very similar to that of **15** but is distinctly different to that of **16**. This difference explains experimental findings obtained by Christl *et al.* who found that a 1-aza-2,3-cyclohexadiene behaves completely different whether it has a three or four coordinated nitrogen center.[203]

The doubly charged anion of **14**, $\mathbf{14}^{2-}$ ($X = BH^{2-}$) is isoelectronic to **16**. The MR-CI+Q computations predict that, in analogy to **16**, $\mathbf{14}^{2-}$ possesses a planar equilibrium structure and a "zwitterionic" ground state. The energy gap to the diradical state is much smaller than for $X = NH$. This may result from the stronger charge concentration in the ring. In the diradical state, the charge is more distributed over the whole molecule.

This qualitative model, which explains the varying planarisation energy and the order of the electronic states at planar structure, is based on orbital energies, while the influence of the geometry, which varies considerably from state to state, and many-electron effects are not taken into account. The question arises to what extent these effects influence the picture. An answer to this question is given in Table 6.10 which depicts the energy differences between the zwitterionic and diradical states of $X = CH_2$, $X = CH^-$, NH and O at various levels of theory. The levels of theory used, reach from simple orbital considerations to the computed MR-CI+Q energies at optimized geometries. The latter were taken as reference (last line of Table 6.10).

The first level that can be used to estimate the energy differences between zwitterionic and diradical states, is based on the orbital energy differences ($\epsilon_{\pi^*} - \epsilon_\sigma$). For $X = NH$ and $X = O$ this simple approach reproduces the correct order of the states but overestimates the energy differences substantially by a constant value of about $70 \text{ kcal} \cdot \text{mol}^{-1}$ ($X = NH$: $97 \text{ kcal} \cdot \text{mol}^{-1}$ vs. $25 \text{ kcal} \cdot \text{mol}^{-1}$, $X = O$: $87 \text{ kcal} \cdot \text{mol}^{-1}$ vs. $14 \text{ kcal} \cdot \text{mol}^{-1}$). For $X = CH_2$ the correct order (diradical state below zwitterionic state) cannot be described since spin

Table 6.10: Analysis of the orbital, geometric and many-electron contributions to the energy difference between **b** and **c** [a].

	X = CH$_2$	X = CH$^-$	X = NH	X = O
$\epsilon_{\pi^*} - \epsilon_\sigma$ [b]	+62	+127	+97	+87
$E^{\text{det}}_{\sigma\pi^*}(\mathbf{c}) - E^{\text{det}}_{\sigma^2}(\mathbf{c})$ [c]	-3	+70	+36	+23
$E^{\text{det}}_{\sigma\pi^*}(\mathbf{b}) - E^{\text{det}}_{\sigma^2}(\mathbf{c})$ [c]	-20	+51	+22	+11
$E^{\text{CI}}(\mathbf{b}) - E^{\text{CI}}(\mathbf{c})$	-15	+51	+25	+14

[a] All energy differences are given in kcal $\cdot$ mol^{-1}. Geometries, that were taken for the calculation (diradical **b** and zwitterion **c**), are denoted in brackets. Orbital energies were taken from CASSCF/cc-pVDZ computations. CI energies are obtained from MR-CI+Q/ano-l//(U)B3LYP/cc-pVDZ computations and energy expectation values were calculated with the DIESEL package[53] utilizing a cc-pVDZ basis set at (U)B3LYP/cc-pVDZ geometries. Note that the CI results in both basis sets are comparable.

[b] Differences of orbital energies were computed with the orbitals of the $^1A''$ state at zwitterionic geometries. Orbital corresponding orbital energy differences obtained for diradical geometries are 35 to 43 kcal $\cdot$ mol^{-1} lower in energy, *i.e.* all trends are kept.

[c] E^{det} corresponds to the energy expectation value of the main configuration state function (CSF) for each individual state. The main configuration for the zwitterionic state is $|...(19a''^2)(20a''^2)(21a'^2)\rangle$ and $|...(19a''^2)(20a''^2)(21a'^1)(22a''^1)\rangle$ for the diradical state, respectively. In C_s symmetry the orbitals of the irreducible representation a$'$ correspond to σ orbitals while those of a$''$ correspond to π orbitals.

pairing effects can only be considered if many-electron effects are taken into account.

The term $E^{\text{det}}_{\sigma\pi^*}(\mathbf{c}) - E^{\text{det}}_{\sigma^2}(\mathbf{c})$ includes many-electron effects in a comparable manner as the Hartree-Fock approach[12] but neglects effects from geometry relaxation (equivalent to the

---

[12]This procedure is comparable to the restricted Hartree-Fock level, which includes many electron effects in an averaged manner.[56, 57] The difference between the restricted Hartree-Fock ansatz and the computation of E^{det} are the orbitals. In this study orbitals are used that were optimized for the diradical state, while

vertical excitation energy). This level describes the order of the electronic states for $X = CH_2$ already correctly (-3 kcal $\cdot$ mol^{-1}). For $X = NH$ and $X = O$ the computed differences between the diradical and the zwitterionic state are uniformly overestimated with respect to the MR-CI+Q values by only about 10 kcal $\cdot$ mol^{-1} (36 kcal $\cdot$ mol^{-1} and 23 kcal $\cdot$ mol^{-1}). If the geometric differences between the diradical state **b** and the zwitterionic state **c** are also considered, the relative energies $E_{\sigma\pi^*}^{det}(\mathbf{b}) - E_{\sigma^2}^{det}(\mathbf{c})$ deviate only by 3 to 5 kcal $\cdot$ mol^{-1} to the MR-CI+Q results that also contain electron correlation effects. The small influence of correlation effects is astonishing since open-shell species are compared to closed-shell species.

6.3.5 Summary

Strained cyclic allenes containing heteroatoms of the second and third row exhibit strong variations in their chemical properties depending on the heteroatomic fragment. In this study, a comprehensive model that explains all these variations including the parent compound 1,2,4-cyclohexatriene is presented. It is based on an orbital diagram in which the orbitals of the fragments (allene moiety, ethylene unit and fragment X) are successively combined to build up the molecular orbitals of the entire system. The variations in the order of the electronic states and the planarisation energy which are both crucial for the reactivity of such compounds[174, 193] are explained by the interplay between the stabilization depending on the degree of orbital interaction and the strain effects in the allene moiety. The model is found to be valid for all experimentally known species and its predictions are in agreement with results obtained from MR-CI+Q computations which were employed to characterize still unknown compounds. Apart from neutral species, the model is also applicable to charged systems. The consideration of $\mathbf{15}^-$ ($X = CH^-$) reveals an unexpected relation between strained cyclic allenes and the phenyl anion. Comparing $X = NH$ with $X = NH_2^+$ the model also answers the open question why three and four coordinated 1-aza-2,3-cyclohexadienes and 1-aza-2,3,5-cyclohexatrienes show different chemical reactivity. In addition, the present work investigates how the various factors (orbital energies, geometry relaxation, many electron and correlation effects) influence the various properties. It is shown that orbital energy differences already describe all trends correctly. For quantitative

---

Hartree-Fock uses orbitals optimized for each state. The choice of the diradical orbitals is necessary to ensure a balanced description of both the σ-type orbitals (σ/σ') and the lowest unoccupied π-type orbital (π_3/π_4')

predictions, the different geometries of the various states and the many electron effects are essential while the remaining electron correlation is less important. This shows that qualitative predictions for an unknown species can be made if its orbital energies, which are already accessible by simple DFT calculations, are compared to the orbital energies of the system with X=NH. For quantitative predictions, however, more sophisticated approaches are necessary.

Chapter 7

Summary

This work encompasses two parts. The first part focuses on the characterization of systems, which complex electronic structures require the application of multi-reference methods. In the second part, algorithms for a Møller-Plesset perturbation theory (MR-MP2) program, designed to perform large-scale computations, were developed and implemented.

The Hartree-Fock method recovers about 99% of the exact energy of a system, however, the effect of electron correlation, is essential to describe chemical processes. The standard *ab initio* methods of todays theoretical chemistry (second-order Møller-Plesset perturbation theory, Coupled Cluster, ...), which encompass electron correlation, employ the Hartree-Fock wavefunction as starting point, *i.e.* distinguish one Slater determinant as reference and are therefore called single-reference approaches. Density Functional Theory uses a somewhat different ansatz, but also represents a single-reference method. Approaches that start from more than one reference (multi-reference) allow a much greater flexibility and accuracy. These approaches are essential for the understanding of the reactivity and the properties of systems with complex electronic and structural features, because the commonly applied single-reference methods are quite unreliable. In the present work three systems were characterized, which were of interest for biologically oriented chemists as well as physical organic chemists.

The anti-tumor efficacy of the natural product Neocarzinostatin (Figure 7.1) is based on the formation of diradicals and causes DNA cleavage by hydrogen abstraction from adenine and thymine. Computations on model systems performed in the present work, which
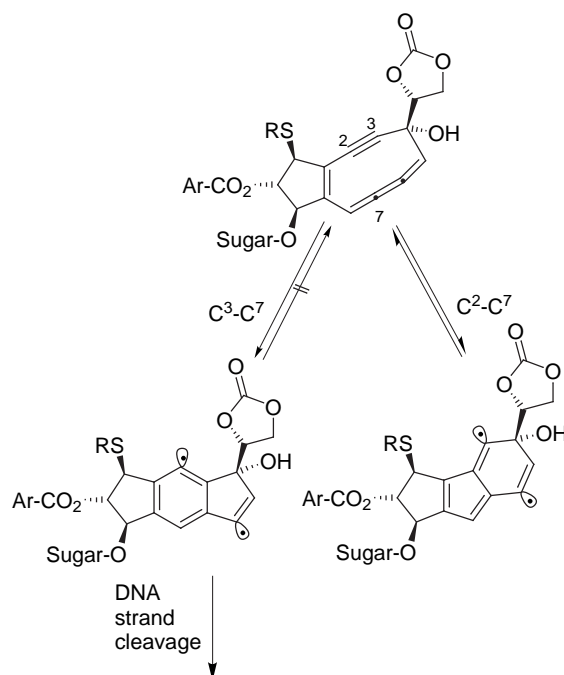
Figure 7.1: Reaction mode of the enyne-cumulene system of the Neocarzinostatin chromophore.

considered the main structural motifs of the Neocarzinostatin chromophore, indicated that the observed mode of action is the most efficient one. Although there is no kinetic preferrence for the formation of the intermediate of the $C^3 - C^7$ cyclization mode or the $C^2 - C^7$ cyclization (Figure 7.1), the reaction via the $C^3 - C^7$ cyclization route is thermodynamically favored. The distinct endothermicity of the $C^2 - C^7$ cyclization results in a small barrier for the reverse reaction and the equilibrium is supposed to prefer the reactant. The reaction then can occur quantitively *via* the $C^3 - C^7$ route. It was even more important that the computations presented in this work also offered an explanation why nature "tailored" the reaction of the Neocarzinostatin chromophore that way. The computed singlet-triplet splitting of the intermediates, which expresses their diradical character, correlates with the hydrogen-abstraction ability. The intermediate formed in the $C^3 - C^7$ cyclization possesses a more pronounced diradical character than the intermediate of the $C^2 - C^7$ pathway, so that only the intermediate of the $C^3 - C^7$ cyclization is able to abstract hydrogens from adenine or thymine which finally causes cytolyis.

The discovery that diradicals occur in biological processes and that molecules with highly unsaturated bonds play a role in biochemistry, led to the development of new compounds with related structural motifs (Figure 7.2). One example of compounds related to Neocarzi-

nostatin are the enyne-allenes and the hetero-anlogue class of enyne-ketenes (Figure 7.2). Both were expected to possess similar reactivities and in this work, their similarities and differences were investigated. Generally, the reactivities are indeed very similar, if the reduced electron density introduced by the ketene moiety with respect to the allene moiety is compensated by appropriate substituents.
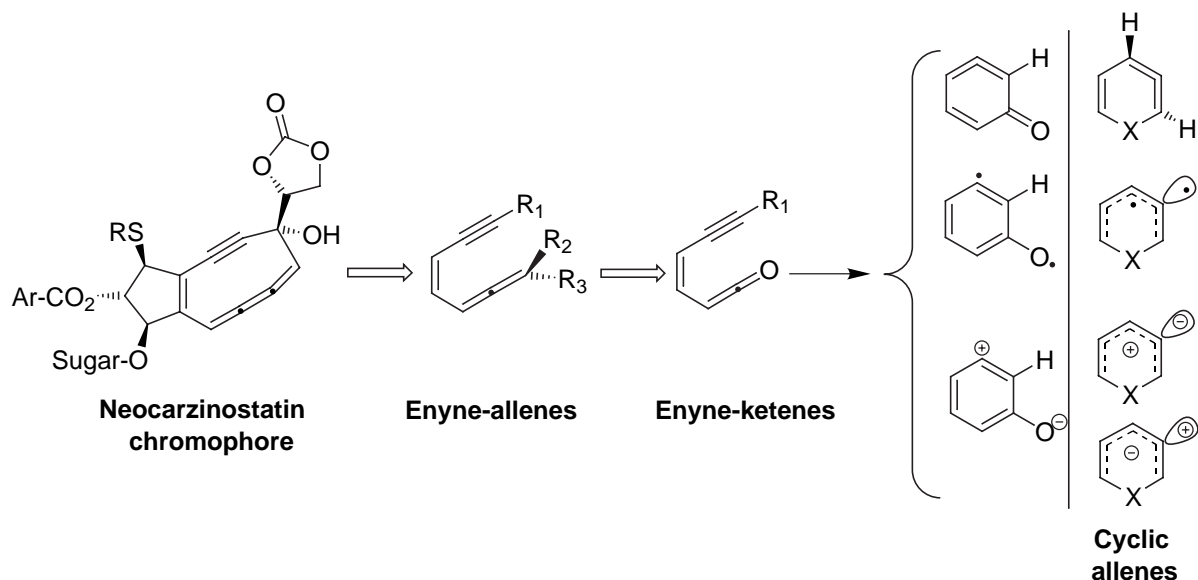
Figure 7.2: Structural relationships between enyne-allenes, enyne-ketenes and cyclic allenes.

In the course of the reaction not only the discussed diradical intermediates can be formed. Others isomeric species, like cyclic allenes and zwitterions, are close in energy to the diradicals and their formation *via* different reaction paths is possible. This work shows, that the preference for a particular electronic structure (allene, diradical, or zwitterion) is governed by the substituent attached to the enyne-ketene framework. The systematic influence of substituents on the preferred electronic structure was evaluated for the class of cyclic allenes substituted in the ring with fragments of the second and third period of the periodic table of elements. For these compounds a comprehensive model on the basis of orbital considerations is proposed. This model explains the experimentally found and/or computed variations on a qualitative level, but also provides a simple tool to predict the electronic structure of unknown compounds. Additionally, the factors influencing the energy differences between the zwitterionic and the diradical states (orbital contributions, geometric considerations, higher-order effects) were analyzed quantitatively on different levels of theory and interpreted.

The presented examples contributed significantly to the understanding of the reactivity

and electronic structure of typical multi-reference cases but also showed up the limits inherently existing in the currently available programs with respect to the size of the molecules.

This experience motivated the second topic of the present work in which an improved MR-MP2 code was developed. The MR-MP2 approach represents the most efficient multireference ansatz. The criterions of an efficient MR-MP2 code encompass a couple of issues. First, an efficient evaluation of the Hamilton matrix is crucial to reduce the computational cost and an intelligent parallelization scheme can be used to further reduce the overall wall time of the computation. As the third feature, the program is supposed to incorporate is a flexible infrastructure with implementations of interfaces to commonly used quantum-chemical programs and a substantial flexibility to extend and to improve the code.

The efficient construction of the Hamilton matrix represents the most crucial part for the calculations because the major fraction of computational cost is consumed in this step. The considered matrices are of the dimension of up to 10^{10}, but are characterized by an tremendous sparsity (99%to 99.999%). An efficient recognition of non-vanishing matrix elements is therefore crucial for a well-performing code. Additionally, the effort to recognize the vanishing blocks of matrix elements has to be as small as possible to ensure that the computational cost is spent on the evaluation of the non-vanishing matrix elements and not on the vanishing ones. For this purpose the configurations are partitioned employing a three-fold separation into inactive (orbitals are doubly occupied in all reference configurations), active (differently occupied orbitals in the reference configurations), and external (orbitals not present in the reference configurations) configuration rests. This division blocks the Hamilton matrix and the inhomogenity and the smaller blocked matrices are treated separately. In these submatrices the structure of even finer interaction classes can be efficiently exploited. Large parts of the $\mathbf{H_0}$ matrix and the inhomogenity can be excluded from the computational procedure, because the separation scheme forces the matrix elements to vanish by virtue of the maximum allowed level of excitation.

The partitioning scheme of the configuration space further allows to enhance the efficiency of the matrix evaluation. In the surviving submatrices a three-level structure exists. The basic level is characterized by the interactions of the inactive configuration rests. For each two interacting inactive configuration rests, sets of active rests exist that also possess a well-defined interaction pattern. Setting up on the active level, the same situation holds

for the external configuration rests. Once at a certain level (*e.g.* inactive) an interaction is computed to vanish, all interaction blocks on the following levels need not to be computed since the matrix elements associated with this matrix block are zero. This scheme allows the early recognition of vanishing blocks of matrix elements. Further on, by explicitely sorting the sets of configuration rests, the interactions at each level represent repeating patterns. These patterns in principal can be precomputed, stored and accessed in each step of the computation. The computed patterns not only contain the position of the non-vanishing interacting configurations but also the characterization of the interactions including excitation order and interacting orbitals.

The evaluation of the interactions at the inactive level, for which the inactive configuration rests are supposed to consist of the majority of occupied orbitals, is not performed in the natural representation of electrons but with inverted configuration rests characterizing the interaction with holes. Since the number of holes is magnitudes smaller than the number of occupied orbitals a significant speed-up is achieved. The use of patterns in the actual process of building the Hamilton matrix guarantees that only those interactions are considered, which contribute non-vanishing matrix elements.

The implementation of a comparable algorithm for the external interactions is impossible because of the too large memory consumption due to the large number of external orbitals. Therefore, a direct scheme utilizing the limited number of possibilities of occurring interactions between the external orbitals was employed. Algorithms for these few different external interaction schemes, which depend of the excitation level of the internal configuration rests and the position of the previously mentioned submatrices, were deduced and implemented. The implementation is completely done without constructing the external configuration rests, but directly sets up the difference configurations needed for the evaluation of the matrix elements. This algorithm is computationally much less expensive than the full evaluation of external configurations and does not need any storage requirements as pattern schemes.

Simultaneous to an efficient implementation of the Hamilton matrix, the parallelization of the MR-MP2 procedure reduces the overall wall time. For this reason emphasis was put on an efficient and well-balanced parallelization. The parallelization scheme employed in the DIESEL-MP program is controlled by the amount of allocatable memory and is

based on the inactive patterns to distribute only the non-vanishing matrix blocks. The implemented parallelization scheme requires a significantly smaller amount of memory than the existing MR-MP2 procedure in the DIESEL package and thus makes large-scale parallel computations feasible. Parallelization causes a certain communication overhead compared to a serial code and the performance of the current program and the old MR-MP2 procedure in the DIESEL program suite are comparable for small test systems (40 electrons, 150 basis functions, 16 inactive and 9 active orbitals). Comparisons for larger systems, for which the new program is designed for, were not possible because of the limitations in the old DIESEL program package.

The infrastructure of the code was designed to be as flexible as possible. The class structure of the DIESEL-MP program is implemented to perform a full MR-MP2 computation but also to incorporate flexible selection schemes. These schemes can either employ an individual selection of configurations or a scheme based on the contributions of particular external orbitals. Basic classes that are needed for an MR-CI approach are also present and can be easily used to implement an efficient MR-CI procedure.

The integral handling was a weak point in the old MR-MP2 code implemented in the DIESEL package. The implemented integral interface of the new program to the TUR-BOMOLE suite possesses a couple of advantages. The integrals generated by a RI integral transformation in conjunction with the on-the-fly generation of the required four-index-integrals from the three-index RI integrals consume less storage resources than the conventional four-index-integrals. The integrals are needed only once for the computation of the inhomogenity vector and the contraction to an effective one-electron multi-reference Fock matrix. For the iterative process, the $\mathbf{H_0}$ matrix is constructed from the effective Fock matrix and the RI integrals are no longer needed. This means that the transformation from the RI-integrals to the four-index integrals has only to be performed once and does not represent the bottleneck of the computation. The orbitals for the integral transformation can be taken from Hartree-Fock, Local-Hartree-Fock, or gradient-corrected density functionals. This offers the flexibility to use appropriate orbitals for computations of reactions paths and excited states. The possibility of the implemented TURBOMOLE interface to easily incorporate the DIESEL-MP program into existing QM/MM environments like the CHEMSHELL[204] package offers the perspective to treat systems like enzymes accurately.

Chapter 8

Zusammenfassung

Die vorliegende Arbeit untergliedert sich in zwei Teile. Der erste Teil beinhaltet die Charakterisierung von Molekülen, deren komplexe elektronische Struktur Multireferenz-Methoden erfordern. Im zweiten Teil werden Algorithmen für die Behandlung großer Systeme mittels der Multireferenz-Møller-Plesset-Störungstheorie (MR-MP2) entwickelt und implementiert.

Die Hartree-Fock Methode erfasst etwa 99% der Gesamtenergie eines Systems, jedoch zeigt es sich, daß die von ihr nicht berücksichtigten Korrelationseffekte von essentieller Bedeutung für die Beschreibung chemischer Reaktionen sind. Die heutigen *ab initio* Standardmethoden der Theoretischen Chemie, die die Elektronenkorrelation berücksichtigen (Møller-Plesset-Störungstheorie zweiter Ordnung, Coupled Cluster, ...), basieren auf der Hartree-Fock Wellenfunktion, d.h. sie zeichnen eine Slater-Determinante als Referenz aus und werden daher als Einreferenzmethoden bezeichnet. Die Dichtefunktionaltheorie besitzt einen etwas anderen Ansatzpunkt, stellt aber ebenfalls eine Einreferenzmethode dar. Ansätze, die von mehreren Referenzen aus starten (Multireferenz) ermöglichen eine wesentliche größere Flexibilität und erreichen damit eine höhere Genauigkeit.

Sie sind für das Verständnis der Reaktivität und der Eigenschaften von Systemen mit komplexen elektronischen und strukturellen Eigenschaften notwendig, da die standardmäßig verwendeten Eindeterminantenansätze für diese Systeme sehr unzuverlässig sind. In dieser Arbeit wurden Systeme charakterisiert, deren Verständnis sowohl für physikalisch-organisch als auch für bioorganisch arbeitende Chemiker von großer Bedeutung ist.

Die Wirkung des natürlich vorkommenden Antitumor-Antibiotikums Neocarzinostatin
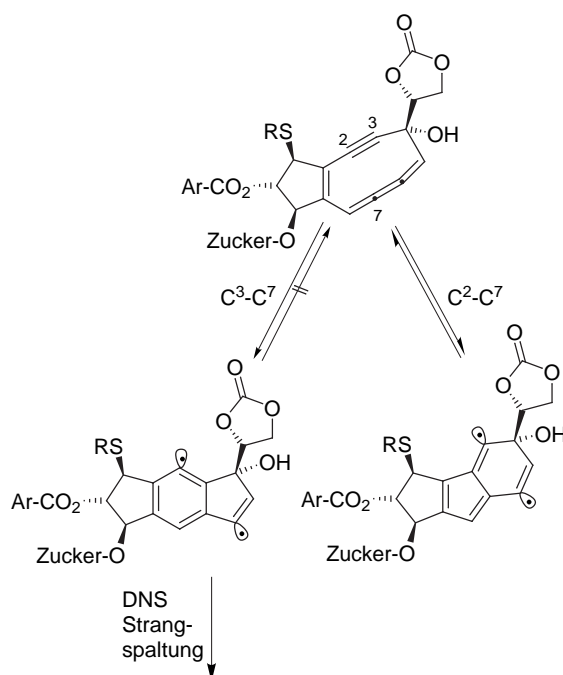
Abbildung 8.1: Reaktionsschema des Neocarzinostatin-Chromophors.

beruht auf der Bildung von diradikalischen Intermediaten (Abbildung 8.1), die durch Abstraktion von Wasserstoffatomen des Adenins und Thymins DNS-Spaltungen hervorrufen. Die in dieser Arbeit durchgeführten Berechnungen an Modellsystemen, welche die dominierenden strukturellen Motive des Neocarzinostatin-Chromophors beinhalten, legen nahe, daß der beobachtete Reaktionsweg der $C^3 - C^7$-Cyclisierung (siehe Abbildung 8.1) auch der Effizienteste ist. Obwohl die $C^3 - C^7$-Cyclisierung gegenüber dem alternativen Reaktionsweg der $C^2 - C^7$-Cyclisierung kinetisch nicht favorisiert ist, ist der Reaktionsverlauf über die $C^3 - C^7$-Cyclisierung thermodynamisch bevorzugt. Die starke Endothermie der $C^2 - C^7$-Cyclisierung bedingt eine sehr kleine Aktivierungsbarriere für die Rückreaktion und es ist davon auszugehen, daß das Gleichgewicht der $C^2 - C^7$-Cyclisierung auf der Seite des Reaktanden liegt und die Reaktion vollständig über die $C^3 - C^7$-Cyclisierung abläuft.

Viel eindrucksvoller als die Verifizierung des ablaufenden Reaktionsweges der $C^3 - C^7$-Cyclisierung ist die, durch Berechnungen erhaltene Erklärung, warum die Reaktion des Neocarzinostatin-Chromophors von der Natur gerade auf diese Weise "maßgeschneidert" worden ist. Sie resultiert aus der berechneten Singulett-Triplett-Aufspaltung der Intermediate, die den diradikalischen Charakter des Systems wiedergibt und korreliert mit der Abstraktionsfähigkeit für Wasserstoffatome. Das Intermediat der $C^3 - C^7$-Cyclisierung besitzt einen deutlich ausgeprägteren diradikalischen Charakter als das der $C^2 - C^7$-Cyclisierung,

sodaß einzig allein das Intermediat der C^3-C^7-Cyclisierung in der Lage ist Wasserstoffatome von Adenin und Thymin zu abstrahieren und die Cytolyse einzuleiten.

Die Entdeckung, daß Diradikale in biologischen Prozessen eine Rolle spielen führte zur Entwicklung neuer Verbindungen mit verwandten Strukturmotiven (Abbildung 8.2). Ein Beispiel zu dem Neocarzinostatin-Chromophor verwandte Verbindungen sind Eninallene und die heteroanalogen Eninketene (Abbildung 8.2). Für beide Verbindungsklassen sind ähnliche Reaktivitäten zu erwarten und die Gemeinsamkeiten und Unterschiede bezüglich der Reaktivität wurden im Rahmen dieser Arbeit untersucht. Es konnte gezeigt werden, daß die Reaktivität der beiden Klassen im Allgemeinen sehr ähnlich ist, falls die, durch die Keteneinheit verursachte, verminderte Elektronendichte durch geeignete Substituenten ausgeglichen wird.
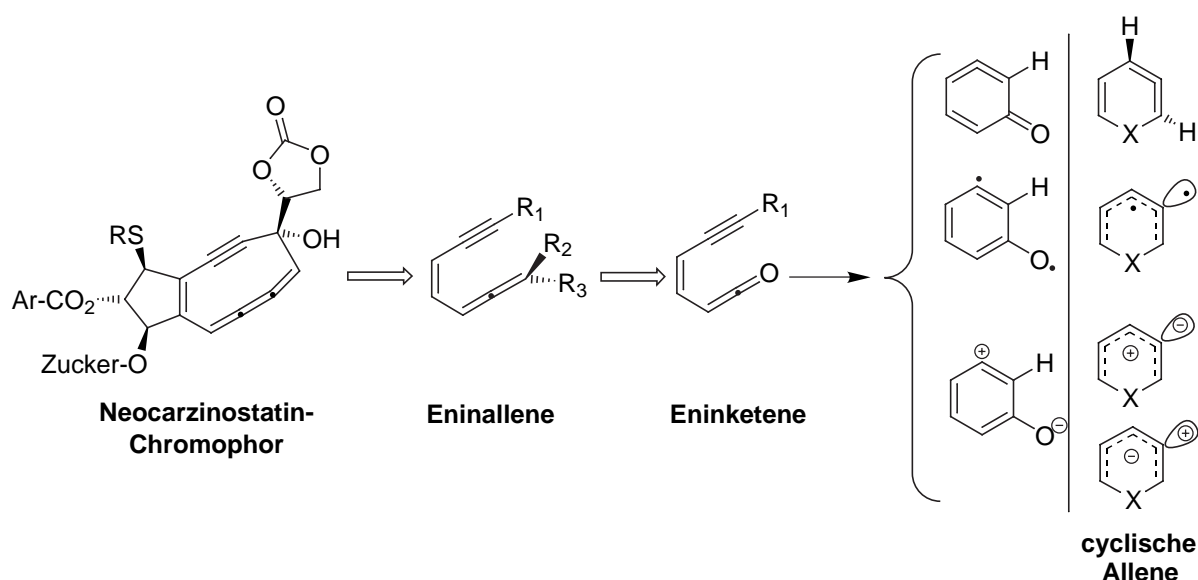
Abbildung 8.2: Strukturverwandtschaft zwischen Eninallenen, Eninketenen und cyclischen Allenen.

Im Verlauf der Cyclisierungsreaktionen können nicht nur die diskutierten diradikalischen Intermediate gebildet werden sondern auch andere Isomere wie cyclische Allene und Zwitterionen, die energetisch so nahe liegen, daß sie über andere Reaktionspfade ebenfalls auftreten können. Diese Arbeit zeigt deutlich, daß die energetisch bevorzugte elektronische Struktur (Allen, Diradikal oder Zwitterion) von den Substituenten am Eninketen-Grundgerüst bestimmt wird. Der systematische Einfluß von Substituenten auf die bevorzugte elektronische Struktur wurde für die Verbindungen der cyclischen Allene evaluiert, bei denen im sechsgliedrigen Ring eine Position durch Fragmente von Elementen aus der zweiten oder dritten

Periode ersetzt wurden. Für diese Verbindungen wurde ein Modell vorgeschlagen, das die experimentell gefundenen oder berechneten Variationen qualitativ erklären kann, aber auch eine einfache Möglichkeit eröffnet die elektronische Struktur unbekannter Verbindungen vorherzusagen. Weiterhin wurden Faktoren, die die Energiedifferenzen zwischen zwitterionischen und diradikalischen Zuständen bestimmen (Orbitalbeiträge, geometrische Faktoren, Effekte höherer Ordnung) auf verschiedenen theoretischen Niveaus quantitativ analysiert und gewichtet.

Die vorgestellten Beispiele tragen zum Verständnis der Reaktivität und der elektronischen Struktur von typischen Multireferenz-Systemen bei, zeigen aber auch noch die Grenzen der aktuell verfügbaren Programme im Hinblick auf große Moleküle auf.

Diese Limitierung der existierenden Programme war die Motivation für das zweite Projekt im Rahmen dieser Arbeit, das sich mit der Entwicklung eines Multireferenz-Møller-Plesset-Störungstheorie (MR-MP2) Programmes beschäftigt. Ihm Rahmen der Multireferenz-*ab initio*-Verfahren stellt der MR-MP2-Ansatz das Verfahren dar, welches den geringsten Aufwand benötigt. Für ein solches MR-MP2-Programm bestehen die Anforderungen aus folgenden Punkten. Zum einen ist eine effiziente Berechnung der Hamiltonmatrix nötig, um die Rechenzeit auf ein vernünftiges Maß zu reduzieren, die durch geschickte Parallelisierung weiter verringert werden kann. Eine flexible Infrastruktur des Programmcodes ist der dritte Aspekt, der Beachtung fand. Dies beinhaltet die Implementierung von Schnittstellen zu quantenchemischen Standardprogrammen und ein offenes Design des Quellcodes, das es erlaubt das Programm auf einfache Art und Weise zu erweitern und zu verbessern.

Der effiziente Aufbau der Hamiltonmatrix stellt im Hinblick auf die Rechnungen den wichtigsten Teil dar, da hier der größte Anteil der Rechenleistung erbracht werden muß. Die betrachteten Matrizen besitzen Dimensionen in der Größenordnung von 10^{10}, sind aber mit sehr vielen Elementen besetzt, die gleich Null sind (etwa $99\%-99.999\%$). Für einen effizienten Programmcode ist daher eine schnelle Erkennung der verschwindenden Blöcke von Matrixelementen bei minimalem Rechenaufwand wesentlich, um sicherzustellen, daß die Rechenzeit durch die Berechnung der nicht-verschwindenden Matrixelemente bestimmt wird. Die Minimierung des Aufwandes zur Bestimmung der verschwindenden Matrixelemente wurde durch eine dreifache Unterteilung der Konfigurationen in inaktive (aus Orbitalen, die in allen Referenzkonfigurationen doppelt besetzt sind), aktive (verschiedene

Besetzung in den Referenzkonfigurationen) und externe (in Referenzkonfigurationen unbesetzt) Konfigurationsreste erreicht. Diese Unterteilung führt zu einer blockweisen Aufteilung sowohl der $\mathbf{H_0}$-Matrix als auch der Inhomogenität, bei der die einzelnen Matrixblöcke getrennt voneinander behandelt werden können. Innerhalb dieser Matrizen existiert wiederum eine Struktur feingliedriger Wechselwirkungsklassen, die effizient ausgenutzt werden können. Aufgrund der vorgenommmen Aufteilung können große Teile der $\mathbf{H_0}$-Matrix bzw. der Inhomogenität von der Berechnung ausgeschlossen werden, da verschwindende Matrixelemente sofort erkannt werden (Überschreitung der maximal zulässigen Anregungsstufe).

Die in dieser Arbeit verwendete Aufteilung des Konfigurationsraumes erlaubt zusätzlich auch eine Behandlung der nicht-verschwindenden Untermatrizen auf drei Ebenen. Auf unterster Ebene stehen die Wechselwirkungen zweier inaktiver Konfigurationsreste. Für jede Kombination aus zwei inaktiven Resten existiert eine Menge von aktiven Resten, die ebenfalls ein wohl-definiertes Wechselwirkungsmuster besitzen. Aufbauend auf die Ebene der aktiven Konfigurationsreste findet man die gleiche Situation für die externen Anteile. Wird nun auf einer Ebene (z.B. inaktiv) eine Wechselwirkung zu Null bestimmt, müssen alle nachfolgenden Ebenen nicht mehr durchlaufen werden, da die zugehörigen Matrixelemente verschwinden. Dies bedeutet, daß die Unterteilung auch hier das frühzeitige Erkennen verschwindender Matrixelemente ermöglicht. Weiterhin können durch eine geschickte Sortierung der Konfigurationsmenge die Wechselwirkungen durch sich wiederholende Muster charakterisiert werden. Auf diese Muster, die im Voraus berechnet und abgespeichert werden, wird in jedem Iterationsschritt zugegriffen. Die implementierten Muster speichern nicht nur Informationen über die Positionen der nicht-verschwindenden wechselwirkenden Konfigurationen ab, sondern auch deren Charakterisierung in Form der auftretenden Anregungsstufe und der wechselwirkenden Orbitale.Die Berechnung der Wechselwirkungen auf der Ebene der inaktiven Konfigurationsreste, für die die Anzahl der besetzten Orbitale in einer Konfiguration am Größten ist, erfolgt nicht in Form der natürlichen Darstellung als besetzte Orbitale, sondern in der invertierten Form, bei der fehlende Elektronen als Löcher berücksichtigt werden. Da die Anzahl der auftretenden Löcher um Größenordnungen kleiner ist als die der besetzten Orbitale, wird eine signifikante Beschleunigung bei der Berechnung der inaktiv-Wechselwirkungen erreicht. Die Ausnutzung der Muster ermöglicht es für den Aufbau der Hamilton-Matrix nur die wechselwirkenden Konfigurationsreste zu berücksichtigen, die zu nicht-verschwindenden Matrixelementen beitragen.

Die Implementierung eines vergleichbaren Algorithmus für die Wechselwirkungen der externen Konfigurationsreste ist aufgrund des immensen Speicherbedarfs, der durch die große Anzahl der externen Orbitale entsteht, nicht möglich. Daher wurden Algorithmen für ein direktes Verfahren, welches die eingeschränkten Möglichkeiten der externen Wechselwirkungen ausnutzt, abgeleitet und implementiert. Diese Algorithmen unterscheiden verschiedene Fälle bezüglich der Position der Matrixblöcke in der Hamilton-Matrix und der Anregungsstufe auf interner Ebene. Die Wechselwirkungen wurden nicht mit Hilfe der externen Konfigurationsreste berechnet, die wiederum aus den externen Orbitalen aufgebaut werden, sondern direkt aus den auftretenden Orbitalen. Dieses Verfahren ist weniger aufwendig als die Berechnung auf Konfigurationsebene und benötigt nur minimalen Speicher verglichen mit einem auf Muster basierenden Verfahren.

Neben einer effizienten Implementierung der Hamilton-Matrix reduziert die Parallelisierung des MR-MP2-Verfahrens die Gesamtrechenzeit. Das im neuen MR-MP2-Program (DIESEL-MP) verwendete Schema richtet sich nach der Menge an verfügbarem Hauptspeicher und beruht auf den Mustern der nicht-verschwindenden inaktiven Konfigurationsreste. Die implementierte Parallelisierung benötigt einen beträchtlich kleineren Hauptspeicherbedarf als das alte DIESEL Programmpaket und erlaubt so die Berechnung größerer Moleküle. Parallelisierung erzeugt jedoch verglichen mit einem seriellen Code einen gewissen Kommunikationsaufwand. Die Geschwindigkeit des neuen Programmes auf einem Knoten, ist bei kleinen Testsystemen (42 korrelierte Elektronen, 150 Basisfunktionen, 16 inaktive und 9 aktive Orbitale) vergleichbar zum MR-MP2-Programmcode im DIESEL. Vergleiche mit größeren Systemen, für die das Programm ausgelegt ist, waren aufgrund der Beschränkungen im alten DIESEL nicht möglich.

Die Infrastruktur des Programms wurde so konzipiert, daß eine möglichst große Flexibilität im Hinblick auf zukünftige Erweiterungen möglich ist. Die Struktur der Klassen des DIESEL-MP-Programmes ermöglicht ein volles MR-MP2-Verfahren, erlaubt aber auch eine einfache Einbeziehung von Selektionsschemata, die sowohl auf individuellen oder orbitalbasierten Ansätzen aufbauen. Die vorhandene Klassenstruktur kann ebenfalls für ein MR-CI-Verfahren verwendet werden.

Den Schwachpunkt des MR-MP2-Algorithmus im alten DIESEL-Programmpaket stellte die Behandlung der Integrale dar. Die implementierte Integralschnittstelle des neuen

DIESEL-MP zum TURBOMOLE-Programmpaket besitzt eine Reihe von Vorteilen. Zum einen wird der Speicherplatz der Zwei-Elektronen-Integrale durch die Verwendung von RI-Integralen und einer "on-the-fly"-Generierung der vollständigen Vier-Index-Integrale deutlich reduziert. Weiterhin werden die RI-Integrale lediglich einmal zur Berechnung der Inhomogenität und der Kontraktion zu effektiven Ein-Elektronen-Integralen benötigt. Im iterativen Prozess sind die RI-Integrale nicht mehr notwendig. Die Transformation der RI- zu Vier-Index-Integralen wird demnach nur einmal durchgeführt und stellt nicht den geschwindigkeitsbestimmenden Schritt im Verfahren dar. Die Möglichkeit verschiedene Orbitale (Hartree-Fock, Lokales-Hartree-Fock oder gradienten-korrigierte Dichtefunktionale) für die Integraltransformation zu verwenden stellt einen weiteren Vorteil des Interfaces zum TURBOMOLE Programm dar. Die Möglichkeit durch dieses Interface das DIESEL-MP-Programm einfach in bestehende QM/MM Umgebungen wie z.B. das CHEMSHELL[204] einzubinden, eröffnet die Perspektive der genauen Behandlung von realen System wie Enzymen.

Bibliography

[1] A. R. Leach, *Molecular Modelling: Principles and Applications*, Pearson Education, Harlow, 2nd edition, **2001**.

[2] P. Sherwood, *Hybrid Quantum Mechanics/Molecular Mechanics Approaches*. In J. Grotendorst (ed.), *Modern Methods and Algorithms of Quantum Chemistry*, John von Neumann Insitute for Computing, Jülich, **2000**, 257.

[3] P. R. Ortiz de Montellano, *Cytochrome P450: Structure, Mechanisms, and Biochemistry*, Volume 2, Plenum Press, New York, 2nd edition, **1995**.

[4] J. C. Schöneboom, H. Lin, N. Reuter, W. Thiel, S. Cohen, F. Ogliaro, S. Shaik, *J. Am. Chem. Soc.* **2002**, *124*, 8142.

[5] P. J. Knowles, H.-J. Werner, M. Schütz. *Ab initio Methods for Electron Correlation in Molecules*. In J. Grotendorst (ed.), *Modern Methods and Algorithms of Quantum Chemistry*, John von Neumann Insitute for Computing, Jülich, **2000**, 69.

[6] M. Schütz, G. Hetzer, H.-J. Werner, *J. Chem. Phys.* **1999**, *111*, 5691.

[7] G. Hetzer, M. Schütz, H. Stoll, H.-J. Werner, *J. Chem. Phys.* **2000**, *111*, 9443.

[8] S. Saebø, P. Pulay, *J. Chem. Phys.* **2001**, *115*, 3975.

[9] D. Walter, A. B. Szilva, K. Niedfeldt, E. A. Carter, *J. Chem. Phys.* **2002**, *117*, 1982.

[10] J. Gauss, *Coupled Cluster Theory*. In P. v. R. Schleyer, N. Allinger, T. Clark, J. Gasteiger, P. Kollman, H. F. Schaefer, III., P. R. Schreiner (eds.), *The Encyclopedia of Computational Chemistry*, John Wiley and Sons, Chichester, **1998**, 485.

[11] M. Schütz, *Phys. Chem. Chem. Phys.* **2002**, *4*, 3941.

[12] M. Schütz, *J. Chem. Phys.* **2000**, *113*, 9986.

[13] M. Schütz, *J. Chem. Phys.* **2002**, *116*, 8772.

[14] R. Ahlrichs, S. Elliot, U. Huniar, *Ab initio Treatment of Large Molecules.* In J. Grotendorst (ed.), *Modern Methods and Algorithms of Quantum Chemistry*, John von Neumann Insitute for Computing, Jülich, **2000**, 1.

[15] C. Ochsenfeld, C. A. White, M. Head-Gordon, *J. Chem. Phys.* **1998**, *109*, 1663.

[16] C. Ochsenfeld, *Chem. Phys. Lett.* **2000**, *327*, 216.

[17] W. Liang, Y. Shao, C. Ochsenfeld, A. T. Bell, M. Head-Gordon, *Chem. Phys. Lett.* **2002**, *358*, 43.

[18] W. T. Borden, *Diradicals.* In P. v. R. Schleyer, N. Allinger, T. Clark, J. Gasteiger, P. Kollman, H. F. Schaefer, III., P. R. Schreiner (eds.), *The Encyclopedia of Computational Chemistry*, John Wiley and Sons, Chichester, **1998**, 708.

[19] G. Frenking, T. Wagner, *Transition Metal Chemistry.* In P. v. R. Schleyer, N. Allinger, T. Clark, J. Gasteiger, P. Kollman, H. F. Schaefer, III., P. R. Schreiner (eds.), *The Encyclopedia of Computational Chemistry*, John Wiley and Sons, Chichester, **1998**, 3073.

[20] C. W. Bauschlicher, Jr., *Transition Metals: Applications.* In P. v. R. Schleyer, N. Allinger, T. Clark, J. Gasteiger, P. Kollman, H. F. Schaefer, III., P. R. Schreiner (eds.), *The Encyclopedia of Computational Chemistry*, John Wiley and Sons, Chichester, **1998**, 3084.

[21] W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley VCH, Weinheim, 2nd edition, **2001**.

[22] M. Torrent, M. Solá, G. Frenking, *Chem. Rev.* **2000**, *100*, 439.

[23] G. Frenking, N. Fröhlich, *Chem. Rev.* **2000**, *100*, 717.

[24] I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Myers, D. A. Benson, R. M. Sweet, D. Ringe, G. A. Petsko, S. G. Sligar, *Science* **2000**, *287*, 1675.

[25] Y.-K. Choe, T. Hashimoto, H. Nakano, K. Hirao, *Chem. Phys. Lett.* **1998**, *295*, 380.

[26] Y.-K. Choe, T. Nakajima, K. Hirao, *J. Chem. Phys.* **1999**, *111*, 3837.

[27] P. Celani, H.-J. Werner, *J. Chem. Phys.* **2000**, *112*, 5546.

[28] M. Parac, S. Grimme, *J. Phys. Chem. A* **2002**, *106*, 6844.

[29] C. Neiss, P. Saalfrank, M. Parac, S. Grimme, *J. Phys. Chem. A* **2003**, *107*, 140.

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

[31] R. J. Buenker, S. D. Peyerimhoff, *Theor. Chim. Acta* **1975**, *39*, 217.

[32] M. Hanrath, B. Engels, *Chemical Physics* **1997**, *225*, 197.

[33] S. Grimme, M Waletzke, *Phys. Chem. Chem. Phys* **2000**, *2*, 2075.

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

[35] K. Andersson, P.-Å. Malmquist, B. O. Roos, A. J. Sadlej, K. Wolinski, **1990**, *94*, 5483.

[36] K. Andersson, P.-Å. Malmquist, B. O. Roos, *J. Chem. Phys.* **1992**, *96*, 1218.

[37] J. Finley, P.-Å. Malmquist, B. O. Roos, L. Serrano-Andrés, *Chem. Phys. Lett.* **1998**, *288*, 299.

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

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

[40] P. Pulay, S. Saebø, *Theor. Chim. Acta* **1986**, *69*, 357.

[41] S. Saebø, P. Pulay, *J. Chem. Phys.* **1987**, *86*, 914.

[42] P. Y. Ayala, G. E. Scuseria, *J. Chem. Phys.* **1999**, *110*, 3660.

[43] M. S. Lee, P. E. Maslen, M. Head-Gordon, *J. Chem. Phys.* **2000**, *112*, 3592.

[44] R. A. Friesner, R. B. Murphy, M. D. Beachy, M. N. Ringnalda, W. T. Pollard, B. D. Dunietz, Y. Cao, *J. Phys. Chem. A* **1999**, *103*, 1913.

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

[46] M. Schütz, H.-J. Werner, *J. Chem. Phys.* **2001**, *114*, 661.

[47] M. Schütz, H.-J. Werner, *Chem. Phys. Lett.* **2000**, *318*, 370.

[48] B. D. Dunietz, R. A. Friesner, *J. Chem. Phys.* **2001**, *115*, 11052.

[49] D. Walter, A. Venkatnathan, E. A. Carter, *J. Chem. Phys.* **2003**, *118*, 8127.

[50] R. B. Murphy, R. A. Friesner, *Chem. Phys. Lett.* **1998**, *288*, 403.

[51] K. Andersson, M. Barysz, A. Bernhardsson, M. R. A. Blomberg, D. L. Cooper, T. Fleig, M. P. Flscher, C. de Graf, B. A. Hess, G. Karlstöm, R. Lindh, P.-Å. Malmquist, P. Neogràdy, J. Olsen, B. O. Roos, A. J. Sadlej, M. Schtz, B. Schimmelpfennig, L. Seijo, L. Serrano-Andr'es, P. E. M. Siegbahn, J. Stålring, T. Thorsteinsson, V. Veryazov, P.-O. Widmark, *MOLCAS5*, **2000**, *Lund University, Sweden.*

[52] MOLPRO is a package of *ab initio* programs designed by H.-J. Werner and P. J. Knowles, Version 2002.1, R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P. J. Knowles, T. Korona, R. A. Lindh, W. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, H.-J. Werner, **2002**.

[53] B. Engels, M. Hanrath, *Universität Bonn, Germany*, **1997**.

[54] M. Hanrath, *Ein individuell selektierendes intern-extern-separiertes Multireferenz-Konfigurationswechselwirkungsverfahren*, PhD thesis, Universität Bonn, **1999**.

[55] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester, **1999**.

[56] A. Szabo, N. S. Ostlund, *Modern Quantum Chemistry*, Dover Publications, Mineola, **1989**.

[57] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester, **1999**, 59.

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

[59] D. Cremer, *Møller-Plesset Perturbation Theory.* In P. v. R. Schleyer, N. Allinger, T. Clark, J. Gasteiger, P. Kollman, H. F. Schaefer, III., P. R. Schreiner (eds.), *The*

Encyclopedia of Computational Chemistry, John Wiley and Sons, Chichester, **1998**, 1706.

[60] K. Wolinski, H. L. Sellers, P. Pulay, *Chem. Phys. Lett.* **1987**, *140*, 225.

[61] K. Wolinski, P. Pulay, *J. Chem. Phys.* **1989**, *90*, 3647.

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

[63] E. R. Davidson, A. A. Jarzecki, *Multi-Reference Perturbation Theory.* In K. Hirao (ed.), *Recent Advances in Multireference Methods*, World Scientific, Singapore, River Edge, London, **1999**, 31.

[64] V. N. Staroverov, E. R. Davidson, *Chem. Phys. Lett.* **1998**, *296*, 435.

[65] K. Hirao, *Chem. Phys. Lett.* **1992**, *196*, 397.

[66] K. Hirao, *Chem. Phys. Lett.* **1992**, *190*, 374.

[67] K. Hirao, *Chem. Phys. Lett.* **1993**, *201*, 59.

[68] Y.-K. Choe, Y. Nakao, K, Hirao, *J. Chem. Phys.* **2001**, *115*, 621.

[69] R. W. A. Havenith, H. J. J. Van Dam, J. H. van Lenthe, L. W. Jenneskens, *Chemical Physics* **1999**, *246*, 49.

[70] H. Nakano, *J. Chem. Phys.* **1993**, *99*, 7983.

[71] H. Umeda, S. Koseki, U. Nagashima, M. W. Schmidt, *J. Comput. Chem.* **2001**, *22*, 1243.

[72] H. A. Witek, H. Nakano, K. Hirao, *J. Chem. Phys.* **2003**, *118*, 8197.

[73] C. Angeli, R. Cimiraglia, J.-P. Malrieu, *Chem. Phys. Lett.* **2000**, *317*, 472.

[74] F. Chen, E. R. Davidson, S. Iwata, *Int. J. Quant. Chem.* **2002**, *86*, 256.

[75] S. Grimme, M. Parac, M. Waletzke, *Chem. Phys. Lett.* **2001**, *334*, 99.

[76] E. R. Davidson, *Chem. Phys. Lett.* **1995**, *241*, 432.

[77] H. J. J. Van Dam, J. H. Van Lenthe, P. Pulay, *Molecular Physics* **1998**, *93*, 431.

[78] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C++*, Cambridge University Press, Cambridge, **2002**.

[79] J. Olsen, M. P. Fülscher, *Chem. Phys. Lett.* **2000**, *326*, 225.

[80] Y.-K. Choe, H. A. Witek, J. P. Finley, K. Hirao, *J. Chem. Phys.* **2001**, *114*, 3913.

[81] H. A. Witek, Y.-K. Choe, J. P. Finley, K. Hirao, *J. Comput. Chem.* **2002**, *23*, 957.

[82] E. R. Davidson, *J. Comput. Phys.* **1975**, *17*, 87.

[83] Y. Saad, *Numerical Methdos for Large Eigenvalue Problems*, Manchester University Press, Manchester, **1992**.

[84] B. Liu. *The Simultaneous Expansion Method for the Iterative Solution of Several of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices*, LBL-8158, Lawrence Berkeley Laboratory, Berkeley, **1978**, 49.

[85] M. Crouzeix, B. Philippe, M. Sadkane, *J. Sci. Comput.* **1994**, *15*, 62.

[86] G. L. G. Sleijpen, H. A. van der Vorst, E. Meijerink, *El. Trans. on Numer. Anal.* **1998**, *7*, 75.

[87] H. M. Bücker, *Iteratively Solving Large Sparse Linear Systems on Parallel Computers.* In J. Grotendorst, D. Marx, A. Muramatsu (eds.), *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, John von Neumann Insitute for Computing, Jülich, **2002**, 521.

[88] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing, New York, **1996**.

[89] G. Alefeld, I. Lenhardt, H. Obermaier, *Parallele numerische Verfahren*, Springer, Berlin, **2002**.

[90] R. J. Buenker, S. Krebs, *The Configuration-Driven Approach for Multi-Reference Configuration Interaction Calculations.* In K. Hirao (ed.), *Recent Advances in Multireference Methods*, World Scientific, Singapore, River Edge, London, **1999**, 1.

[91] V. Pless, *Ein direktes, individuell selektierendes Multireferenz-Konfigurationswechselwirkungsverfahren*, PhD thesis, Universität Bonn, **1994**.

[92] R. Pauncz, *Spin Eigenfunctions, Construction and Use*, Plenum Press, New York, **1979**.

[93] W. Duch, J. Karwowski, *Symmetric Group Approach to Configuration Interaction Methods*, North Holland, Amsterdam, **1985**.

[94] R. J. Buenker, *Prcoeedings of the Workshop on Quantum Chemistry and Molecular Physics*, Wollongong, **1980**.

[95] R. J. Buenker, *Current Aspects of Quantum Chemistry: Studies in Physical and Theoretical Chemistry, Vol. 21*, Elsevier, Amsterdam, **1981**.

[96] R. J. Buenker, R. A. Philips, *J. Mol. Struct.* **1985**, *123*, 291.

[97] *http://www.openmp.org*

[98] R. Chandra, R. Menon, L. Dagum, D. Kohr, D. Maydan, J. McDonald, *Parallel Programming in OpenMP*, Morgan Kaufmann, San Francisco, **2000**.

[99] X. Tian, A. Bik, M. Girkar, P. Grey, H. Saito, E. Su, *Intel Technology Journal Q1* **2002**, *6*, 1.

[100] *http://www-unix.mcs.anl.gov/mpi/*

[101] W. Gropp, E. Lusk, A. Skjellum, *Using MPI. Portable Parallel Programming with the Message-Passing Interface*, MIT Press, Cambridge, Massachusetts, 2nd edition, **1999**.

[102] P. Sanders, T. Worsch, *Parallele Programmierung mit MPI - ein Praktikum*, Logos Verlag, Berlin, **1998**.

[103] R. R. Sedgewick, *Algorithmen in C++*, Pearson Studium, München, 3. überarbeitete Auflage, **2002**.

[104] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Chem. Phys. Lett.* **1998**, *294*, 143.

[105] M. Feyereisen, G. Fitzgerald, A. Komornicki, *Chem. Phys. Lett.* **1993**, *208*, 359.

[106] D. E. Bernholdt, R. J. Harrison, *Chem. Phys. Lett.* **1996**, *250*, 477.

[107] TURBOMOLE, Version. 5.6, Quantum Chemistry Group, University Karlsruhe, Germany, **2002**.

[108] P. W. Musch, B. Engels, *Angew. Chem.* **2001**, *113*, 3951; *Angew. Chem. Int. Ed.* **2001**, *40*, 3833.

[109] P. W. Musch, C. Remenyi, H. Helten, B. Engels, *J. Am. Chem. Soc.* **2002**, *124*, 1823.

[110] P. W. Musch, D. Scheidel, B. Engels, *J. Phys. Chem. A* **2003**, in press.

[111] K. C. Nicolaou, W.-M. Dai, *Angew. Chem.* **1991**, *103*, 1453; *Angew. Chem. Int. Ed.* **1991**, 30, 1387.

[112] K. Edo, Y. Koide, *Neocarzinostatin chromophore: structure and mechanism of DNA cleavage.* In Hiroshi Maeda,Kiyoto Edo, Nakao Ishida (eds.), *Neocarzinostatin: The Past, the Present and the Future of an Anticancer Drug*, Springer, Tokyo, **1997**, 23.

[113] N. Ishida, K. Miyazaki, K. Kumagai, M. Rikimaru, *J. Antibiot.* **1965**, *18*, 68.

[114] K. Edo, M. Mizugaki, Y. Koide, H. Seto, K. Furihata, N. Otake, N. Ishida, *Tetrahedron Lett.* **1985**, *26*, 331.

[115] K. C. Nicolaou, A. L. Smith, *Acc. Chem. Res.* **1992**, *25*, 297.

[116] K. H. Kim, B. M. Kwon, A. G. Myers, D. C. Rees, *Science* **1993**, *262*, 1042.

[117] A. G. Myers, *Tetrahedron Lett.* **1987**, *28*, 4493.

[118] O. D. Hensens, G. L. Helms, D. L. Zink, D.-H. Chin, L. S. Kappen, I. H. Goldberg, *J. Am. Chem. Soc.* **1993**, *115*, 11030.

[119] A. G. Myers, P. J. Proteau, *J. Am. Chem. Soc.* **1989**, *111*, 1146.

[120] T. Mastuura, I. Saito, H. Sugiyama, T. Morii, H. Kawabata, *Pure & Appl. Chem* **1989**, *61*, 473.

[121] A. G. Myers, S. B. Cohen, B.-M. Kwon, *J. Am. Chem. Soc.* **1994**, *116*, 1670.

[122] H. Sugiyama, K. Yamashita, T. Fujiwara, I. Saito, *Tetrahedron* **1994**, *50*, 1311.

[123] A. G. Myers, S. P. Arvedson, R. W. Lee, *J. Am. Chem. Soc.* **1996**, *118*, 4725.

[124] D. H. Chin, I. H. Goldberg, *J. Am. Chem. Soc.* **1992**, *114*, 1914.

[125] R. R. Jones, R. G. Bergman, *J. Am. Chem. Soc.* **1972**, *84*, 660.

[126] T. P. Lockhart, P. B. Comita, R. G. Bergman, *J. Am. Chem. Soc.* **1981**, *103*, 4082.

[127] A. G. Myers, E. Y. Kuo, N. S. Finney, *J. Am. Chem. Soc.* **1989**, *111*, 8057.

[128] R. Nagata, H. Yamanaka, E. Okazaki, I. Saito, *Tetrahedron Lett.* **1989**, *30*, 4995.

[129] Y. Araki, T. Konoike, *Tetrahedron Lett.* **1998**, *39*, 5549.

[130] M. Tokuda, K. Fujiwara, T. Gomibuchi, M. Hirama, M. Uesugi, Y. Sugiura, *Tetrahedron Lett.* **1992**, *34*, 669.

[131] F. Ferri, R. Brückner, R. Herges, *New J. Chem.* **1998**, 531.

[132] K. K. Wang, H.-R. Zhang, J. L. Petersen, *J. Org. Chem.* **1999**, *64*, 1650.

[133] K. C. Nicolaou, P. Maligres, J. Shin, E. de Leon, D. Rideout, *J. Am. Chem. Soc.* **1990**, *112*, 7825.

[134] Y. W. Andemichael, Y. G. Gu, K. K. Wang, *J. Org. Chem.* **1992**, *57*, 794.

[135] A. G. Myers, P. S. Dragovich, E. Y. Kuo, *J. Am. Chem. Soc.* **1992**, *114*, 9369.

[136] P. G. Wenthold, S. G. Wierschke, J. J. Nash, R. R. Squires, *J. Am. Chem. Soc.* **1993**, *115*, 12611.

[137] K. K. Wang, *Chem. Rev.* **1996**, *96*, 207.

[138] M. Schmittel, M. Strittmatter, S. Kiau, *Angew. Chem.* **1996**, *108*, 1952; *Angew. Chem. Int. Ed.* **1996**, *35*, 1843.

[139] B. Engels, Ch. Lennartz, M. Hanrath, M. Schmittel, M. Strittmatter, *Angew. Chem.* **1998**, *110*, 2067; *Angew. Chem. Int. Ed.* **1998**, *37*, 1960.

[140] P. R. Schreiner, M. Prall, *J. Am. Chem. Soc.* **1999**, *121*, 8615.

[141] J. Wisniewski Grissom, G. U. Gunawardena, D. Klingberg, D. Huang, *Tetrahedron* **1996**, *52*, 6453.

[142] P. G. Wenthold, M. A. Lipton, *J. Am. Chem. Soc.* **2000**, *122*, 9265.

[143] C. J. Cramer, B. L. Kormos, M. Seierstad, E. C. Sherer, P. Winget, *Org. Lett.* **2001**, *3*, 1881.

[144] S. R. Brunette, M. A. Lipton, *J. Org. Chem.* **2000**, *65*, 5114.

[145] R. W. Saalfrank, M. Haubner, C. Deutscher, W. Bauer, T. Clark, *J. Org. Chem.* **1999**, *64*, 6166.

[146] M. Schmittel, J.-P. Steffen, M. Maywald, B. Engels, H. Helten, P. Musch, *J. Chem. Soc., Perkin Trans. 2* **2001**, 1331.

[147] B. Engels, M. Hanrath, *J. Am. Chem. Soc.* **1998**, *120*, 6356.

[148] B. Engels, M. Hanrath, Ch. Lennartz, *Computers and Chemistry* **2001**, *25*, 15.

[149] M. Prall, A. Wittkopp, P. R. Schreiner, *J. Phys. Chem. A* **2001**, *105*, 9265.

[150] T. S. Hughes, B. K. Carpenter, *J. Chem. Soc., Perkin Trans. 2* **1999**, 2291.

[151] P. W. Musch, B. Engels, *J. Am. Chem. Soc.* **2001**, *123*, 5557.

[152] J. Gabriel Garcia, B. Ramos, L. M. Pratt, A. Rodríguez, *Tetrahedron Lett.* **1995**, *36*, 7391.

[153] C. J. Cramer, R. R. Squires, *Org. Lett.* **1999**, *1*, 215.

[154] K. K. Wang, B. Liu, Y. de Lu, *Tetrahedron Lett.* **1995**, *36*, 3785.

[155] K. Fujiwara, H. Sakai, M. Hirama, *J. Org. Chem.* **1991**, *56*, 1688.

[156] J. Suffert, E. Abraham, S. Raeppel, R. Brückner, *Liebigs Ann.* **1996**, 447.

[157] S. W. Scheuplein, R. Machinek, J. Suffert, R. Brückner, *Tetrahedron Lett.* **1993**, *34*, 6549.

[158] T. Doi, T. Takahashi, *J. Org. Chem.* **1991**, *56*, 3465.

[159] *Gaussian98, Rev. A.7*, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, Jr. J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski,

J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gompertsand R: L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, **1998**, *Gaussian Inc., Pittsburgh PA.*

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

[161] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.

[162] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257.

[163] A. Klamt, G. Schürmann, *J. Chem. Soc., Perkin Trans. 2* **1993**, 799.

[164] Thomas J. Dunning, Jr., *J. Chem. Phys.* **1989**, *90*, 1007.

[165] MOLPRO is a package of *ab initio* programs designed by H.-J. Werner and P. J. Knowles, Version 2000.1, R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, P. J. Knowles, T. Korona, R. A. Lindh, W. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, H.-J. Werner, **2000**.

[166] M. Schmittel, J.-P. Steffen, D. Auer, M. Maywald, *Tetrahedron Lett.* **1997**, *38*, 6177.

[167] H. W. Moore, M. Yerxa, *Chemtracts* **1992**, 273.

[168] R. W. Sullivan, V. M. Coghlan, S. A. Munk, M. W. Reed, H. W. Moore, *J. Org. Chem.* **1994**, *59*, 2276.

[169] A. Padwa, D. Austin, U. Chiacchio, J. M. Kassir, A. Rescifina, S. L. Xu, *Tetrahedron Lett.* **1991**, 5923.

[170] L. D. Foland, J. O. Karlsson, S. T. Perri, R. Schwabe, S. L. Xu, S. Patil, H. W. Moore, *J. Am. Chem. Soc.* **1989**, *111*, 975.

[171] M. Schmittel, M. Strittmatter, S. Kiau, *Tetrahedron Lett.* **1995**, *36*, 4995.

[172] M. Schmittel, S. Kiau, *Liebigs Ann.* **1997**, 733.

[173] M. Schmittel, M Keller, S. Kiau, *Chem. Eur. J.* **1997**, *3*, 807.

[174] B. Engels, J. C. Schöneboom, A. Münster, S. Groetsch, M. Christl, *J. Am. Chem. Soc.* **2002**, *124*, 287.

[175] P. Musch, *Diplomarbeit, Universität Würzburg* **2000**.

[176] R. Seeger, J. A. Pople, *J. Chem. Phys.* **1977**, *66*, 3045.

[177] R. Bauernschmitt, R. Ahlrichs, *J. Chem. Phys.* **1996**, *104*, 9047.

[178] G. D. Purvis, R. J. Bartlett, *J. Chem. Phys.* **1982**, *76*, 1910.

[179] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon, *Chem. Phys. Lett.* **1989**, *157*, 479.

[180] P.-O.Widmark, P.-Å. Malmquist, B. O. Roos, *Theor. Chim. Acta* **1990**, *77*, 291.

[181] J. Gräfenstein, A. M. Hjerpe, E. Kraka, D. Cremer, *J. Phys. Chem. A* **2000**, *104*, 1748.

[182] J. Gräfenstein, D. Cremer, *Phys. Chem. Chem. Phys.* **2000**, *2*, 2091.

[183] E. D. Glendering, A. E. Reed, J. E. Carpenter, F. Weinhold, *NBO 4.0*, **1998**.

[184] M. Prall, A. Krüger, P. R. Schreiner, H. Hopf, *Chem. Eur. J.* **2001**, *7*, 4386.

[185] W. Pan, M. Balci, P. B. Shevlin, *J. Am. Chem. Soc.* **1997**, *119*, 5035.

[186] W. Pan, P. B. Shevlin, *J. Am. Chem. Soc.* **1997**, *119*, 5091.

[187] C. J. Emanuel, P. B. Shevlin, *J. Am. Chem. Soc.* **1994**, *116*, 5991.

[188] M. L. McKee, P. B. Shevlin, M. J. Zottola, *J. Am. Chem. Soc.* **2001**, *123*, 9418.

[189] M. Christl, M. Braun, G. Mueller, *Angew. Chem.* **1992**, *104*, 471; *Angew. Chem. Int. Ed.* **1992**, *31*, 473.

[190] I. Yavari, F. Nourmohammadian, D. Tahmassebi, *J. Mol. Struct.* **2001**, *542*, 199.

[191] M. A. Hofmann, U. Bergsträßer, G. J. Reiß, L. Nyulászi, M. Regitz, *Angew. Chem.* **2000**, *112*, 1318; *Angew. Chem. Int. Ed.* **2000**, *39*, 1261.

[192] M. Christl, S. Groetsch, *Eur. J. Org. Chem.* **2000**, 1871.

[193] J. Schöneboom, S. Groetsch, M. Christl, B. Engels, *Chem. Eur. J.* **2003**, *9*, 4641.

[194] M. Christl, S. Drinkuth, *Eur. J. Org. Chem.* **1998**, 273.

[195] R. P. Johnson, *Chem. Rev.* **1989**, *89*, 1111.

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

[197] P.-O.Widmark, B. J. Persson, B. O. Roos, *Theor. Chim. Acta* **1991**, *79*, 419.

[198] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.

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

[200] J. B. Lambert, Y. Zhao, R. W. Emblidge, L. A. Salvador, X. Liu, J.-H. So, E. C. Chelius, *Acc. Chem. Res.* **1999**, *32*, 183.

[201] G. E. Davico, V. M. Bierbaum, C. H. DePuy, G. B. Ellison, R. R. Squires, *J. Am. Chem. Soc.* **1995**, *117*, 2590.

[202] A. Nicolaides, D. M. Smith, F. Jensen, L. J. Radom, *J. Am. Chem. Soc.* **1997**, *119*, 8032.

[203] S. Drinkuth, S. Groetsch, E.-M. Peters, K. Peters, M. Christl, *Eur. J. Org. Chem.* **2001**, 2665.

[204] *http://www.cse.clrc.ac.uk/qcg/chemshell/*

Erklärung

Hiermit erkläre ich an Eides statt, daß ich die Dissertation

"Large-Scale Applications of Multi-Reference Methods in Chemistry

and

Development of a Multi-Reference Møller-Plesset Perturbation Theory Program"

selbständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfs-
mittel benutzt habe.

Ich erkläre weiterhin, daß diese Dissertation weder in gleicher oder anderer Form bereits
in einem anderen Prüfungsverfahren vorgelegen hat.

Ich habe früher außer den mit dem Zulassungsgesuch urkundlich vorgelegten Graden keine
weiteren akademischen Grade erworben oder zu erwerben versucht.

Würzburg, 15.12.2003

Patrick Musch

Lebenslauf

Name:	Patrick Werner Musch
Geburt:	28. April 1975 in Erlenbach am Main
Staatsangehörigkeit:	deutsch
Familienstand:	ledig

Schulbildung

Sept. 1981 – Juli 1985	Besuch der Johannes-Obernburger Grundschule in Obernburg am Main
Sept. 1985 – Juli 1994	Besuch des Hermann-Staudinger-Gymnasiums in Erlenbach am Main und Erwerb der Allgemeinen Hochschulreife

Zivildienst

Juli 1994 – Okt. 1995	Ambulanter Pflegedienst bei der Caritas Sozialstation Sankt Johannes in Klingenberg am Main

Hochschulstudium

Nov. 1995 – Sept. 2000	Studium des Faches Chemie (Diplom)an der Bayerischen Julius-Maximilians-Universität Würzburg
Oktober 1997	Absolvieren der Prüfungen zum Vordiplom (Fakultätspreis)
März 1999 – Juli 1999	Praktikum bei Dr. John Simmie an der National University of Ireland, Galway über die Modellierung von Verbrennungsprozessen in einer "Rapid-Compression-Machine"
Januar 2000	Absolvieren der Prüfungen zum Diplom-Hauptexamen
Feb. 2000 – Sept. 2000	Diplomarbeit im Institut für Organische Chemie der Universität Würzburg unter Leitung von Prof. Dr. Bernd Engels und Abschluß des Diploms (mit Auszeichnung); Fakultätspreis
Okt. 2000 – Dez. 2003	Durchführung der Arbeiten und Anfertigung der Dissertation im Institut für Organische Chemie der Universität Würzburg unter Leitung von Prof. Dr. Bernd Engels gefördert durch ein Doktorandenstipendium der Stiftung Stipendien-Fonds im Verband der Chemischen Industrie im Zeitraum von Jan. 2001 bis Dez. 2002

Berufliche Tätigkeiten

Jan. 2003 – Dez. 2003	Wissenschaftlicher Mitarbeiter am Institut für Organische Chemie der Universität Würzburg (Administration des Unix Netzwerkes)