Quantum Chemical Modeling of Phosphoesterase Mimics and Chemistry in Confined Spaces

Henrik Daver

Academic dissertation for the Degree of Doctor of Philosophy in Organic Chemistry at Stockholm University to be publicly defended on Friday 1 December 2017 at 14.00 in Magnélisalen, Kemiska övningslaboratoriet, Svante Arrhenius väg 16 B.

Abstract

In this thesis, density functional theory is employed in the study of two kinds of systems that can be considered to be biomimetic in their own ways. First, three binuclear metal complexes, synthesized by the group of Prof. Ebbe Nordlander, have been investigated. The complexes are designed to resemble the active sites of phosphatase enzymes and have been examined in complexes where either two Zn(II) ions or one Fe(III) and one Mn(II) ion are bound. These dinuclear compounds were studied as catalysts for the hydrolysis of bis(2,4-dinitrophenyl) phosphate and the transesterification of 2-hydroxypropyl *p*-nitrophenyl phosphate, which are model systems for the same reactions occurring in DNA or RNA. It was found that the two reactions take place in similar ways: a hydroxide ion that is terminally bound to one of the metal centers acts either as a nucleophile in the hydrolysis reaction or as a base in the transesterification. The leaving groups depart in an effectively concerted manner, and the formed catalyst-product complexes are predicted to be the resting states of the catalytic cycles. The rate-determining free energy barriers are identified from the catalyst-product complex in one catalytic cycle to the transition state of nucleophilic attack in the next.

Another type of biomimetic modeling is made with an aim of imitating the conceptual features of selective binding of guests and screening them from solute-solvent interactions. Such features are found in so-called nanocontainers, and this thesis is concerned with studies of two capsules synthesized by the group of Prof. Julius Rebek, Jr. First, the cycloaddition of phenyl acetylene and phenyl azide has experimentally been observed to be accelerated in the presence of a capsule. Computational studies were herein performed on this system, and a previously unrecognized structure of the capsule is discovered. Two main factors are then identified as sources of the rate acceleration compared to the uncatalyzed reaction, namely the reduction of the entropic component and the selective destabilization of the reactant supercomplex over the transition state.

In the second capsule study, the alkane binding trends of a water-soluble cavitand was studied. It is found that implicit solvation models fail severely in reproducing the experimental equilibrium observed between binding of *n*-decane by the cavitand monomer and encapsulation in the capsule dimer. A mixed explicit/implicit solvation protocol is developed to better quantify the effect of hydrating the cavitand, and a simple correction to the hydration free energy of a single water molecule is proposed to remedy this. The resulting scheme is used to predict new hydration free energies of the cavitand complexes, resulting in significant improvement *vis-à-vis* experiments.

The computational results presented in this thesis show the usefulness of the quantum chemical calculations to develop understanding of experimental trends observed for substrate binding and catalysis. In particular, the methodology is shown to be versatile enough such that experimental observations can be reproduced for such diverse systems as studied herein.

Keywords: *density functional theory, catalysis, phosphoester hydrolysis, transesterification, supramolecular chemistry, inclusion complex, host-guest chemistry, cycloaddition.*

Stockholm 2017
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-148259

ISBN 978-91-7797-016-3
ISBN 978-91-7797-017-0

Department of Organic Chemistry

Stockholm University, 106 91 Stockholm

QUANTUM CHEMICAL MODELING OF PHOSPHOESTERASE MIMICS AND CHEMISTRY IN CONFINED SPACES

Henrik Daver

Stockholm
University

Quantum Chemical Modeling of Phosphoesterase Mimics and Chemistry in Confined Spaces

Henrik Daver

ISBN print 978-91-7797-016-3
ISBN PDF 978-91-7797-017-0

Cover picture: "In the prebio(mime)tic puddle", © the author, 2017
Printed in Sweden by Universitetsservice US-AB, Stockholm 2017
Distributor: Department of Organic Chemistry

– Vi är naturligtvis väldigt glada att kunna offentliggöra resultatet av många års forskningar, som nu resulterat i siffran sju.

Doktor Eva i Fem myror är fler än fyra elefanter

List of Publications

This thesis is based on the following papers.

I **Theoretical Study of Phosphodiester Hydrolysis and Transesterification Catalyzed by an Unsymmetric Biomimetic Dizinc Complex**
Henrik Daver, Biswanath Das, Ebbe Nordlander and Fahmi Himo
Inorg. Chem. **2016**, *55*, 1872-1882
DOI: 10.1021/acs.inorgchem.5b02733

II **A dinuclear zinc(II) complex of a new unsymmetric ligand with an N_5O_2 donor set; A structural and functional model for the active site of zinc phosphoesterases**
Biswanath Das, Henrik Daver, Monika Pyrkosz-Bulska, Elke Persch, Suman K. Barman, Rabindranath Mukherjee, Elzbieta Gumienna-Kontecka, Martin Jarenmark, Fahmi Himo and Ebbe Nordlander
J. Inorg. Biochem. **2014**, *132*, 6-17
DOI: 10.1016/j.jinorgbio.2013.08.001

III **A Heterobimetallic $Fe^{III}Mn^{II}$ Complex of an Unsymmetrical Dinucleating Ligand: A Structural and Functional Model Complex for the Active Site of Purple Acid Phosphatase of Sweet Potato**
Biswanath Das, Henrik Daver, Amrendra Singh, Reena Singh, Matti Haukka, Serhiy Demeshko, Franc Meyer, George Lisensky, Martin Jarenmark, Fahmi Himo and Ebbe Nordlander
Eur. J. Inorg. Chem. **2014**, *13*, 2204-2212
DOI: 10.1002/ejic.201301375

IV **Quantum Chemical Modeling of Cycloaddition Reaction in a Self-Assembled Capsule**
Henrik Daver, Jeremy N. Harvey, Julius Rebek, Jr. and Fahmi Himo
J. Am. Chem. Soc. **2017**, *In press.*
DOI: 10.1021/jacs.7b09102

V **Toward Accurate Quantum Chemical Modeling of Water-Soluble Self-Assembled Capsules**
Henrik Daver, Andrés Algarra, Julius Rebek, Jr., Jeremy N. Harvey and Fahmi Himo
Manuscript

In the papers, the author of this thesis performed all the calculations (Papers I-IV) or a major part of the calculations (Paper V), analyzed the results, wrote the first draft of the articles (Paper I, IV and V) or the theoretical parts thereof (Paper II and III), and took part in the subsequent revision of the drafts.

Contents

Abbreviations and atom colors

B3LYP	Becke's three-parameter functional with LYP correlation
BJ	Becke-Johnson, referring to the damping function used in DFT-D3
BP86	Functional with Becke's exchange functional and Perdew's correlation functional from 1986
BSSE	Basis set superposition error
C-PCM	Conductor-like polarizable continuum method
COSMO	Conductor-like solvation model
COSMO-RS	Conductor-like solvation model for real systems
DFT	Density functional theory
DFT-D	Density functional theory corrected with Grimme's dispersion corrections
DNA	Deoxyribonucleic acid
DOF	Degree of freedom
HF	Hartree-Fock, referring to the theory
LANL2DZ	Double-zeta basis set with pseudopotential core from the Los Alamos National Laboratories
LDA	Local density approximation
LYP	Lee, Yang and Parr, referring to the correlation functional developed by those authors
M06	Minnesota type functional published in 2006
NMR	Nuclear magnetic resonance spectroscopy
PAP	Purple acid phosphatase
PCM	Polarizable continuum method
PTE	Phosphotriesterase
RNA	Ribonucleic acid
RRHO	Rigid-rotor harmonic-oscillator model
SMD	Solvation model based on the solute electron density
TS	Transition state
TST	Transition state theory
TZVP	Ahlrich's triple-zeta basis set with valence and polarization functions
UEG	Uniform electron gas
VWN	Vosko, Wilks and Nusair, referring to the correlation functional developed by those authors

In the figures, all energies are given in kcal/mol and all distances are given in Ångströms. Atoms are colored according to the following scheme:

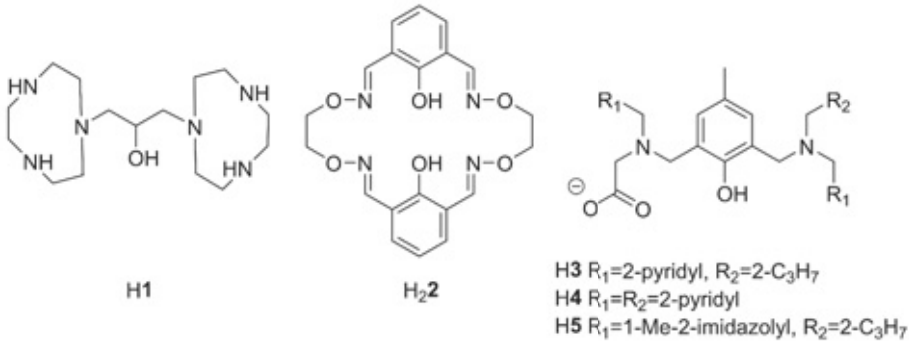
- Carbon (in ligand or capsule)
- Carbon (in substrate)
- Hydrogen
- Nitrogen
- Oxygen
- Phosphorus
- Manganese
- Iron
- Zinc
- Gallium

1 Introduction

Enzymes are large biological molecules that catalyze chemical reactions. They do so by selective binding of reactants to a so-called active site, where amino acid residues are arranged in such a way that the binding is favored and the reaction is accelerated.

In this thesis two kinds of systems are studied that can be considered to mimic enzyme active sites in different ways. The systems are examined by means of quantum chemistry, which allows that detailed free energies are calculated in order to reproduce the experimental findings and provide insights into the observed chemistry.

A direct biomimetic approach is to synthesize complexes that resemble the active sites of enzymes. Several biomimetic complexes have been synthesized that resemble the active site of phosphoesterases, i.e. enzymes that catalyze the hydrolysis of phosphates.[1-6] These enzymes typically bind two or three metal ions in their active sites,[1,7-10] so the complexes are designed to provide binding pockets for these metals. A number of ligands designed to bind two metal ions are shown in Figure 1. While **1**[11] and **2**[12] are symmetric, complexes **3**,[13] **4**[14] and **5**[15] are unsymmetric with respect to the coordination of the metal ions.



H1 H_2**2**

H3 R_1=2-pyridyl, R_2=2-C_3H_7
H4 R_1=R_2=2-pyridyl
H5 R_1=1-Me-2-imidazolyl, R_2=2-C_3H_7

Figure 1. Examples of water-soluble dinucleating ligands used in biomimetic phosphoesterase complexes.

One of the goals with designing such complexes is to gain insight into the workings of the model enzymes. Typically, though, the enzyme is a many times more efficient catalyst than the model complexes.[2,6] However, by

optimizing the ligand for activity, insights can be gained into the importance of certain factors for the catalysis.

The biomimetic complexes are also interesting on their own, i.e. as agents that can perform the same tasks as the enzymes but under different conditions or with modified reactivity. For phosphoesterase mimics, possible applications include the selective cleavage of DNA or RNA strings, where the mimic complex can have different sequence specificity than the enzymes, or as degrading agents towards organophosphate pesticides.[3]

Different complexes of ligands **3**, **4** and **5**, that bind either two Zn(II) ions or one Fe(III) and one Mn(II) ion are studied in this thesis. The complexes were experimentally studied as catalysts for hydrolysis and transesterification of two phosphate diesters. The aim of the computational investigations is to elucidate the catalytic mechanisms of the reactions, which can then be analyzed to propose modifications in order to increase the catalytic efficiency. The results of the computational investigations are presented in Chapter 3.

Next, the results of two studies of molecular capsules will be presented, i.e. studies of host molecules that more or less completely surround their guests.[16] A selection of such hosts is shown in Figure 2.[17-20] Selective binding of guests to the capsule host occurs when there is shape complementarity between the guest and the interior of the host.[21] The guests are screened from interactions with the solvent, and thus species that are unstable in solution can be stabilized in the presence of such hosts.[16,22] Additionally, several observations have been made of two or more guests being bound simultaneously.[23] The container can act as a nanoreactor, bringing two substrates together, and the system can even be catalytic.[22,24-29] The systems can be called biomimetic because the selective binding of substrates, the shielding of guests from solvent and the catalytic features conceptually resemble the situation in active sites of enzymes.[26,28,30,31]

One specific reaction that was experimentally observed to be accelerated in the presence of capsule $\mathbf{8}_2$ (Figure 2) is the cycloaddition reaction between phenyl acetylene and phenyl azide.[32] This was studied computationally, with the aim to determine the specific origins of rate acceleration in this system. The results are presented in Chapter 4.

The cycloaddition reaction was experimentally studied in a nonpolar solvent. To properly mimic an enzyme, the chemistry should be possible in the same medium as the enzymes are active in, i.e. water.[33] Capsule $\mathbf{9}_2$ which is very similar to $\mathbf{8}_2$ was made water-soluble by the choice of appropriate substituents at the R_4 position indicated in Figure 2.[20] When studied experimentally for the binding of alkanes of different lengths, some interesting trends were observed.[20,34,35] The system is studied computationally in Chapter 5. Standard methodology, as presented in Section 2.5, fails to reproduce the experimental findings, and therefore a new protocol is developed to properly model the solvation effects.

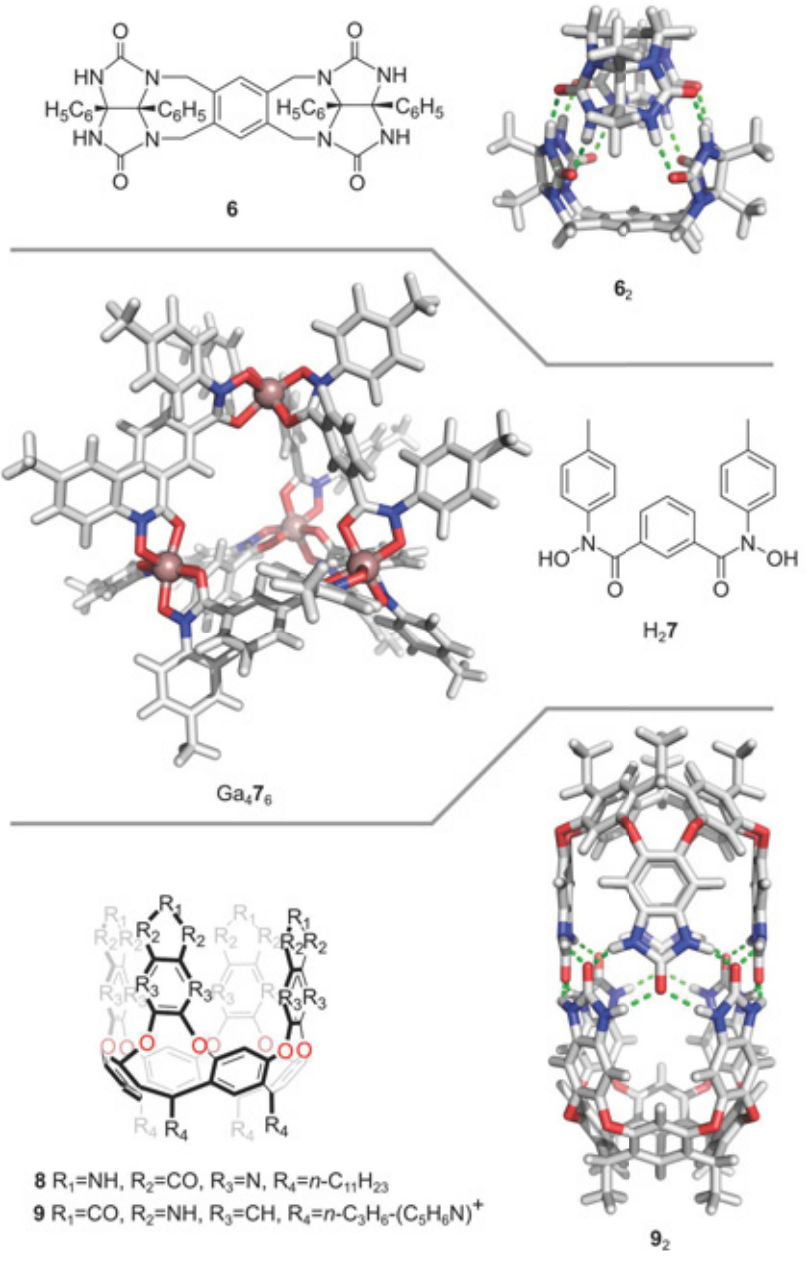


Figure 2. Examples of self-assembled molecular containers: 'tennis ball' $\mathbf{6}_2$, tetrahedron $Ga_4\mathbf{7}_6$, and capsules $\mathbf{8}_2$ and $\mathbf{9}_2$. Alkyl substituents are replaced with methyls in the graphical representations.

The thesis is organized as follows. In Chapter 2, the computational methodology and aspects of the calculations are briefly introduced. Then in Chapter 3 the dizinc complexes of **3** and **4** and the Fe(III)Mn(II) complex of

5 are studied as catalysts for hydrolysis and transesterification of phosphodiester substrates. Next, the study of capsule-accelerated cycloaddition is presented in Chapter 4. Finally, the study of water-soluble host **9** and its alkane-binding properties are discussed in Chapter 5, and conclusions and future outlook are presented in Chapter 6.

2 Concepts and computational methodology

In this chapter, the computational methodology used in the thesis is briefly introduced. First, the concepts of energy profiles, rate constants, and transition state theory will be presented. Then, basic density functional theory will be introduced, together with various corrections used to calculate internal energies of molecules. To go from the microscopic molecular quantities to the macroscopic world of chemistry, corrections must also be added for solvation and entropic effects. These will also be discussed before the computational protocol employed in Chapter 3 to 5 is summarized in the last section of this chapter.

2.1 Energy profiles, rate constants and free energies

A chemical reaction can be visualized as a series of geometrical changes resulting in the transformation of reactants into products. Since every molecular geometry is associated with an energy, proceeding from reactants to products can be likened to moving along a trajectory on a (multidimensional) energy landscape, on which every change in molecular geometry is related to a change on the energy coordinate.

In Figure 3, a slice of a hypothetical three-dimensional energy surface is shown for the case of hydrolysis of a phosphate substrate, named **S**. In this reaction, the energy first increases as the $H_2O\cdots P$ bond is formed, and then decreases somewhat as a pentacoordinate intermediate **I** is generated. It then increases again as the P···O bond to the leaving group dissociates, but after some distance along the reaction coordinate the energy decreases and the products **P** and ROH are formed.

In Figure 3, the path from **S** to **P** with the lowest increase in energy along the way is indicated in blue. On this two-dimensional *energy profile*, the stationary points are indicated with horizontal lines. Several concepts can be deduced from this representation. Reactants, intermediates and products are minima on the energy profile, while the maxima are called *transition states* (TSs). The minima thus possess Hessians of the energy with respect to the molecular coordinates which have only positive eigenvalues, while the TSs, which are saddle points on the multidimensional energy surface, have one negative eigenvalue.

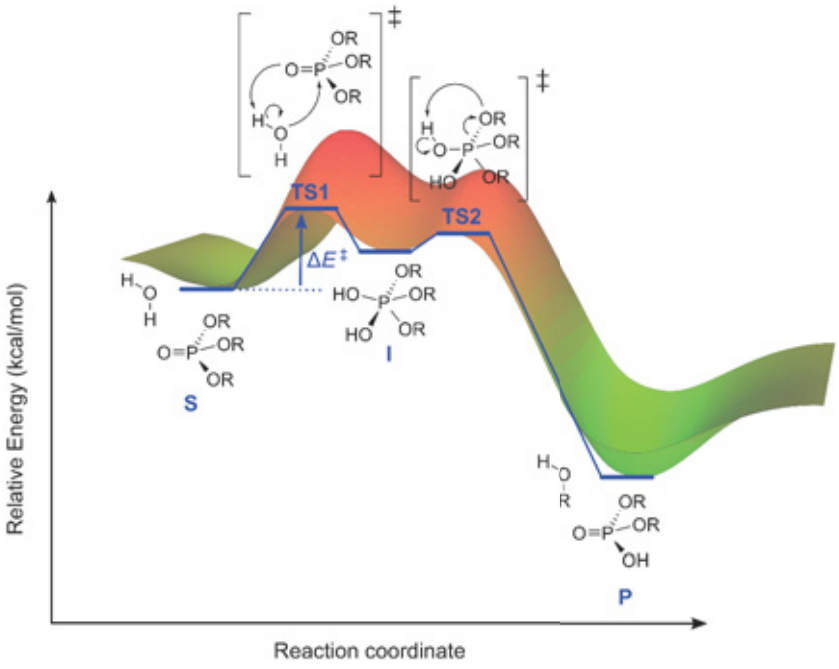


Figure 3. Hypothetical energy profile of a phosphoester hydrolysis reaction, shown as a two-dimensional cross-section of a three-dimensional molecular energy surface.

The highest *energy barrier* in the left-to-right direction is associated with going from the reactant to **TS1**. The barrier is denoted as $\Delta E^{\ddagger}$ and is the energy difference between reactant and TS. (Likewise, for the reverse reaction, going from right to left in Figure 3, the highest barrier would be associated with going from the product to **TS1**.)

In a bimolecular reaction, such as the one in Figure 3, the rate v of the reaction is proportional to the intrinsic rate constant k of the reaction and the concentrations of the reacting species, Equation (i). According to transition state theory (TST), k is related to the highest free energy barrier $\Delta G^{\ddagger}$ that must be surmounted for the reaction to go. This is therefore referred to as the rate-determining barrier.

$$v = k[\mathbf{S}]\left[\mathrm{H_2O}\right] \tag{i}$$

$$k = \frac{k_{\mathrm{B}}T}{h} e^{-\frac{\Delta G^{\ddagger}}{RT}} \tag{ii}$$

The quantity $\Delta G^{\ddagger}$ is called the Gibbs free energy barrier, and is related to the energy barrier $\Delta E^{\ddagger}$ according to Equation (iii).

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger} = \Delta E^{\ddagger} + P\Delta V^{\ddagger} - T\Delta S^{\ddagger} \quad \text{(iii)}$$

The quantity $\Delta H^{\ddagger} = \Delta E^{\ddagger} + P\Delta V^{\ddagger}$ is called the enthalpy. The volume change is usually negligible for reactions taking place in solution, and the energy barrier $\Delta E^{\ddagger}$ is a good estimate of $\Delta H^{\ddagger}$. $T\Delta S^{\ddagger}$ is the temperature times the change in entropy between reactant and transition state. This term is related to the number of degrees of freedom (DOFs) in the system, i.e. the number of different states (spatial and electronic) that are occupied under the reaction conditions.

Many observable quantities can be converted into free energy differences that are not necessarily connected to the transition state energy. For example, the equilibrium constant K of a reaction is related to the free energy difference between reactants and products. For the model reaction in Figure 3, K is defined in Equation (iv). From this constant, the free energy difference between products and reactants can be determined according to Equation (v).

$$K = \frac{[\mathbf{P}][\mathrm{ROH}]}{[\mathbf{S}]\left[\mathrm{H_2O}\right]} \quad \text{(iv)}$$

$$\Delta G = -RT\ln(K) \quad \text{(v)}$$

Molecular energies can be calculated with various computational methods. With suitable models for the interaction with solvent and for the molecular degrees of freedom, corrections can be added to these internal energies to yield free energies. Calculated free energy differences can hence be converted to rate or equilibrium constants, which can be compared to experimental observations.

2.2 Density functional theory

In the present thesis, density functional theory (DFT) has been used to calculate molecular energies. It is chosen because it provides a good balance between accuracy and computational cost for the systems studied here, being up to 300 atoms in size. DFT is based on the Hohenberg-Kohn theorems, the first of which states that the energy of a system depends in a unique fashion on its electron density.[36] In DFT, the molecular energy is therefore expressed in terms of the electron density. The energy expression $E[\rho]$ is a *functional* of the electron density ρ and is commonly formulated as in Equation (vi).[37,38]

$$E[\rho] = T_{\mathrm{S}}[\rho] + J[\rho] + V_{\mathrm{NN}} + E_{\mathrm{Ne}}[\rho] + E_{\mathrm{xc}}[\rho] \quad \text{(vi)}$$

In the above expression, the first four terms on the right-hand side are known. $T_S[\rho]$ is the kinetic energy for a system of non-interacting electrons. $J[\rho]$ and V_{NN} are repulsive terms; they represent the Coulomb interaction between the electrons and between the nuclei, respectively. $E_{Ne}[\rho]$ is the (attractive) nucleus-electron interactions. Finally, the $E_{xc}[\rho]$ term is the difference between the exact energy and the energy calculated by summing up the first four terms. This is called the exchange-correlation functional, since these are the expressions that are missing in the Coulomb description of the electrons: the exchange energy is the energy change that stems from exchange of two electrons, and the correlation energy is the part of the total energy which cannot be accounted for by treating the electrons in a mean-field manner.

Unfortunately, the exact expression for $E_{xc}[\rho]$ is not known. Over the years, a large number of suggestions have been put forward for this expression. One of the most commonly used ones is B3LYP,[39] which uses a linear combination of the exchange energy derived from the local density approximation (LDA) used to model the uniform electron gas (UEG),[40] the exchange energy calculated by Hartree-Fock (HF) theory,[41] and Becke's exchange energy functional called B88.[42] For the correlation energy, a combination of the UEG correlation functional derived by Vosko, Wilks and Nusair (VWN)[43] and the correlation energy developed by Lee, Yang and Parr (LYP)[44] is used. Three parameters are fitted in order to reproduce a set of experimental energies.

$$E_{xc}^{B3LYP} = (1-a_0)E_x^{LDA} + a_0 E_x^{HF} + a_x E_x^{B88} + a_c E_c^{LYP} + (1-a_c)E_c^{VWN} \quad \text{(vii)}$$

The fitted values of a_0, a_x and a_c are 0.20, 0.72 and 0.81, respectively.[39] In a slightly modified functional, B3LYP*, the a_0 value is changed to 0.15. This functional is commonly used for transition metal systems, since it has been found to better reproduce the energy differences between spin states in such complexes.[45] Importantly, corrections have also been developed for B3LYP and other functionals to better describe nonuniformities in the long-range dispersion interactions, which are inherently missing in the LDA and VWN descriptions that were developed for the UEG. The most popular implementation of is DFT-D3,[46] which corrects the exchange-correlation functional with a new expression.

$$E_{xc}^{B3LYP\text{-}D3} = E_{xc}^{B3LYP} + \sum_{AB}\sum_{n=6,8} s_n \frac{C_n^{AB}}{r_n^{AB}} f_{d,n}(r_{AB}) + \sum_{ABC} f_{d,(3)}\left(\bar{r}_{ABC}\right) E^{ABC} \quad \text{(viii)}$$

In this extension of B3LYP, the atom-pair coefficients C_n^{AB} for atoms of elements A and B are computed by time-dependent DFT, s_6 is set to unity and s_8 is parameterized to fit experimental data.[46] The f_d functions are damping functions that are used to avoid singularities for $r = 0$. The choice of Becke-Johnson (BJ) damping is recommended,[47,48] which includes two more parameters to be fitted for each pair of atom elements. Finally, the three-body energy expression E^{ABC} also depends on the atom-pairwise coefficients of the elements involved. The DFT-D3 and DFT-D3(BJ) corrections have been parameterized for a number of popular density functionals.[46,48] The DFT-D3(BJ)-corrected B3LYP functional, called B3LYP-D3(BJ), has been shown to perform well for both binding energies and energy barriers in benchmark tests.[48-50]

Another popular set of density functionals is called the Minnesota family.[51,52] In M06, the exchange energy is a combination of HF exchange and E_x^{DFT}, a density-weighted combination of LDA exchange and the exchange functional of Perdew, Burke and Ernzerhof.[53,54] For the correlation energy a modified version of the UEG expression is used, here named E_c^{DFT}. In contrast to B3LYP, the M06 type functionals already in their definitions contain terms that are supposed to account for dispersion interactions.

$$E_{xc}^{M06} = b_0 E_x^{HF} + (1 - b_0) E_x^{DFT} + E_c^{DFT} \qquad \text{(ix)}$$

Despite its simple appearance, a large number of parameters are used in the expressions for E_x^{DFT} and E_c^{DFT}. In total 35 parameters are fitted in order to reproduce experimental data. Three functionals are parameterized by setting b_0 to 0, 0.27 and 0.54, respectively, and fitting the other 34 parameters; they are named M06-L,[51] M06[52] and M06-2X.[52] They thus differ in the influence of the HF exchange. M06-L is suggested for systems with transition metals, M06-2X is proposed for non-covalent interactions, and M06 is suggested to be a balanced representation for all kinds of systems.[51,52]

The performances of the B3LYP and M06 families of functionals have been extensively tested in many benchmarks.[49,50,52,55-61] Based on these tests, one can make informed selections of which functional to use in the particular studies.

Apart from the choice of DFT functional, one must also choose the basis set. Linear combinations of basis functions from this set are used to represent the molecular orbitals. The larger the basis set, i.e. the more complex shape of the corresponding basis functions, the better can the orbitals be modeled. However, no basis set is complete, and there will be inadequacies. One such problem is basis set superposition error (BSSE), where basis functions on one molecular fragment delocalize such that also another fragment is being described by it. The result is an artificial lowering of the energy of the molecular complex. One means of correcting for this is the counterpoise method.[62,63] In this approach, for a two-fragment supermolecule AB with

fragments A and B, the molecular energy is calculated for each fragment in the presence of the basis functions from itself and from the other fragments. The overlapping parts are then subtracted to give a molecular energy that is corrected for BSSE. In Equation (x), the subscript after each energy denotes from which species the basis functions are used.

$$\Delta E_{\text{counterpoise}} = E(A)_{AB} - E(A)_A + E(B)_{AB} - E(B)_B \tag{x}$$

DFT methods are single-determinantal, which means that the electrons are treated in a mean-field manner and only one electronic configuration is considered when the energy of a compound is calculated. One shortcoming of such a treatment is that the representations of specific spin states, in which the electrons can pair up in different ways, become erroneous. For the case of antiferromagnetic coupling ($S = 0$) the $\langle S^2 \rangle$ value calculated with standard DFT deviates from the value that is expected from a pure spin state, $S(S + 1) = 0$. The calculated spin state is therefore said to be contaminated.

One way to correct for this is to use the broken symmetry approach developed by Noodleman.[64] If the spin contaminated system is denoted S', the difference in energy between this state and the high-spin state S_{max} is

$$E(S_{\text{max}}) - E(S') = (S_{\text{max}})^2 \frac{J}{2} \tag{xi}$$

The coupling constant J is calculated from this expression. It is then used to calculate the energy correction for the true antiferromagnetic state according to

$$E(0) = E(S_{\text{max}}) + S_{\text{max}}(S_{\text{max}} + 1)\frac{J}{2} \tag{xii}$$

2.3 Solvation models

All systems studied in the present thesis were experimentally investigated in the solution phase, where the molecule is surrounded by a huge number of solvent molecules. The solvent and the solute can interact in a variety of ways. For example, in the case of water solvent, it can form hydrogen bonds to the polar parts of the solute, coordinate to metal centers in organometallic compounds, form hydrophobic interactions with the nonpolar parts of the solute, and even react with the solute to form new species (like in Reaction (1)). Nonpolar solvents can interact with the solute via for example dispersive interactions.

Taking these possibilities into account in the computational modeling explicitly requires an extensive amount of work. Just positioning a few

solvent molecules around a solute in an optimal way can be a tedious task. Fortunately, there are simpler approaches to this problem that in many cases work satisfactorily.

In polarizable continuum models (PCMs),[65,66] the solvent is implicitly modeled as a homogeneous dielectric continuum that is characterized by a dielectric constant ε. The charges on the molecular surface of the solute (Figure 4) are calculated and are then allowed to interact with the dielectric medium. The solute and the dielectric medium are iteratively allowed to polarize each other in a self-consistent manner until convergence is reached. The interaction free energy between the solute and the solvent is estimated from this process.

The PCM methods are parameterized against experimental solvation free energies. Two of the most widely used implicit models are C-PCM,[67,68] which is based on a scaled-conductor treatment of the dielectric (called COSMO),[69,70] and SMD,[71] which relies on a similar approximation and also includes a term that represents the non-electrostatic interactions between solvent and solute. They both predict hydration free energies of neutral species with accuracies on the order of 0.6-1.6 kcal/mol.[72,73]

A slightly different model is COSMO-RS,[70,74] where RS stands for 'real solvents'. In this approach, the surface charges are calculated for both the solvent and the solute, and for both species a histogram is constructed, representing the number of surface segments of a certain charge. These so-called σ-profiles (Figure 4) are then used to give the solvation free energy of the solute. The agreement with experiments is of the same order as those obtained with C-PCM and SMD. For example, for the hydration free energies of a set of 284 neutral species the mean absolute deviation from experiments was 0.6 kcal/mol.[75]

The implicit models usually perform well when the solvent interacts with the solute through weak interactions. However, sometimes strong hydrogen bonds are made or the solvent binds to a coordination site on the solute. The implicit solvent models discussed above do typically not account in an ideal way for such interactions. One way to treat this situation is to include a few explicit solvent molecules in the model of the solute, and then describe the rest of the solvent with an implicit model. Several computational protocols have been derived for how to compare models with a different number of coordinated solvent molecules.[76-78] We deal more with such a situation in Chapter 5, which is concerned with the study of a water-soluble capsule.

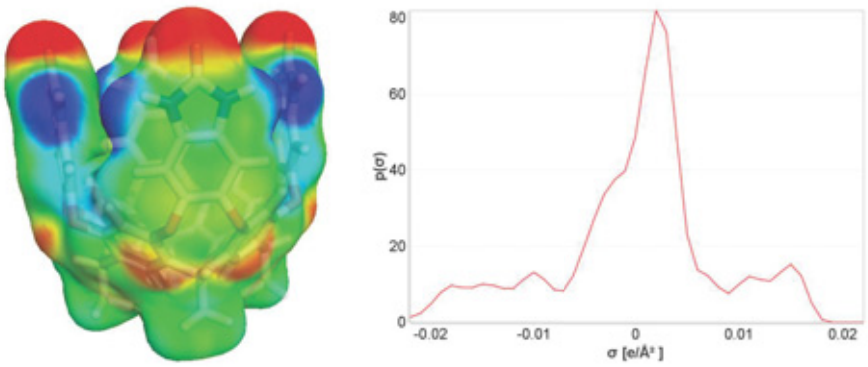


Figure 4. Molecular surface of ***9****, calculated with COSMO-RS, and the corresponding σ-profile.*

Apart from the electrostatic solvent-solute interactions, another contribution to the solvation free energy comes from the difference in standard state in the gas phase and in the solution. The solvation free energy is usually defined as the free energy gain (or cost) when a molecule is taken from an ideal gas phase of that compound to solution, yielding an infinitely diluted compound.[79,80] In the ideal gas, each mol of solute occupies a volume of 24.5 liters. In solution, the standard state is 1 mol/l of solute. For a solute X, the free energy difference that comes from this change in concentration is

$$-RT \ln\left(\frac{[X]_{\text{solv}}}{[X]_{\text{gas}}}\right) = RT \ln(1 \text{ mol/l} \times 24.5 \text{ l/mol}) = +1.9 \text{ kcal/mol} \qquad \text{(xiii)}$$

For the solvent itself, the standard state is the concentration of the pure solvent. For water, the state changes from 24.5 l/mol to 55.4 mol/l, which is associated with a change in free energy of +4.3 kcal/mol.

2.4 Entropic corrections

As stated above, the molecular entropies depend on the number of accessible states. In the current thesis, the rigid-rotor harmonic-oscillator (RRHO)[81] approach is used, in which the molecular partition function is divided into four types of DOFs: translational, electronic, rotational and vibrational. The thermal corrections corresponding to each DOF, i.e. the correction to the energies to give free energies, are then calculated according to equations derived for these DOFs.[81] Three of the DOFs are translational and for a nonlinear molecule there are three rotational DOFs as well. The remaining $3N - 6$ DOFs are vibrational.

For large molecules, the vibrational DOFs give the largest contributions to the free energy corrections. The vibration modes have corresponding vibrational frequencies, which are related to the square root of the second derivative of the forces on each atom with respect to the atomic positions. Hence from this analysis, it can be determined whether the molecule is a minimum (only real vibrational frequencies) or a saddle point (one imaginary vibrational frequency) on the free energy landscape of molecular conformation space.

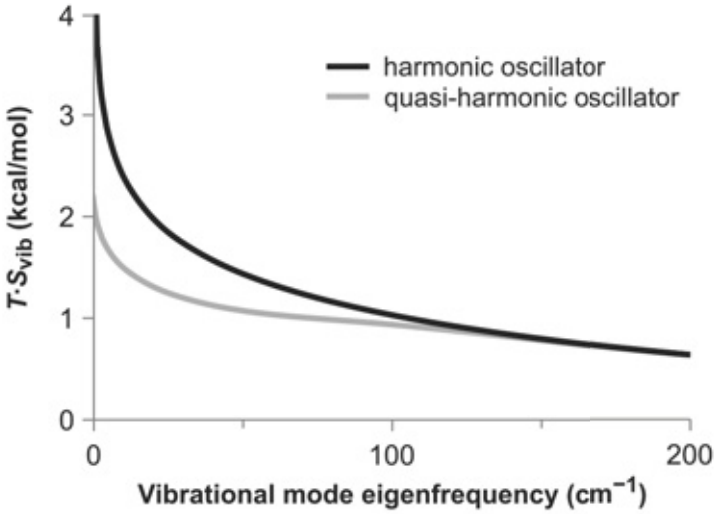


Figure 5. The free energy contribution $T \cdot S_{vib}$ at room temperature calculated with the harmonic-oscillator (black) or quasi-harmonic-oscillator (light gray) approaches. A cutoff value of 100 cm^{-1} and an average moment of inertia of 10^{8} amu cm^{2} is used for the latter calculations.

One shortcoming of the RRHO protocol is that the entropic contributions of vibrational DOFs with small frequencies approach infinity as the corresponding eigenfrequencies approach zero, and are thus overestimated (Figure 5). One way to correct for this is to use the so-called quasi-RRHO approach, in which vibrational modes with low-lying frequencies are treated as rigid rotors with the same frequency.[82] A weighting function $w(\bar{\nu})$ is used to enforce a smooth transition from vibrational to rotational treatment of the low-frequency modes. Typically, a value of $\bar{\nu}_0 = 100\ \mathrm{cm}^{-1}$ is used.[82]

$$S(\bar{\nu}) = w(\bar{\nu}) S_V(\bar{\nu}) + \left(1 - w(\bar{\nu})\right) S_R(\bar{\nu}), \qquad w(\bar{\nu}) = \frac{1}{1 + \left(\frac{\bar{\nu}_0}{\bar{\nu}}\right)^4} \qquad \text{(xiv)}$$

2.5 Computational details

The calculations in the present thesis were conducted with the Gaussian 09 software suite.[83] Geometries were optimized with B3LYP (Section 3.1 and 3.2), B3LYP* (Section 3.3), or B3LYP-D3(BJ) (Chapter 4 and 5). In the optimizations, the 6-31G(d,p) basis set was used for all atoms except phosphorus and oxygens bound to phosphorus, for which 6-311+G(d) was used, and Zn, Fe and Mn, for which the LANL2DZ pseudopotential was used.[84] For all compounds, thorough conformational searches were performed in order to identify the most stable geometries. On the basis of these geometries, single-point calculations were done with the same functional as the geometry optimizations, using the 6-311+G(2d,2p) basis set for main-group atoms and LANL2DZ for metals. In Chapter 3, the two-body DFT-D3 corrections for B3LYP (the second term in Equation (viii)) were added separately to the energies. For the supramolecular complexes studied in Chapter 4 and 5, the three-body DFT-D3 term was also added to all energies, since it is not included in the Gaussian 09 implementation of the B3LYP-D3(BJ) functional. In Section 3.3, the antiferromagnetic energies are corrected using the broken symmetry approach. In Chapter 5, some of the obtained energies were also corrected according to the counterpoise scheme, Equation (x).

Solvation effects were modeled with the C-PCM model (Chapter 3), at the same level of theory as the geometry optimizations, or the COSMO-RS model (Chapters 4 and 5), calculated at $T = 298.15$ K and using the BP86/TZVP level of theory[42,85,86] with the COSMO*therm* software.[87,88] Concentration corrections, Equation (xiii), were added for all compounds. At the same level of theory as the geometry optimizations, frequency calculations were carried out to confirm the identity of the structure (minimum or transition state) and to calculate the thermal free energy corrections at 298.15 K according to the quasi-RRHO protocol. In Chapter 4 and 5, the quasi-RRHO calculations were done using rotational symmetry number $\sigma_R = 1$, and the molecular point groups were determined using GaussView.[89] The corresponding corrections to the rigid-rotor entropies, $RT \ln(\sigma_R)$, were then added to the quasi-RRHO corrections.

3 Phosphodiester hydrolysis and transesterification with phosphoesterase complexes (Papers I, II and III)

Phosphotriesterase (PTE), or aryldialkylphosphatase, is an interesting enzyme for which no natural substrate has been identified.[90-92] It catalyzes the hydrolysis of organophosphates, that are rest products of pesticides and nerve gases.[93] Both experimental and theoretical studies have been conducted to gain knowledge on how the enzyme, and its likes, function.[1-4,7,9,94-99] The active site of PTE is shown in Figure 6.[96]

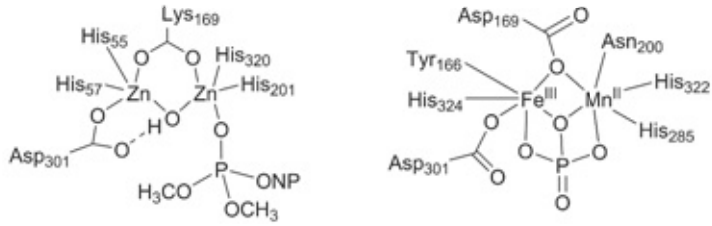
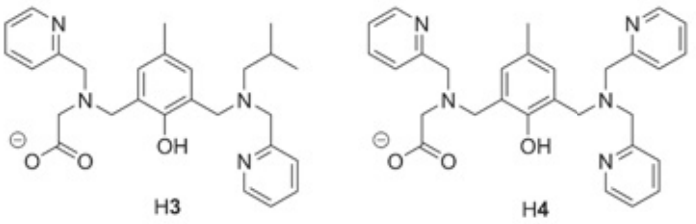
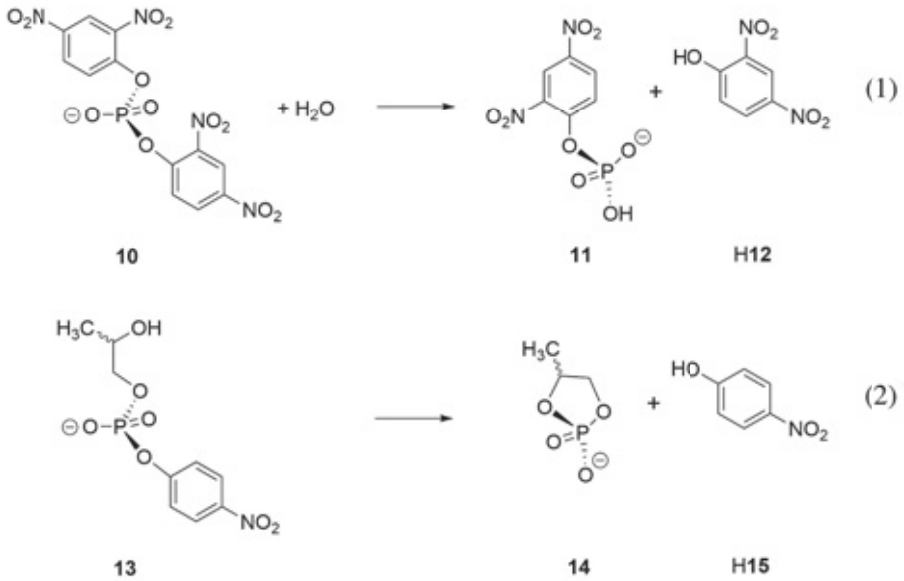


Figure 6. Schematic representations of the active sites of phosphotriesterase (left) and purple acid phosphatase from sweet potato (right), both with phosphate substrates bound.

In this chapter, the hydrolytic effect of dinuclear complexes of two ligands that have been designed to mimic the active site of PTE and similar enzymes is investigated by means of computations. Both ligands **3** and **4** provide binding pockets for two metal ions and have been studied experimentally with two Zn(II) bound, similarly to the active site of PTE. Both complexes are unsymmetric, as they provide different coordination environments for the two metal ions. One of the metal ions, here called Zn1, is coordinated by a phenyl oxide, a carboxylate, a tertiary amine and a pyridyl nitrogen, while the other one, Zn2, is coordinated by the same phenyl oxide, a tertiary amine and one (**3**) or two (**4**) pyridyl nitrogens. This unsymmetric arrangement resembles the active sites of dinuclear phosphoesterase enzymes, which typically provide different coordination environments for the two metal ions in the active site.[7,9]



The dizinc complexes were experimentally examined for activity towards Reaction (1),[14,100] in which bis(2,4)-dinitrophenyl phosphate (**10**) is hydrolyzed to form arylphosphate **11** and phenol **H12**. The $(Zn(II)_2\mathbf{3})^{2+}$ complex has also been considered towards Reaction (2),[100] a transesterification reaction in which a hydroxide unit of 2-hydroxypropyl *p*-nitrophenyl phosphate (**13**) binds to the phosphorus center to yield cyclic phosphate **14** and phenol **H15**. Both reactions are shown in Scheme 1.



*Scheme 1. Reactions catalyzed by dimetal complexes of **3**, **4** and **5**.*

Apart from the coordination environment provided by the ligands, the Zn(II) ions can also coordinate water molecules from the 50:50 water:acetonitrile solution that the experiments were performed in. Upon coordination, the pK_a of the water molecules decreases significantly such that the coordinated water molecules can be deprotonated.[4,100] The reactions were experimentally observed to show a dependency on pH. The complexes were characterized in the absence of substrate at the pH for which the highest initial rates were measured. Under these conditions, the $(Zn(II)_2\mathbf{3})^{2+}$ complex was found to have two hydroxide ions bound,[98] while $(Zn(II)_2\mathbf{4})^{2+}$, that has

one more coordination site occupied by the ligand, was observed to bind one hydroxide ion.[14]

The aim of the calculations is to characterize the active forms of the catalysts in Reaction (1) and (2) and to elucidate the reaction mechanisms. For $(Zn(II)_2\mathbf{3})^{2+}$, the possibilities of one or two hydroxide ions being coordinated to the catalyst must be considered, and the coordination of the substrate to the catalyst must be determined. An X-ray structure was previously obtained for the tetranuclear complex $[(Zn(II)_2\mathbf{3}\cdot OAc)_2]^{2+}$, i.e. a dimer of the $(Zn(II)_2\mathbf{3})^{2+}$ complex with one acetate bridging the two zinc ions in each monomer.[100] This structure is used as a starting point for the calculations, in which one of the monomer structures is extracted from the X-ray geometry and the acetate is manually substituted for one or two hydroxide ions. In the case of $(Zn(II)_2\mathbf{4})^{2+}$, the isopropyl moiety is also replaced with a pyridyl. The substrate is coordinated in a number of ways to identify the most stable catalyst-substrate complex, and from there different mechanistic scenarios were assessed.

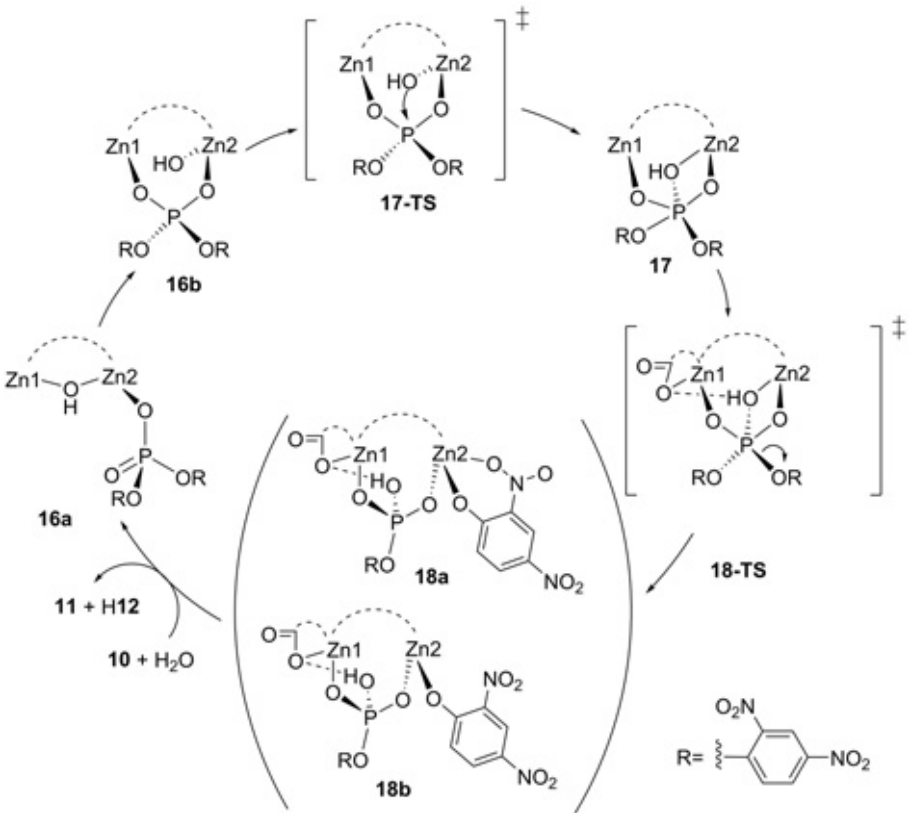
The results of the calculations will now be presented, first for Reaction (1) and then for Reaction (2).

3.1 Hydrolysis reaction with dizinc complexes

First, the reaction is studied with one hydroxide ion bound to the catalyst in a position bridging the two zinc ions. From here the reaction is found to take place *via* the same steps regardless of whether the ligand is **3** or **4**. The obtained reaction mechanism, catalyzed with $(Zn(II)_2\mathbf{3}\cdot OH)^+$ and $(Zn(II)_2\mathbf{4}\cdot OH)^+$, is given in Scheme 2.

Several binding modes of substrate **6** to the catalysts are evaluated, and in the most stable one (**16a**) the substrate coordinates to Zn2 with one of the phosphate oxygens. However, for the reaction to proceed, it is found that the other phosphate oxygen must bind to Zn1. The hydroxide now loses its bridging role and instead takes a terminal position on Zn2 to form the **16b** complex. This conformation is calculated to be 7.2 kcal/mol less stable than **16a** with ligand **3** and 9.0 kcal/mol with ligand **4**.

From this position, the hydroxide performs a nucleophilic attack on the phosphorus center of **10**. In the transition state **17-TS**, the distance between the hydroxide and the phosphorus is 2.13 Å with **3** and 2.08 Å with **4** (Figure 7). Compared to **16a**, the TSs are calculated to be 10.8 and 11.7 kcal/mol in energy with **3** and **4**, respectively. The result of the nucleophilic attack is a pentacoordinate phosphate species bound to the catalyst, complex **17**, with energies similar to the TSs. With **3**, it is calculated to be +9.4 kcal/mol, and with **4** +11.0 kcal/mol.



*Scheme 2. Mechanism for hydrolysis of **10** with dizinc complexes of **3** and **4**.*

From here, the leaving group can dissociate from the phosphorus, while the phosphate-bound hydroxide forms a hydrogen bond to the carboxylate oxygen coordinated to Zn1. Without the solvation corrections, this is calculated to occur with barriers of 6.1 and 0.2 kcal/mol with **3** and **4** respectively, but when the solvation effects were added, this second TS is calculated to be somewhat lower in free energy than **17** (see Figure 8). Thus, the calculations suggest that the dissociation of the leaving group dissociation is essentially occurring in a concerted fashion with the nucleophilic attack.

Complex **18a** is now formed with **3** and **18b** is formed with **4**. In these complexes, both hydrolysis products coordinate to the catalyst. Arylphosphate **11** bridges the zinc ions via two oxygens, while the hydroxyl group of the phosphate has lost its coordination to the zinc ions and only hydrogen-binds to the carboxylate moiety. With the **3** ligand, the phenoxide product **12** coordinates to Zn2 with the phenolic oxygen and one of the nitro groups, such that Zn2 is now hexacoordinate. The formation of these catalyst-product complexes is calculated to be quite exergonic; compared to

16a, this species is found to be −20.2 kcal/mol with **3** and −23.3 kcal/mol with **4**.

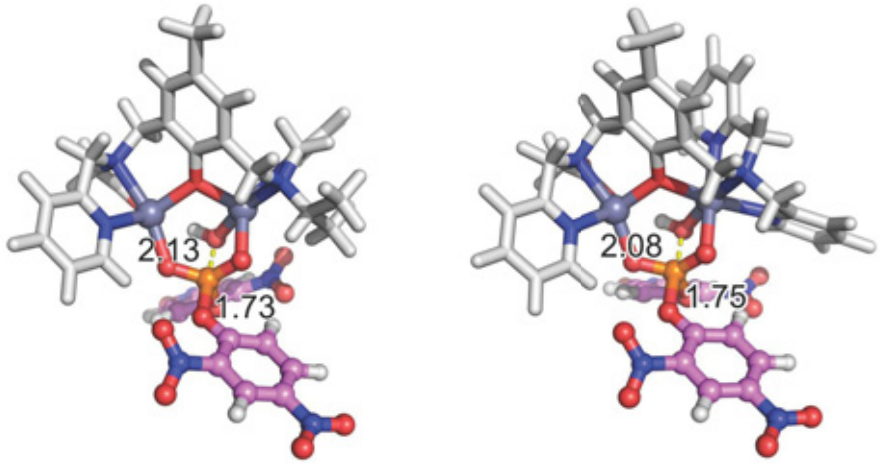


Figure 7. Optimized geometries of ***17-TS*** *with dizinc complexes of* ***3*** *(left) and* ***4*** *(right). Important distances (Å) are indicated.*

To regenerate the catalyst-substrate complex **16a** for another hydrolysis reaction to take place, a new substrate **10** must bind and replace the **11** and **12** products of the reaction, and one water molecule from solution must bind and become deprotonated. The net reaction is identical to Reaction (1). The free energy of this reaction is calculated to be −10.0 kcal/mol, and the catalyst does not change this value. Thus, in going from **16a** in catalytic cycle 1 to the same step in cycle 2, 10.0 kcal/mol is calculated to be released to the surroundings. This means that the regeneration step, i.e. going from **18a** in cycle 1 to **16a** in cycle 2, is calculated to be endergonic by $-10.0-(-20.2)=+10.2$ kcal/mol with **3**. Analogously, the energy from **18b** to **16a** is calculated to be $-10.0-(-23.3)=+13.3$ kcal/mol with **4**. The full free energy profiles of the reactions thus obtained are shown in Figure 8.

Complexes **18a/18b** are thus calculated to be the lowest energy species along the reaction coordinate, constituting the resting state of the reaction. This result is supported by recent experimental findings for a dinuclear Fe(III)Fe(II) complex, catalyzing the hydrolysis of **10**, for which the catalyst-product complex proved to be stable enough that it could be characterized by X-ray crystallography.[101] In the same study, the authors used Mössbauer spectroscopy to identify the nucleophile in the reaction as a hydroxide, coordinated to one of the metal ions, that becomes terminal prior to the nucleophilic attack on the phosphorus center.[101] This is in agreement with the mechanism presented above with dizinc complexes of **3** and **4**.

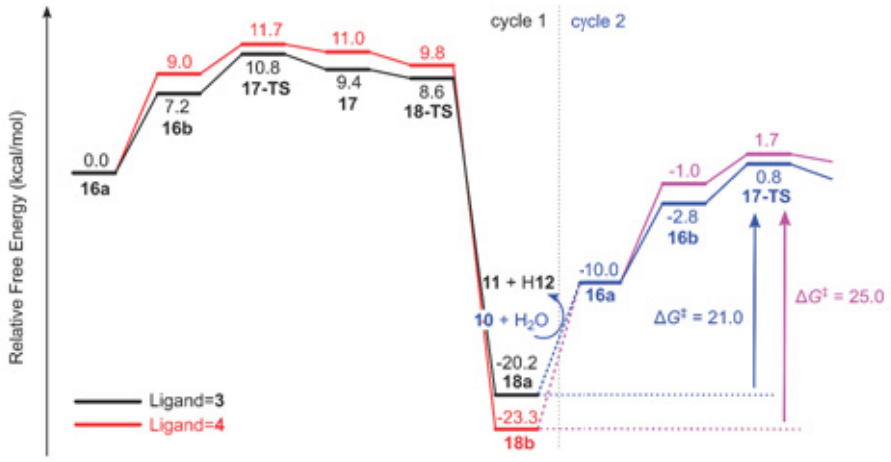


*Figure 8. Free energy profile for hydrolysis of **10** with dizinc complexes of **3** (black/blue curve) and **4** (red/magenta curve).*

From **16a**, a second hydrolysis reaction can take place with the same barrier as in cycle 1. The overall barrier from **18a** in one catalytic cycle to **17-TS** in the next is calculated to be 10.2 + 10.8 = 21.0 kcal/mol with **3**, and the barrier from **18b** to **17-TS** is calculated to 13.3 + 11.7 = 25.0 kcal/mol with **4**. This is in excellent agreement with the experimentally measured rates,[14,100] which can be converted to free energy barriers of approximately 22 and 26 kcal/mol, respectively.

In an alternative reaction mechanism, the catalyst-bound hydroxide ion can act as a Brønsted base instead of as a nucleophile, thus activating a water molecule from the solution which in turn performs the nucleophilic attack on phosphorus. The corresponding transition states are calculated to be 17.2 and 13.9 kcal/mol higher in energy than **16a** with **3** and **4**, respectively. Hence, with **3** this general-base mechanism is calculated to have a significantly higher barrier (+6.4 kcal/mol relative to **17-TS**), while it is calculated to be of similar magnitude (+2.2 kcal/mol) as **17-TS** with **4**. With the latter ligand, such a general-base mechanism can thus not be ruled out on the basis of the calculated free energies.

Finally, the barrier calculated with **3** can be compared to the same barrier with two hydroxides present in the catalyst-substrate complex. The overall free energy barrier is then found to increase to 28.8 kcal/mol, i.e. 7.8 kcal/mol higher than what was calculated above with the monohydroxide catalyst. Such a mechanism can thus be discarded.

To conclude this section, on the basis of calculations the dizinc complexes of **3** and **4** are suggested to catalyze the hydrolysis of **10** via an essentially S_N2-like reaction with a catalyst-bound hydroxide as the nucleophilic species. Importantly, for the identification of the overall free energy barrier it is found to be essential to consider two consecutive catalytic

cycles in the calculations. This insight is now acknowledged by other researchers who later studied similar reactions computationally.[102]

3.2 Transesterification reaction with dizinc complex

Next, Reaction (2), i.e. transesterification of substrate **13**, is studied with the $(Zn(II)_2\mathbf{3})^{2+}$ complex. Like in the previous section, the first step in the computational study is to identify the most stable catalyst-substrate complex. This analysis proves to be harder than with the substrate **10** since **13** harbors an extra potentially coordinating group in its isopropanol moiety. The binding mode calculated to be the most stable with one hydroxide present in the catalyst turns out to be one that is very similar to what was previously suggested for a symmetric dizinc complex.[103,104] In this binding mode, named **19a** (Scheme 3 and Figure 10a), the phosphate oxygens coordinate to one zinc ion each. Furthermore, the isopropanol moiety binds to Zn1, forming a hydrogen bond to a hydroxide ion that is terminally bound to Zn2.

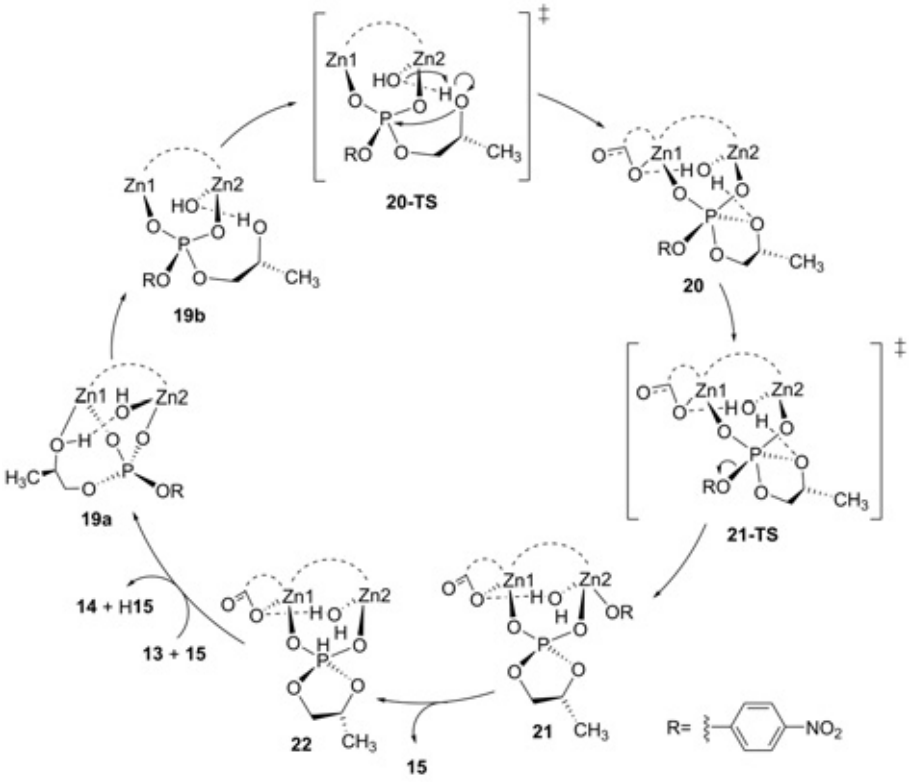
In order to form the transesterification products of Reaction (2), a bond must be formed between the isopropanol oxygen and the phosphorus center, and the alcohol group must be deprotonated. These events can occur in different orders: The deprotonation can occur before, concertedly with, or after the nucleophilic attack. These events will be taken into consideration in the present section.

By consideration of the **19a** binding mode, it is natural to assume that the Zn2-bound hydroxide is the most likely species to act as Brønsted base in the mechanism. However, although many attempts are made to identify transition states for the deprotonation of the substrate or for the nucleophilic attack of the Zn1-bound alcohol group, no such TSs are located.

Instead, it is found that in order for the nucleophilic attack to occur, the isopropanol group must decoordinate from Zn1 while preserving the hydrogen bond to the Zn2-bound hydroxide. This catalyst-substrate binding mode is named **19b** and is calculated to be 7.9 kcal/mol higher in energy than **19a**. The reaction mechanism starting from here is shown in Scheme 3. From its position in **19b**, the isopropanol can attack the phosphorus center of **13**, and it is calculated that in the corresponding TS, named **20-TS**, this occurs in concert with deprotonation of the alcohol by the Zn2-coordinated hydroxide. The free energy of **20-TS** is calculated to be +16.4 kcal/mol with respect to **19a**.

A pentacoordinate phosphate species named **20** is now formed. This species is calculated to be +13.2 kcal/mol compared to **19a**. The bond between phosphorus and leaving group **15** is then cleaved via **21-TS**. In analogy with the catalyzed hydrolysis of **10** presented above, this TS is located at the level of theory used for the geometry optimizations, but when

all effects were added it turns out to be lower in energy than **20**. Thus, also this reaction is found to be of essentially concerted nature.



*Scheme 3. Mechanism for transesterification of **13** with dizinc complex of **3**.*

The catalyst-product complex **21** is then formed, and the most stable geometry of it is calculated to be −9.0 kcal/mol in energy compared to **19a**. When compared to complex **22**, in which product **15** has been released to solution, the latter complex is calculated to be marginally more stable, −0.2 kcal/mol. Thus the calculations suggest this to be the resting state of the reaction.

From here, the energetics of the catalyst regeneration process is estimated in the same way as was done in the study of Reaction (1) above, and it is calculated to be endergonic by 3.1 kcal/mol. Adding this value to the 16.4 kcal/mol calculated for accessing **20-TS** from **19a**, an overall barrier of 19.5 kcal/mol is obtained. This is in good agreement with the barrier obtained from the experimentally-measured rate constant, which is 23 kcal/mol. The free energy profile of the complete reaction is shown in Figure 9.

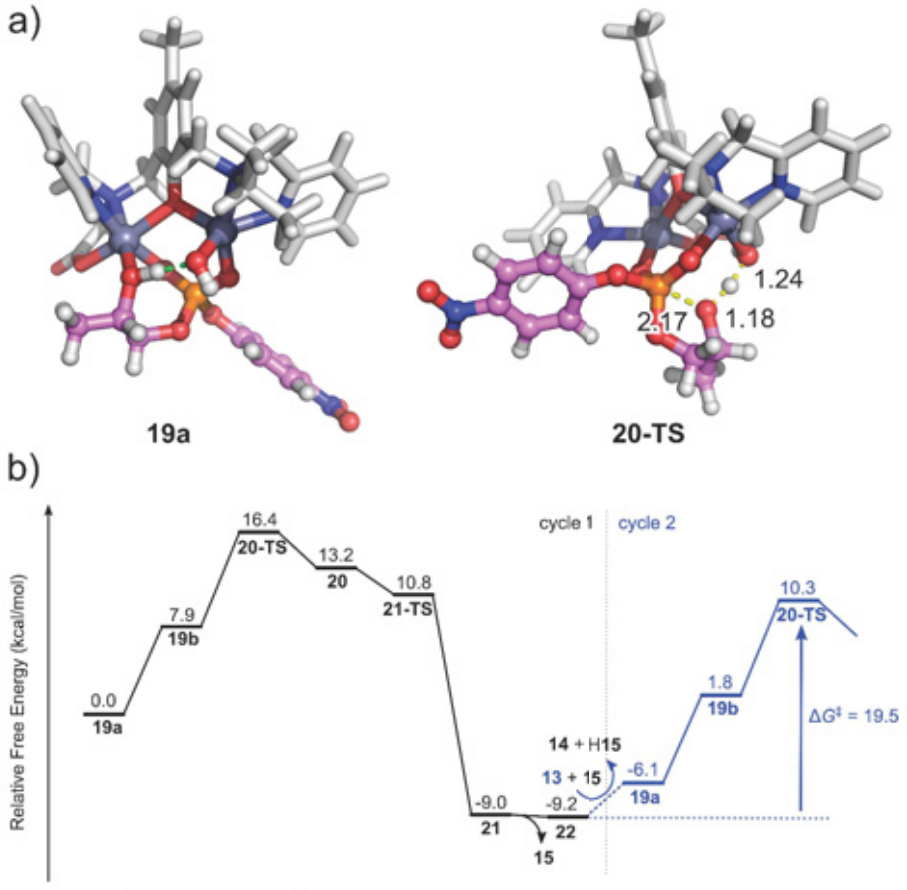


*Figure 9. (a) Optimized geometries of **19a** and **20-TS**. (b) Free energy profile for transesterification of **13** with dizinc complex of **3**.*

In the reaction mechanism presented above, the nucleophilic attack on the phosphorus center was calculated to occur concertedly with the deprotonation of the nucleophilic propanol moiety. The deprotonation can also be thought to occur prior to the nucleophilic attack, an event which was suggested on the basis of kinetic experiments, where a pre-equilibrium event was observed to occur before the rate-determining step.[100] While no such TS could be located in spite of many attempts, the computational study could start with a deprotonated substrate being bound to the catalyst. However, such a reaction is calculated to have an overall barrier of 29.3 kcal/mol, significantly higher than the barrier calculated for the concerted reaction mechanism above. Alternative mechanisms with a catalyst species in which two hydroxides are present are also considered, but these are calculated to give higher free energy barriers than above.

To summarize, for the $(Zn(II)_2\mathbf{3})^{2+}$-catalyzed transesterification of **13**, a general-base reaction mechanism is calculated to give the lowest overall free

energy barrier. No theoretical support is found for the existence of a pre-equilibrium before the rate-determining step. However, an unproductive **19a** binding mode is identified as the most stable catalyst-substrate complex. Since dissociation of this complex is required for the reaction to proceed, this might hypothetically be involved in the experimentally observed pre-equilibrium.

3.3 Hydrolysis reaction with iron-manganese complex

Purple acid phosphatase (PAP) is another phosphatase enzyme, which has its name from the purple color arising from a charge transfer from an Fe(III) ion to a coordinated tyrosine residue.[9,105] Apart from the Fe(III) ion, the active site also contains another metal ion which could be either of Fe(II), Zn(II), Mg(II) or Mn(II).[9,105] A schematic representation of the active site of PAP from sweet potato is shown in Figure 6 above.[106] This enzyme has an Fe(III) ion and an Mn(II) ion in its active site.[107]

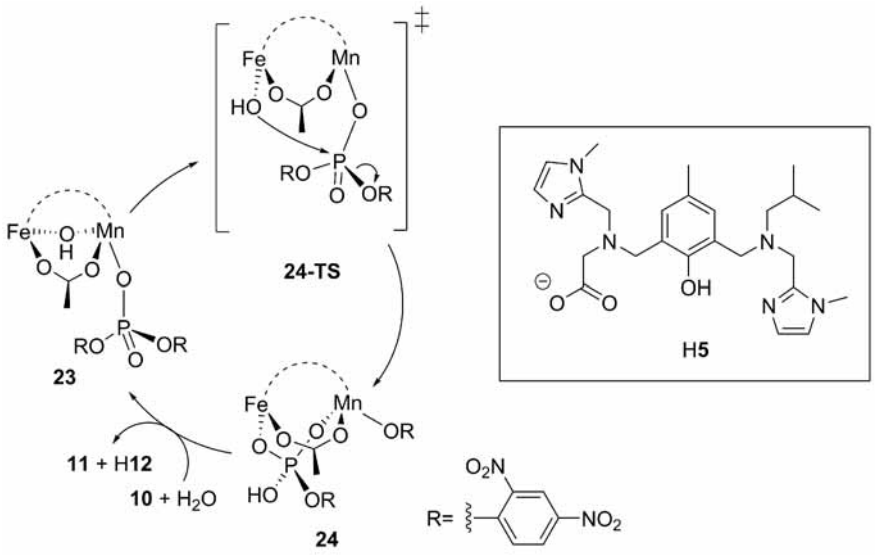
A biomimetic complex, based on ligand **5**, was synthesized[15] and was shown to bind an Fe(III)Mn(II) motif.[108] The ligand is identical to **3** except that the pyridyl moieties of **3** have been replaced with methyl imidazoles. It is thus also an unsymmetric ligand, in which the two metal ions bind in one binding pocket each; Fe(III) binds to the carboxylate site, providing four coordinations from the ligand, and Mn(II) binds to the other, three-coordinated, site.[108]

An X-ray structure was obtained for the compound.[108] In this structure, two acetates are bridging the two metal ions, and a chloride ion binds to Mn(II). Both metal ions are thus hexacoordinate. The complex was experimentally monitored for activity towards hydrolysis of **10**, and it was found that the reaction proceeds at similar rates as was measured for the dizinc complexes studied in the previous sections.[108]

Computational studies are conducted in order to characterize the active catalyst and to identify the nucleophile in the reaction. Several possibilities exist: the acetates and the chloride can be substituted for water (or hydroxides) and the substrate. First, the results will be presented with a complex in which the chloride ion has been replaced by the substrate, binding in a monodentate fashion to Mn(II), and one of the acetates has been replaced by a hydroxide, bridging the metal ions. The resulting complex is named **23** (Scheme 4).

This complex is characterized from an electronic point of view. While the ferromagnetic ($S = 11/2$) and antiferromagnetic ($S = 0$) cases are calculated to be similar in energy, other possible spin configurations are found to be significantly higher. The antiferromagnetic singlet state is calculated to be slightly more stable than the ferromagnetic one, so this species was used in

the calculations. This is also in agreement with experiments, as the coupling constant measured for the system (-15 cm^{-1})[108] is small and negative.



*Scheme 4. Mechanism for hydrolysis of **10** with FeMn complex of **5**.*

In the transition state for hydrolysis of **10**, the hydroxide then attacks the phosphorus center of the substrate. Concomitantly, the hydroxide decoordinates from Mn(II) and the bond between phosphate and leaving group **12** is broken. This TS, called **24-TS** (Figure 10a), is calculated to be 13.1 kcal/mol higher in energy than **23**.

In the most stable product state, **12** is found to coordinate to Mn(II) with its phenyl oxygen. The phosphate product **11** is now bound to both metal ions in a bidentate fashion. This species, named **24**, is calculated to be -16.2 kcal/mol in free energy compared to **23**.

Following the same scheme as above to estimate the energetics of the catalyst regeneration process, it is calculated that the cost of forming **23** from **24** is 6.4 kcal/mol. Hence, the overall barrier for the reaction, going from **24** in one catalytic cycle to **24-TS** in the next, is calculated to be 19.5 kcal/mol (Figure 10b). This is similar to what was calculated for the dizinc catalysts above.

Several other mechanistic scenarios are tried, varying both the identity and the position of the nucleophile, the binding mode of the catalyst, the number and identity of coordinated species, the spin configuration, and whether the hydroxide acts as a nucleophile or as a Brønsted base. However, all these attempts give higher species and/or higher energy barriers. Thus, the calculations suggest complex **23** to be the active catalyst.

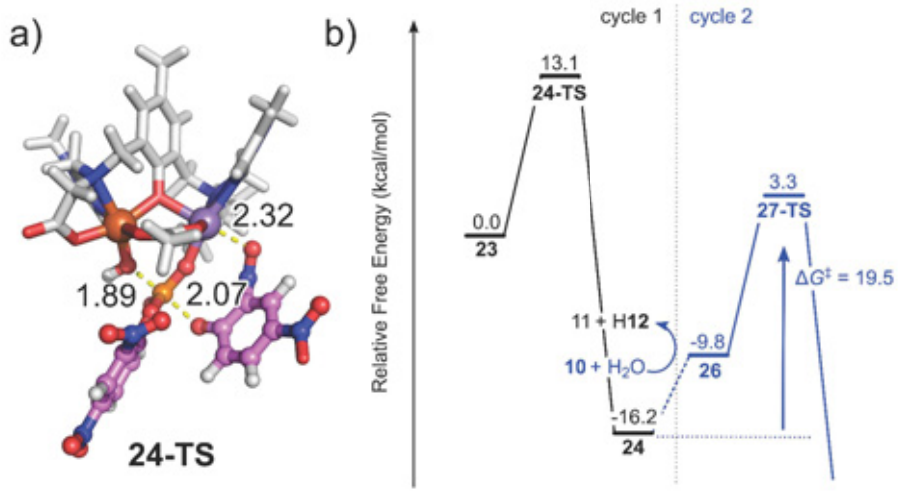


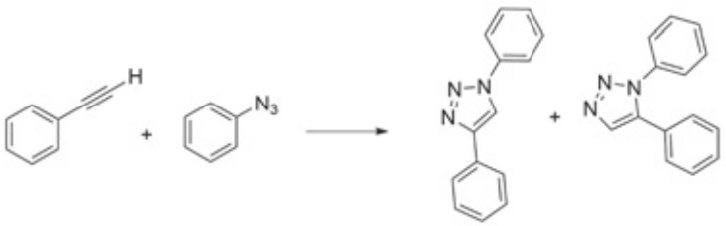
*Figure 10. (a) Optimized geometry of **24-TS**. (b) Calculated free energy profile for hydrolysis of **10** with FeMn complex of **5**. Note that the values reported here differ from the ones in Paper III; this difference is due to that the values in Paper III were calculated with the RRHO protocol, but here the values are recalculated with quasi-RRHO.*

3.4 Conclusions

There are several similarities in the calculated mechanisms of Reaction (1) and (2) catalyzed by dinuclear complexes of **3**, **4** and **5**. First, one hydroxide ion is calculated to be present in the active catalyst-substrate complexes. This hydroxide plays an active role in the catalytic reaction mechanism, either as a hydroxide (in Reaction (1)) or as a Brønsted base (in Reaction (2)). Second, the reactions are calculated to take place by essentially concerted mechanisms, and the resulting catalyst-product complexes are found to be the resting states in the catalytic cycles. The overall free energy barriers are identified as the progression from there to the TS of nucleophilic attack in the next cycle.

4 Cycloaddition reaction inside a self-assembled capsule (Paper IV)

In the previous chapter, we studied compounds that were synthesized with the aim to resemble the active sites of enzymes. Another interesting biomimetic approach is to design cavities that mimic the fact that many enzymes provide binding pockets that shield the substrates from interactions with solvent while prearranging them for reactions. Such features are found in so-called nanocontainers.[16,26,28,30,31] These molecules are synthesized to provide a cavity where one or a few substrate molecules can bind.[23] Size and shape complementarity between the cavity and the substrate(s) is found to be crucial for selective binding to take place.[21] Bonded inside the nanocontainer, unwanted contaminations can be extracted, unstable reaction intermediates can be recognized, and reactions that usually do not take place in solution can occur.[16,22] In the latter case, the reaction can be accelerated and even catalyzed.[22,24-29]



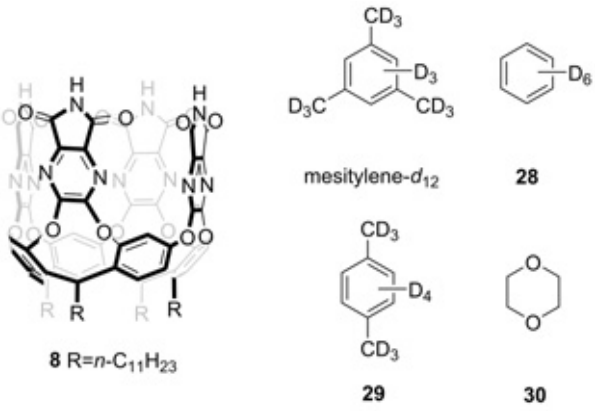
(3)

25 **26** **27a** **27b**

*Scheme 5. Cycloaddition reaction between acetylene **25** and azide **26**.*

Experimentally, the cycloaddition of acetylene **25** and azide **26**, as shown in Scheme 5, was found to be accelerated in the presence of self-assembled capsule $\mathbf{8}_2$ in mesitylene-d_{12} solvent.[32] The rim of this capsule is lined with imide units, and two cavitand monomers were observed to self-assemble around suitably sized guests *via* hydrogen-bonding between imides.[19] Furthermore, the capsule imposes a certain regioselectivity to the reaction. In the absence of capsule, the two regioisomers **27a** and **27b** were formed in roughly equivalent amounts, but in the presence of $\mathbf{8}_2$, only the encapsulated regioisomer complex **27a**@$\mathbf{8}_2$ was observed.[32] While the regioselectivity can be explained by the different shape complementarities between **27a** and **27b**

and the host, it is of fundamental interest to determine what factors are causing the rate acceleration.



To this end, quantum chemical calculations are performed here, and an energy decomposition analysis of the results is carried out to gain insight into the acceleration effect of the capsule. Another implicit aim of the calculations is, as usual, to evaluate the adopted quantum chemical methodology. Hitherto, only a few theoretical studies of rate acceleration or catalysis have been performed for reactions occurring inside supramolecular host-guest complexes,[109-120] so there is not yet a well-established methodology to be used in studies of host-guest catalysis.

4.1 Characterization of capsule

The experimental reaction was performed in off-the-shelf mesitylene-d_{12} solvent, in which the two impurities benzene-d_6 (**28**) and *p*-xylene-d_{10} (**29**) were present in millimolar amounts. Despite the low concentrations of the impurities, when **8** was dissolved in mesitylene-d_{12} the only complex that was observed was **28·29@8**$_2$. Therefore, this complex is chosen as the starting point of the calculations.

In the NMR experiments, the **28·29@8**$_2$ complex was observed to give signals corresponding to a C_{4v} symmetry.[19,32,121] In such a complex the N-H groups at the rim of each pyrazine-imide unit makes bifurcated hydrogen bonds to two carbonyl oxygens in the other **8** subunit so that in total eight bifurcated hydrogen bonds are formed. The monomeric unit in this arrangement is from now on called **8a**, and the geometry of the **28·29@8a**$_2$ complex is shown in Figure 11. As it turns out, when the geometry of **28·29@8a**$_2$ is optimized, another arrangement is found in which the **8** monomers were of approximate C_4 symmetry. In this cavitand geometry, named **8b**, each pyrazine-imide unit twists such that one of the carbonyl

oxygens faces the solvent, while the other accepts a hydrogen bond from an N-H group in the other **8b** monomer (Figure 11).

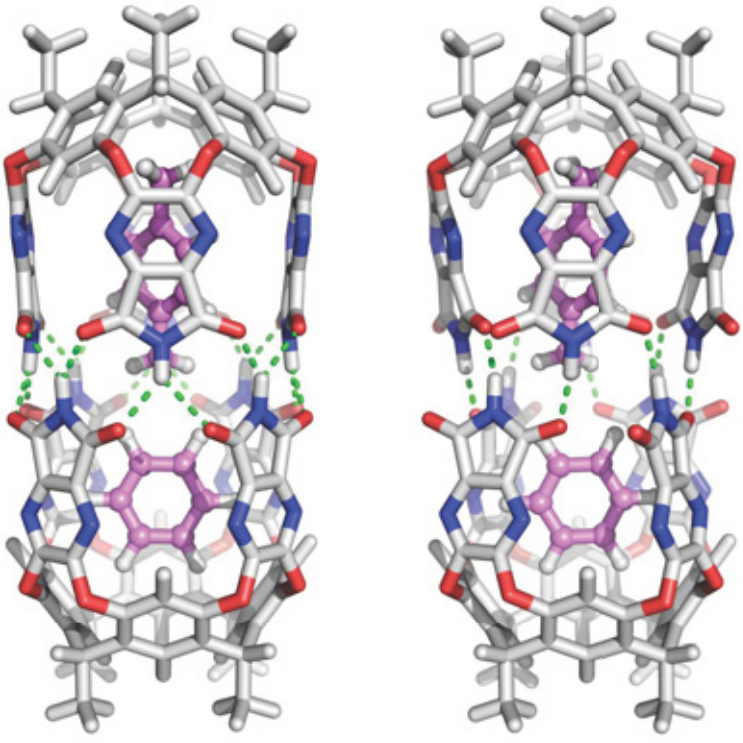


*Figure 11. Optimized geometries of inclusion complexes between cavitands **8a** and **8b** and guests **28** and **29**.*

Thus, the number of hydrogen bonds is the same in the two geometries, but the more linear hydrogen bonds in the latter structure result in shorter hydrogen bonds (NH···O distance 1.80 Å, compared to 2.12 Å for the bifurcated hydrogen bonds) and an overall significantly more stable structure (−12.1 kcal/mol, calculated at the B3LYP-D3(BJ)/6-311+G(2d,2p) level of theory). A reasonable suggestion is that the C_4-symmetric **8b** units are constantly interconverting between its two isomers, in which different carbonyl oxygens are facing the solvent. The TS for such a rotation would be similar to **8a,** and the 12.1 kcal/mol energy difference corresponds to a rate constant of about 9000 s^{-1}. This is much faster than the NMR timescale, so in the experiments it is likely that a time-averaged C_{4v} signal is measured.

4.2 Binding free energies inside the capsule

Experimentally, when reactants **25** and **26** were added to the mixture of cavitand **8** in mesitylene-d_{12}, the host-guest complexes observed by NMR (in order of abundance) were **25·26@8**$_2$, **26·30@8**$_2$, **26·26@8**$_2$, **25·25@8**$_2$ and **25·30@8**$_2$.[32] Note that of these complexes, two include dioxane guest **30** which is an impurity in the commercial solution of **25**. To see whether the binding trend could be reproduced by the calculations, all possible host-guest combinations are evaluated with the methodology presented in Section 2.5. The results are given in Table 1, where the relative stabilities of selected complexes are compared to **28·29@8b**$_2$.

*Table 1. Calculated stabilities of selected inclusion complexes, relative to **28·29@8b**$_2$.*

Complex	ΔG_0	$\Delta\delta G_{\text{free}}^{\text{conc}}$	$\Delta G_0 + \Delta\delta G_{\text{free}}^{\text{conc}}$
28·29@8b$_2$	0.0	0.0	0.0
25·25@8b$_2$	+0.5	−2.7	−2.2
25·26@8b$_2$	−1.5	−2.3	−3.8
25·30@8b$_2$	−1.0	0.0	−1.0
26·26@8b$_2$	−3.2	−1.9	−5.1
26·30@8b$_2$	−2.7	+0.4	−2.3

As seen from Table 1, all experimentally observed compounds are calculated to be of similar or lower free energy than **28·29@8b**$_2$, but the experimental order of abundance is not fully reproduced. However, when the experimental concentrations are taken into consideration via Equation (xv), where *X* and *Y* are different guests, the agreement with experiments is significantly improved.

$$\delta G_{\text{free}}^{\text{conc}} = RT \ln\left(\frac{[X][Y]}{[\mathbf{28}][\mathbf{29}]}\right) \qquad \text{(xv)}$$

Corrected in this way, the calculations predict **26·26@8b**$_2$ to be the most stable host-guest complex, being 1.3 kcal/mol more stable than **25·26@8b**$_2$, i.e. slightly different from the experimental observation. This is a relatively small error, and in the following sections, the **25·26@8b**$_2$ complex (shown in Figure 12) will be used as the ground state species, according to the experiments.

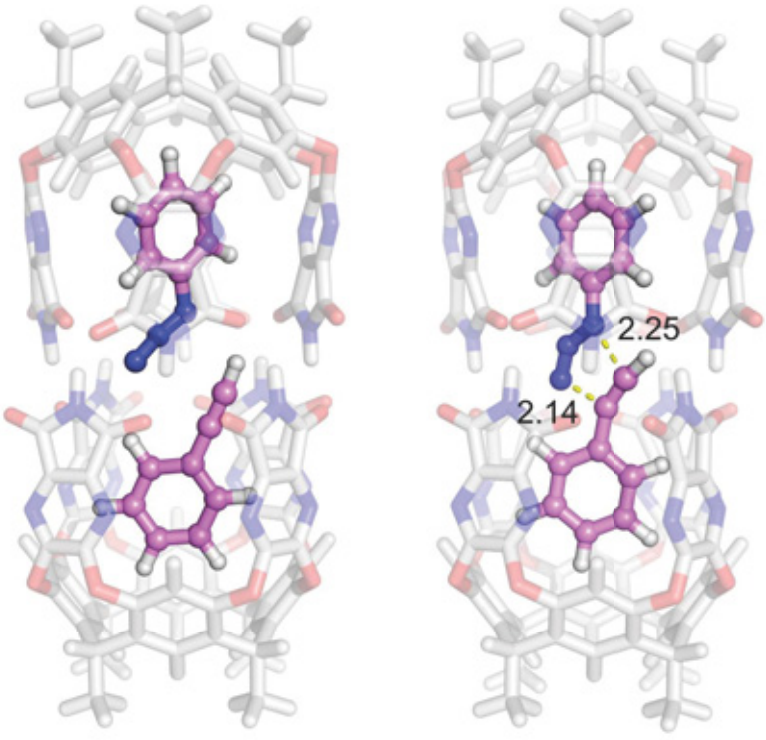


Figure 12. Optimized geometries of inclusion complexes with reactant and TS for the encapsulated cycloaddition reaction.

4.3 Acceleration of cycloaddition reaction

From the optimized geometry of **25·26@8b**$_2$, it can be observed that the two reactants are well arranged for reacting inside the capsule. The transition state for formation of the **27a** product, named **27a-TS@8b**$_2$, was optimized. This TS is calculated to be 21.6 kcal/mol higher in free energy than **25·26@8b**$_2$ and is shown in Figure 12. Compared to the TS optimized in the absence of capsule, the two phenyl rings of the reactants are more coplanar in the encapsulated TS, such that the whole complex is more planar.

The **27a** product is then formed inside the capsule, and this complex is calculated to be −56.8 kcal/mol in free energy compared to **25·26@8b**$_2$. This can be compared to the exergonicity of Reaction (3) in the absence of capsule, which is calculated to be −52.6 kcal/mol. The fact that the reaction becomes more exergonic in the presence of **8**$_2$ means that **27a** is calculated to be a better guest to the capsule than the combination of **25** and **26**. To regenerate the **25·26@8b**$_2$ complex for a second round of cycloaddition, one

25 and one **26** molecule must bind and one **27a** product must leave the capsule. The net reaction is thus identical to a regioselective variant of Reaction (3), and using the same protocol as in Chapter 3 to estimate the free energy of the catalyst regeneration process, this is calculated to be endergonic by $-52.6 - (-56.8) = 4.2$ kcal/mol. The free energy barrier for performing a second cycloaddition reaction in the capsule becomes thus $4.2 + 21.6 = 25.8$ kcal/mol. The free energy profile of the reaction is shown in Figure 13.

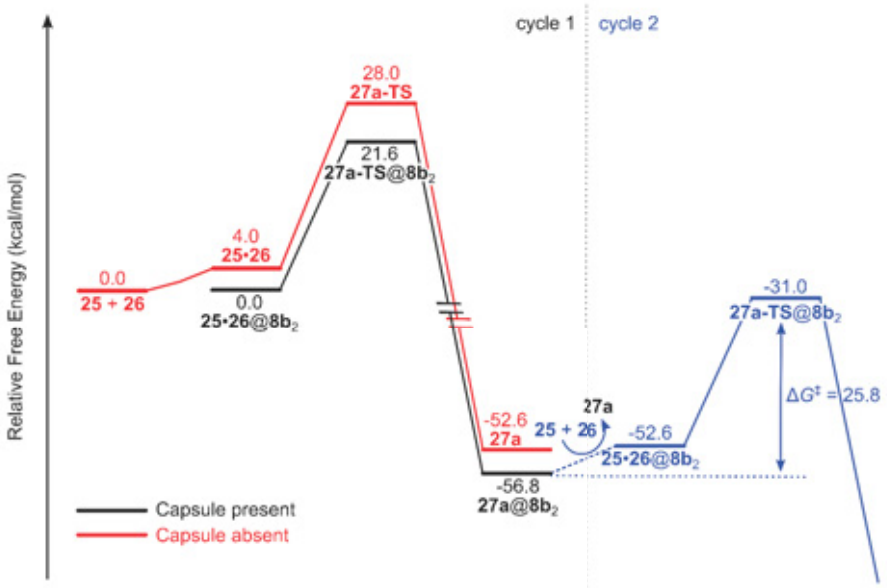


*Figure 13. Free energy profile for cycloaddition of **25** and **26** in presence (black/blue curve) or absence (red curve) of capsule (Figure altered from Paper IV).*

For comparison, in the uncatalyzed reaction, **27a-TS** is calculated to occur with a free energy barrier of 28.0 kcal/mol with respect to the separated reactants. Hence the calculations predict the capsule to lower the barrier by 6.4 kcal/mol in the first turnover and by 2.2 kcal/mol in succeeding catalytic cycles.

In the experiments, the formation of **27a@8**$_2$ was monitored for a week, and in this time only one turnover of Reaction (3) was observed.[32] The calculations predict the free energy barrier to increase in the second catalytic cycle compared to the first, by 4.2 kcal/mol. Using TST, this can be converted to a decrease in reaction rate by a factor of more than 1000. Hence, it can be expected that even if the capsule indeed catalyzes the reaction, the effect will not be measurable on the timescale of weeks. The rate measured for the cycloaddition reaction can be converted to a free

energy barrier of approximately 26 kcal/mol.[122] The calculations thus underestimate this value by roughly 4 kcal/mol.

The TS towards formation of the **27b** regioisomer is also calculated inside the capsule but is found to have a significantly higher free energy barrier, about 40 kcal/mol compared to **25·26@8b**. This pathway can thus be discarded.

It is now of interest to decipher how the capsule accelerates the reaction. To do so, the reaction barriers calculated in the presence and in absence of capsule are compared (Figure 14). First, the results are compared without the free energy corrections, which roughly constitute the entropic effects. At this level of theory, the barrier for the uncatalyzed reaction is calculated to be 23.7 kcal/mol. The entropic effect on the uncatalyzed barrier is thus approximately +4.3 kcal/mol. Then, the reactants and the TS are calculated with the same geometries as they adopt when bound inside the capsule. The calculated effect of this strain, imposed on the molecules by the incarceration, is a lowering of the barrier by 2.4 kcal/mol. The cause of this lowering is due to the geometry of the reactant supercomplex being more distorted by the capsule than the TS is. This leads to a favoring of the latter with respect to the former complex.

Finally, the distorted geometries of the reactants and the TS are placed inside the capsule, and it is calculated how the interaction between the capsule and the encapsulated species affects the free energy barrier. It is found that this effect reduces the barrier by only 0.7 kcal/mol. Hence, only taking into account the enthalpic component of the energies, the barrier for the reaction is calculated to be $23.7 - 2.4 - 0.7 = 20.6$ kcal/mol. Comparing this value to the free energy barrier calculated above, 21.6 kcal/mol, it is seen that the entropic component only affects the barrier by +1.0 kcal/mol. On the other hand, for the uncatalyzed reaction the effect of including the entropic corrections was +4.3 kcal/mol. Thus, compared to the outside reaction, this is calculated to reduce the barrier by $4.3 - 1.0 = 3.3$ kcal/mol. This is in part due to the fact that in absence of capsule, formation of the reactant supercomplex from the separated reactants is calculated to be an endergonic process, while in the presence of $\mathbf{8}_2$ the encapsulated reactant supercomplex becomes the ground state. In other words, it is costly to bring the reactants together in solution, but in the presence of capsule this process becomes downhill in free energy.

Taken together, this energy decomposition has identified three components that lower the free energy barrier of the cycloaddition reaction in the presence of capsule compared to in its absence. The entropic effect, discussed above, reduces the barrier by 3.3 kcal/mol. The imposed strain, disfavouring the reactant supercomplex more than the TS, lowers the barrier by an additional 2.4 kcal/mol. Finally, the capsule is calculated to interact slightly better with the TS than with the reactants, causing a further reduction of the barrier by 0.7 kcal/mol.

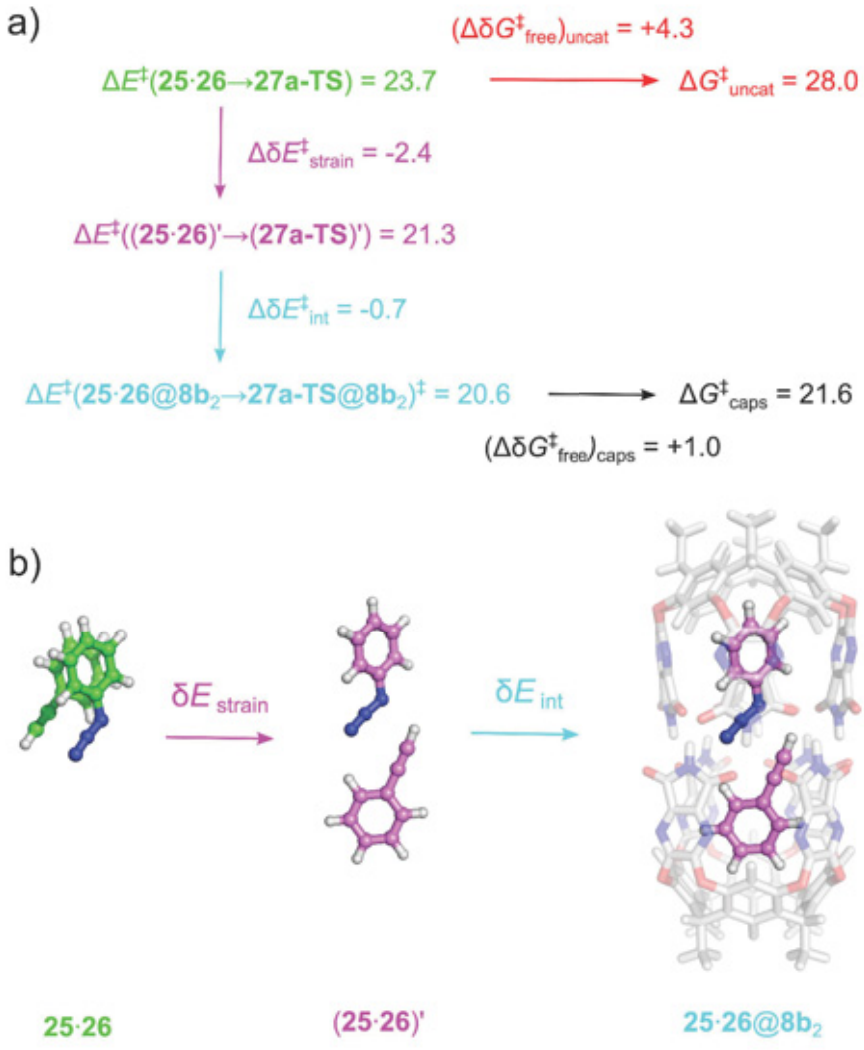


Figure 14. a) Energy decomposition scheme for the cycloaddition barrier. b) The reactant supercomplex geometry optimized in the absence of capsule (left) is distorted such that its arrangement in the presence of capsule is obtained (middle) before the interactions between reactants and capsule are considered (right).

4.4 Conclusions

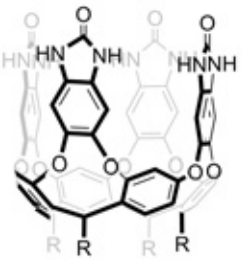
In this chapter, it is found that the methodology presented in Section 2.5 is appropriate to reproduce the lowering of the free energy barrier of the cycloaddition reaction in presence of capsule host. Furthermore, with the possibility to do an energy decomposition analysis and compare the

catalyzed and the uncatalyzed reactions, the origins of the rate acceleration can be quantified. In this case, it is shown that two main components contribute to the lowering of the barrier: the reduced entropic penalty and the fact that the reactants are more distorted than the TS by the encapsulation. These insights can potentially be used to design more efficient catalysts in the absence of electrostatic stabilization of the TS: the host should provide better shape complementarity with the TS than with the reactants and the products. The reactants should, however, be bound strongly enough for the entropic penalty of association to be overcome, while the binding energy of the products should ideally be low in order to facilitate turnover.

5 Alkane binding to water-soluble cavitand and capsule (Paper V)

In the previous chapter, capsule $\mathbf{8}_2$ was studied in nonpolar solvent mesitylene-d_{12}. For a host to be truly biomimetic, it should be able to imitate its template under biological conditions – which means in water.[33] This does, however, make the feat of self-assembly harder to accomplish, especially if this process utilizes polar contacts like hydrogen bonds. The water molecules surrounding the cavitands will compete for these interactions. Despite this challenge, several self-assembling supramolecular hosts have been synthesized for the use in aqueous environments.[33,123-126]

One such host molecule is **9**, which is similar to **8** but differs in two aspects. Namely, the *n*-undecyl substituents are replaced by alkyl pyridinium ions, and the pyrazine-imide units (the 'side panels') are replaced by benzimidazolone units, such that the imide groups at the rim of the vase-shaped monomers are replaced by urea moieties.[20] The cavitand has been observed to self-assemble to encapsulate hydrophobic guests in water.[20,34,35]



9 R=*n*-C_3H_6-(C_5H_6N)$^+$

31 n=1
32 n=2
33 n=3

An interesting trend was seen when these cavitands were mixed with *n*-alkanes.[34] For alkanes of length up to and including *n*-nonane (**31**), the cavitand formed 1:1 complexes with the alkane, which was observed to bind to vase monomer **9** in a coiled conformation (i.e. making *gauche* interactions along all of the alkane chain).[34] For *n*-alkanes from *n*-undecane (**33**) and up, the cavitands were observed to self-assemble around the alkane, forming 2:1 complexes.[20,34] An equilibrium was observed between the 1:1 and 2:1 host-guest complexes when the guest was *n*-decane (**32**), see Figure 15. For this

n-alkane, NMR signals were obtained for both cavitand **9** and capsule $\mathbf{9}_2$, so Equilibrium (4) is observed.[34]

$$2 \times \mathbf{32@9} \rightleftharpoons \mathbf{32@9}_2 + \mathbf{32} \tag{4}$$

By careful examination of the NMR spectra, Equilibrium (4) can be deduced to be exergonic by approximately 5 kcal/mol under standard state concentrations. In the present chapter, this system is studied with the primary aim to develop a computational protocol that can reproduce the energy of Equilibrium (4). As it turns out, the computational methodology presented in Section 2.5 is found to fail considerably in reproducing the energy of the equilibrium: the reaction is calculated to be exergonic by as much as 46 kcal/mol. To find a working protocol for the quantum chemical study of these species, each step of the quantum chemical calculations (as presented in Chapter 2) is examined, and corrections to the computational protocol will be suggested with the aim to reproduce the experimental observations summarized above.

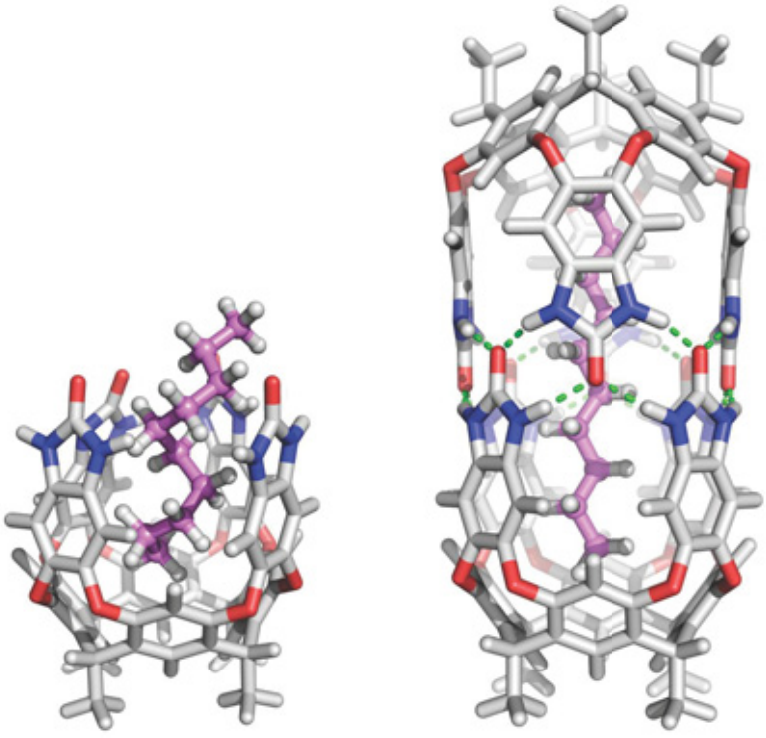


Figure 15. Optimized geometries of inclusion complexes between ***9*** *and* ***32****.*

5.1 Effect of using a truncated model

Before any changes are made to the computational protocol presented in Section 2.5, it is tested whether the truncation of the model, i.e. the substitution of the alkyl pyridinium moieties for methyls, affects the computational results. It turns out that they do not if one counterion per positive charge is included in the model. For the full (untruncated) system, both the gas-phase potential energies (E_{gas}) and the implicit solvation corrections (δG_{solv}) are found to be very similar to the values calculated with the truncated model. Optimized geometries of the truncated models of **32@9** and **32@9**$_2$ are shown in Figure 15, and computational results are shown in Table 2.

Table 2. Energy (kcal/mol) of Equilibrium (4) calculated with different computational approaches. × = The value is taken from the 'reference protocol'.[a]

Method	ΔE_{gas}	$\Delta\delta G_{free}$	$\Delta\delta G_{solv}$	ΔG_{tot}
Experimental value				−5
Reference protocol[a]	−115.9	+7.5	+62.0	−46.4
Cavitand model				
$[\text{Full model}]^{4+}$	+118.1	×	−99.7	+25.9
$[\text{Full model}]^{4+}\cdot(Cl^-)_4$	−115.5	×	+58.4	−49.6
Functional				
M06-2X	−95.7	+6.2	+60.1	−29.4
B3LYP (no dispersion correction)	−105.9	+19.2	+66.3	−20.5
B3LYP-D2[b]	−102.3	×	×	−32.8
Basis set				
6-311+G(2d,2p) + Counterpoise	−114.7	×	×	−45.2
Free energy protocol				
RRHO	×	+8.6	×	−45.3
Implicit solvation model				
C-PCM	×	×	+44.6	−63.8
SMD	×	×	+61.1	−47.3

[a] The computational protocol described in Section 2.5.
[b] Calculated with geometries optimized with B3LYP-D3(BJ).

5.2 Variations of computational methodology

Next, the dependency on the methodological choices is assessed. That is, the density functional, the dispersion model, the basis set, the protocol used to calculate thermal effects and the implicit solvation model are varied and the effects on the free energy of the reaction are discussed (Table 2).

In general, the free energy is calculated to be quite insensitive to basis set superposition error and the method to calculate the free energy contribution but proves to be quite sensitive to the choice of density functional and solvation model. The choice of density functional will be discussed first.

The largest effect is obtained if the calculations are carried out without dispersion corrections at all, which is an expected result since complex **32**@**9**$_2$ should be quite stabilized by these interactions. Calculating the free energies with the B3LYP functional, i.e. using the original expression (vii) and not correcting it for dispersion, the exergonicity of Equilibrium (4) is calculated to decrease significantly: the reaction free energy changed from −46.4 to −20.5 kcal/mol. However, an auxiliary constraint can be introduced here: that **32** should be predicted to be bound by the cavitand, i.e. that Equilibrium (5) should be exergonic.

$$\mathbf{32} + \mathbf{9} \rightleftharpoons \mathbf{32}@\mathbf{9} \qquad (5)$$

Calculated with B3LYP, the energy of this reaction is calculated to be *endergonic* by as much as 22.8 kcal/mol. This can be compared to the results obtained with B3LYP-D3(BJ), which actually also predicts Equilibrium (5) to be endergonic, but only by 2.5 kcal/mol. It is apparent that correcting for dispersion is necessary to properly represent the binding of **32** to **9**.

Changing B3LYP-D3(BJ) to B3LYP-D2, in which the predecessor of DFT-D3 is used to model the dispersion,[127] or to M06-2X, also causes a significant change in the predicted free energy of Equilibrium (4). With M06-2X, the free energy is calculated to be −29.4 kcal/mol, and B3LYP-D2 predicts it to be −32.8 kcal/mol. At the same time, with these approaches the free energy of Equilibrium (5) is calculated to be rather close to the experimental observation: M06-2X predicts the reaction to be endergonic by 1.2 kcal/mol while B3LYP-D2 predicts it to be quite exergonic, −7.5 kcal/mol.

Apart from the density functional, the results are also found to be sensitive to the choice of implicit solvent model. While COSMO-RS and SMD give similar results (−46.4 and −47.3 kcal/mol, respectively), C-PCM predicts the reaction to become even more exergonic: −63.8 kcal/mol. This discrepancy can be related to how the solvation models treat urea, which is present at the rim of the vase monomers. The hydration free energy of the urea molecule is calculated to be about −11 kcal/mol with COSMO-RS and SMD, but −8 kcal/mol with C-PCM. As a comparison, the hydration free energy of urea has been calculated to about −14 kcal/mol with Monte Carlo simulations.[128] Compared to this value, all three implicit approaches underestimate the interactions between urea and water, but C-PCM does so more than COSMO-RS and SMD. As an effect, the hydration of vase-shaped monomer **9** is expected to also be underestimated by C-PCM, and Equilibrium (4) is calculated to be more endergonic.

As mentioned in Chapter 2, implicit solvent models fall short when it comes to representing systems that make strong interactions with the solvent. In these cases, mixed explicit/implicit models can be used to model the compound of interest, in which a limited number of explicit solvent molecules are included in the model.[76-78,129,130] This approach will now be tested to further assess the sensitivity of Equilibrium (4) to the choice of solvation model. For now, the B3LYP-D3(BJ) functional will be kept, but the results will in the end also be presented with M06-2X and B3LYP-D2.

5.3 Mixed explicit/implicit solvation

In this section, the solvation effect on the free energy of Equilibrium (4) will be examined with a mixed explicit/implicit scheme. Since we will be dealing with systems containing many hydrogen-bonded entities, both counterpoise corrections and free energy corrections $\delta G_{\text{free}}^{\text{exch}}$ for systems with N identical species are included.

$$\delta G_{\text{free}}^{\text{exch}}(N) = RT \ln(N!) \qquad \text{(xvi)}$$

The most common approach when using mixed explicit/implicit methods is to include the same number of solvent molecules on each side of the equation. In the case of Equilibrium (4), such an approach is not viable, since one can imagine many more positions to add water molecules in the **32@9** structure than in **32@9**$_2$. Hence, a model is needed in which different numbers of water molecules can be added on each side of the reaction arrow.

One such approach was described by Goddard et. al,[77] where it was proposed that the surplus water molecules on either side of the reaction arrow should be added as water clusters.[77] This way, there is a cancellation of systematic errors since similar clusters are compared – on one side of the equation, the water molecules are bound to a solute in a cluster-like formation, and on the other side they are bound to other water molecules in a cluster. A caveat of this method, however, is that water clusters of different sizes should not be compared. For example, a 12-water molecule cluster is calculated to be 7.1 kcal/mol less stable than two 6-water clusters.

This affects Equilibrium (4): if for example N water molecules are used in the model of **32@9** and no waters in the model of **32@9**$_2$, then a cluster of $2N$ water molecules should be added on the right-hand side of the equation. One approach to remedy this problem is to add two N-sized water clusters on the right-hand side. Another, more systematic way, is presented below.

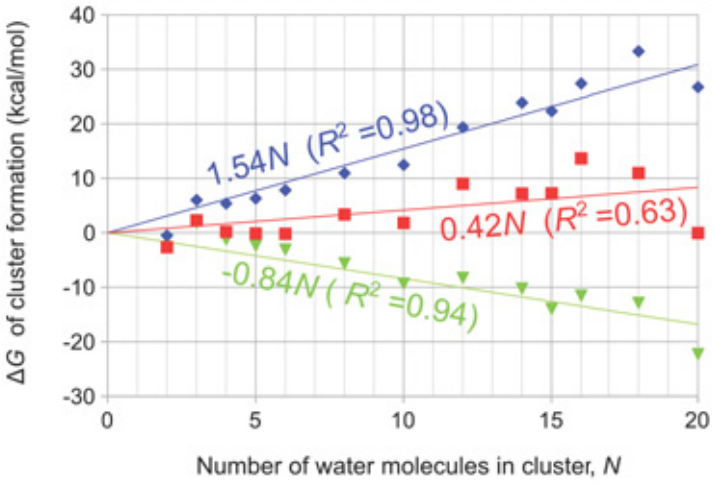


Figure 16. Free energies of water cluster formation, calculated with COSMO-RS (red), SMD (blue), and C-PCM (green).

We assume that the main reason why water cluster formation in water is calculated to not be thermoneutral is that the hydration free energy calculated for a single water molecule is slightly wrong. This value can then be corrected to make the formation of water cluster thermoneutral; the correction is obtained by fitting a straight line to the cluster formation free energies as a function of the cluster size, as done in Figure 16. With B3LYP-D3(BJ) and SMD, which is the solvation model that shows the most linear dependency of the cluster formation free energies on the number of waters (Figure 16), the correction to $\delta G_{solv}(H_2O)$ is calculated to be +1.54 kcal/mol.

To apply this method to Equilibrium (4), water molecules are added to **32@9** and **32@9**$_2$ and their binding free energies are calculated with the corrected value for hydration of H_2O. If these energies are negative, the water molecules are included in the model, and the solvation free energies are appended with the binding free energies of the waters. As it turns out, at least 16 water molecules are needed in the model of **32@9** in order for the calculated solvation free energy to converge (Figure 17a). With the approach described above, they are calculated to bind by in total −15.4 kcal/mol. For **32@9**$_2$, no water molecules are calculated to be bound. The geometry of the complex of **32@9** with 16 waters is shown in Figure 17b.

The effect on the free energy of Equilibrium (4) is thus that the left-hand side is stabilized by as much as $2 \times (-15.4) = -30.8$ kcal/mol, which means that the calculated exergonicity of the reaction decreases: the free energy changes from −47.3 to −16.5 kcal/mol. The effect of this treatment thus corrects the free energy of Equilibrium (4) a great deal towards the value of −5 kcal/mol that was derived from experiments.

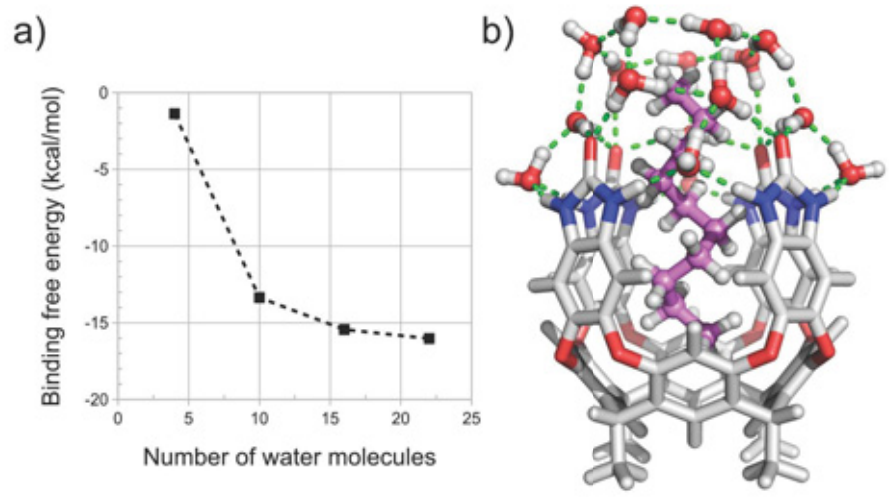


Figure 17. (a) Binding free energies of water molecules to ***32@9****, calculated with the revised method. (b) Optimized geometry of* ***32@9*** *with 16 water molecules bound.*

The revised methodology is thus: Geometries are optimized and quasi-RRHO free energy corrections are calculated at the 6-31G(d,p) level of theory. On the basis of the optimized geometries, solvation corrections are calculated with the same functional and basis set, but adding a correction to the calculated hydration free energy of a single water molecule such that water cluster formation in water becomes thermoneutral. Counterpoise-corrected energies are calculated on the same geometries with the 6-311+G(2d,2p) basis set. Standard state corrections and corrections for the number of exchangeable molecules are also added, as described in Equations (xiii) and (xvi).

The sensitivity of this revised method to the choices of functional and solvation model is now assessed. In Table 3, the results are presented when the solvation model is changed from SMD to C-PCM and COSMO-RS, and when the functional is changed from B3LYP-D3(BJ) to M06-2X and B3LYP-D2. Of the possibilities evaluated here, the combinations of functional and solvation model that comes closest to the experimental value of −5 kcal/mol are M06-2X/SMD and B3LYP-D3(BJ)/SMD, which when corrected with the explicit-implicit model yield free energies of Equilibrium (4) of −0.8 kcal/mol and −16.5 kcal/mol, respectively.

The method is now put to test for calculating the energetics of a few more encapsulation reactions with **9**, for which approximate free energies can be deduced from experiments.

Table 3. Free energy (kcal/mol) of Equilibrium (4), calculated with different functionals and solvation models.

Functional	Solvation model	Reference protocol[a]	Revised protocol
B3LYP-D3(BJ)	SMD	−47.3	−16.5
B3LYP-D3(BJ)	C-PCM	−63.8	−28.4
B3LYP-D3(BJ)	COSMO-RS	−46.4	−27.8
M06-2X	SMD	−30.9	−0.8
B3LYP-D2[b]	SMD	−32.5	+12.8
Experimental value			−5

[a] The computational protocol described in Section 2.5, with functional and solvation model as defined in this table.
[b] Calculated with geometries optimized with B3LYP-D3(BJ).

5.4 Tests of the mixed explicit-implicit method

As stated above, nonane **31** was found to be preferentially bound by the cavitand and undecane **33** by the capsule.[20,34] This gives two more equilibria to evaluate:

$$2 \times \mathbf{31@9} \rightleftharpoons \mathbf{31@9_2} + \mathbf{31} \quad (6)$$

$$2 \times \mathbf{33@9} \rightleftharpoons \mathbf{33@9_2} + \mathbf{33} \quad (7)$$

The free energies of these reactions are related to that of Equilibrium (4). Equilibrium (6) should be less exergonic (or more endergonic), and Equilibrium (7) should be more exergonic than Equilibrium (4). The free energies of Equilibria (4), (6) and (7), calculated with the revised method, are given in Table 4. Here, B3LYP-D3(BJ) and M06-2X are employed in conjunction with the SMD solvation model.

The experimental trends are well-reproduced with the two methods considered in Table 4. B3LYP-D3(BJ) predicts a steady increase in the affinity of the *n*-alkane for binding to dimer $\mathbf{9}_2$ *vis-à-vis* monomer **9**: For **32**, this affinity is 12.5 kcal/mol higher than with **31**, and with **33** the affinity further increases by 11.5 kcal/mol. With M06-2X the three equilibria in Table 4 are calculated to be closer in energy. Equilibrium (4) and (10) are predicted to have very similar free energies, which disagrees with experiments. However, the free energy of Equilibrium (4) is significantly closer to the experimental value of −5 kcal/mol than when calculated with B3LYP-D3(BJ).

Table 4. Free energies (kcal/mol) of Equilibria (4), (6) and (7) calculated with the reference protocol and with the revised computational protocol.

Reaction	B3LYP-D3(BJ)		M06-2X	
	Reference protocol[a]	Revised protocol	Reference protocol[a]	Revised protocol
(6)	−43.3	−3.5	−28.0	+6.1
(4)	−47.3	−16.5	−30.9	−1.1
(7)	−52.5	−28.0	−34.8	−1.2

[a] The computational protocol described in Section 2.5, with SMD solvation.

A final test here is a revisit of Equilibrium (5), i.e. the binding of **32** to **9**. The reference protocol with SMD solvation predicts **32** to be slightly unbound by the cavitand, +2.5 kcal/mol with B3LYP-D3(BJ) and +1.2 kcal/mol with M06-2X. It turns out that the new method actually predicts **32** to be *more* unbound by **9**: the free energy of Equilibrium (5) is calculated to be +19.7 and +10.5 kcal/mol with B3LYP-D3(BJ) and M06-2X, respectively. As many as 28 water molecules can be added to cavitand **9** with decreasing energies (Figure 18a).

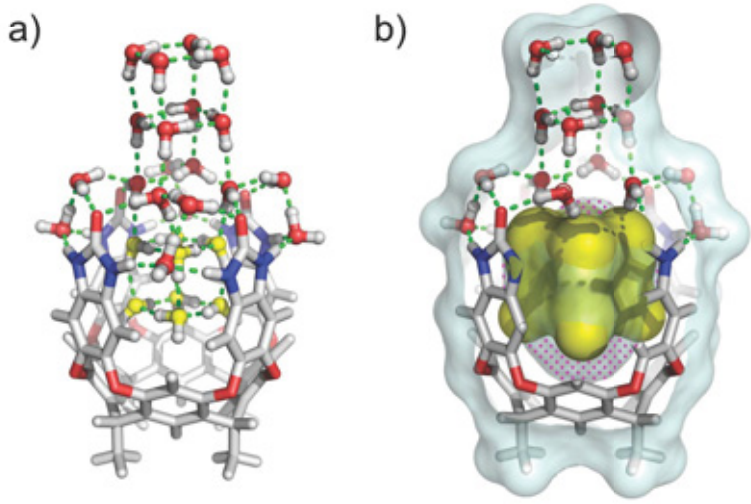


*Figure 18. (a) Optimized geometry of **9** with 28 water molecules bound. Eight water molecules that reside in the interior of the cavitand are shown in yellow. (b) Cross-section of the molecular cavities around the interior waters (yellow) and around the rest of the complex (transparent). Pockets between the two cavities are dotted.*

We realized that the water molecules bound inside the cavitand are artificially stabilized by some of the implicit dielectric medium, which is found to extend to the interior of **9** (Figure 18b). This is due to the fact that in the PCM calculations, the molecular cavity is generated from atom-

centered spheres, and when they are merged some empty pockets still reside between guest and host. They are treated as parts of the dielectric continuum, which means that the guests are modeled to be directly interacting with the solvent. To quantify the effect of this interaction, the empty pockets between host and guests are eliminated by manual addition of cavity spheres. This treatment reduces the calculated endergonicity of Equilibrium (5) to +14.1 and +6.1 kcal/mol with B3LYP-D3(BJ) and M06-2X, respectively.

Hence, while the preference between the formation of 1:1 and 2:1 host:guest complexes with host **9** and *n*-alkane guests is better reproduced with the mixed explicit-implicit solvent method introduced above, there is still uncertainty involved when water is treated as a guest, as in Figure 18.

5.5 Conclusions

To summarize, a new computational protocol is here developed for more accurate quantum chemical calculation of stabilities of models including explicit solvent molecules, compared to when the solvent is only modeled implicitly. The simple protocol, in which the hydration free energy of water is corrected such that the formation of water clusters is predicted to be thermoneutral, is not at all limited to cases like the one presented above. It can be used in calculations of all kinds of systems where it is of interest to determine how many solvent molecules that should be included in the model of the system. For the system studied above, the success of the applied method shows that the key to calculating the free energy of Equilibrium (4) correctly with DFT methods is to correct for the underestimation of the solvation energy of complex **32@9**.

6 Conclusions and outlook

In this thesis, the catalytic and substrate-binding properties of five biomimetic complexes were studied using quantum chemical methods. The calculations have provided insights into the observed trends that would not be easily available otherwise.

In the case of the dinuclear complexes of **3**, **4** and **5**, the identity of the nucleophile in the studied hydrolysis and transesterification reactions was identified as a hydroxide terminally bound to one of the metal centers. The calculated free energy profiles, spanning two consecutive catalytic cycles, predicted catalyst-product complexes to be the resting state in the reaction mechanism, a prediction that has been confirmed by experiments on similar complexes. The rate-determining free energy barriers were thus found to be associated with proceeding from the catalyst-product complex in one cycle to the transition state of nucleophilic attack in the next. For the design of future complexes, the computational results indicate that efforts should be concentrated on disfavouring the binding of the reaction products.

Next, a cycloaddition reaction was studied inside capsule-shaped host $\mathbf{8}_2$, for which a previously unrecognized geometry was identified. In agreement with experiments, the free energy barrier for the reaction was calculated to be lowered with respect to the uncatalyzed reaction, but to a larger extent in the first reaction cycle than in following turnovers. An energy decomposition analysis predicted that the rate acceleration can be mainly attributed to two effects: a lowering of the entropic cost of bringing the substrates together, and a larger destabilization of the reactant supercomplex than of the transition state caused by the encapsulation. These conclusions imply that shape complementarity between the supramolecular host and the transition state is of high importance in the development of more efficient catalytic host-guest systems.

Finally, the alkane binding trends of capsule $\mathbf{9}_2$ were studied, and the necessity of incorporating explicit solvent molecules to properly reproduce the experimental observations was identified. A new computational protocol was introduced for calculation of the solvation free energies of mixed explicit/implicit models that can be compared to the results obtained with implicit models. In this protocol, high-level quantum chemical calculations were combined with a simple correction for the hydration free energy of a single water molecule. The method was put to the test and with the choice of the appropriate density functional, it well reproduced the alkane binding

trends observed by experiments. A future application is to use this protocol to study the binding of longer alkanes, which has been proposed to lead to capsule widening via interactions between the capsule rim and water molecules from the solution.[35] The method is however not in any way restricted to the modeling of supramolecular complexes but could be used to give estimates of binding energies of water molecules to any system.

Populärvetenskaplig sammanfattning

Enzymer är stora biologiska molekyler som gör kemiska reaktioner möjliga som annars inte skulle ske. Flera processer som är nödvändiga för livet som vi känner det beror av att det finns enzymer som kan snabba på dem utan att själva förgås – det vill säga *katalysera* dem. En ökad förståelse av hur enzymkatalys fungerar kan ge nya möjligheter till att till exempel utforma läkemedel och göra kemiska processer i industrin mer miljövänliga.

Det finns flera sätt för oss att lära känna enzymer närmare. Ett är givetvis att studera enzymerna direkt i sin naturliga miljö, celler. Sådana studier är emellertid ofta svåra då celler är väldigt stökiga miljöer. Kompletterande tillvägagångssätt behövs, och ett sådant är att skapa och studera modeller som liknar enzymet på ett eller flera sätt. Hur sådana så kallade *biomimetiska* modeller lyckas reproducera enzymets verkningar ger insikt i vilka faktorer som är viktiga för enzymets beteende.

I den här avhandlingen har två sorters system studerats, som försöker efterlikna enzymer på olika sätt. Systemen har tidigare undersökts noggrant i laboratorier. I den här avhandlingen används teoretiska metoder, där beräkningskemiska datorprogram används för att fördjupa kunskapen om de biomimetiska modellerna. Insikter som inte är lätta att komma fram till med laborativa mätmetoder kan på så vis nås. Ett centralt begrepp i det här fallet är energi: varje steg i en kemisk reaktion motsvaras av att molekyler är arrangerade på ett särskilt vis, vilket i sin tur motsvaras av en energi. Med så kallad *täthetsfunktionalteori* räknas energier ut för olika molekylarrangemang och när de jämförs kan slutsatser dras om varför reaktionerna sker.

I de första studierna undersöks tre komplex som är gjorda för att efterlikna den lilla del av det stora enzymet där den katalyserade kemiska reaktionen sker, det så kallade *aktiva sätet*. De modellkomplex som studerats i den här avhandlingen har tillverkats för att efterlikna det aktiva sätet i en klass enzymer som bryter fosfoesterbindingar, starka bindningar som finns i bland annat DNA. Dessa enzymer har ofta två metalljoner i sitt aktiva säte, och de komplex som undersökts här binder antingen två zinkjoner eller en järn- och en manganjon. Med beräkningarnas hjälp ser vi att en vattenmolekyl, som sitter mellan de två metalljonerna, spelar en aktiv roll i att fosfoesterbindningarna bryts (Bild 1). När reaktionen har gått så måste produkten släppas loss så en ny fosfoestermolekyl kan binda in och reaktionen upprepas. I det här fallet visar beräkningarna dock att produkten

binder väldigt hårt till modellkomplexet, vilket är hämmande för reaktiviteten då det krävs energi för att få produkten att släppa.

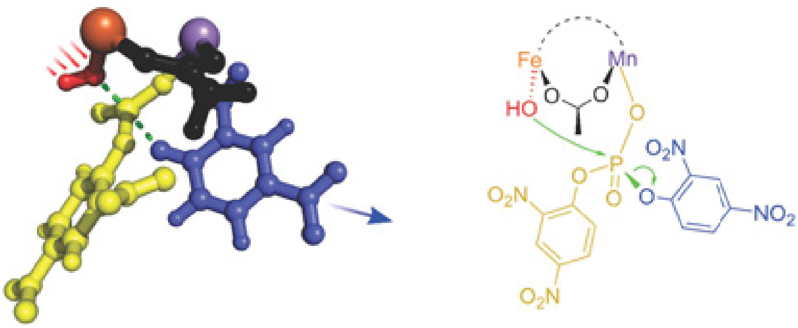


Bild 1. Reaktioner har studerats där vattenmolekyler som är bundna till metalljoner är aktiva i att bryta fosfoesterbindningar. Till vänster: kul- och pinnmodell av skeendet. Fosfoestermolekylen visas i gult och blått, och en vattenmolekyl som här är bunden till en järnatom och har tappat ett av sina väten visas i rött. Gröna streckade linjer indikerar kemiska bindningar som bildas eller bryts. Till höger: schematisk bild av samma reaktion.

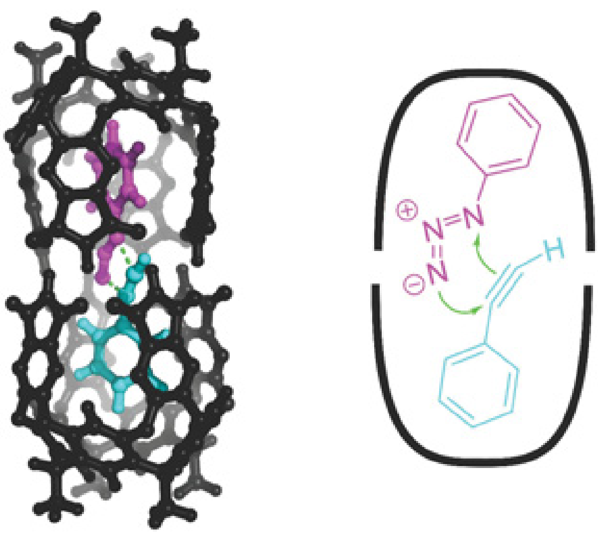


Bild 2. Reaktionen mellan fenylazid (magentafärgad) och fenylacetylen (cyan) inuti kapseln. Till vänster: kul- och pinnmodell av reaktionen. Till höger: schematisk bild av densamma.

Den andra sortens biomimetiska molekylär som här studerats är så kallade nanokapslar. Som namnet antyder kan dessa molekylkomplex kapsla in andra molekyler och därigenom skydda dem från påverkningar utifrån, liknande hur enzymer selektivt kan binda molekyler från lösningen för att använda på olika sätt. Ett kapselkomplex som undersöks i avhandlingen fångar två molekyler som kan reagera med varandra (Bild 2). Inuti kapseln

går reaktionen snabbare än utanför, och bara en av två möjliga produkter fås. Vi har undersökt vilka faktorer som är grunden till att reaktionen går fortare: är kapselns roll bara att vara "match-maker" och föra ihop reaktanterna, eller gör den också något med molekylerna för att underlätta reaktionens gång? Beräkningarna indikerar att kapselns huvudsakliga verkan är dels att föra samman de två reagenterna, dels att rikta dem rätt så att reaktionen kan ske utan större omorganisationer. Beräkningarna visade även att kapseln inte ser ut riktigt som det tidigare troddes, utan sitter ihop på ett sätt som inte hade föreslagits förr.

En annan kapsel som framställts har visat sig binda olika långa kolväten på olika sätt. Kapseln består av två halvor, och beroende på kolvätets längd kan det antingen forma komplex med bara en vasformad halva eller kapslas in av två halvor (Bild 3). Ett specialfall är linjära kolväten som är tio kolatomer långa, och har observerats inuti både vas och kapsel. När våra beräkningsmetoder använts för att försöka efterlikna detta och ge insikt i balansen mellan växelverkningar med vas och kapsel så misslyckas de rejält. Det huvudsakliga syftet i studien här är att uppdatera beräkningsmetodologin för att korrekt kunna återskapa den experimentellt observerade trenden. Det visar sig att beräkningarnas tillkortakommanden till största delen beror på att vattnet som omger vasen (och kapseln) inte modelleras tillräckligt bra med den metod som vi vanligtvis använder, där vattnet beskrivs som ett homogent elektriskt fält. För att resonabla resultat ska uppnås behöver ett antal explicita vattenmolekyler inkluderas i den beräkningskemiska modellen.

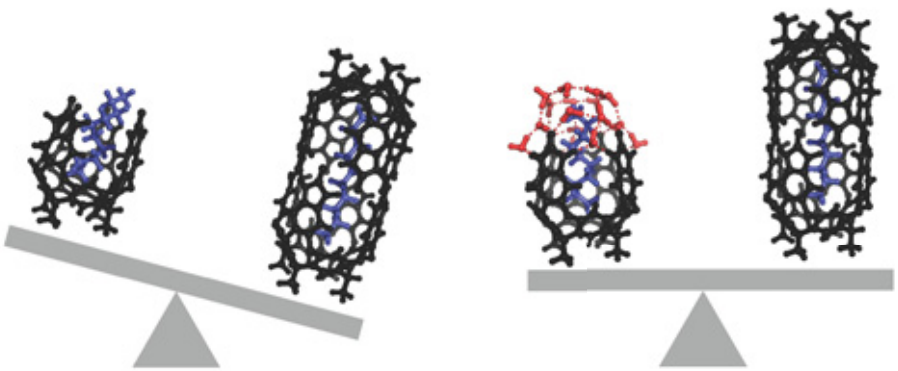


Bild 3. Till vänster: Om standardmetoden används i våra beräkningar så förutspås det inkapslade kolvätet (blått) vara mycket mer stabilt än det som binder i vasen. Till höger: Genom att modellera växelverkningar mellan vaskomplexet och omkringvarande vattenmolekyler (röda) fås en mer balanserad beskrivning.

Allt som allt visar studierna här att den beräkningskemiska metoden kan användas för att, med goda resultat, tillämpas på vitt skilda typer av komplex och för att studera både bindning och reaktivitet. Beräkningarna har givit nya

insikter och därmed kompletterat de bilder som målats upp med laborativa observationer. Både fosfoesterkomplexen och kapslarna har möjliga tillämpningar inom läkemedelsutveckling: utvecklingar av de förra skulle kunna användas för att till exempel bryta ner infekterat DNA, och de senare skulle kunna användas för att transportera läkemedel till rätt plats i kroppen eller binda skadliga ämnen och förhindra dem från att reagera. Förhoppningsvis kan kunskapen som utvecklats i denna avhandling bidra till utvecklingen av sådana hjälpmedel.

Acknowledgements

First and foremost, I want to thank my supervisor *Fahmi Himo*. Thanks for having me in your group for the last seven years, and special thanks for your tenacity in improving my scientific thinking and writing.

Ebbe Nordlander and *Biswanath Das*, for the collaborations on phosphoesterase mimics. Thanks for good discussions and nice visits.

Julius Rebek, Jr., for the collaboration on the capsules. I appreciate you taking time to meet with us while in Europe, and I am grateful for your encouraging comments.

Andrés, for the collaboration on the water-soluble capsule. I apologize for not naming the correction appropriately in the manuscript, but you know that it will always in my heart be the 'Algarra factor'.

Pher G. Andersson, for showing interest in this thesis, and *Oriana* for proofreading it.

C. F. Liljevalch J:ors stipendiefond and *Knut and Alice Wallenberg Foundation* for travel grants during my Ph.D. studies.

Per Siegbahn and *Margareta Blomberg* for sharing your knowledge.

All past and present members of the FH group, for making my work here so fun. *Sara*, thanks for sharing office space, and for co-preparing the stjärnanis schnaps that time. *Maria*, thanks for your playful attitude. *Bianca*, thanks for always showing appreciation. *Micke*, thanks for lifting my mood with hip-hop lyrics. *Stefano*, thanks for encouraging me to do a Ph.D. in quantum chemistry, and *Oscar*, likewise. *Ferran* and *Oriana*, thanks for stepping into my shoes. *Xiang*, thanks for giving me a birthday buddy. *Liao*, *Masoud*, *Binh*, *Genping*, *Marcin*, *Karim*, *Jiji*, *Jacob*, *Helena* and *Lynn*, thanks for making the group such a nice place during my years here.

Maria Rudbeck, for introducing me to the world of quantum chemistry.

All the people at the *Department of Organic Chemistry*. I will happily remember the nice atmosphere and all the nice people I have met here. Special thanks to *Louise* and *Jenny* for helping with all kinds of issues, and to *Carin* and *Sigrid* for the assistance in buying the computer cluster, which has been used to calculate most of the results herein.

Mamma och *pappa*, tack för all hjälp med barnpassande och för att ni tog er tid att läsa den populärvetenskapliga sammanfattningen.

Sofia och *Sally*, tack mina hjärtan för att ni ser till att mitt fokus är på det som är viktigt.

Sara, min älskling. Tack för allt.

References

[1] Weston, J. *Chem. Rev.* **2005**, *105*, 2151-2174.
[2] Brown, R. S. *Pure Appl. Chem.* **2015**, *87*, 601-614.
[3] Mancin, F.; Tecilla, P. *New J. Chem.* **2007**, 31, 800-817.
[4] Christianson, D. W.; Cox, J. D. *Annu. Rev. Biochem.* **1999**, *68*, 33-57.
[5] Dalle, K. E.; Meyer, F. *Eur. J. Inorg. Chem.* **2015**, *21*, 3391-3405.
[6] Bonfá, L.; Gatos, M.; Mancin, F.; Tecilla, P.; Tonellato, U. *Inorg. Chem.* **2003**, *42*, 3943-3949.
[7] Sträter, N.; Lipscomb, W. N.; Klabunde, T.; Krebs, B. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2024-2055.
[8] Mitić, N.; Smith, S. J.; Neves, A.; Guddat, L. W.; Gahan, L. R.; Schenk, G. *Chem. Rev.* **2006**, *106*, 3338-3363.
[9[Lipscomb, W. N.; Sträter, N. *Chem. Rev.* **1996**, *96*, 2375-2433.
[10] Carboni, M.; Latour, J.-M. *Coord. Chem. Rev.* **2011**, *255*, 186-202.
[11] Sessler, J. L.; Sibert, J. W.; Burrell, A. K.; Lynch, V.; Markert, J. T.; Woorten, C. L. *Inorg. Chem.* **1993**, *32*, 4277-4283.
[12] Anbu, S.; Kamalraj, S.; Varghese, B.; Muthumary, J.; Kandaswamy, M. *Inorg. Chem.* **2012**, *51*, 5580-5592.
[13] Jarenmark, M.; Kappen, S.; Haukka, M.; Nordlander, E. *Dalton Trans.* **2008**, 993-996.
[14] Das, B.; Daver, H.; Pyrkosz-Bulska, M.; Persch, E.; Barman, S. K.; Mukherjee, R.; Gumienna-Kontecka, E.; Jarenmark, M.; Himo, F.; Nordlander, E. *J. Inorg. Biochem.* **2014**, *132*, 6-17.
[15] Carlsson, H.; Haukka, M.; Nordlander, E. *Inorg. Chem.* **2004**, *43*, 5681-5687.
[16] Hof, F.; Craig, S. L.; Nuckolls, C.; Rebek, J., Jr. *Angew. Chem. Int. Ed.* **2002**, *41*, 1488-1508.
[17] Wyler, R.; de Mendoza, J.; Rebek, J., Jr. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1699-1700.
[18] Beisell, T.; Powers, R. E.; Raymond, K. N. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1084-1086.
[19] Heinz, T.; Rudkevich, D. M.; Rebek, J., Jr. *Nature* **1998**, *394*, 764-766.
[20] Zhang, K.-D.; Ajami, D.; Rebek, J., Jr. *J. Am. Chem. Soc.* **2013**, *135*, 18064-18066.
[21] Mecozzi, S.; Rebek, J., Jr. *Chem. Eur. J.* **1998**, *4*, 1016-1022.
[22] Yoshizawa, M.; Klosterman, J. K.; Fujita, M. *Angew, Chem. Int. Ed.* **2009**, *48*, 3418-3438.
[23] Rebek, J., Jr. *Angew. Chem. Int. Ed.* **2005**, *44*, 2068-2078.
[24] Deraedt, C.; Astruc, D. *Coord. Chem. Rev.* **2016**, *324*, 106-122.
[25] Hooley, R. J., Rebek, J., Jr. *Chem. Biol.* **2009**, *16*, 255-264.
[26] Raynal, M.; Ballester, P.; Vidal-Ferran, A.; van Leeuwen, P. W. N. M. *Chem. Soc. Rev.* **2014**, *43*, 1734-1787.
[27] Koblenz, T. S.; Wassenaar, J.; Reek, J. N. H. *Chem. Soc. Rev.* **2008**, *37*, 247-262.

[28] Marchetti, L.; Levine, M. *ACS Catal.* **2011**, *1*, 1090-1118.
[29] Breiner, B.; Clegg, J. K.; Nitschke, J. R. *Chem. Sci.* **2011**, *2*, 51-56.
[30] Wiester, M. J.; Ulmann, P. A.; Mirkin, C. A. *Angew. Chem. Int. Ed.* **2011**, *50*, 114-137.
[31] Dong, Z.; Luo, Q.; Liu, J. *Chem. Soc. Rev.* **2012**, *41*, 7890-7908.
[32] Chen, J.; Rebek, J., Jr. *Org. Lett.* **2002**, *4*, 327-329.
[33] Jordan, J. H.; Gibb, B. C. *Chem. Soc. Rev.* **2015**, *44*, 547-585.
[34] Zhang, K.-D.; Ajami, D.; Gavette, J. V.; Rebek, J., Jr. *Chem. Commun.* **2014**, *50*, 4895-4897.
[35] Gavette, J. V.; Zhang, K.-D.; Ajami, D.; Rebek, J., Jr. *Org. Biomol. Chem.* **2014**, *12*, 6561-6563.
[36] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864-B871.
[37] Jensen, F. *Introduction to Computational Chemistry*, John Wiley & Sons Ltd. **2007**, p. 236.
[38] Leach, A. R. *Molecular Modelling: principles and applications*, Pearson Education Ltd. **2001**, p. 128.
[39] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–52.
[40] Dirac, P. A. M. *Math. Proc. Cambridge Philos. Soc.* **1930**, *26*, 376-385.
[41] Jensen, F. *Introduction to Computational Chemistry*, John Wiley & Sons Ltd. **2007**, pp. 93-98.
[42] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100.
[43] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200-1211.
[44] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B.* **1988**, *37*, 785-789.
[45] Reiher, M.; Salomon, O.; Hess, B. A. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
[46] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104–22.
[47] Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2005**, *123*, 154101.
[48] Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456-1465.
[49] Mládek, A.; Krepl, M.; Svozil, D.; Čech, P.; Otyepka, M.; Banáš, P.; Zgarbová, M.; Jurečka, P.; Šponer, J. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7295-7310.
[50] Steinmetz, M.; Grimme, S. *ChemistryOpen* **2013**, *2*, 115-124.
[51] Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.
[52] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
[53] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865-3868.
[54] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.
[55] Burns, L. A.; Vázquez-Mayagoitia, Á.; Sumpter, B. G.; Sherrill, C. D. *J. Chem. Phys.* **2011**, *134*, 084107.
[56] Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C. T.; Goddard, W. A. III *J. Chem. Theory. Comput.* **2009**, *5*, 1016-1026.
[57] Mardirossian, N.; Parkhill, J. A.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19325-19337.
[58] Goerigk, L.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670-6688.
[59] Weymuth, T.; Couzijn, E. P. A.; Chen, P.; Reiher, M. *J. Chem. Theory Comput.* **2014**, *10*, 3092-3103.
[60] Kazaryan, A.; Baerends, E. J. *J. Comput. Chem.* **2013**, *34*, 870-878.
[61] Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *J. Phys. Chem. A* **2007**, *111*, 10439-10452.
[62] Boys, S. F.; Bernardi, F.; *Mol. Phys.* **1970**, *19*, 553.
[63] Simon, S.; Duran, M.; Dannenberg, J. J. *J. Chem. Phys.* **1996**, *105*, 11024-11031.
[64] Noodleman, L. *J. Chem. Phys.* **1981**, *74*, 5737-5743.
[65] Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027-2094.

[66] Mennucci, B. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 386-404.
[67] Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995-2001.
[68] Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comp. Chem.* **2003**, *24*, 669-681.
[69] Klamt, A.; Schüürman, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799-805.
[70] Klamt, A. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 699-709.
[71] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378-6396.
[72] Takano, Y.; Houk, K. N. *J. Chem. Theory. Comput.* **2005**, *1*, 70-77.
[73] Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378-6396.
[74] Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224-2235.
[75] Klamt, A.; Mennucci, B.; Tomasi, J.; Barone, V.; Curutchet, C.; Orozco, M.; Luque, F. J. *Acc. Chem. Res.* **2009**, *42*, 489-492.
[76] Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2001**, *105*, 7241-7247.
[77]. Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A. III *J. Phys. Chem. B* **2008**, *112*, 9709-9719.
[78] Fennell, C. J.; Kehoe, C. W.; Dill, K. A. *Proc. Nat. Acad. Sci.* **2011**, *108*, 3234-3239.
[79] Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792-803.
[80] Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016-2027.
[81] Jensen, F. *Introduction to Computational Chemistry*, John Wiley & Sons Ltd. **2007**, pp. 429-439.
[82] Grimme, S. *Chem. Eur. J.* **2012**, *18*, 9955-9964.
[83] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision D.01, Wallingford CT, **2013**.
[84] Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270-283.
[85] Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822-8824.
[86] Schaefer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829-5835.
[87] COSMO*therm*, C3.0, release 1601, revision 2587, COSMOlogic GmbH & Co KG.
[88] Eckert, F.; Klamt, A. *AlChE J.* **2002**, *48*, 369-385.
[89] Dennington, R.; Keith, T. A.; Millam, J. M. *GaussView, Version 6*, Semichem Inc., Shawnee Mission, KS, **2008**.
[90] Aldridge, W. N. *Biochem. J.* **1953**, *53*, 110-117.
[91] Dumas, D. P.; Caldwell, S. R.; Wild, J. R.; Raushel, F. M. Biochem. *J. Biol. Chem.* **1989**, *264*, 19659-19665.
[92] Scanlan, T. S.; Reid, R. C. *Chem. Biol.* **1995**, *2*, 71-75.

[93] Ramalho, T. C.; de Castro, A. A.; Silva, D. R.; Silva, M. C.; França, T. C. C.; Bennion, B. J.; Kuca, K. *Curr. Med. Chem.* **2016**, *23*, 1041-1061.
[94] Chen, S.-L.; Fang, W.-H.; Himo, F. *J. Phys. Chem. B* **2007**, *111*, 1253-1255.
[95] Kim, J.; Tsai, P.-C.; Chen, S.-L.; Himo, F.; Almo, S. C.; Raushel, F. M. *Biochem.* **2008**, *47*, 9497-9504.
[96] Omburo, G. A.; Kuo, J. M.; Mullins, L. S.; Raushel, F. M. J. Biol. Chem. 1992,
[97] Benning, M. M.; Shim, H.; Raushel, F. M.; Holden, H. M. *Biochem.* **2001**, *40*, 2712-2722.
[98] Hill, C. M.; Li, W.-S.; Thoden, J. B.; Holden, H. M.; Raushel, F. M. *J. Am. Chem. Soc.* **2003**, *125*, 8890-8891.
[99] Ribeiro, A. J. M.; Alberto, M. E.; Ramos, M. J.; Fernandes, P. A.; Russo, N. *Chem. Eur. J.* **2013**, *19*, 14081-14089.
[100] Jarenmark, M.; Csapó, E.; Singh, J.; Wöckel, S.; Farkas, E.; Meyer, F.; Haukka, M.; Nordlander, E. *Dalton Trans.* **2010**, *39*, 8183-8194.
[101] Gouré, E.; Carboni, M.; Troussier, A.; Lebrun, C.; Pécaut, J.; Jacquot, J.-F.; Dubourdeaux, P.; Clémancey, M.; Blondin, G.; Latour, J.-M. *Chem. Eur. J.* **2015**, *21*, 8064-8068.
[102] Barman, S. K.; Mondal, T.; Koley, D.; Lloret, F.; Mukherjee, R. *Dalton Trans.* **2017**, *46*, 4038.
[103] Fan, Y.-B.; Gao, Y.-Q. *Acta Phys.-Chim. Sin.* **2010**, *26*, 1034-1042.
[104] Gao, H.; Ke, Z.; DeYonker, N. J.; Wang, J.; Xu, H.; Mao, Z.-W.; Phillips, D. L.; Zhao, C. *J. Am. Chem. Soc.* **2011**, *133*, 2904-2915.
[105] Schenk, G.; Mitić, N.; Hanson, G. R.; Comba, P. *Coord. Chem. Rev.* **2013**, *257*, 473-482.
[106] Schenk, G.; Gahan, L. R.; Carrington, N.; Mitić, N.; Valizadeh, M.; Hamilton, S. E.; de Jersey, J.; Guddat, L. W. *Proc. Nat. Acad. Sci.* **2005**, *102*, 273-278.
[107] Schenk, G.; Ge, Y.; Carrington, L. E.; Wynne, C. J.; Searle, I. R.; Carrol, B. J.; Hamilton, S.; de Jersey, J. *Arch. Biochem. Biophys.* **1999**, *370*, 183-189.
[108] Das, B.; Daver, H.; Singh, A.; Singh, R.; Haukka, M.; Demeshko, S.; Meyer, F.; Lisensky, G.; Jarenmark, M.; Himo, F.; Nordlander, E. *Eur. J. Inorg. Chem.* **2014**, *13*, 2204-2212.
[109] Alvira, E.; Cativiela, C.; García, J. I.; Mayoral, J. A. *Tetrahedron Lett.* **1995**, *36*, 2129-2132.
[110] Kim. S. P.; Leach, A. G.; Houk, K. N. *J. Org. Chem.* **2002**, *67*, 4250-4260.
[111] Chen, W.; Sun, L.; Tang, Z.; Chang, C.-E.; *bioRxiv* **2017**, 169565.
[112] Xu, L.; Hua, W.; Hua, S.; Li, J.; Li, S. *J. Org. Chem.* **2013**, *78*, 3577-3582.
[113] Furuki, T.; Hosokawa, F.; Sakurai, M.; Inoue, Y.; Chûjô, R. *J. Am. Chem. Soc.* **1993**, *115*, 2903-2911.
[114] Luzhkov, V.; Åqvist, J. *Chem. Phys. Lett.* **1999**, *302*, 267-272.
[115] Karnes, J. J.; Benjamin, I. *J. Phys. Chem. C* **2017**, *121*, 19209-19217.
[116] Fruschicheva, M. P.; Mukherjee, S.; Warshel, A. *J. Phys. Chem. B*, **2012**, *116*, 13353-13360.
[117] Ootani, Y.; Akinaga, Y.; Nakajima, T. *J. Comput. Chem.* **2015**, *36*, 459-466.
[118] Carlqvist, P.; Maseras, F. *Chem. Commun.* **2007**, 748-750.
[119] Goehry, C.; Besora, M.; Maseras, F. *ACS Catal.* **2015**, *5*, 2445-2451.
[120] Cantillo, D.; Ávalos, M.; Babiano, R.; Cintas, P.; Jiménez, J. L.; Palacios, J. *Org. Biomol. Chem.* **2011**, *9*, 7638-7642.
[121] Moran. J. R.; Ericson, J. L.; Dalcanale, E.; Bryant, J. A.; Knobler, C. B.; Cram, D. J. *J. Am. Chem. Soc.* **1991**, *113*, 5707-5714.
[122] Cacciapaglia, R.; Di Stefano, S.; Mandolini, L. *Acc. Chem. Res.* **2004**, *37*, 113-122.
[123] Biros, S. M.; Rebek, J., Jr. *Chem. Soc. Rev.* **2007**, *36*, 93-104.

[124] Laughrey, Z.; Gibb, B. C. *Chem. Soc. Rev.* **2011**, *40*, 363-386.
[125] Oshovsky, G. V.; Reinhoudt, D. N.; Verboom. W. *Angew. Chem. Int. Ed.* **2007**, *46*, 2366-2393.
[126] Murray, J.; Kim, K.; Ogoshi, T.; Yao, W.; Gibb, B. C. *Chem. Soc. Rev.* **2017**, *46*, 2479-2496.
[127] Grimme, S. *J. Comp. Chem.* **2006**, *27*, 1787-1799.
[128] Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *Isr. J. Chem.* **1993**, *33*, 323-330.
[129] da Silva, C. O.; Mennucci, B.; Vreven, T. *J. Phys. Chem. A* **2003**, *107*, 6630-6637.
[130] Groenenboom, M. C.; Keith, J. A. *J. Phys. Chem. B* **2016**, *120*, 10797-10807.