BERGISCHE
UNIVERSITÄT
WUPPERTAL

Investigation of Biologically Active Vanadium-Containing Complexes using DFT-computed NMR parameters

PhD Thesis

This thesis is submitted as required for the fulfilment of a
Doctorate of Philosophy at the University of Wuppertal
Presented by
Geethalakshmi K. Rangaswamy

Inaugural- Dissertation

Zu Erlangung des Doktorgrades der Naturwissenschaften
dem Fachbereich der Bergischen Universität Wuppertal
vorgelegt von
Geethalakshmi K. Rangaswamy

MAX-PLANCK-INSTITUT FÜR KOHLENFORSCHUNG · MÜLHEIM-RUHR

Mülheim an der Ruhr

2008

Referent: Prof. Dr. Michael Bühl

Co-Referent: Prof. Dr. Per Jensen

Date of Verbal Defence**: 05-02-2009**

Dedicated To My Beloved Parents

" When a mother hears her son heralded a good and learned man, Her joy exceeds that of his joyous birth "

Thirukural -Verse 69

Acknowledgments

It is my pleasure and privilege to thank many persons who made this thesis possible.

Supervisory panel:

A large thanks to Prof. Dr. Michael Bühl for being an excellent supervisor, his support and generosity throughout my candidature are highly appreciated. He has invested so much time and energy throughout my PhD, a simple thank you will never be able to repay his kindness. I would like to thank and acknowledge the continued support of Prof. Dr. W. Thiel, especially for making me welcome in his laboratory.

Financial support:

I would like to thank DFG and Max Planck Institute for Kohlenforschung for the financial support during the period of my Ph.D. programme.

Personal Support:

I would like to appreciate the past and present members of the theory group at the MPI in Mülheim, especially Dr. Mark P.Waller, Dr. Dongqi Wang, for their fruitful collaboration, valuable assistance, interesting discussions and support during the period of my work. It has been my pleasure to work in this team amongst such a pleasant atmosphere and I would like to offer my sincere thanks and gratitude to one and all. I would also like to thank computer department of the MPI-Mülheim, especially Horst Lenk for essential technical support in hardware and software related problems.

I would like to take this opportunity to thank my school teachers and my professors who taught me and encouraged me in all aspects to come to this position.

I am very thankful to many of my close friends and colleagues especially, for their wealth of advice and support to complete this work successfully.

I wish to thank my entire extended family: My sister, my brother in law, my little nephew-Aathi, uncles, aunties, cousins and my fiancé for their eternal support and providing a loving environment for me.

Lastly, and most importantly, I am indebted and would like to thank my beloved parents, Rangaswamy and Sarasvathi for their inspiration and never ending love, care, support and encouragement.

Table of Contents

Chapter 1

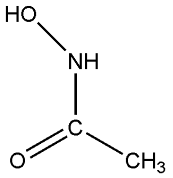
1 Introduction

^{51}V NMR spectroscopy is an important analytical tool for diamagnetic vanadium complexes.[1] Among transition metal nuclei, ^{51}V is one of the most abundant natural isotopes with characteristics that are very favorable for NMR spectroscopy, affording a sensitive probe for electronic and steric effects of the coordinating ligands. From the large body of data available, spans of ^{51}V chemical shifts can be extracted that are characteristic for a given ligand environment and, to a lesser extent, for the oxidation state. ^{51}V NMR can also be a potential tool to study vanadium compounds in the biosphere. Bioinorganic chemistry of vanadium[2] is less well developed than that of, e.g., iron. Besides nitrogenases, haloperoxidases is an important area of vanadium chemistry, that can be studied with ^{51}V NMR.[1b] Another fascinating aspect of vanadium chemistry is that of vanadium accumulation by living organisms, for example to provide a sufficient concentration for the use of the vanadium in one or more specific biochemical functions and/or as a means of protection against the toxicity that arises as an excess of vanadium.[3] An intriguing example of such biologically active vanadium complexes that may be used for this purpose is amavadin[8] which, in oxidised form, is also amenable to ^{51}V NMR spectroscopy.

Density functional theory (DFT) computations of ^{51}V chemical shifts are well advanced[1b,4] and have been applied so far mostly for small molecules, usually in context with their structures, reactivities (catalytic activities) and the relation between molecular dynamics and chemical shifts.[1b,5,125] One of the remaining challenges of such applications is the size of the complexes when it comes to very large systems such as biomolecules. To treat such large systems, quantum mechanics/molecular-mechanics (QM/MM) approaches have become the method of choice. QM/MM methods have been mostly applied for the modelling of the enzymatic reactions and other biomolecular processes that involve changes in the electronic structure, such as charge transfer or electronic excitation.[6] However, properties such as chemical shifts are only rarely addressed.[48a,d] The main goal of this thesis is to validate a suitable QM/MM method for computing ^{51}V NMR parameters of VHPO's (vanadium containing haloperoxidases), to apply this technique to specific structural problems and to assist in the interpretation of the observed results. Also included are applications of ^{51}V NMR computations for a smaller system, oxidized amavadin, aiming to rationalize its unusual ^{51}V chemical shift.

1.1 Amavadin

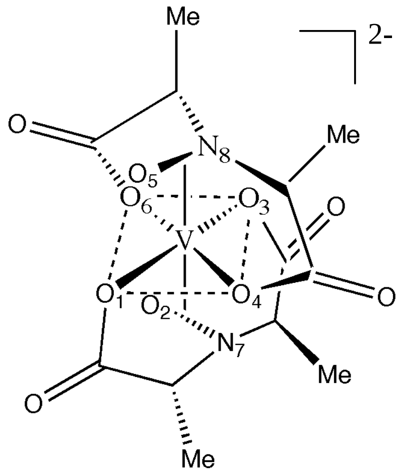
Occasionally deviations from "normal" chemical-shift ranges can suggest unusual electronic structures. For instance, exceptionally strong downfield shifts are observed for ^{51}V nuclei in vanadium (V) complexes containing catecholate or hydroxamate-based ligands (see Scheme 1.1.1). These downfield shifts can exceed 1000 ppm with respect to corresponding vanadates without these former ligands. This observation was first made by Pecoraro et al.[7] who labelled these ligands as "non-innocent" and who reported close relationships between ^{51}V chemical shifts and λ_{max} from low energy electronic excitations in the near IR spectra at around 900 nm.



Acetohydroxamic acid

Scheme 1.1.1

Oxidized amavadin is a species with an unusually high ^{51}V chemical shift, which, as had been speculated,[1b] could arise from non-innocent behavior of the ligands in this compound. Amavadin is a vanadium (IV) complex with the composition $[\Delta\text{-}V^{IV}\{(S,S)\text{-hidpa}\}_2]^{2-}$,where H_3hidpa is 2,2´-hydroxyiminodipropionic acid. Amavadin is found in the fungal genus Amanita and has an unusual structure and metal-ligand bonding mode as shown in Figure 1.1.1. Since the isolation of this compound from *A. muscaria*[8] there have been discordant reports in the literature with respect to its structure and function.[9,10] Initially a V=O species had been postulated[11,12], whereas later studies indicated that amavadin is a non-oxo octacoordinated vanadium complex (see Figure 1.1.1).[13] Since amavadin is the only naturally occurring compound with such a structure, the apparent stability of this complex is very unusual.[14] The hydroxylamido moieties in the amavadin complexes have fascinated chemists for some time, and various model systems and spectroscopic characterizations have been carried out, including EPR measurements.[15,16,17,18] Previous studies of vanadium hydroxylamido complexes have shown that the hydroxylamido ligand coordinates in a side on manner as does the peroxo functionality.[19,20,21]



1.

Figure 1.1.1 Amavadin, $(V^{IV}\text{-hidpa})_2)^{2-}$, hidpa=2,2´((hydroxyimino)diproponoic acid).

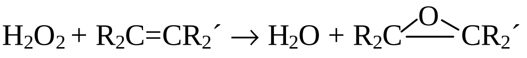
Amavadin and its derivatives are readily oxidized to afford diamagnetic vanadium (V) complexes, which are amenable to ^{51}V NMR spectroscopy. Oxidized amavadin $[\Delta\text{-}V^{V}\{(S,S)\text{-hidpa}\}_2]^{-}$ and $[\Delta\text{-}V^{V}(\text{hida})_2]^{-}$ compounds (H_3hida = 2,2´-hydroxyiminodiacetic acid) showed single ^{51}V resonances at δ -281 and -263, respectively.[22] Further studies on model amavadin complexes obtained with (R,S)-H_3hidpa revealed, after oxidation, three signals at $\delta(^{51}V) = -250, -270$, and -280, due to the presence of three diastereomers.[23] In a recent study on the self-exchange electron transfer in amavadin-type complexes synthesized from the racemic hidpa ligand, Lenhardt et al also reported three ^{51}V NMR signals for the oxidized form, at $\delta = -217, -234$, and -252,[24] again due to the presence of three diastereomers. All these δ values display an unusually low shielding of the ^{51}V nucleus, given that hydroxylamine ligands, much like peroxo moieties, generally induce high field shifts, up to $\delta = -860$ in oxo-vanadium complexes.[25,26] Whether the low shielding in the oxidized amavadin derivatives would originate from to a non-innocent nature of the specific hydroxylamido ligands is addressed in the present work for the parent hida complex and selected stereoisomers of hidpa derivatives, and the results are discussed in section 4.1.

1.2 Vanadium containing haloperoxidases

Vanadium-containing haloperoxidases (VHPOs) are efficient at catalyzing the oxidation of a number of halides, olefins, and organic sulfides using hydrogen peroxide as an oxygen source (see Scheme 1.2.1). [27,28,29,30,31] The controlled partial oxidation of such substrates to well-defined products is potentially very useful.

a) $H_2O_2 + X^- + H^+ \rightarrow H_2O + HOX$ oxidation of halide to hypohalous acid

b) $H_2O_2 + RSR' \rightarrow H_2O + RS(O)R'$ oxidation of sulfide to sulfoxide

c) 
$H_2O_2 + R_2C{=}CR_2' \rightarrow H_2O + R_2C\overset{O}{\frown}CR_2'$ oxidation of olefin to epoxide

Scheme 1.2.1 The oxidation of a) halide b) sulfide and c) olefin in the presence of hydrogen peroxide.

In the biosphere, VHPOs are thought to be responsible for the majority of halogenated natural products. These enzymes are especially abundant in the marine environment.[32] As the catalytic turnover of VHPO far exceeds any synthetic catalyst for halide oxidation known to date[30] this family of proteins has aroused interest for bio-inspired catalyst design.[33] The observed stability of VHPO's in the presence of high concentrations of strong oxidants or organic solvents and their stability at elevated temperatures make them attractive as industrial biocatalysts. Pharmaceutical applications have recently emerged based upon structural analogues of the active site in VHPO.[34] The pharmacotherapeutic activity of these model systems has prompted further research into the structure and function of these enzymes. Another potential commercial application of VHPO's is their use in environmentally friendly antifouling paints. A haloperoxidase has already been used in the coating on the outer walls of marine vehicles in order to prevent these walls from biofouling, making use of the natural presence of hydrogen peroxide in sea water.[35]

Experimentally VHPOs have been shown to contain vanadium in the oxidation state (V)[36] which is believed to be constant throughout the catalytic cycle.[37] The first stage of the catalytic cycle is thought to involve an initial proton transfer to one of the oxygen atoms directly bound to the vanadium.[38]

Two classes of vanadium haloperoxidase enzymes have been isolated, namely vanadium chloroperoxidase (VCPO) and vanadium bromoperoxiase (VBPO). Haloperoxidases are named after the most electronegative halide that they are able to oxidize; hence chloroperoxidases can oxidize chloride, bromide, or iodide, whereas bromoperoxidases are incapable of using chloride as a substrate. Both VCPO and VBPO bind vanadates formally derived from a VO_4^{3-} core as a prosthetic group.

Vanadium containing chloroperoxidase

The X-ray structure was reported for the vanadium dependent chloroperoxidase (VCPO) extracted from *Curularia Inaequalis* in the native form by Messerschmidt et al. in 1996.[39] The resolution of protein X-ray crystallography precludes definite conclusions on protonation states and hydrogen bonding interactions, which may be crucial for the action of this enzyme. The X-ray structure of the peroxo form of VCPO has also been resolved.[39]

A number of theoretical studies have investigated small vanadium complexes as models of the vanadium chloroperoxidase protein. Zampella et al.[40] performed a systematic survey of a large number of small active site mimics and found the doubly protonated monoanionic vanadate to be the most energetically stable using density functional theory (DFT) in the gas phase. A time dependent density functional study by Bangesh and Plass[41] also advocated the doubly protonated monoanionic vanadate to be the resting state. However the omission of the protein environment may lead to large geometric and/or electronic structure changes, which can qualitatively affect the results of these two studies.

There have been two previous QM/MM investigations of this protein (VCPO) to date. [42,43] Carlson and co-workers[42] identified the triply protonated, neutral vanadate moiety with an axially coordinated water and one hydroxyl group in the equatorial position as the lowest energy model and therefore considered it likely to be the resting state in the naturally occurring enzyme. They used a truncated protein model, designed to capture the electrostatic effects acting on the active site of VCPO. More recently Raugei and Carloni[43] reported on a series of CPMD/MM simulations. Free energy calculations and geometric similarity to the original X-ray structure of the native VCPO provided support for the doubly protonated, monoanionic vanadate form.

There now exists a multitude of studies on ^{51}V NMR spectra of vanadium-containing peptides recorded in solution. Butler et al. [44] reported a ^{51}V NMR spectrum of human transferrin (Tf) δ = -529.5 / -531.5 ppm. Rehder et al.[45] measured ^{51}V NMR chemical shifts of bovine apo-transferrin(Tf) δ = -515 ppm and of bovine prostatic acid phosphatase (Pp) δ= –542 ppm in solution. It is reasonable to assume that the ^{51}V NMR chemical shifts might be in a similar region for VCPO. On the other hand, an unusually strong shielding ($\delta \approx -930$ ppm) was found for a related vanadium-containing bromoperoxidase.[45]

A recent paper by Pooransingh-Margolis et al. [46] was a major impetus for the current study. The authors published the first experimental ^{51}V solid-state magic angle spinning (MAS) NMR spectrum of VCPO. They determined the isotropic chemical shift to be -507.5 ppm in the solid state (corrected for the second-order quadrupole shift), which is in fair agreement with the solution studies referred to above for other biological systems. The nuclear quadrupole coupling constant C_Q and asymmetry parameter η_Q of the electric field gradient (EFG) tensor were both determined along with the reduced anisotropy δ_σ and the asymmetry η_σ of the chemical shift anisotropy (CSA) tensor. Gas-phase DFT calculations were used to investigate a large number of small vanadium-containing complexes as models of VCPO. The theoretical work did not include the protein environment, but aimed instead at the identification of small model systems that would resemble the experimentally observed ^{51}V NMR chemical shifts.

It has been shown recently that ^{51}V chemical shifts can be computed with modern DFT methods and are quite often sensitive to structural details. [47] Herein chemical shifts are calculated from QM/MM models that incorporate the protein environment using a fully solvated and equilibrated system starting from the X-ray structure. The values refined from the experimental ^{51}V NMR spectrum [46] are used to evaluate the QM/MM models of the resting state of VCPO.

Computations of NMR chemical shifts in a QM/MM framework are becoming increasingly popular,[48] and the potential usefulness of such schemes for the treatment of a transition metal nucleus in a model for an active site of an enzyme has been demonstrated early on. [48a] To the best of our knowledge, the first application of such a QM/MM scheme to study NMR properties of a transition metal in a metalloenzyme are reported here. The results and elaborate discussions are given in section 4.2.

Vanadium containing bromoperoxidase

The native form of vanadium dependent bromoperoxidase (VBPO), extracted from *Ascophyllum nodosum*, was characterized by X-ray crystallography in 1999 (2.0 Å resolution).[49] Despite little similarity in the peptide backbones, the vanadium-binding sites of VBPO and VCPO show a high degree of homology[50], see Figure 1.2.1. The key difference on going from VCPO to VBPO is the replacement of a hydrophobic PHE residue close to the vanadate with a hydrophilic HIS residue. Also, VBPO is found to exist as a homodimer and is roughly twice the size of VCPO.

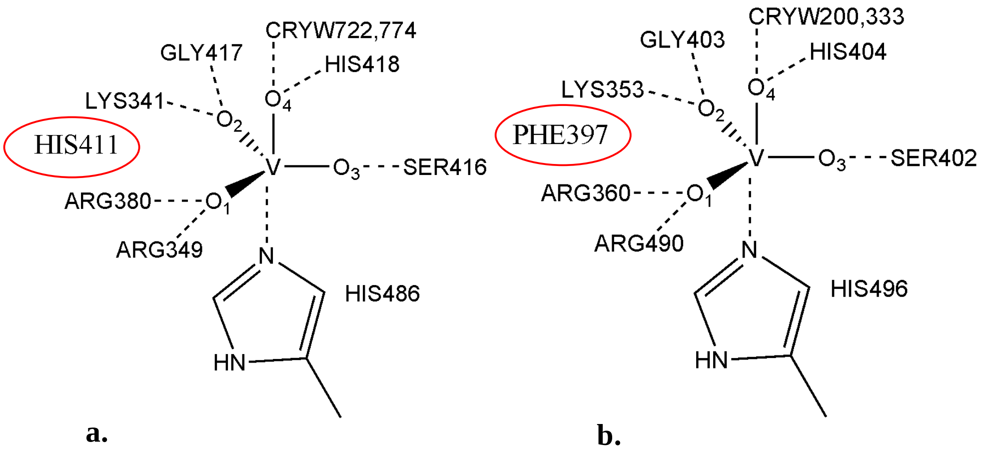


Figure 1.2.1 a. VBPO- Vanadate moiety and b. VCPO- Vanadate moiety with the key difference circled in red. The vanadium and four O atoms with the coordinated imidazole moiety, show the oxygen labeling used throughout this study and potential hydrogen bonding interactions with the protein environment (all labels are consistent with those in the 1QI9.pdb and 1IDQ.pdb file for VBPO and VCPO respectively).

The ^{51}V isotropic chemical shift for VBPO recorded by Rehder et al.[45] in solution is ca. -931ppm, and is some 400 ppm more shielded compared to typical vanadium complexes in solution, and also to the isotropic chemical shift extracted from the solid state for VCPO.[46] The reason for this large difference is certainly interesting giving the high degree of similarity in the active sites of these two enzymes. Preliminary studies on the solid-state MAS spectrum of VBPO[51] suggest an isotropic chemical shift of around -687 ppm. This isotropic value in the solid-state is therefore also significantly shielded with respect to that in solid VCPO, but only by ca. 170 ppm. The reasons for this apparent quantitative discrepancy between solution and solid-state ^{51}V NMR of VBPO is unclear at present.

Site directed mutagenesis was used by Wever et al.[52] to investigate the notable difference in the primary sequence between the active sites of VBPO and VCPO. They mutated PHE397 into a HIS effectively converting the VCPO to a VBPO active site mimic. They noted an increase in bromonation activity and a pH dependant inactivation of the enzyme. This clearly shows the residue has some important role in the relative activity of VBPO and VCPO. This primary structural difference is herein studied using QM/MM methods in combination with ^{51}V chemical shift calculations.

A bromine K-edge EXAFS study by Dau et al.[53] on VBPO from *Ascophyllum nodosum* proposed that the SER416 may be protonated during the catalytic cycle. A more recent study by Feiters et al.[54] also suggested some possible bromination within the active site. To investigate the importance of the brominated SER402, Wever et al. performed site directed mutagenesis (SER402ALA) studies on the VCPO and found that the extent of this mutation was only marginal. The VBPO protein resolved by X-ray crystallography was found to crystallize as a homodimer.[49] Whether the two different active sites are indeed fully isostructural, also in terms of protonation state and H-bond network, is an important question that has yet to be addressed.

To date, there appears to be no explicit computational studies of VBPO in the literature. In contrast there exists quite a substantial body of literature for the VCPO enzyme. Due to the high homology of the active sites of the two enzymes, conclusions based on VCPO are often transferred to the VBPO enzyme and the remainder of the VHPO family. Zampella et al. published a gas-phase DFT study on the reactivity of the peroxo form of VCPO.[55] The study of the peroxo form does not include the primary difference between VCPO and VBPO, namely the distal HIS. However, the experimental observation that different members of the VHPO family can oxidize halogens of different electronegativities does suggest that some (perhaps subtle) differences are present. Furthermore the experimental EXAFS and site directed mutagenesis studies above further discourage the transferability of conclusions from VBPO and VCPO.

There have been no QM/MM studies of VBPO to date, presumably as the homodimer is significantly larger in size than the closely related VCPO. Such study is now presented in section 4.3, calling special attention to the computation of ^{51}V NMR chemical shifts and a detailed comparison of both geometries and chemical shifts between VBPO and VCPO .

The peroxo intermediate

Vanadium peroxides have the potential to generate a number of biological and biochemical responses[56], and are utilized as insulin-mimetic agents in the treatment of human diabetes.[57] The X-ray-derived structure of the peroxide-form of VCPO (2.24 Å resolution),[39] denoted *p*-VCPO, reveals a distorted tetragonal bipyramidal vanadium site, see Figure 1.2.2. The peroxide ligand is coordinated in a side-on manner, and the apical oxygen ligand is no longer present. In addition to the two peroxo oxygen atoms bound to vanadium, two equatorial oxo ligands and a nitrogen ($N^{\varepsilon 2}$) from HIS496 in the basal plane completes the coordination scheme. One of the peroxo oxygens O_4, is in a pseudo-axial position, and the other is in an equatorial position, therefore giving an overall distorted pyramidal coordination geometry. One of the equatorial oxo ligands is hydrogen-bonded to ARG490, the coordinated peroxide is hydrogen-bonded to the neighboring glycine amide backbone and also to LYS353, and the remaining oxo ligand is hydrogen-bonded to ARG360, see Figure 1.2.2. It has been shown that the VCPO protein has a higher affinity for peroxovanadate than for vanadate.[58]

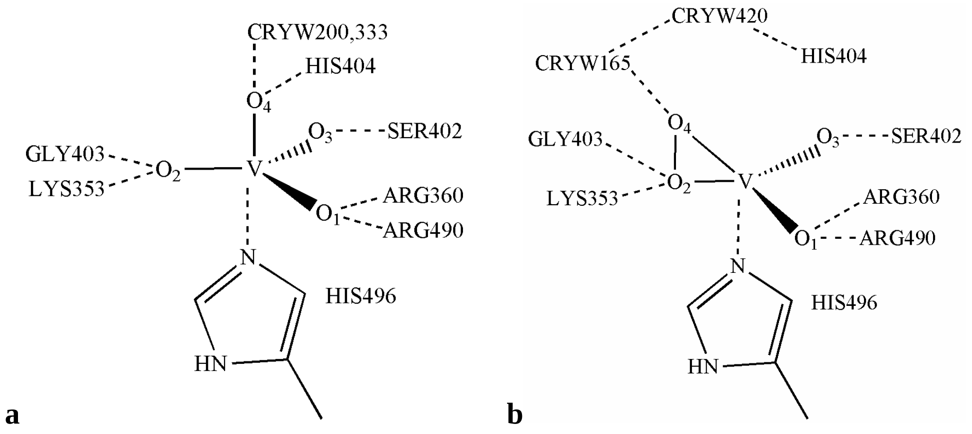
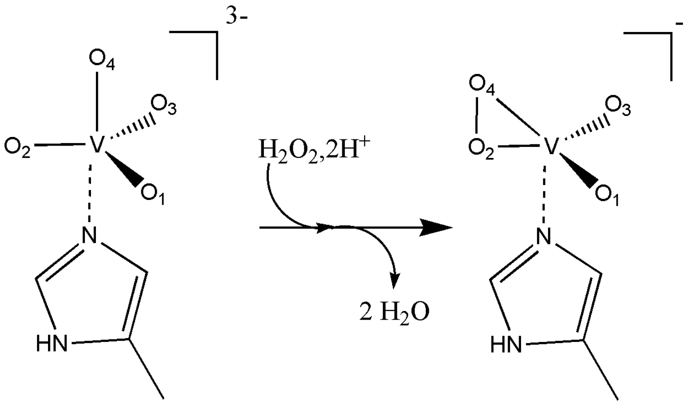


Figure 1.2.2 a. Vanadate moiety and coordinated imidazole in native VCPO. b. Peroxovanadate moiety and coordinated imidazole. The labels used throughout this study for potential hydrogen bonding interactions with the protein environment are consistent with those in the 1IDU.pdb file. The oxygen labeling scheme of the vanadate has been edited to be consistent with the *p*-VCPO labeling scheme within the 1IDU.pdb file.

A possible pathway for the formation of *p*-VCPO from VCPO has been outlined by Zampella et al.:[55] Assuming the axial ligand in native VCPO (O_4 in Figure 1.2.2a) is present as OH^-, which is made more basic by its H-bond to HIS404, this OH^- unit would deprotonate the approaching H_2O_2, thereby generating a HOO^- species. The weakly ligated axial water molecule dissociates from the vanadate and a side-on bound peroxide intermediate is formed after the departure of another water molecule, see Scheme 1.2.2. Furthermore, Zampella et al.[55] suggested that an attack of a chloride ion on one of the peroxo atoms, and the uptake of a proton from one of the surrounding water molecules, leads to the generation of the hypochlorous acid (HOCl) and the restoration of the native state. Wever et al.[59] has shown that the vanadium ion plays the role of a strong Lewis acid which activates the peroxide, as seen by its inability to undergo redox cycling during catalysis. Kinetic studies showed that the protonation of the bound peroxo group is a crucial step in the heterolytic cleavage of the O-O bond.[60] These observations prompted the proposal of the protonated oxygen being transferred to the halide according to an oxo-transfer mechanism. A subsequent computational study disagrees, [61] instead suggesting that the non-protonated peroxo oxygen is the atom transferred to the substrate.



Scheme 1.2.2 Proposed formation of *p*-VHPOs. The actual protonation states of vanadate and peroxovanadate moieties are not known with certainty.

The protonation state of the peroxovandate cofactor affects its reactivity and this might be a crucial factor in tuning the selectivity profile of these enzymes.[62,63] Determining the protonation state of the intermediate peroxo-forms of the VHPO enzymes is experimentally challenging, as a VHPO enzyme shuttles between a trigonal bipyramidal structure (native-form) and a distorted tetragonal structure (peroxo-intermediate) during the catalytic cycle. The X-ray structure solved by Messerschmidt et al.[39] failed to resolve the protonation state of the peroxovanadate active site within *p*-VCPO. Furthermore, there currently exists no X-ray crystal structure for the peroxo-form of VBPO (hereafter labeled *p*-VBPO). A gas-phase DFT study on the active site of VCPO[55] addressed the protonation state of the intermediate peroxo-form of the enzyme, suggesting a singly protonated species. Protonation of the peroxovanadate was previously suggested[64] to be a possible explanation for the difference between VCPO and VBPO, therefore further studies that incorporate the intrinsic environmental differences might be necessary.

Wever et al.58 used stopped-flow UV-vis spectroscopy to investigate the formation of the peroxo-form of VCPO from the native-form after the addition of H_2O_2. The native-form had a characteristic peak at 315nm and upon formation of the peroxo-form of VCPO a peak appeared at 385nm. The pH dependence on the activity was also investigated experimentally within this study, revealing maximum stability at a pH value of 8.3, while maximum activity was observed at pH 5.0. These findings highlight the importance of understanding the protonation scheme within the active site.

The very center of the active site has been probed with ^{51}V NMR spectroscopy. The ^{51}V isotropic chemical shift for VBPO recorded by Rehder et al.[45] in solution is -931ppm for the native form, and -1135ppm for the peroxo-form . Interestingly, the peroxo-resonance is shifted upfield by ca. -200ppm from that of the native-form, in line with observations for small vanadates and peroxovanadates derived thereof. The experimental isotropic chemical shift of VPBO in the native-form is 400ppm more shielded compared to typical vanadium complexes in solution, in particular VCPO, and also to the isotropic chemical shift extracted from the solid-state for VCPO.[46]

Quantum-chemical computations can be an important structural tool for investigating the active sites of vanadium haloperoxidases. Zampella et al.[55] performed an active-site model study, using DFT, for both the native and peroxo-form of the VHPOs and considered the reaction energies combined with TD-DFT results when suggesting a singly protonated peroxovanadate as being the most likely candidate for the intermediate peroxo-form, as it had the best agreement with the UV-vis spectral band at 385nm. Identifying which of the two equatorial oxo ligands would be protonated within VHPO, either O^1 or O^3, is a question that remains open at present.

The resolution of the solid-state structure of *p*-VCPO is not high enough to locate hydrogen atoms, and the uncertainty in the atom positions of heavier atoms makes it very speculative to differentiate between oxo and hydroxy bond lengths for the vanadate cofactor. Apparently, the crystallization of the VBPO enzyme is as yet unsuccessful for the peroxo-intermediate. In view of these current limitations in the experimental data, computational methods were applied to probe and assess structural differences between these proteins, differences that may have important implications for their reactivities. The results are summarized in section 4.4, where protonation states of the intermediate peroxo-form of the VHPOs are critically compared to the previously investigated native-forms.

Chapter 2

2 Theoretical background

Since NMR properties can be very sensitive to the molecular structure, accurate geometries need to be used as input in the calculations. One choice, in order to obtain accurate geometries, is to use structures optimized at appropriate quantum-mechanical levels. However, such methods scale unfavourably with the size of the system, so that they become very CPU-intensive, if not impossible at all, for large systems. A popular way to overcome this limitation is to use hybrid QM/MM theory, a new class of emerging method that combines the advantages of QM and MM calculations, namely accuracy and speed, respectively. In this approach, a small part of the system is treated quantum-mechanically (typically active-site of an enzyme) and the remaining system is treated classically.[65] For systems containing transition metals, density functional theory (DFT) is the QM method of choice. As the systems get larger, additional problems arise due to the very large number of degrees of freedom. Frequently the systems are characterized by a conformational flexibility that is reflected in many local minima, which makes it difficult to find the relevant ones (including the global minimum). The molecular dynamics technique is one strategy that allows detailed sampling of a representative phase space. The present chapter briefly outlines the formal foundations on which computational methods of this thesis are based, providing descriptions of the molecular dynamics simulations, the specific QM/MM methodology, DFT in static and time-dependent forms, as well as molecular properties.

2.1 Molecular Dynamics

Molecular dynamics (MD) uses computed forces to propagate the atomic positions in time applying the laws of Newtonian (as opposed to quantum) mechanics. The simulation is performed by numerically integrating Newton's equations of motion over small time steps (usually 10-15 secs or 1 fsec). The Verlet algorithm is used to compute the velocities of the atoms from the forces and atom locations. Once the velocities are computed, new atom locations and the temperature of the assembly can be calculated. These values then are used to calculate trajectories, or time dependent locations, for each atom. Over a period of time, these values can be stored on disk and played back after the simulation has completed to produce a "movie" of the dynamic nature of the molecule.

The simplest way to compute the necessary forces is by way of a classical empirical force field (see sec. 2.2). With this approach, MD simulations have been used in a variety of biomolecular applications.[66] The technique, when combined with data derived from Nuclear Magnetic Resonance (NMR) studies, has been used to derive 3D structures for peptides and small proteins in cases where X-ray crystallography was not practical.[67] Additionally, structural, dynamic and thermodynamic data from molecular dynamics has provided insights into the structure-function relationships, binding affinities, mobility and stability of proteins, nucleic acids and other macromolecules that cannot be obtained from static models. MD can be used to disturb an optimized structure to generate a new starting point for an optimization, in order to discover multiple minima for molecules with a complex energy surface. The time evolution gives us the time dependence of molecular motion and so allows access to transport properties, such as the self diffusion coefficient. Molecular dynamics can be used to supply configurations for averaging properties (i.e. chemical shifts) as an alternative to Monte Carlo.

To set the momenta we need to give each atom an initial velocity such that the overall kinetic energy of the system is consistent with the desired temperature. From the Boltzmann distribution we know that the expectation value for the molecular velocity is:

Eqn 2.1.1

$$\left\langle v_i^2 \right\rangle = \frac{kT}{m}$$

The temperature of the system gives the atomic speeds to use but not their direction. The directions can be chosen at random but with the proviso that the system as a whole has zero momentum. This ensures that the entire system does not drift off during the simulation.

Molecular dynamics is a simulation technique that uses Newton's Laws of motion to integrate the trajectory of atoms in a simulation. The forces on the atoms may be derived from a force field or from quantum mechanics but the method of generating the trajectory will be the same.

In a simulation we calculate the force $\underline{F}$ on any atom from the geometric arrangement of the atom centers, and, knowing the atom mass, *m,* this equation is used to find out the acceleration, $\underline{a}$, on a given atom:

Eqn. 2.1.2 $$\underline{a} = \frac{\underline{F}}{m}$$

The acceleration is the rate of change of velocity with time, i.e.:

Eqn. 2.1.3 $$\frac{d\underline{v}}{dt} = \frac{\underline{F}}{m}$$

If a particle has an initial velocity, $\underline{u}$, and moves under the action of this force for a time, τ, its velocity will be given by integrating this equation:

Eqn. 2.1.4 $$\underline{v} = \underline{u} + \int_0^{\tau} \frac{d\underline{v}}{dt} dt = \underline{u} + \int_0^{\tau} \frac{\underline{F}}{m} dt$$

Provided the force on the atom does not change during the time of the integration, $\underline{F}$ will be constant and so:

Eqn. 2.1.5 $$\underline{v} = \underline{u} + \frac{\underline{F}}{m}\tau$$

The velocity itself is a rate of change of atom position, $\underline{s}$, with time,

Eqn. 2.1.6 $$\underline{v} = \frac{d\underline{s}}{dt}$$

and so the position of the particle after a certain time will again be an integration:

Eqn. 2.1.7 $$\underline{s} = \underline{s}_o + \int_0^{\tau} \frac{d\underline{s}}{dt} dt$$

Remembering that the velocity is always changing even under uniform acceleration:

Eqn. 2.1.8 $$\underline{s} = \underline{s}_o + \int_0^{\tau} \left(\underline{u} + \frac{\underline{F}}{m} t \right) dt = \underline{s}_o + \underline{u}\tau + \frac{1}{2}\frac{\underline{F}}{m}\tau^2$$

So we work out the forces on a given atom from its disposition with respect to all other atoms in the simulation. This allows us to work out the acceleration on that atom, then its velocity and finally its position after a time interval. The forces can also be computed quantum mechanically with the molecular dynamics techniques such as Born-Oppenheimer (BOMD) or Car-Parrinello MD (CPMD).[5,68] However, such techniques have the restriction of quite limited simulation times.

2.2 Molecular Mechanics

Molecular mechanics is a mathematical formalism which attempts to reproduce molecular geometries, energies and other features by adjusting bond lengths, bond angles and torsion angles to equilibrium values that are dependent on the hybridization of an atom and its bonding scheme (this atom description is referred to as the atom type). Rather than utilizing quantum physics, the method relies on the laws of classical Newtonian physics and experimentally derived parameters to calculate geometry as a function of steric energy. The general form of the force field equation is

Eqn 2.2.1 $$E_{pot} = \Sigma\, E_{bnd} + \Sigma\, E_{ang} + \Sigma\, E_{tor} + \Sigma\, E_{oop} + \Sigma\, E_{nb} + \Sigma\, E_{el}$$

E_{pot} is the total steric energy which is defined as the difference in energy between a real molecule and an ideal molecule. E_{bnd}, the energy resulting from deforming a bond length from its natural value, is calculated using Hooke's equation for the deformation of a spring ($E = 1/2\ K_b(b - b_0)2$ where K_b is the force constant for the bond, b_0 is the equilibrium bond length and b is the current bond length). E_{ang}, the energy resulting from deforming a bond angle from its natural value, is also calculated from Hooke's Law. E_{tor} is the energy which results from deforming the torsion or dihedral angle. E_{oop} is the out-of-plane bending component of the steric energy. E_{nb} is the energy arising from non-bonded interactions and E_{el} is the energy arising from coulombic forces.

When the terms shown in the general form of the force field are expanded, the equation becomes

Eqn 2.2.1

$$E_{pot} = \Sigma 1/2\ K_b(b - b_0)^2 + \Sigma\ 1/2\ K_\theta(\theta - \theta_0)^2 + \Sigma\ 1/2\ K_\Phi\ (1 + \cos N\Phi)^2$$

$$+ \Sigma\ 1/2\ K\chi(\chi - \chi_0)^2\ \ \Sigma\ ((B/r)^{1/12} - (A/r)^{1/6}) + \Sigma\ (qq/r)$$

The manner in which these terms are utilized to build a model is referred to as the functional form of the force field. The force constants K_b, K_θ, K_Φ, $K\chi$ and equilibrium values b_0, θ_0, K_Φ, $K\chi$ are atomic parameters which are experimentally derived from X-ray, NMR, IR, microwave, Raman spectroscopy, thermochemical (calorimetric data) and ab initio calculations on a given class of molecules (alkanes, alcohols, etc). The energy of the atoms in a molecule is calculated and minimized using a variety of directional derivative techniques.

In contrast to ab initio methods, molecular mechanics is used to compute molecular properties which do not depend on electronic effects. These include geometry, rotational barriers, vibrational spectra, heat of formation and the relative stability of conformers. Since the calculations are fast and efficient, molecular mechanics can be used to examine systems containing thousands of atoms and to explore their conformational flexibility (conformational search).[69]. What is inherently difficult to describe with this technique is the breaking and formation of chemical bonds. Also, unlike ab initio methods, molecular mechanics relies on experimentally derived parameters so that calculations on new molecular structures may be misleading.

2.3 Quantum mechanics:

One-electron systems

Electrons display both wavelike and particle behaviour, and it is necessary to describe their motions and distributions using a probabilistic theory (quantum mechanics).

At the heart of this 'wave mechanics' is the state function or "wavefunction", $\psi(\mathbf{r},t)$, associated with an electron. This function describes the time evolution of the electron's spatial distributions and also its motion (momentum and/or kinetic energy). In principle it also depends on spin of the electron; but since the calculations presented in this thesis concern solely closed-shell, non-radical species, we will refer to the spin-free formulation of the background theory.

At low (non-relativistic) speeds, the wavefunction of the electron evolves according to the time-dependent Schrödinger equation

Eqn 2.3.1
$$\hat{H}(\mathbf{r},t)\ \psi(\mathbf{r},t) = i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t}$$

where $\hat{H}$ is the Hamilton operator. If the Hamiltonian is independent of time then the equation simplifies to an *eigenvalue equation*:

Eqn 2.3.2
$$\hat{H}(\mathbf{r})\ \psi_n(\mathbf{r}) = E_n\ \psi_n(\mathbf{r})$$

Here the different electronic states are labelled by the subscript n; the lowest energy state is the *ground state* (all subsequent calculations presented are ground state calculations). One of the most useful properties which may be obtained from the wavefunction is the *electron density* (often also called the *charge density*), and in this one-electron case it is given by $\rho(\mathbf{r}) = \psi(\mathbf{r})^*\psi(\mathbf{r}) = \psi(\mathbf{r})^2$ if the wavefunction is real. $\rho(\mathbf{r})d\mathbf{r}$ represents the probability of finding the electron in the volume element $d\mathbf{r}$. The commonest method for estimating ground state properties is to write a trial wavefunction which is a linear combination of atomic orbitals (the LCAO approximation).

Eqn 2.3.3
$$\psi = \sum_{i=1}^{N} a_i \phi_i$$

Then we can apply the variational principle to obtain the ground-state wavefunction with the associated energy E_0

Eqn 2.3.4
$$\frac{\int \psi \; \hat{H} \; \psi \; d\mathbf{r}}{\int \psi^2 d\mathbf{r}} = E_0$$

Many-electron systems

In a molecule with N electrons and M nuclei, the Schrodinger equation becomes

Eqn 2.3.5
$$\hat{H}(\{\mathbf{r}\},\{\mathbf{R}\})\,\Psi_n(\{\mathbf{r}\},\{\mathbf{R}\}) \;=\; E_n\,\Psi_n(\{\mathbf{r}\},\{\mathbf{R}\})$$

Where **R** is a nuclear coordinate. For almost any system of interest, this equation is intractable and a number of approximations are introduced. The most fundamental of these is where the Born-Oppenheimer approximation, therefore the total wavefunction is written as a product of *electronic* and *nuclear* wavefunctions:

Eqn 2.3.6
$$\Psi_{nm}(\{\mathbf{r}\},\{\mathbf{R}\}) \;=\; \Phi_n(\{\mathbf{r};\mathbf{R}\})\,\chi_{nm}(\{\mathbf{R}\})$$

Where Φ_n is the electronic wavefunction and χ_{nm} is the nuclear wavefunction.
The electronic wavefunction is therefore only parametrically dependant on the nuclear coordinates and is an eigenfunction of an electronic Schrödinger equation

Eqn 2.3.7
$$\hat{H}_{el}(\{\mathbf{r};\mathbf{R}\})\,\Phi_n(\{\mathbf{r};\mathbf{R}\}) \;=\; E_n(\{\mathbf{R}\})\,\Phi_n(\{\mathbf{r};\mathbf{R}\})$$

Note that the eigenvalue of the electronic equation is a function of the nuclear coordinates – this is usually called the potential energy surface (PES). Focusing on the electronic wavefunction, the electron density is now obtained by integrating (in effect averaging) over all electronic coordinates except one.

Eqn 2.3.8 $$\rho(\mathbf{r}) = \int \Phi_n(\{\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3 \ldots \mathbf{r}_n\};\{\mathbf{R}\})^* \, \Phi_n(\{\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3 \ldots \mathbf{r}_n\};\{\mathbf{R}\}) \, d\mathbf{r}_2 d\mathbf{r}_3 \ldots . d\mathbf{r}_n$$

The forms of the electron kinetic energy, electron-nuclear attraction energy and electron-electron repulsion energy operators in the electronic Hamiltonian are as follows (in a.u.):

Eqn 2.3.9 $$\hat{H}_{el} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee}$$

where

$$\hat{T}_e = \sum_i -\frac{1}{2}\nabla^2{}_{e,i}$$

$$\hat{V}_{ne} = \sum_i \sum_j -\frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}$$

$$\hat{V}_{ee} = \sum_i \sum_{j>i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

$\hat{T}_e$ is the electronic kinetic energy operator; $\hat{V}_{ne}$ is the nuclear-electron potential operator, $\hat{V}_{ee}$ is the electron-electron potential operator, Z_j is the charge of the j-th nucleus, $\mathbf{R}_j$ is the position of the j-th nucleus, $\mathbf{r}_i$ is the position vector of the i-th electron, and

$$\nabla^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)$$

2.4 Density Functional Theory

Once the Born-Oppenheimer approximation is applied, the electronic part of the total energy

$E = E_e + E_{nn}$ can be related to the electron density $\rho(\mathbf{r})$ of a system of electrons and nuclei using density functional theory (DFT):[70]

Eqn 2.4.1 $$E_e = V_{ne} + T_e + V_{ee}$$

Where: V_{ne} is the electron-nuclear energy

T_e is the electronic kinetic energy

V_{ee} is the electron-electron interaction energy.

The potential energy terms in DFT are given by

Eqn 2.4.2 $$\mathrm{V}_{\mathrm{ne}} = \int \rho(\mathbf{r})\, \hat{V}_{ne}\, d\mathbf{r}$$

Eqn 2.4.3 $$\mathrm{V}_{\mathrm{ee}} = \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + \int \rho(\mathbf{r})\hat{V}_{xc}[\rho(\mathbf{r})]\, d\mathbf{r}$$

where the electron-nuclear operator $\hat{V}_{ne}$ was already defined in equation 1.19 and $\hat{V}_{xc}$ is the exchange-correlation potential. In other words, the potential energy can in principle be expressed as a function of the density, providing we know the form of the exchange-correlation potential $V_{xc}[\rho(\mathbf{r})]$. Only crude approximations to the kinetic energy functional $T_e[\rho(\mathbf{r})]$ are known. Modern density functional theory circumvents this problem by employing Kohn-Sham[17] theory outlined below.

Kohn-Sham Theory

In the Kohn-Sham formalism,[71] the total energy functional is partitioned in the following way:

Eqn 2.4.4 $$E_e = T_{ni} + V_{ne} + V_{ee} + \Delta T + \Delta V_{ee}$$

where T_{ni} is the kinetic energy of a system of non-interacting electrons, and ΔT is the correction to the kinetic energy arising from instantaneous electron-electron interactions.

Similarly ΔV_{ee} is the exchange energy plus that part of the electron-electron interaction energy due to their instantaneous interactions (i.e. exchange plus the so-called electron correlation energy). In DFT the terms $\Delta V_{ee,}$ and ΔT are lumped together in a single term that is known as the exchange-correlation functional. Expressing the density as a linear combination of molecular orbital products

Eqn 2.4.5 $$\rho(\mathbf{r}) \;=\; \sum_{j} n_j \, \phi_j(r)\, \phi_j{}^*(r)$$

the Kohn-Sham orbital total energy expressed in terms of these orbitals is

Eqn 2.4.6

$$E\,[\rho(\mathbf{r})] \;=\; \sum_{i}^{N} \left\langle \phi_i \middle| -\frac{1}{2}\nabla_i^{\,2} \middle| \phi_i \right\rangle \;-\; \sum_{k}^{\text{nuclei}} \left\langle \phi_i \middle| \frac{Z_k}{|\mathbf{r}_i - \mathbf{r}_k|} \middle| \phi_i \right\rangle \;+\; \frac{1}{2} \int \frac{\rho(\mathbf{r}_i)\rho(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|}\, d\mathbf{r}_i\, d\mathbf{r}_i \;+\; \int \rho(\mathbf{r})\, V_{xc}\,[\rho(\mathbf{r})]\, d\mathbf{r}$$

The exact form of the exchange-correlation functional is unknown and has been the source of much research; a large number have been proposed. Minimisation of E[ρ] is achieved in the usual way by expanding the molecular orbitals in an LCAO basis and determining the coefficients variationally. This results in a pseudo one-electron equation (the Kohn-Sham equation) for the orbitals with eigenvalues ε_i, which is analogous to the Hartree Fock(HF) SCF equation:

Eqn 2.4.7 $$[\; -\frac{1}{2}\nabla^2 \;+\; \hat{V}_{\text{ne}}(\mathbf{r}) \;+\; \int \frac{\rho(\mathbf{r}')}{|r - r'|}\text{dr}' \;+\; \hat{V}_{\text{xc}}(\mathbf{r}) \;]\; \phi_{\text{i}}(\mathbf{r}) \;=\; \varepsilon_{\text{i}}\, \phi_{\text{i}}(\mathbf{r})$$

These equations are solved iteratively using methods similar to those applied in HF calculations (the difference is the requirement for numerical integration methods for evaluating the complicated exchange-correlation potential terms).[72]

Density functional theory in its original Hohenberg-Kohn context is essentially a ground state theory and not applicable a priori to excited states.

2.5 Time-Dependent Density Functional Theory

TDDFT is a quantum mechanical method used to investigate the properties of many-body systems beyond the electronic ground state structure. TDDFT extends the concept of stationary DFT to time-dependent situations by describing the interaction of a quantum many-particle system with a given time-dependent potential, e.g. a periodic electromagnetic field. Using additional simplifications in the linear response regime,[73] the main area of TDDFT applications till now has been the calculation of energies of electronically excited states. The excitation energies can be computed as the poles of the response function of the system which can be computed using a Dyson equation. The key ingredients become the Kohn-Sham non-interacting response function and the Hartree plus exchange-correlation kernel, which is the functional derivative of the effective potential with respect to the density.

In TDDFT, excitation energies are essentially computed from ground-state densities, without recourse to the actual excited states (which are not covered by the original Hohenberg-Kohn theorems). In many cases TDDFT rivals sophisticated MR based *ab* initio methods in terms of accuracy of the excited energies. Problems can arise with certain types of excitations, e.g. involving high-lying diffuse orbitals (Rydberg states) and, most notably, for long-range CT states, the energies of which are notoriously underestimated with present day Xc functionals.[74]

2.6 Basis Sets

In LCAO orbital methods the orbitals are expanded using predefined basis set -a collection of mathematical functions used to model the spatial variation of the orbitals. Historically Slater-type orbitals were used due to their geometric similarity to the hydrogenic orbitals. However they do have a significant drawback in that there are no analytical solutions for the many-centre integrals involving Slater-type orbitals. Consequently most practitioners of quantum chemistry turned to Gaussian type orbitals.

The general form for a normalised Gaussian-type orbital (GTO) within a cartesian coordinate system is

Eqn 2.6.1

$$\phi(x,y,z:\alpha,i,j,k)=\left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}}\left[\frac{(8\alpha)^{i+j+k}\,i!\,j!\,k!}{(2i)!(2j)!(2k)!}\right]^{\frac{1}{2}}x^{i}y^{j}z^{k}e^{-\alpha(r^{2})}$$

GTO's have an atom-centered gaussian form (*i.e.* decay exponentially with the square of the distance from the atom). A major problem with GTOs is the dissimilarity to hydrogenic orbitals in the radial portion of the function: they have no "cusp" at the nucleus. This issue is partly resolved by using a fixed linear combination of GTO's (known as primitives) to better approximate the shape of an STO closer to the nucleus (it doesn't rectify the problem of there being no cusp actually *at* the nucleus). A basis function that is composed of a fixed linear combination of primitive Gaussians is commonly known as a contracted basis function. These effectively combine the computational advantage of the Gaussian functions while providing a good approximation to the shapes of Slater-type orbitals.

Eqn 2.6.2

$$\phi(x,y,z:\{\alpha\},i,j,k)=\sum_{a=1}^{M}C_{a}\phi(x,y,z,\alpha_{a},i,j,k)$$

Larger basis sets give rise to more accurate representations of the MOs. It is mathematically evident that an infinite basis is required for exact solutions of the Schrödinger equation. However, practical considerations have dictated the need for optimisation of finite compact basis sets to best described molecular properties within sensible computational expense. This has led to the development of various basis sets each with differing characteristics reflecting their intended applications. The use of such optimised finite basis sets does produce reasonably accurate geometries (as basis sets do not strongly influence structure). However energies are much more basis set dependent.

Minimal basis sets (e.g. STO-3G[75]) assign a single basis function to each atomic orbital in the ground state. This provides a fairly inflexible description of the molecular orbitals, particularly with regard to expansion/contraction of the function as a response to differing molecular environment.

Extended basis sets have been developed with the addition of extra functions for all orbitals which would be occupied in the ground state of the associated atom (or in some cases only valence orbitals are assigned these extra functions). Two or more contracted GTOs per atomic orbital, as in the so-called double-zeta basis sets[76], provide much greater flexibility. Split-valance basis sets partitions the atomic orbitals into core and valence regions. The core AO's are assigned a minimal basis whilst the valence orbitals are described at the double-zeta level. For example the 6-31G basis set[77] describes the core AO's with a single contracted Gaussian composed of primitive Gaussian functions. The valence AO's are described using two functions: a contracted GTO with 3 primitives and a separate single diffuse GTO.

Polarisation functions are frequently added to improve the flexibility: this is a function with higher angular momentum than is found in the occupied ground state atomic orbitals. The basis set 6-31G*[78a] adds 5 or 6 *d*-type GTOs to first and second-row atoms; the basis set 6-31G**[78b,c] also adds three *p*-type GTO's to hydrogen atoms.

2.7 QM/MM methodology

The basic idea is to identify a small region, the Inner Region in Figure 2.7.1 to which the higher (QM) level of theory is to be applied. As in an extended covalent system, the excision of this region involves bonds, it will be necessary to compute the QM wavefunction of the Inner Region in the presence of at least a layer of atoms which serve to provide a satisfactory termination for the QM calculation.

In many methods, traditionally those used for covalent systems, a number of additional atoms, which have no counterparts in the real system, are added to the QM region to terminate the broken bonds, these are typically called link atoms. These are hereafter referred to as link-atom methods**.** When handling solid-state systems, the models, often referred to as embedding models, handle the termination of the QM region by defining a boundary region in which the atoms have some quantum character, as well as being described fully by the classical force field. No new centres are needed. These models are called boundary region methods.

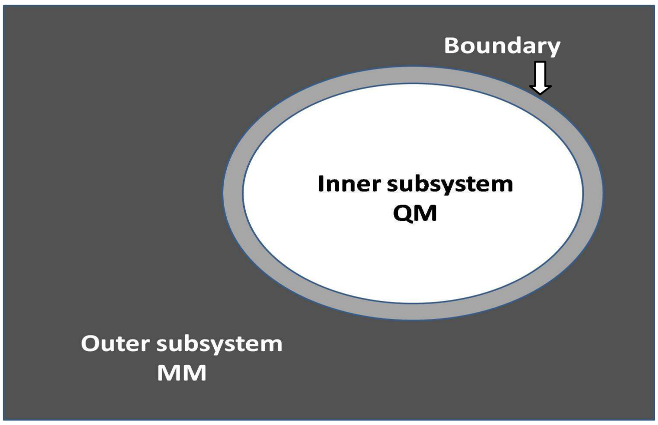


Figure 2.7.1 Inner, outer and boundary regions

QM/MM schemes may also be classified by the nature of the total energy expression. At one extreme, a QM/MM energy expression can be obtained by adding the QM and MM contributions, together with the appropriate coupling terms

Eqn 2.7.1. $$E = E(\text{Inner,QM}) + E(\text{Outer, MM}) + E(\text{Couple, QM/MM})$$

E(Couple,QM/MM) is a QM/MM coupling term that includes all interactions between the two regions, for example, classically handled bonding and van der Waals interactions and modifications to the QM Hamiltonian to reflect the electrostatic influence of some or all of the atoms in the outer region (*vide infra*).

This approach is referred as the additive QM/MM model. The method can be applied easily to systems where there are no boundary atoms (as in studies of solvation) but as presented above the approach takes no account of the handling of link atoms at the interface.

Coupling of the QM and MM regions

The QM and MM regions are coupled by bonded and non-bonded interactions. The bonded terms are generally handled by terms in the classical force field or (for boundary region methods) by a combination of quantum and classical terms between QM and boundary atoms. The non-bonded terms comprise van der Waals and electrostatic interactions. The former are handled in the same way as they are in the MM region (perhaps with modified parameters). Bakowies and Thiel[79] defined three ways of treating the QM/MM electrostatic interaction, labelled A–C, as follows.

A. ***Mechanical embedding:*** The QM calculation is essentially performed in the gas phase, without electronic coupling to the environment. The electrostatic interaction between QM and MM regions is either omitted or performed by the MM code, using a classical point charge model for the QM charge distribution (e.g. a potential derived charge model).

B. ***Electrostatic embedding:*** The classical partition appears as an external charge distribution (e.g. a set of point charges) in the QM Hamiltonian. The polarisation of the QM region by the MM charge distribution thus occurs as part of the QM electronic structure calculation. The partial charges used to describe the MM distribution are frequently taken to be those used in the force field [80,81,82], relying on the use of electrostatic properties in the force field charge derivation. In ab initio schemes it is clear that the electrostatic embedding scheme should be implemented, at least at long range, by adding the contribution of the MM point charges to the 1-electron Hamiltonian.

C. ***Polarised embedding*:** The polarisation of the MM region in response to the QM charge distribution is also included. Intuitively this makes most sense when the force field incorporates polarisation as unpolarised force-fields implicitly incorporate MM polarisation in their parameterisation, and care must be taken to ensure such implicit contributions do not occur in the QM/MM potential. A variety of models for classical polarisation are possible, including the shell model [83], and coupled distributed atomic polarizabilities.[84] Polarisation of MM atoms close to the QM region (e.g. those connected by link-atom terminated bonds) was found to be unphysically large, leading to the suggestion that these atoms be treated as unpolarizable.[79]

2.8 NMR chemical shifts

Nuclear magnetic resonance (NMR) is an extremely popular experimental technique. The information that can be extracted from such spectra is ever increasing. Structural information such as distances and angles are now routinely extracted from elegant experiments on systems ranging in size up to large bio-molecules. Coordination numbers, bond types and even some insight into the electronic structure can be interpreted from NMR spectra. Evolution in the field of quantum chemistry has made possible calculations of useful accuracy and reliability on fragments of biomolecules and transition metal complexes, containing dozens of atoms.[85] Here, a brief explanation of the chemical shifts is given.

NMR active nuclei have a non-zero spin I, which can take 2I+1 orientations. In an external magnetic field the degeneracy of these orientations is lifted due to the splitting of the energy levels. For I=1/2 the energy splitting is :

Eqn 2.8.1 $$\Delta E = \frac{ghB}{2\pi}$$

Where g is the gyrometric ratio, an intrinsic constant of every nucleus, and h is the Planck constant. B is the magnetic field at the nuclei.

The effective magnetic field at the nucleus is not equal to the applied magnetic field since the latter is shielded by the electrons moving around the nucleus. The resulting effective field is :

Eqn 2.8.2 $$\vec{B}^{eff} = \vec{B}^{ext}(1-\vec{\sigma})$$

Where $\vec{\sigma}$ is the nuclear shielding tensor, which in the isotropic case simplifies to the isotropic shielding constant Tr$\vec{\sigma}$/3. The shielding tensor depends on the chemical environment of the nucleus. Nuclei of the same kind placed in a different chemical context are shielded differently and can therefore be distinguished. The chemical shift is then defined as:

Eqn 2.8.3 $$\delta / ppm = 10^6 \frac{\nu_L - \nu_{ref}}{\nu_{ref}}$$

Where ν_L is the Larmor frequency of the sample and ν_{ref} is the Larmor frequency of a reference compound. The Larmor frequency is related to the shielding constant as follows:

Eqn 2.8.4 $$\nu_{\mathrm{L}} = (1-\sigma)\frac{g \cdot B^{eff}}{2\pi}$$

The shielding constants can be calculated from the mixed second-order derivative of the energy with respect to the nuclear magnetic moment and the external field assuming that in lowest order the energy of the system is linear to both $\vec{B}^{ext}$ and μ:

Eqn 2.8.5 $$\sigma_{st} = \left.\frac{\partial^2 E}{\partial B_s^{ext} \partial \mu_t}\right|_{\vec{B}ext=\vec{\mu}=0}$$

The calculation of the shielding constant requires an accurate description of the wavefunction close to the nuclei. Thus, the nuclear magnetic shielding (absolute shielding) is the molecular electronic property. The isotropic chemical shift δ_{iso} is defined as $\delta_{iso} = \sigma_{ref} - \sigma$ where σ_{iso} is the isotropic shielding of the vanadium nucleus and σ_{ref} is the isotropic chemical shielding of the reference compound.

Almost all modern approaches to the shielding use some method with the distributed origin for the gauge of the magnetic field. This is necessary to avoid the nonphysical-dependence of calculated properties on the coordinate origin or, on more generally, on the gauge of the magnetic vector potential. This dependence vanishes, of course, if the quantum mechanical equations have been solved exactly. It can, however, have serious consequences if this is not the case, for instance for finite basis sets. One of the most popular methods for avoiding this gauge dependence is the so-called ‘gauge including atomic orbitals (GIAO, originally known as gauge invariant atomic orbitals) method. [86,87] The GIAO/DFT approach[88] is known to give satisfactory chemical shifts for different nuclei with larger molecules.

NMR chemical shift tensors

The NMR chemical shift is a tensor quantity. Depending on the local symmetry at the nuclear site, the magnitude of the chemical shift will vary as a function of the orientation of the molecule with respect to the external magnetic field. This orientation dependence of the chemical shift is referred to as chemical shift anisotropy (CSA). Mathematically, the chemical shift anisotropy is described by a second-rank tensor (a 3 by 3 matrix), which in the case of the symmetric part of the chemical shift (CS) tensor consists of six independent components. Generally, one is able to express the chemical shift tensor in a coordinate frame where all off-diagonal elements vanish. In this principal axis system, the chemical shift tensor is fully described by the three diagonal elements - the principal components - and the three eigenvectors or Euler angles describing the orientation of the principal axes with respect to an arbitrary frame. In addition, various combinations of the principal components (and their orientations) are in use to describe the chemical shift tensor.

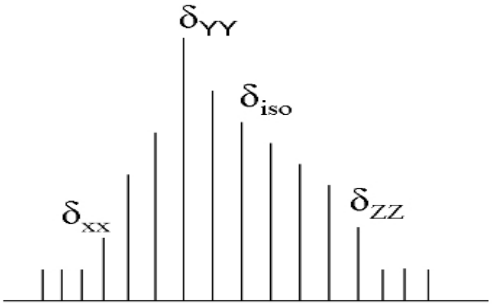


Figure 2.8.1 Principal components and isotropic chemical shift representation

In the Haeberlen-Mehring-Spiess convention, the principal axes are defined as follows:

$$\left|\delta_{zz} - \delta_{iso}\right| \geq \left|\delta_{xx} - \delta_{iso}\right| \geq \left|\delta_{yy} - \delta_{iso}\right|$$

where $\hat{\delta} = \hat{I}\sigma_{iso} - \hat{\sigma}$

The centre of gravity of the line shape is described by the isotropic value (δ_{iso}), which is the average value of the principal components and is represented by,

Eqn 2.8.6

$$\delta_{iso} = \frac{1}{3}(\delta_{xx} + \delta_{yy} + \delta_{zz})$$

The reduced anisotropy (δ_{σ}) describes the largest separation from the centre of gravity which is shown below.

Eqn 2.8.7

$$\delta_{\sigma} = \delta_{zz} - \delta_{iso}$$

The asymmetry parameter (η_{σ}) indicates by how much the line shape deviates from that of an axially symmetric tensor.

Eqn 2.8.8

$$\eta_{\sigma} = \left(\frac{\delta_{yy} - \delta_{xx}}{\delta_{\sigma}}\right)$$

EFG tensors and NQC

Quadrupole nuclei with a half-integer spin larger than 1/2 possess a quadrupole moment Q, which interacts with the electric-field gradient (EFG) generated by their surroundings. The coupling of Q (a property of the nucleus) with an EFG (a property of a sample) is called the quadrupole interaction.

Consider a free nucleus in a uniform space, that is, the three coordinate axes x, y, and z are equivalent. The Hamiltonian representing the quadrupole interaction of this nucleus, independent of the Cartesian coordinate frame, is defined by:

Eqn 2.8.9

$$\hbar H_Q = \frac{eQ}{6I(2I-1)} \sum_{j,k=x,y,z} V_{jk}\left[\frac{3}{2}(I_j I_k + I_k I_j) - \delta_{jk} I(I+1)\right]$$

where V_{jk} are the Cartesian components of V, the EFG at the origin, which is a second-rank symmetrical tensor, and δ_{jk} is the Kronecker delta. In the principal-axis system of the EFG (X_{PAS}, Y_{PAS}, Z_{PAS}), V is diagonal:

Eqn 2.8.10

$$V = \begin{pmatrix} V_{XX} & 0 & 0 \\ 0 & V_{YY} & 0 \\ 0 & 0 & V_{ZZ} \end{pmatrix}$$

with the convention

$$|V_{ZZ}| \geq |V_{YY}| \geq |V_{XX}|$$

Furthermore, the Laplace equation, $V_{XX} + V_{YY} + V_{ZZ} = 0$, holds for V, because the electric field at the nucleus is produced by charges wholly external to the nucleus.

Thus, only two independent quadrupole parameters are required:

Eqn 2.8.11

$$eq = V_{zz}$$

Eqn 2.8.12

$$\eta_Q = \frac{(V_{xx} - V_{yy})}{V_{zz}} \qquad \textbf{with } 1 \geq \eta_Q \geq 0$$

the largest component eq and the asymmetry parameter η_Q, respectively.

In practice, the asymmetry parameter η defines the powder lineshape and eq is related to the linewidth.

The product of eq with eQ divided by Planck's constant is called the quadrupole coupling constant C_Q

During line shape analysis of the solid-state NMR spectra, values for the quadrupole coupling constant C_Q (MHz) and asymmetry parameter η_Q are refined, which are related to the EFG tensor by the following formulas:

$$C_Q = \frac{e^2 Q V_{zz}}{h} \qquad \eta_Q = \frac{(V_{xx} - V_{yy})}{V_{zz}}$$

where V_{ii} is the EFG tensor at nucleus in atomic units, ordered according to $|V_{zz}| > |V_{yy}| \geq |V_{xx}|$;

NMR Chemical shift with QM/MM approach

NMR is a widely used tool for studying the structure and dynamics of chemical and biological systems ranging from small molecules to proteins and nucleic acids.[89] Although shielding (chemical shifts) tensors are determined and assigned in an essential step of any NMR study, it has been shown only recently that they can provide useful information for protein structure determination.[90]

Quantum chemical calculations of NMR parameters, often have to be limited to isolated (gas-phase) molecules and to some preferred (optimized) structures while experimental NMR spectra are commonly statical averages affected by dynamic processes such as conformational equilibria as well as intra and/or intermolecular interactions. Since DFT is the most useful quantum mechanical technique for larger systems to about 100 atoms,[91] particularly where metal ions are involved, implementation for calculating magnetic properties such as chemical shielding tensors within the DFT framework are of considerable importance. For macromolecules, including proteins[92] and nucleic acids[93], useful empirical and semiempirical methods for chemical shift estimates have been developed. They are based on measured values for different systems, augmented by semiempirical calculations. The recent increase in the number of protein structures determined by NMR has made possible a simple, yet useful, empirical description for proton chemical shifts. The model includes electric field effects, the contribution of the magnetic susceptibility of other atoms or groups, and closed contact interactions.[94]

To complement these empirical approaches and to treat nuclei other than the proton, it is important to be able to calculate chemical shielding with high accuracy for specific groups in a large molecule. This can be achieved with a method that treats the shielding of only a part of the systems in detail and describes the environment by a simpler and fast approach. QM/MM methodologies are well suited for this purpose where the important part is treated quantum mechanical methods, and the environment part is described with molecular mechanics method. The chemical shielding tensor in a large molecule like a protein is usefully divided into terms that arise from the electronic structure of the atoms of the interest perturbed by the environment and additive contribution from the other parts of the system. The former includes primarily the electrostatic effects arising from the partial charges on the MM atoms, and the later includes the contribution of anisotropic magnetic susceptibility tensors of aromatic and carbonyl groups.[95]

2.9 Molecular Properties

Population Analysis

The electron density can be arbitrarily partitioned and assigned to nuclei in a molecule, giving rise to the concept of the *electron population* of an atom. The atomic charge is therefore the sum of this electron population and the nuclear charge. A number of schemes exist for this type of partitioning. Population analysis is a scheme that follows naturally from the LCAO form of most wavefunctions. One of the most popular schemes is due to Mulliken.[96] Mulliken analysis also employs the (fixed) AO basis function overlap matrix *S* to give the following expression for the electron population of an atom:

Eqn 2.9.1

$$P_{\mathrm{A}} = \sum_{i \in \mathrm{A}} \mathrm{P}_{ii} - \frac{1}{2} \sum_{i \in \mathrm{A}} \sum_{j \in \mathrm{B}; i \neq j} \mathrm{P}_{ij} \, \mathrm{S}_{ij}$$

The use of diffuse functions in quantum mechanical calculations presents a challenge for Mulliken analysis. This is due to the delocalised nature of the orbitals arising for such functions, which may no longer be assigned to a single atomic centre. Mulliken charges can thus be quite basis-set dependent. The most preferable method in that respect is Natural Population Analysis (NPA) introduced by Weinhold and co-workers in the mid 1980s.[97] Certainly, NPA method seems to be excellently positioned for the formidable task of substituting the Mulliken method for the assignment of electronic populations to atomic orbitals. On the other hand, it overcomes some of its most serious limitations, namely the basis set dependence and the appearance of physically meaningless results when non-covalent bonds are considered.

Chapter 3

3 Computational Details

3.1 Gas-phase optimisation

Stationary points of amavadin-derived species were optimized with the Gaussian 03 program[98] at the BP86/AE1 level, i.e. employing the exchange and correlation functionals of Becke[99] and Perdew,[100] respectively, together with a fine integration grid (75 radial shells with 302 angular points per shell), the augmented Wachters' basis[101]on V (8s7p4d, full contraction scheme 62111111/3311111/3111), and 6-31G* basis[102] on all other elements. For the amavadin-based model complexes the resolution of identity was used with fitting of the density employing suitable auxiliary basis sets generated automatically in Gaussian 03 (denoted RI-BP86). This and comparable DFT levels have proven quite successful for transition-metal compounds and are well suited for the description of structures, energies, barriers, etc.[103] The nature of the stationary points was verified by computations of the harmonic frequencies at that level. In selected cases, geometries were also optimized at the RI-BP86 level as implemented in the TURBOMOLE program,[104] using suitable auxiliary bases[105] and a medium-sized grid (m3).

3.2 QM/MM optimisation

The pdb file VCPO (1IDQ.pdb) and VBPO (1QI9.pdb) was obtained from the RCSB website.[106] Protonation states of the titratable residues were assigned based on calculated pKa values using the Propka program.[107] The overall charge of the protein model after protonation, e.g. $-19e$ for VCPO, was then neutralized by selectively protonating residues on the surface of the protein which were > 5 Å from the active centre (residues with higher pKa values were preferentially protonated). The InsightII[108] software package was employed to add a 14 Å solvent shell of TIP3 water iteratively (with intermediate relaxation) to the VCPO protein X-ray structure. A frozen crust of solvent molecules (outer 7 Å) was imposed to prevent solvent ‘escaping’ during the model relaxation and equilibration phase. The solvation of the VBPO homodimer was achieved by placing a 30 Å solvent sphere of TIP3 water centered on one of the vanadate moieties (resid VO4). The solvation was performed iteratively (with intermediate relaxation) to the protein X-ray structure of VBPO. A boundary potential was imposed to prevent solvent ‘escaping’ during the model relaxation and equilibration phase. Due to the large size of the VBPO homodimer, the previous setup, which immerses the enzyme totally in the solvent, would have resulted in a very large number of solvent molecules. A large majority of these solvent molecules would be located very far from the active site in the non-active QM/MM region, and would thus affect the results only marginally.

The initial geometry of the *p*-VCPO enzyme was taken from the pdb file (1IDU.pdb), and the initial geometry of the *p*-VBPO enzyme was constructed from the pdb file of the *n*-VBPO (1QI9.pdb), obtained from the RCSB website.[106] As the experimental structure of *p*-VBPO is not available till date, we produced a ‘homology-type’ model. The VBPO file was modified by deleting the vanadium cofactor, and then the coordinates from the peroxo-form of the VCPO cofactor were transposed into the emptied active site of the *n*-VBPO.

The vanadate moiety was also fixed at X-ray coordinates with an additional rigid constraint being applied to the vanadium - histidine 496 N^{ε} bond during equilibration. The solvated protein was relaxed via several cycles of molecular mechanics (MM) minimization and molecular dynamics (MD) simulations. Once the system was equilibrated (~200-400 picoseconds of MD simulation) six snapshots were taken at random intervals along the MD trajectory for QM/MM optimizations, in order to generate a reasonable sample size with different protein configurations.[109]

ChemShell[110] was utilized as the QM/MM software suite. Turbomole[104] with the BP86[99] functional and the AE1 basis set was applied to the QM regions using the resolution of identity approximation, RI-BP86/AE1. The AE1 basis comprises the

Wachters basis[101] augmented with two diffuse *p* and one diffuse *d* sets for metal centers (8s7p4d, full contraction scheme 62111111/3311111/3111), and 6-31G*[102] for all other atoms in the QM region, together with suitable auxiliary fitting functions from the Turbomole library.[104] DL_POLY[111] provided the MM energy and gradients using the CHARMM27 force field.[112] An electrostatic embedding scheme was applied.[113] Partial atomic charges for the vanadate moiety were created using a Mulliken population analysis from gas-phase models optimized at the RI-BP86/AE1 level of theory. Partial atomic charges for the vanadate moiety were created using Weinhold natural population analysis (NPA)[97] from gas-phase models optimized at the RI-BP86/AE1 level of theory and are shown in Tables 3.2.1. and 3.2.2.

Table 3.2.1 Partial atomic charges that are calculated by NPA[97] for selected vanadate moieties of native form of VHPOs

	VHPO-T	VHPO-D	VHPO-S
V	1.48	1.14	1.01
O-Terminal(O_1)	-0.43	-0.63	-1.00
O-Hydroxyl(O_2)	-0.84	-0.63	-0.81
O-Hydroxyl (O_3)	-0.84	-0.89	-0.81
O-Hydroxyl (O_4)	-0.84	-0.89	-0.80
H1	-	-	-
H2	0.49	-	-
H3	0.49	0.45	0.41
H4	0.49	0.45	-
Total Charge	0	-1	-2

Table 3.2.2 Partial atomic charges that are calculated by NPA[97] for selected vanadate moieties of peroxo form of VHPOs

	p-VHPO-Z	*p*-VHPO-S1	*p*-VHPO-D13
V	1.09	1.19	1.32
O-Terminal	-0.58	-0.48	-
O-Hydroxyl	-	-0.74	-0.63
O-Peroxo (O^2)	-0.51	-0.27	-0.11
O-Peroxo (O^4)	-0.42	-0.18	-0.17
H	-	0.48	0.62
Total Charge	-1	0	1

3.3 Molecular dynamics simulations

For amavadin-derived species, Born-Oppenheimer MD simulations (denoted BOMD) were performed at the RI-BP86/AE1 level, using the ChemShell program[110] for NVE ensembles at ca. 300 K for 3000 ps, with a time step of 1 fs. In these simulations, Turbomole was used as QM program,[104] and the C-H distances were frozen with the SHAKE algorithm. Data and snapshot sampling was started after the first picosecond, which was taken for equilibration.

3.4 NMR chemical shifts

The magnetic shielding tensors $\hat{\sigma}$ were computed using the Gaussian 03[98] package with the B3LYP[99,114] functional and the AE1+ basis (AE1 basis augmented with additional diffuse functions[115] on all non-hydrogen atoms). The protein and solvent surrounding the QM region were treated as fixed point charges.

A single $VOCl_3$ molecule was used as a reference for the conversion to relative chemical shifts with a calculated isotropic magnetic shielding of -2294.4 ppm (which is bracketed by the reported B3LYP values,[46] -2317.2 and -2279.4 ppm for the TZV and 6-311+G basis, respectively). The isotropic chemical shift δ_{iso} is defined as $\delta_{iso} = \sigma_{iso} - \sigma_{VOCl_3}$ where σ_{iso} is the isotropic shielding of the vanadium nucleus and σ_{VOCl3} is the reference isotropic chemical shielding of $VOCl_3$.

There are a number of conventions in the literature with respect to the definitions of the principal components of the CSA and EFG tensors, which may lead to confusion. The Haeberlen-Mehring-Spiess[116] convention is adopted here (see chapter 2.8) for consistency with the previous work by Pooransingh-Margolis et al.[46]

Magnetic shieldings were computed for equilibrium structures and for snapshots along the trajectories employing the B3LYP[114] hybrid functional, together with AE1+ basis, i.e. AE1 augmented with diffuse functions on C, N, and O.[115] Snapshots were taken every 20 fs during 1000-2000 ps of the total runs. Chemical shifts are reported relative to $VOCl_3$, optimized or simulated at the same respective level (σ values -2294 ppm and -2320 ppm employing BP86/AE1, and RI-BP86/AE1 geometries, respectively, and -2382 ppm for a BOMD simulation averaged over 1 ps).

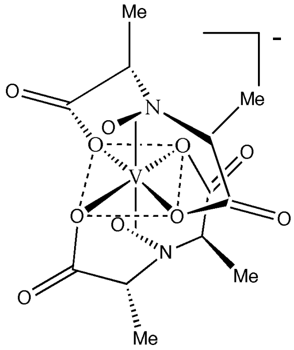
In selected cases, magnetic shieldings were evaluated at the GIAO-B3LYP/AE1+ level in conjunction with the polarizable continuum model (PCM) of Tomasi and coworkers[117] (employing UFF radii and the parameters of water, both for geometries optimized in the gas phase and the continuum at the BP86/AE1 level.

Chapter 4

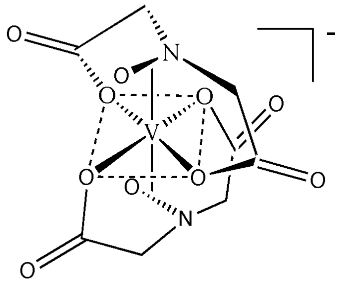
4 Results

4.1 A DFT directed verdict on oxidized amavadin

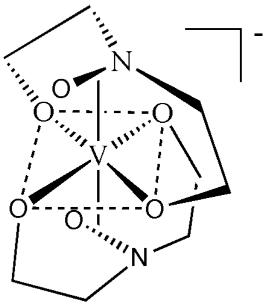
The oxidized amavadin complex [Δ-V^V{(S,S)-hidpa}$_2$]$^-$ (**2a**) and its hida analogue (**3**) are known in the solid state (PPh_4^+ counterion).[22] Their structures, as well as hypothetical (**4 – 6**) with the same donor atoms in the first coordination sphere are shown in Figure 4.1.1. These model species were constructed from **3** by deleting the appropriate atoms and saturating the dangling bonds with H atoms.



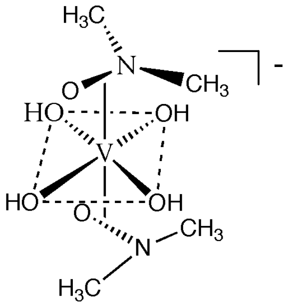
Compound 2a. hidpa (-382)



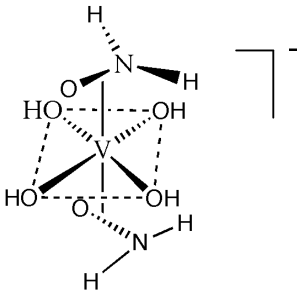
Compound 3. hida (-341)



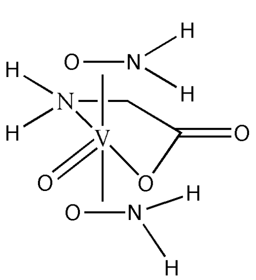
Compound 4. (-415)



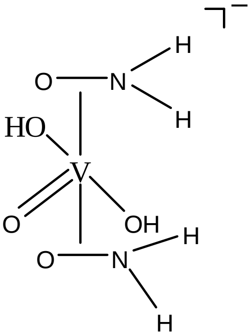
Compound 5. (-578)



Compound 6. (-633)



Compound 7. (-863)



Compound 8. (-866)

Figure 4.1.1 Oxidized amavadin (2a), its parent (3), and model complexes (4 – 6), Oxo-vanadium(V) hydroxylamido complex [$VO(NH_2O)_2Gly$] (7) and model complex [$VO(NH_2O)_2(OH)_2$]$^-$ (8), together with computed gas-phase equilibrium ^{51}V chemical shifts (GIAO-B3LYP/AE1+ level, in parentheses).

Geometries:

As mentioned in the introduction, model complexes **2** can exists in the form of several stereoisomers, as it contains four stereogenic carbon centers and one axis of symmetry giving rise to axial chirality. Gas-phase and PCM-optimized geometrical parameters for **2a** and selected stereoisomers, as well as for **3 - 6** are collected in Table 4.1.1 and Table 4.1.2, respectively, together with experimental data for 2a and 3 in the solid-state.[22] As expected, the different diastereomers of the oxidized amavadin **2a** have quite similar gas phase optimized bond lengths. The computed values for the all-S form **2a** and for **3** tend to be longer than the corresponding experimental ones, which may in part be due to packing effects in the solid, and may in part be related to the particular density functional employed.[118] Typically, distances are overestimated by a few picometers in the computations, with a maximum deviation of 0.046 Å for a V-N bond in **2a**.

For the discussion of substituent effects on the 51V chemical shifts (see below), several model complexes **4 - 6** were constructed that preserve the general ligand environment of amavadin. Turning the carboxylate donors into alkoxy groups introduces only minor changes in the geometrical parameters (compare **3** and **4** in Table 4.1.2). Dissecting the tetradentate ligand backbone in **4** to afford **5** effectively leads to a dissociation of the N atoms from the metal (with fully optimized V-N distances of 2.978 Å in **5**). Apparently, the chelate is necessary to counteract the steric bulk of 8 donor atoms about the metal. In order to evaluate substituent effects for the same ligand environment, partial optimizations were performed for **5**, in which the two V-N and V-O distances to the hydrozylimido moiety were fixed to their respective values in **4**. It is the data from these constrained optimizations that are included in Table 4.1.1. and Table 4.1.2. Relative to the corresponding fully optimized minimum, this partially optimized structure of 5 is higher in energy by 23.7 kcal/mol (RI-BP86/AE1 level).[119]

Table 4.1.1 RI-BP86/AE1 optimized bond lengths (Å) of the diastereomeric oxidized amavadin complexes [Δ-V^V(hidpa}$_2$]$^-$ (2).

	2a SS,SS *Expt.*[a]	**2a** SS,SS (PCM)	**2a** SS,SS	**2b** RR,RR	**2c** RR,SS	**2d** SR,RS	**2e** RS,SR	**2f** RS,RS
V-O_1	*1.993*	2.025	2.038	2.035	2.037	2.029	2.031	2.029
V-O_2	*1.923*	1.951	1.958	1.957	1.956	1.968	1.950	1.968
V-O_3	*1.959*	1.987	1.976	1.977	1.978	1.972	1.968	1.972
V-O_4	*1.977*	1.987	1.976	1.977	1.976	1.972	1.968	1.967
V-O_5	*1.940*	1.951	1.958	1.957	1.958	1.968	1.950	1.950
V-O_6	*1.972*	2.025	2.038	2.035	2.037	2.029	2.031	2.029
V-N_7	*2.018*	2.047	2.045	2.046	2.046	2.050	2.063	2.053
V-N_8	*1.999*	2.047	2.045	2.046	2.046	2.050	2.063	2.053

[a] Solid state, PPh_4^+ counterions, from reference 22; see Figure 1 for numbering of atoms.

Table 4.1.2 RI-BP86/AE1 optimized bond lengths (Å) of the oxidized amavadin model complexes [Δ-V^V(hida}$_2$]$^-$ (3), and for selected models thereof.

	3 *Expt.*[a]	**3**	**4**	**5**	**6**
V-O_1	*1.991*	2.041	2.036	2.037	2.109
V-O_2	*1.963*	1.961	1.953	1.953[b]	1.952
V-O_3	*1.955*	1.979	2.016	1.950	1.952
V-O_4	*1.920*	1.979	2.016	1.950	1.952
V-O_5	*1.936*	1.961	1.963	1.963[b]	1.952
V-O_6	*1.978*	2.041	2.036	2.036	2.109
V-N_7	*2.016*	2.038	2.040	2.040[b]	2.048
V-N_8	*2.028*	2.038	2.040	2.040[b]	2.048

[a] Solid state, PPh_4^+ counterions, from reference 22. [b] Fixed to the corresponding value of **4**.

Chemical shifts

Equilibrium ^{51}V chemical shifts computed in the gas phase are included in Figure 4.1.1. For **2a** and **3**, the static values δ = -382 and -341, respectively, are considerably more shielded than the corresponding experimental numbers, δ = -281 and -263, respectively, a frequent observation for this particular level of theory (see below for further results on **2a - 2h**). Figure 4.1.1 shows a typical oxovanadium(V) complex with two hydroxylamido ligands, [VO(NH_2O)$_2$Gly] (**7**, Gly = glycine anion).[26a] Even though some geometrical parameters observed for **7** in the solid-state are not very well reproduced by the gas-phase optimization (see, Table 4.1.3), the computed ^{51}V chemical shift for **7**, δ = -863, compares favorably to experiment, δ = -830 (aqueous solution value). Clearly, the substantial deshielding on going from **7** to **2a** is well captured by the computations.

Table 4.1.3 Geometrical parameters for gas phase optimized (RI-BP86/AE1) oxo-vanadium(V) hydroxylamido complexes.

Bond length Å	**7** *Exp.*[a]	**7**	**8**
V_1-O2	----	------	2.007
V_1-N2	*2.121(2)*	2.253	-----
V_1-N3	*2.008(2)*	2.068	2.068
V_1-N4	*2.021(2)*	2.065	2.068
V_1=O5	*1.603(2)*	1.626	1.662
V_1-O6	*2.188(14)*	2.064	1.992
V_1-O7	*1.898(2)*	1.898	1.960
V_1-O8	*1.901(2)*	1.900	1.960

[a] From solid-state X-ray diffraction, see Keramidas, A.D.; Miller, S.M.; Anderson, O.P.; Crans, D.C. *J. Am. Chem. Soc.*, **1997**,*119*, 8901.

Starting from **7**, model complex **8** was constructed by replacing the carboxylate and amine donors with hydroxy substituents. The vanadium(V) centers in both complexes are seven-coordinate in a pentagonal bipyramidal geometry with the oxo ligand and an amino or hydroxy group in apical positions, and they contain two bidentate hydroxyamido ligands, one oxo ligand, and two monodentate N or O donors. The two hydroxylamido ligands are in an equatorial plane perpendicular to the V=O. Despite different overall charges and some variations in the metal-ligand bond distances (cf.Table 4.1.3), the computed chemical shifts for **7** and **8** are remarkably similar (Figure 2). The latter is closely related to the amavadin model **6** in Figure 4.1.2, via the equation

$$\mathbf{6} \rightarrow \mathbf{8} + H_2O, \quad \Delta E = -7.8 \text{ kcal/mol (RI-BP86/AE1)} \quad (1)$$

On going from **8** to **6**, a substantial deshielding is computed for the ^{51}V nucleus, by $\Delta\delta$ = +233 ppm. This large effect is difficult to analyze, because the mutual orientation of the ligands differs in both complexes. For instance, the X-V-X angle (where X denotes the midpoint of the side-on NO group) is 128° and 179° in **8** and **6**, respectively. In any event, it appears that the transition from a heptacoordinate oxo complex to an octacoordinate non-oxo species entails a large deshielding of the metal.

Replacing the hydroxylamido moieties in **6** with N,N´-dimethylhydroxylamido ligands (constraining the latter to prevent detachment) results in a further, slight downfield shift, $\Delta\delta$ = +55 ppm (compare **6** and **5** in Figure 4.1.1). Closing the tetradentate ligand backbone by going from **5** to **4** produces another large deshielding, by $\Delta\delta$ = +163 ppm. This rather large effect is mostly steric in nature, as the general electronic structure of both complexes should be very similar. One factor that arguably will contribute to this effect is the considerable elongation of the two $V\text{-}O_3$ and $V\text{-}O_4$ bonds by almost 0.07 Å in the course of this transformation (Table 4.1.1.).

Mutating the alkoxy moieties in **4** into the carboxylate donors of the hidpa ligands in **3** produces a final, small downfield shift of $\Delta\delta$ = +74 ppm. All these transformations just discussed have a deshielding effect on the central metal, and add up to $\Delta\delta$ = +525 ppm. Thus, the change in $\delta(^{51}V)$ between a typical oxovanadium complex **8** (or **7**) and the oxidized amavadin parent **3** can be broken down into a series of increments, the two most important of which stem from the difference between oxo- and non-oxo complexes (i.e. an electronic effect), and from those between mono-/bidentate and tetradentate ligands (i.e. a steric effect). There is thus no need to invoke a non-innocent nature of the hidpa and related ligands in order to explain the observed unusual ^{51}V chemical shifts of oxidized amavadin and its derivatives.

When the electronic structure of **2a** is analyzed, no evidence for low-lying ligand-to-metal charge charge transfer (LMCT) excitations are found. For instance, a notable HOMO-LUMO gap and singlet-triplet splitting are computed (3.75 and 1.97 eV, respectively, at the B3LYP level). These are much larger than those in vanadium catecholate complexes, where very low-lying LMCT bands and, possibly, thermal population of triplet states are indicated to be responsible for the unusual deshielding of the ^{51}V resonances in the complexes.[120] This result further corroborates the innocent nature of the ligands in amavadin.

Stereoisomers

Natural amavadin **1** (see Figure 1.1.1) has been shown to consist of two diastereomers with Δ-V{(S,S)-hidpa}$_2$ and Λ-V{(S,S)-hidpa}$_2$ configuration. The interconversion between both is rapid in the native vanadium(IV) form, but is kinetically hindered upon oxidation to the vanadium(V) species. Depending on the ligands employed, and on their optical purity, synthetic amavadin derivatives can consist of complex equilibria between stereoisomers, which can give rise to multiple ^{51}V NMR signals in the oxidized forms. Depending on the source of the hidpa ligand, ^{51}V chemical shifts between δ = -217 and δ = -280 have been reported, a quite substantial variation for complexes with identical composition and connectivity that differ essentially in the stereochemistry at C atoms three bonds away from the metal. Even within the same probe, where experimental conditions are certainly exactly the same, variations of up to $\Delta\delta$ = 35 ppm have been noted.

The stereoisomers considered in this study are shown schematically inFigure 4.1.2. They consist of the pairwise combinations of (S,S) and (R,R) forms, as well as pairs of the meso-(R,S) forms among themselves. Only Δ-isomers have been computed, as most of the corresponding Λ-isomers are implicitly included (e.g., Λ-V{(S,S)-hidpa}$_2$ and Δ-V{(R,R)-hidpa}$_2$ are enantiomeric pairs with identical energies and NMR parameters). In the study by Lenhardt et al,[24] the racemic form of the hidpa ligand has been used to synthesize the vanadium complexes in situ. This racemic form has been shown to consist of a mixture of (S,S)-, (R,R)-, and meso-(R,S)-isomers.[121] Thus, not only the pairwise combinations between each of these ligands are possible in the vanadium complexes, but also mixed combinations such as (S,S) with (R,S). Such mixed forms have not been included in this study.

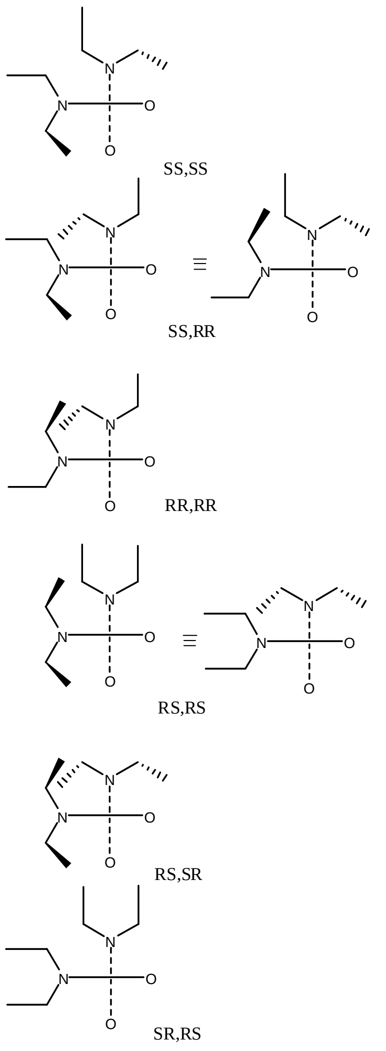


Figure 4.1.2 Amavadin diastereomers selected in this study (Δ-forms); schematic views of the (Me-C)$_2$-N-O fragments along the X-V-X axis (X = midpoint of NO groups).

The equilibrium $\delta(^{51}V)$ values computed for a number of salient stereoisomers are all closely spaced around two values (around ca. -380 ppm and -350 ppm for complexes with asymmetric and meso ligands, respectively, see entry 2 in Table 4.1.4. No further differentiation is apparent between the isomers within each of the two groups upon inclusion of bulk solvent effects via a simple PCM approach, neither in terms of single-point computations on the gas-phase geometries (entry 3), nor upon relaxation in the continuum (entry 4). For selected isomers (**2a-d**) we have performed BOMD simulations in the gas phase and averaged the chemical shifts along the respective trajectories. This or similar procedures to model thermal averaging of chemical shifts has recently been applied to other transition-metal complexes.[122] The BOMD simulations were done at the same level of theory as the Gaussian 03 optimizations (RI-BP86/AE1), but with another program with a slightly different scheme for the numerical integration, resulting in slightly different equilibrium distances and chemical shifts (compare entries 2 and 5 in Table 4.1.4). Dynamical (thermal) averaging results in a noticeable general increase of the δ values, i.e. deshielding, and in a somewhat larger spread of the individual resonances (compare entries 5 and 6). For instance, while the equilibrium δ_e values of **2a** and **2b** are almost identical, the corresponding thermally averaged δ_{av} values differ by ca. 30 ppm. For the individual snapshots, fluctuations in the instantaneous magnetic shieldings are much larger than this value, but the running average over 50 or more such snapshots is well converged within this range, without showing a noticeable drift after 1 ps (see Figure 4.1.3). Note that the computed deshielding due to thermal motion brings the δ_{av} values somewhat closer to experiment than the raw δ_e data.

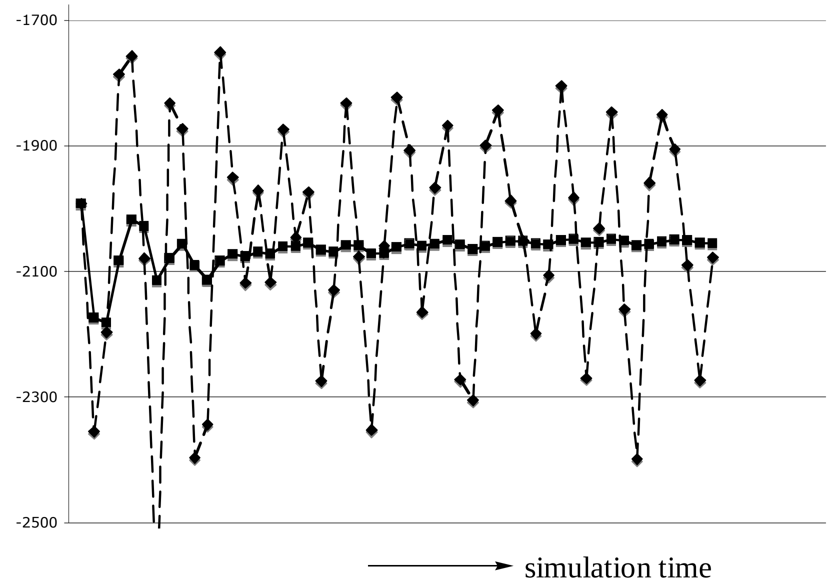


Figure 4.1.3 Convergence of the ^{51}V isotropic magnetic shielding for 2b, from the first to second picosecond of BOMD simulation. Dashed line and diamonds is for individual snapshots, and the solid line with squares represents the cumulative running average (i.e. the average over all preceding points).

Table 4.1.4 Relative energies (in kcal/mol) and ^{51}V NMR chemical shifts δ (in ppm)[a] for selected amavadin diastereomers.

Entry	Isomer	2a	2b	2c	2d	2e	2f
		SS,SS	RR,RR	RR,SS	SR,RS	RS,SR	RS,RS
1	E_{rel}, RI-BP86	0.00	0.17	0.10	0.4	11.37	5.96
	(B3LYP)[a]	0.00	0.47	0.19	1.79	11.77	6.91
2	δ_e (gas)[b]	-382	-381	-379	-357	-349	-354
3	δ (PCM//gas)[c]	-313	-311	-310	-302	-314	-308
4	δ (PCM// PCM)[d]	-353	-352	-349	-350	----	---
5	δ_e (gas)[e]	-394	-393	-397	-382	----	----
6	δ_{av} (BOMD)[f]	-355	-327	-340	-325	----	----
7	Expt. δ [g]	-252	-217	-234	----	----	----
8	Expt. δ [h]	-----	-----	----	-280 / -270/ -250		

[a] B3LYP/AE1+ level. [b] RI-BP86/AE1 optimized using Gaussian 03. [c] PCM single point on gas-phase optimized geometries. [d] Geometries optimized with PCM. [e] RI-BP86/AE1 optimized using Turbomole. [f] Mean value over snapshots from BOMD trajectory. [g] Racemic ligand, reference 24. [h] meso-ligand, reference [23].

As expected, this deshielding is related to an increase of the mean bond distances over the trajectories[123] with respect to the corresponding equilibrium values. On average, this elongation of the vanadium-ligand bonds amounts to 0.01 Å (see Table 4.1.5)

Table 4.1.5 Equilibrium (r_e) and averaged (r_{av}) bond-lengths (in Å) at the RI-BP86/AE1 optimized geometry (from the snapshots selected for NMR computation) in amavadin diastereomers (in parentheses: standard deviation during 1000-2000 ps), all distances are in Å.

	2a		**2b**		**2c**		**2d**	
	(S,S)(S,S)		(R,R)(R,R)		(R,R)(S,S)		(S,R)(R,S)	
	r_e	r_{av}	r_e	r_{av}	r_e	r_{av}	r_e	r_{av}
V-O_1	2.037	2.048(10)	2.034	2.050(20)	2.035	2.051(7)	2.029	2.048(6)
V-O_2	1.956	1.986(21)	1.955	1.962(12)	1.953	1.949(7)	1.968	1.963(8)
V-O_3	1.975	1.980(8)	1.976	1.982(10)	1.975	1.988(8)	1.972	1.994(6)
V-O_4	1.975	1.972(10)	1.976	1.992(16)	1.974	1.986(4)	1.972	1.994(18)
V-O_5	1.956	1.964(13)	1.955	1.967(18)	1.956	1.975(17)	1.968	1.966(7)
V-O_6	2.037	2.040(9)	2.034	2.044(10)	2.035	2.051(10)	2.029	2.041(5)
V-N_7	2.043	2.055(8)	2.045	2.068(6)	2.043	2.057(4)	2.050	2.056(5)
V-N_8	2.043	2.056(5)	2.045	2.058(13)	2.044	2.055(9)	2.050	2.066(7)

For the complexes formed from the racemic ligand (**2a-c**), the computed sequence of the signals is in qualitative agreement with the observed pattern (assuming that the rr,ss combination has double intensity), and the isomers are tentatively assigned correspondingly. For the complexes with the meso ligand, MD simulations have been performed only for one representative (**2d**), because performing them for all possible components of the mixture would be a formidable task, and because confident assignment of experimental numbers to specific stereoisomers based on these results will probably be difficult.

Conclusions

^{51}V chemical shifts of oxidized amavadin derivatives have been computed and the reasons for the unusually deshielded ^{51}V chemical shifts in these species have been analyzed. This deshielding can be broken down into a number of increments from regular substituent effects, the most important of which arises from the non-oxo nature of these complexes. No evidence is found for low-lying LMCT excitations. There is thus no need to invoke a non-innocent nature of the ligand in order to explain the unusual chemical shifts in this case. These findings constitute an important contribution from theory to the interpretation of experimental findings. $\delta(^{51}V)$ values computed for static equilibrium geometries show only little variation for diastereomeric isomers. The observed noticeable discrimination in case of diastereomeric mixtures is indicated to be dynamic in origin. Even though chemical shifts averaged over MD trajectories apparently can reproduce certain experimental trends, definite assignment of individual ^{51}V resonances to specific isomers is difficult at this point.

4.2 ^{51}V NMR Chemical Shifts Calculated from QM/MM Models of Vanadium Chloroperoxidase – Native form

Preface

The investigation on the native VCPO enzyme that is presented in this chapter was done in collaboration with Dr. Dongi Wang and Dr. Mark P.Waller.The calculations on the triply protonated model VOT234 and the singly protonated models VOS1, VOS3 and VOS4 were done exclusively by me. I was involved in selecting the other models and evaluation of all results. Thus, the discussion in this chapter is based on all the selected models.

QM/MM setup

The vanadate moiety has four crystallographically resolved oxygen atoms (one axial and three equatorial) as shown in Figure 4.2.1. This work is based on considering triply, doubly and singly protonated models, which are hereafter referred to as VOT, VOD and VOS respectively. The oxygen labels from the X-ray study (1IDQ.pdb) are used throughout this work. The labeling of the equatorial oxo ligands is not uniform in the literature, since some of the authors[42,43] use different conventions.

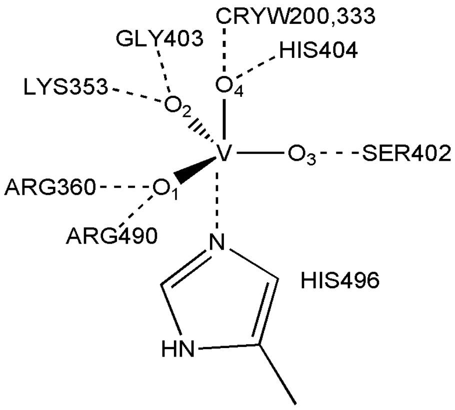


Figure 4.2.1 Vanadate moiety (vanadium and four O atoms) and coordinated imidazole moiety, showing the oxygen labeling used throughout this study and potential hydrogen bonding interactions with the protein environment (all labels are consistent with those in the 1IDQ.pdb file).

Protonation at different vanadate oxygen atoms generates positional isomers. Those that are studied presently are listed in Table 4.2.1, which also introduces the chosen notation. An exhaustive study on all possible permutations of positional isomers for each protonation state was not performed. Visual inspection of the immediate protein environment surrounding O^2 does not indicate any available residues that may act as significant hydrogen bond acceptor. Therefore protonation of O^2 appears less likely than that of the other equatorial oxygen atoms. We thus do not consider the singly protonated VOS2, the doubly protonated VOD24, and the triply protonated VOT124 and VOT244 models as candidate structures and include only the VOT234 model to check the validity of this assumption. Previously Carlson and co-workers[42] have noted that an axial water may represent a low energy conformation with one additional equatorial hydroxo group, and therefore we also studied a number of models that contain water in the axial position.

Table 4.2.1 Selected models labeled according to the protonation state (VOT,VOD and VOS) and the oxygen atoms that are protonated (1,2,3,4).

Triply protonated	Doubly protonated	Singly protonated
VOT134	VOD14	VOS1
VOT234	VOD34	VOS3
VOT144	VOD44	VOS4
VOT344		

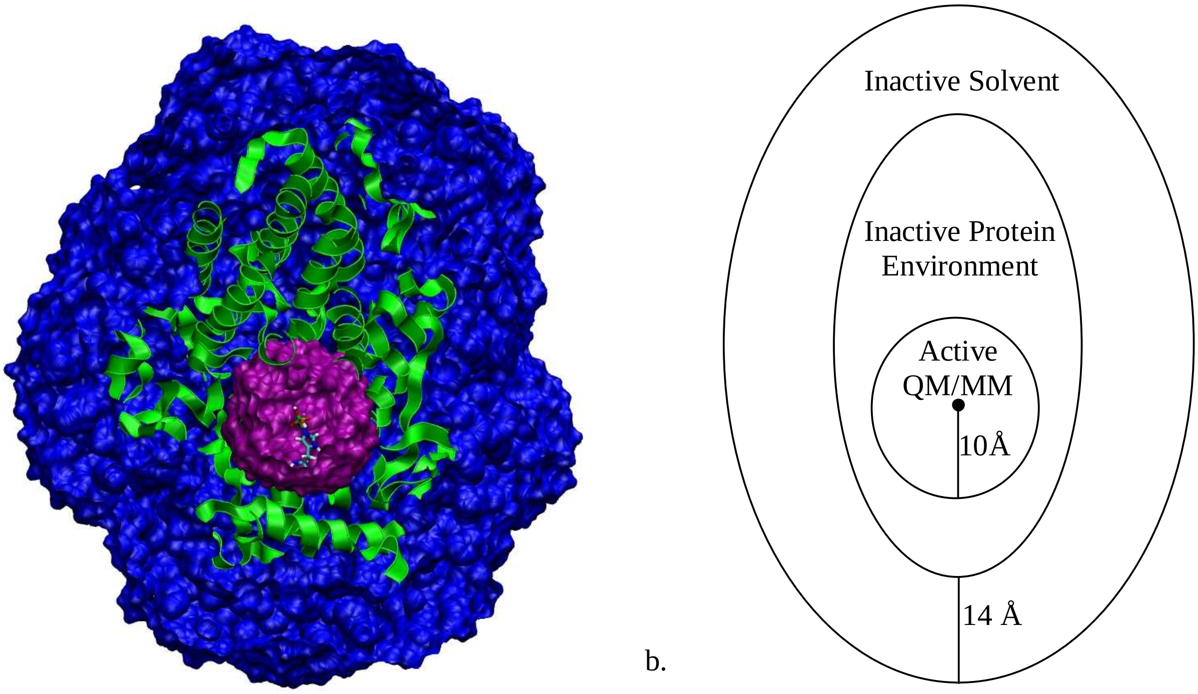


Figure 4.2.2 a. A cross section through the solvated protein: the outer region in blue is solvent (14 Å layer from the protein surface), the fixed protein environment during the QM/MM optimization (residues > 10 Å from vanadate) is displayed as a green ribbon. The purple sphere centered on the vanadate moiety is the active region subjected to the QM/MM optimization (including all residues < 10 Å from vanadate). The residues in the active region (~1000 atoms) are shown in Table 4.2.2. .The vanadate moiety and the bound HIS496 are shown as a visual guide. b. Schematic system partitioning with the filled black circle (roughly in the centre) representing the vanadate moiety.

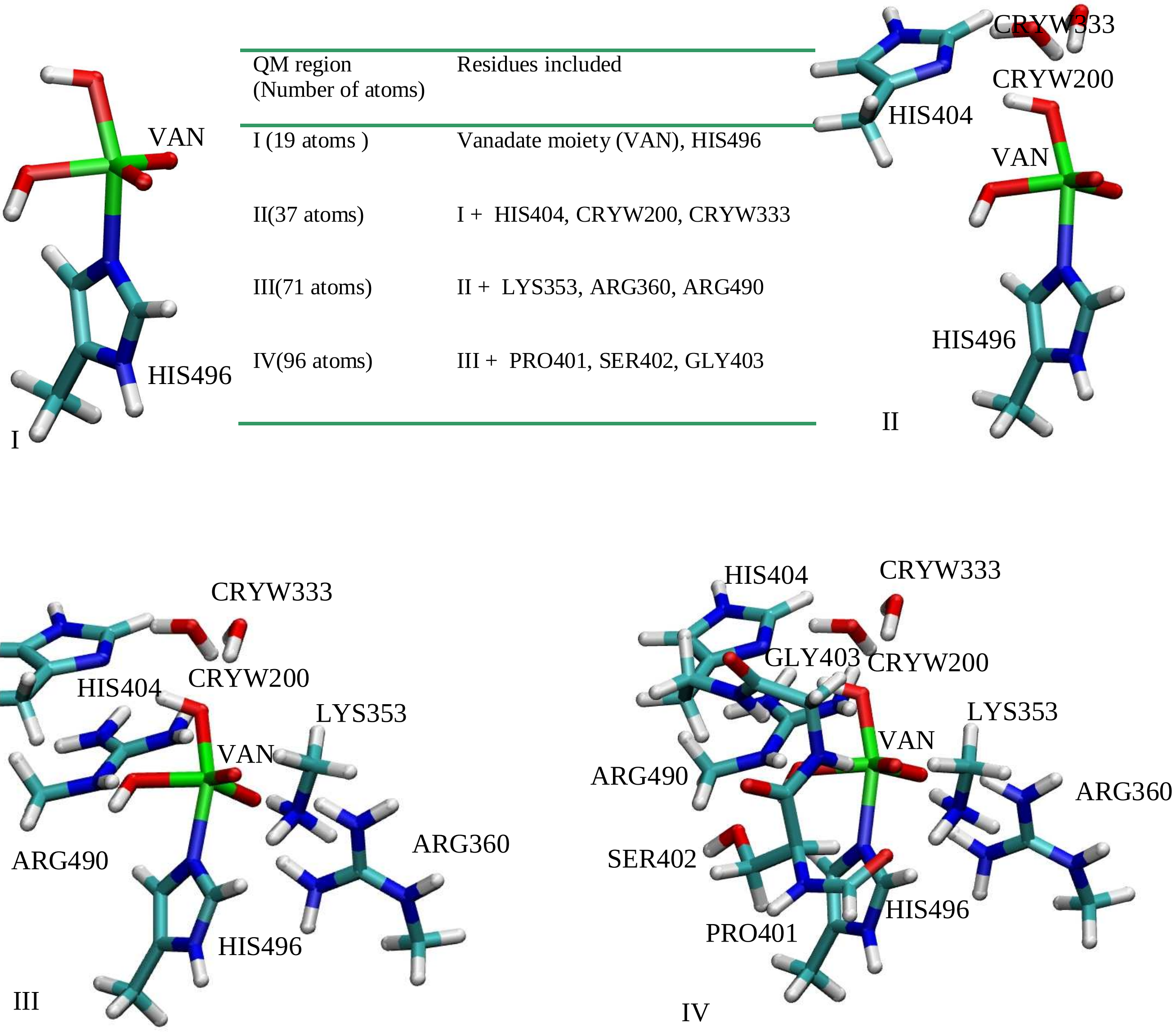
Table 4.2.2 Example of the residues included in the active region for VOT234 (995 atoms, residue numbering according to that in 1IDQ.pdb)

GLY46 PRO47 PRO48 PHE193 HSP222 GLY286 TRP289 ALA290 TYR291 ASP292
GLY293 THR299 PRO300 PRO301 PHE303 TYR304 ALA345 GLY346 SER349 TRP350
LYS353 TRP354 PHE358 ARG360 PRO361 GLY382 ALA383 PRO384 PHE393 LYS394
PRO395 PRO396 PHE397 PRO398 ALA399 TYR400 PRO401 SER402 GLY403 HSE404
ALA405 THR406 PHE407 GLY408 GLY409 GLU485 ASN486 ALA487 ILE488 SER489
ARG490 ILE491 PHE492 LEU493 GLY494 VAL495 HSD496 TRP497 ARG498 PHE499
ASP500 ALA501 ALA502 GLY551 GLY552 VAL553

Crystal water molecules in the active region (48 atoms)

Solvent water molecules in the active region (6 atoms)

Four QM regions were investigated for each of the selected models in



QM region (Number of atoms)	Residues included
I (19 atoms)	Vanadate moiety (VAN), HIS496
II(37 atoms)	I + HIS404, CRYW200, CRYW333
III(71 atoms)	II + LYS353, ARG360, ARG490
IV(96 atoms)	III + PRO401, SER402, GLY403

Figure 4.2.3 QM regions (I-IV), the insert lists the residues that are included in each of the QM regions (for the VOD34 example).

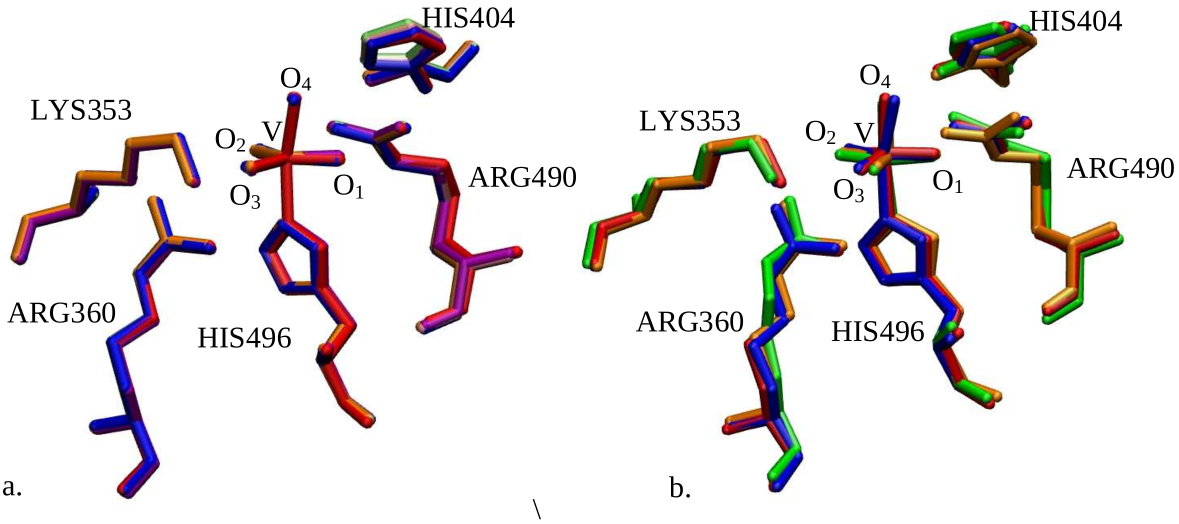
General Considerations

QM/MM results are available for each of the 10 selected models (Table 4.2.1) with four different QM regions I – IV of increasing size (Figure 4.2.3). The results for the largest QM region IV should be the most reliable and will therefore be discussed preferentially, while those for the smaller QM regions I – III will be presented primarily in the context of establishing their convergence (or sensitivity) with regard to the size of the QM region. In each system, we have generated data for 6 individual snapshots (see section 3.2) which will not be discussed individually, but only in terms of their mean values and standard deviations. The latter gives an indication of how much the result may vary due to fluctuations in the protein environment (as encountered along an MD trajectory).

In the following we shall focus on geometries and NMR properties. We shall not address energetics, mainly for two reasons: Firstly, it is extremely demanding to compute reliable relative QM/MM energies for models (Table 4.2.1) that differ in the protonation state of the central vanadate moiety, because of long-range electrostatic interactions and the associated need for extensive sampling. Secondly, in our current computational setup, the various models and snapshots are generated independently, and hence the active (optimized) regions will generally be different so that the energies are no longer directly comparable. We have not attempted to remedy this latter issue, because our primary interest is on geometries and NMR properties where one can compare the computational results with experimental data from X-ray crystallography and ^{51}V NMR spectroscopy. Active-site geometries and chemical shifts are expected to be essentially local properties that should not depend too much on the more distant protein environment (certainly less so than the total energy).

Geometries

In general there appears to be excellent geometric convergence (for individual models) across the six snapshots after QM/MM optimization. Figure 4.2.4a displays an overlay of the heavy atoms in QM region III for the 6 snapshots of VOT144 to illustrate this small deviation. Figure 4.2.4b shows an overlay of the heavy atoms for QM region III for three different models and highlights that the location of the protons can significantly perturb the geometry around the vanadate moiety.



\

Figure 4.2.4 a). Overlay of the six optimized snapshots (VOT144 model, QM region III) showing small variation in the geometry crystallographic water is omitted for clarity. **b)**. Overlay of the heavy atoms in the active site for X-ray[39] (orange), VOD144(blue), VOD14(red) and VOD34(green) for QM region III with noticeable variation amongst different models. Note that the complete residues are plotted here; see Figure 4.2.3 for the atoms included in the actual QM regions.

The X-ray structure of VCPO solved by Messerschmidt et al.[39] has three almost identical equatorial VO bond lengths (within ~0.04 Å), see top of Table 4.2.3. The axial VO distance is reported to be longer than the equatorial ones by around 0.26 Å. This would suggest a protonated axial oxygen and three deprotonated equatorial oxygen atoms. However due to the current resolution limitations of X-ray crystallography, we consider a wider range of possible protonation states. The average bond lengths and standard deviations for the atoms coordinated to vanadium are given in Table 4.2.3 for QM regions I to IV.

A first glance at the standard deviations in Table 4.2.3 (values in parentheses) reveals that the geometrical parameters for each model and QM region are generally very similar within the 6 selected snapshots. For most of the cases included in the table, this standard deviation is no larger than 0.02 Å. This finding indicates that in each case, there is predominant population of minima that are closely related as far as the geometry of the active site is concerned, and that there are only small perturbations in the latter due to the protein environment.

Overall the results for each of the models are fairly insensitive toward extension of the QM regions from I to IV. Apparently, the effects of the local environment on the vanadate geometry are reasonably well described already by the smallest QM region and the electrostatic embedding scheme. There are two notable exceptions, however, namely models VOT134 and VOT234: In these cases, a proton is transferred from O^4 to HIS404 after the QM region is extended to include the two crystal water and HIS404 moieties. This proton transfer is observed in all 6 snapshots during optimization, effectively converting these triply protonated models into the respective doubly protonated ones, with H atoms remaining at O^1,O^3 and O^2,O^3. Concomitantly, the V-O^4 bond length decreases by ca. 0.13 Å as the OH ligand is transformed into a terminal oxo atom (compare QM I and IV data for VOT134 and VOT234 models in Table 4.2.3). Even in these cases, the distances involving the equatorial oxygen atoms O1 – O3 are not strongly affected by increasing the QM region.

Interestingly, in those cases where a water molecule is placed at the axial position (VOT144, VOT344, and VOD44), the V-O^4 bond length also decreases noticeably (by 0.04 – 0.13 Å) upon going from QM region I to IV. In these cases, however, no proton transfer is observed and the water molecules stay intact throughout, forming a persistent hydrogen bond to HIS404. Arguably, the bond between vanadium and water is weaker than that involving an hydroxy or oxo ligand, and is thus more sensitive to specific hydrogen-bonding interactions and the way these are treated computationally.

Table 4.2.3 Bond lengths involving atoms coordinated to vanadium for QM/MM regions I - IV. Mean bond lengths and standard deviations are reported in Å. X-ray refined experimental values[39] are shown for comparison.

		V-N^{HIS496}	V-O^{1}	V-O^{2}	V-O^{3}	V-O^{4}
X-ray[39]		2.08	1.60	1.61	1.64	1.88
VOT134	I	2.15 (0.02)	1.87 (0.02)	1.63 (0.00)	1.84 (0.01)	1.80 (0.01)
	II	2.34 (0.09)	1.87 (0.03)	1.66 (0.01)	1.89 (0.01)	1.69 (0.05)
	III	2.35 (0.11)	1.88 (0.04)	1.66 (0.01)	1.89 (0.01)	1.69 (0.06)
	IV	2.36 (0.11)	1.88 (0.03)	1.66 (0.01)	1.88 (0.02)	1.69 (0.06)
VOT234	I	2.14 (0.01)	1.63 (0.00)	1.90 (0.00)	1.80 (0.01)	1.83 (0.00)
	II	2.38 (0.23)	1.66 (0.00)	1.91 (0.00)	1.88 (0.00)	1.67 (0.00)
	III	2.32 (0.01)	1.66 (0.00)	1.92 (0.00)	1.90 (0.00)	1.67 (0.01)
	IV	2.32 (0.01)	1.66 (0.00)	1.93 (0.00)	1.89 (0.00)	1.67 (0.00)
VOT234[a]	I	2.14 (0.01)	1.63 (0.00)	1.89 (0.00)	1.81 (0.00)	1.82 (0.00)
	II	2.28 (0.02)	1.67 (0.00)	1.92 (0.00)	1.90 (0.01)	1.66 (0.00)
	III	2.29 (0.03)	1.67 (0.00)	1.93 (0.02)	1.89 (0.01)	1.65 (0.00)
	IV	2.28 (0.02)	1.67 (0.01)	1.92 (0.02)	1.90 (0.02)	1.66 (0.00)
VOT144	I	2.04 (0.02)	1.87 (0.00)	1.65 (0.00)	1.65 (0.00)	2.18 (0.01)
	II	2.09 (0.01)	1.90 (0.00)	1.65 (0.00)	1.66 (0.00)	2.05 (0.01)
	III	2.10 (0.01)	1.90 (0.01)	1.66 (0.00)	1.66 (0.01)	2.06 (0.01)
	IV	2.11 (0.00)	1.90 (0.00)	1.66 (0.00)	1.66 (0.00)	2.05 (0.01)
VOT344	I	2.07 (0.00)	1.66 (0.00)	1.66 (0.00)	1.87 (0.00)	2.11 (0.00)
	II	2.10 (0.00)	1.66 (0.00)	1.66 (0.00)	1.87 (0.00)	2.06 (0.00)
	III	2.10 (0.00)	1.66 (0.00)	1.66 (0.00)	1.88 (0.00)	2.06 (0.00)
	IV	2.11 (0.00)	1.66 (0.01)	1.66 (0.00)	1.90 (0.01)	2.07 (0.00)
VOD14	I	2.21 (0.02)	1.91 (0.01)	1.66 (0.01)	1.66 (0.01)	1.89 (0.01)
	II	2.25 (0.01)	1.90 (0.01)	1.66 (0.02)	1.66 (0.02)	1.89 (0.01)
	III	2.24 (0.01)	1.91 (0.01)	1.67 (0.00)	1.67 (0.00)	1.90 (0.02)
	IV	2.24 (0.01)	1.91 (0.01)	1.67 (0.00)	1.67 (0.00)	1.90 (0.02)
VOD34	I	2.20 (0.02)	1.67 (0.00)	1.66 (0.01)	1.91 (0.01)	1.90 (0.01)

	II	2.20 (0.02)	1.67 (0.01)	1.67 (0.01)	1.92 (0.01)	1.91 (0.01)
	III	2.19 (0.01)	1.67 (0.01)	1.67 (0.01)	1.93 (0.01)	1.91 (0.02)
	IV	2.19 (0.01)	1.67 (0.01)	1.67 (0.00)	1.93 (0.01)	1.91 (0.02)
VOD44	I	2.10 (0.00)	1.70 (0.00)	1.69 (0.00)	1.69 (0.00)	2.24 (0.01)
	II	2.13 (0.01)	1.70 (0.00)	1.69 (0.00)	1.70 (0.00)	2.16 (0.04)
	III	2.13 (0.01)	1.70 (0.00)	1.70 (0.00)	1.70 (0.00)	2.16 (0.05)
	IV	2.14 (0.00)	1.70 (0.00)	1.71 (0.00)	1.71 (0.00)	2.14 (0.05)
VOS1	I	2.15 (0.02)	1.97 (0.00)	1.69 (0.01)	1.70 (0.00)	1.72 (0.00)
	II	2.32 (0.02)	1.95 (0.03)	1.68 (0.01)	1.68 (0.01)	1.79 (0.05)
	III	2.34 (0.04)	1.98 (0.02)	1.69 (0.01)	1.68 (0.02)	1.78 (0.06)
	IV	2.29 (0.05)	1.94 (0.02)	1.67 (0.01)	1.67 (0.01)	1.67 (0.01)
VOS3	I	2.54 (0.04)	1.71 (0.00)	1.69 (0.00)	1.95 (0.01)	1.70 (0.00)
	II	2.41 (0.02)	1.70 (0.00)	1.68 (0.00)	1.95 (0.00)	1.75 (0.00)
	III	2.44 (0.02)	1.70 (0.00)	1.69 (0.00)	1.96 (0.01)	1.74 (0.00)
	IV	2.42 (0.01)	1.70 (0.00)	1.69 (0.00)	1.98 (0.00)	1.74 (0.00)
VOS4	I	2.39 (0.03)	1.71 (0.00)	1.70 (0.00)	1.71 (0.00)	1.92 (0.00)
	II	2.37 (0.01)	1.72 (0.00)	1.70 (0.00)	1.71 (0.00)	1.95 (0.00)
	III	2.38 (0.01)	1.73 (0.00)	1.72 (0.00)	1.71 (0.00)	1.93 (0.00)
	IV	2.32 (0.01)	1.72 (0.00)	1.72 (0.00)	1.71 (0.00)	1.93 (0.00)

[a] VOT234 model with no neutrality imposed. (Total system charge of -19e)

Optimized V-O bond distances invariably adopt values that are typical for the particular ligand involved, around ca. 1.7 Å, 1.8-2.0 Å, and 2.1 Å for oxo, hydroxy, and water ligands, respectively, with a slight tendency to increase with the amount of negative charge on the central vanadate unit (e.g. note the elongation of the V-O^3 distance from 1.90 to 1.98 Å on going from VOT344 via VOD34 to VOS3, i.e. as the overall charge on vanadate successively increases from zero to –2).

The V-N bond length involving HIS496 is somewhat more variable across models, ranging from ca. 2.1 – 2.4 Å. Closest V-N contacts are observed when the other axial ligand trans to HIS96 is a water molecule (as in VOT144, VOT344, and VOD44), longer values are found when this is just an oxo atom (as in VOS1, VOS3, or in VOT134 and VOT234 after the aforementioned proton transfer).

Table 4.2.4 Root-mean-square deviations (RMSD, in Å) for all heavy atoms included in the QM region IV, relative to the X-ray-derived structure (1IDQ.pdb).[39]

	I	IV
VOT134	0.31 ± 0.01	0.30 ± 0.02
VOT234	0.33 ± 0.01	0.41 ± 0.01
VOT144	0.26 ± 0.01	0.26 ± 0.01
VOT344	0.32 ± 0.03	0.32 ± 0.01
VOD14	0.30 ± 0.03	0.32 ± 0.02
VOD34	0.35 ± 0.02	0.36 ± 0.04
VOD44	0.22 ± 0.00	0.22 ± 0.01
VOS1	0.28 ± 0.01	0.32 ± 0.02
VOS3	0.24 ± 0.01	0.26 ± 0.01
VOS4	0.28 ± 0.02	0.30 ± 0.02

Unfortunately, a direct comparison of the optimized and observed distances appears difficult due to the limitations set by the resolution achievable experimentally (in this case, 2.03 Å resolution with R = 19.7%, affording mean positional errors of the atoms of ±0.24 Å).[39] Models VOS1 and VOS3 appear to be less likely candidates because the computed V-O distances to the equatorial O atoms are much longer than that involving the axial O^4 (by up to 0.27 Å for the protonated O atoms), whereas the opposite trend is observed experimentally.[39] However, assessment of individual models based on such singular deviations may not be very reliable. To quantify the overall difference between optimized QM/MM models and the X-ray structure the RMSD values for non-hydrogen atoms are reported in. These RMSD values show a rather modest variation and make it difficult to advocate or exclude any model as a possible candidate for the resting state of VCPO based solely upon the RMSD. Interestingly, the RMSD values obtained by Carloni[43] show significantly more variation amongst models and are also much larger in magnitude than observed here. The RMSD values do not significantly differ between the QM I and QM IV regions for individual models (compare columns 1 and 2 of Table 4.2.4).

Hydrogen Bonding

Hydrogen bonding interactions between the vanadate moiety and the surrounding protein matrix are important for active-site properties. They are best studied by QM/MM calculations with QM region IV, which by design contains all such possible direct hydrogen bonding interactions. The corresponding hydrogen-bond lengths around the vanadate moiety are presented in Table 4.2.5. The generally rather small standard deviations across the six snapshots (values in parentheses) indicate that the individual models of VCPO have a dominant hydrogen-bonding network and not a multitude of different networks. As expected, the hydrogen bonds formed with crystallographically resolved water molecules exhibit much larger variations due to the inherent mobility of water.

The O^1-ARG360 hydrogen bond is found to be around 2 Å for all models. This hydrogen bond is stable across the six snapshots (cf. the low standard deviation). The O^1-ARG490 distance shows greater variation and is rather longer (ranging up to 2.9 Å) than the O^1-ARG360 hydrogen bond. The O^2-LYS353 and O^2-GLY403 distances generally mirror each other and range from medium (1.8 Å) to large values (3.0 Å for O^2-LYS353). The O^3-SER402 follows the same basic pattern with stability across snapshots and variation between models. The H^4 atom (when present) is hydrogen bonded to HIS404 and generally has the shortest distance of all possible hydrogen bonds considered around the vanadate moiety. The initially (in QM region I) triply protonated models VOT134 and VOT234 lose the axial H^4 atom to the HIS404 moiety as mentioned previously, and the resulting H^4-HIS404 distances clearly indicate N-H single bonds, see final column of Table 4.2.5. In summary, the hydrogen bonding around the vanadate moiety is generally similar for different snapshots of a given model, but there is some variation between the models.

Table 4.2.5 Possible hydrogen bonds in QM/MM calculations using QM region IV. Mean bond lengths and standard deviations are reported in angstroms. CHARMM atom types are given in brackets and superscripted residue labels are consistent with the X-ray pdb file.

	O^1 - H $[HH11]^{ARG360}$	O^1- H $[HH22]^{ARG490}$	O^2 - H $[HZ]^{LYS353}$	O^2 - H $[HN]^{GLY403}$	O^3 - HG $[HG]^{SER402}$	H^4 - N $[ND1]^{HIS404}$
VOT124	2.15 (0.21)	2.40 (0.44)	1.81 (0.04)	1.83 (0.02)	1.90 (0.05)	1.15 (0.25)
VOT234	2.00 (0.02)	2.04 (0.04)	1.79 (0.01)	1.93 (0.01)	2.09 (0.03)	1.04(0.00)
VOT144	1.84 (0.02)	2.90 (0.06)	2.97 (0.81)	1.88 (0.02)	2.13(0.08)	1.58 (0.02)
VOT344	1.95 (0.02)	2.27 (0.07)	1.82 (0.00)	1.96 (0.01)	1.88 (0.01)	1.68 (0.01)
VOD14	1.92 (0.08)	2.88 (0.11)	1.75 (0.01)	1.78 (0.00)	2.20 (0.38)	1.75 (0.01)
VOD34	1.93 (0.08)	2.79 (0.23)	1.79 (0.08)	1.89 (0.15)	1.99 (0.13)	1.93 (0.06)
VOD44	1.95 (0.08)	2.01 (0.02)	1.56 (0.03)	1.74 (0.00)	1.79 (0.02)	1.78 (0.01)
VOS1	2.07 (0.02)	1.74 (0.01)	1.72 (0.02)	2.00 (0.21)	2.52 (0.17)	-
VOS3	1.99 (0.02)	1.84 (0.01)	1.95 (0.01)	1.99 (0.01)	1.67 (0.00)	-
VOS4	1.84 (0.04)	1.84 (0.04)	1.62 (0.01)	1.74 (0.01)	1.73 (0.01)	1.87 (0.00)

Isotropic ^{51}V NMR Chemical Shifts

Computed isotropic chemical shifts are obtained as the difference between magnetic shieldings and can thus benefit from error cancellation. This applies e.g. for the neglect of relativistic effects, which are especially important for core orbitals, but tend to be quite transferable for 3d transition metals in different environments.[124] The computed isotropic ^{51}V chemical shifts (δ_{iso}) are displayed in Table 4.2.6.

Table 4.2.6 ^{51}V isotropic chemical shifts (ppm) averaged over 6 snapshots, together with the corresponding standard deviations. For each of the QM/MM models considered, results are given for QM regions I – IV. VOS4* indicates HIS404 being doubly protonated. VOS4$^+$ indicates model where 30 Å solvent sphere of TIP3 centered on vanadate moiety.

	I	**II**	**III**	**IV**
VOT124	-550 ± 6	-635 ± 2	-623 ± 8	-618 ± 11
VOT234	-540 ± 4	-587 ± 3	-585 ± 3	-583 ± 3
VOT344	-610 ± 1	-620 ± 2	-610 ± 2	-602 ± 2
VOT144	-606 ± 3	-617 ± 2	-609 ± 3	-602 ± 2
VOD14	-615 ± 8	-620 ± 11	-596 ± 5	-580 ± 13
VOD34	-623 ± 6	-628 ± 11	-618 ± 2	-607 ± 3
VOD44	-541 ± 3	-558 ± 9	-547 ± 10	-541 ± 10
VOS1	-645 ± 2	-627 ± 18	-621 ± 16	-615 ± 18
VOS3	-651 ± 1	-629 ± 2	-619 ± 1	-610 ± 1
VOS4	-560 ± 3	-570 ± 2	-564 ± 2	-553 ± 2
VOS4$^+$	-572	-573	-	558
VOS4*	-554 ± 3	-564 ± 3	-553 ± 3	-546 ± 4

A cursory glance at Table 4.2.6 shows that the δ_{iso} values of all models are rather stable across the QM regions II – IV, where variations typically amount to less than 20 ppm. Fluctuations across QM regions I – IV can be somewhat larger, in particular for the first two triply protonated models, for which deprotonation of O^4 has occurred upon inclusion of HIS404 into the QM region. Concomitant with this proton transfer, a noticeable shielding of the ^{51}V nucleus is found, by ca. –70 ppm and –45 ppm for VOT124 and VOT234, respectively.

Starting from the equilibrated VOS4 model, a new model was generated in which an additional proton was placed on HIS404, labeled VOS4* in Table 4.2.6. Interestingly, the additional proton on HIS404 did not transfer back to the singly protonated vanadate moiety. There is no large difference in the isotropic chemical shift due to the additional proton. Therefore the protonation state of the HIS404 has no direct effect on the isotropic chemical shifts. However, an indirect effect is observed when the deprotonated HIS404 accepts the axial proton in the case of VOT124 and VOT234.

The isotropic NMR shifts range from around -550 ppm to -620 ppm for QM region IV, without showing a clear dependence on the protonation state. This range is dominated by values of the VOS4 and VOD44 outliers, which are considerably more deshielded than the other models. The range for the remaining models is only around 40 ppm, which implies that all models are potential candidates for the experimentally observed structure. The experimentally measured isotropic chemical shift was reported to be -507.5 ppm[46] which is noticeably less shielded than all models shown in Table 5. This is consistent with previous computational results at the same or comparable QM levels for other vanadates,[47,125] where the computed isotropic value is generally too strongly shielded by roughly 100ppmThus, while VOS4 and VOD44 would seem to be the most likely candidates based on the best agreement between computed and experimental isotropic shifts, this assignment cannot be upheld when the systematic errors of the QM method are taken into consideration.

A test calculation was performed for one of the singly protonated models, VOS4 (denoted VOS4^{+} in Table 4.2.6), where a 30 Å solvent sphere of TIP3 water was centered on vanadates moiety, instead of solvating the full enzyme (cf. Figure 4.2.2).

There is no significant variation in isotropic chemical shifts (see Table 4.2.6), thus confirming that ^{51}V isotropic chemical shifts are independent of the types of solvation scheme and the number of solvent molecules far away from the active region.

In summary, the isotropic ^{51}V chemical shifts are found to be a poor discriminator of likely protonation configurations for QM/MM models of VCPO, due to the small variation resulting from changes in the proximal proton environment.

Anisotropic ^{51}V NMR Chemical Shifts: The CSA and EFG Tensors

There are four parameters that characterize the experimental ^{51}V NMR anisotropic spectra, namely the reduced anisotropy δ_σ, asymmetry η_σ, nuclear quadrupole coupling constant C_Q(MHz), and asymmetry parameter η_Q. Their experimentally refined numerical values[46] are given at the top of Table 4.2.7. These same parameters ($\delta_\sigma, \eta_\sigma, C_Q, \eta_Q$) were computed from the CSA and EFG tensors obtained from the QM/MM optimized models and these values are collated in Table 4.2.7. The mean absolute percent deviation (MAPD) of all four parameters for an individual model with the associated standard deviation (SD) is also included in Table 4.2.7. A detailed discussion of the deviations between individual parameters computed from QM/MM models and the experimentally reported values is not given in the interest of brevity.

In general, the computed δ_σ values can be somewhat more sensitive to the size of the QM region than the isotropic chemical shifts. As expected, changes between QM region I and II are particularly pronounced for the initially triply protonated models VOT134 and VOT234, due to the concomitant proton transfer to HIS496. Further extension of the QM region usually produces relatively small changes in δ_σ; the same trend is observed for most of the other models (with some exceptions, notably VOD14 and VOD34), and in the following we will only discuss results obtained with the largest QM region IV.

The δ_σ values computed for five of the models (VOT144, VOT344, VOD34, VOD14, and VOS1) fall into the range of ±100 ppm from experiment, while the remaining models show larger shifts to more positive values. Very large deviations are found for those models that bear no protons on the equatorial oxygen atoms, VOD44 and VOS4. For these models positive δ_σ values are predicted, clearly incompatible with experiment, where a negative value is found for this property. It is noteworthy that these are the very models that might have been anticipated based on the raw structural data (VO distances) and from an uncritical assessment of the isotropic ^{51}V chemical shifts (see above). That these models can now be safely excluded based on the anisotropic chemical shifts is testimony to the usefulness of solid-state NMR in combination with quantum-chemical calculations.

Only the absolute value of C_Q can be determined experimentally, not its sign. There are two models that miss the experimental absolute value, C_Q = 10.5 MHz, by an uncomfortably large margin, namely VOT134 and VOD44, which under- and overestimate C_Q by nearly 4 and 8 MHz, respectively (QM IV results in Table 4.2.7). The asymmetry parameters can only assume values between zero and 1, and both η_σ and η_Q have been refined experimentally to an intermediate value. The VOT134 and VOT344 models (together with VOD44 and VOS4, which are already excluded based on their δ_σ values) afford η values that are significantly too low compared to experiment.

A critical assessment of the models would require precise knowledge of the inherent errors of the applied QM method for each of the CSA and EFG tensor elements. A conservative estimate for these errors is ca. ±20%,[46] that is, computed values are considered satisfactory when they are within this range from experiment. The absolute percent deviation, averaged over the four tensor parameters, is included in Table 4.2.7 (MAPD values, last column). In a ranking based on this value, three models emerge as the most promising candidates, namely VOT144, VOD14, and VOD34, which all have absolute percent errors as low as 11-12%. One of the singly protonated models, VOS1, also shows a satisfactory agreement with experiment (MAPD 18%). However, this model has not been identified previously as a low energy structure (neither from gas-phase or QM/MM studies) and is therefore not thought to represent the resting state of VCPO.

Table 4.2.7 Reduced anisotropy δ_σ, asymmetry η_σ, nuclear quadrupole coupling constant C_Q(MHz), and asymmetry parameter η_Q with the associated standard deviations for six snapshots. The mean absolute percent deviation (MAPD) is also tabulated. Results are given for QM regions I to IV.

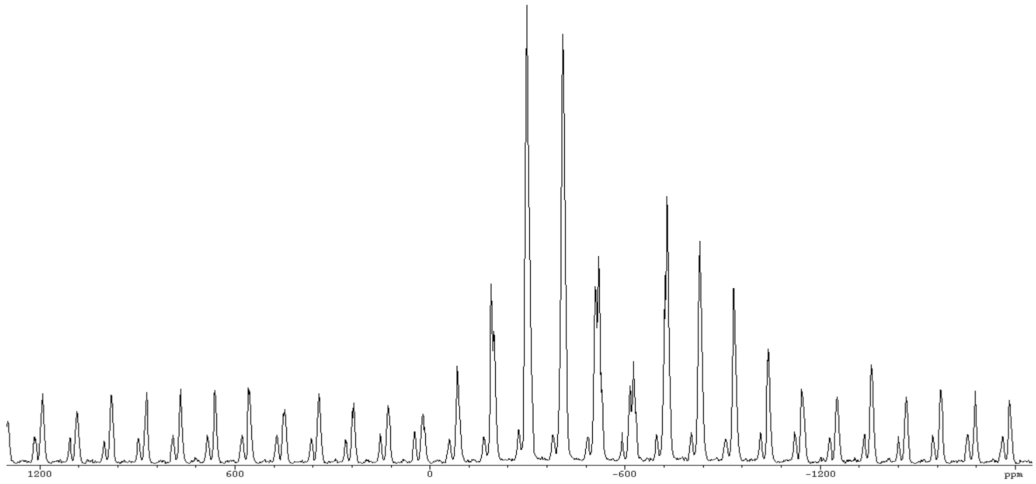
		δ_σ	η_σ	C_Q	η_Q	MAPD(SD)
EXP[46]		-520 ± 13	0.4 ± 0.05	10.5 ± 1.5	0.55 ± 0.15	
VOT134	I	-414 ± 53	0.6 ± 0.13	-6.0 ± 0.5	0.24 ± 0.06	42 ± 16
	II	-252 ± 73	0.8 ± 0.19	-7.5 ± 0.7	0.15 ± 0.06	63 ± 30
	III	-283 ± 92	0.7 ± 0.20	-7.5 ± 0.8	0.13 ± 0.10	56 ± 23
	IV	-280 ± 91	0.2 ± 0.21	-6.9 ± 0.9	0.16 ± 0.11	50 ± 15
VOT234	I	-568 ± 3	0.2 ± 0.01	4.6 ± 0.1	0.63 ± 0.04	32 ± 24
	II	-375 ± 10	0.5 ± 0.02	-8.4 ± 0.1	0.25 ± 0.02	31 ± 15
	III	-369 ± 10	0.5 ± 0.03	-8.5 ± 0.1	0.24 ± 0.03	32 ± 16
	IV	-354 ± 8	0.5 ± 0.02	-8.6 ± 0.1	0.36 ± 0.02	27 ± 7
VOT234[a]	I	-563 ± 14	0.2 ± 0.05	5.83 ± 0.7	0.39 ± 0.11	33 ± 19
	II	-371 ± 25	0.7 ± 0.05	-8.6 ± 0.5	0.43 ± 0.11	35 ± 26
	III	-369 ± 20	0.6 ± 0.02	-9.1 ± 0.5	0.45 ± 0.15	27 ± 16
	IV	-343 ± 15	0.6 ± 0.04	-9.0 ± 0.5	0.47 ± 0.13	28 ± 17
VOT144	I	-605 ± 15	0.3 ± 0.03	14.0 ± 0.4	0.44 ± 0.03	23 ± 7
	II	-648 ± 19	0.4 ± 0.04	12.3 ± 0.4	0.68 ± 0.02	16 ± 11
	III	-619 ± 15	0.4 ± 0.04	12.0 ± 0.4	0.51 ± 0.10	10 ± 8
	IV	-600 ± 19	0.4 ± 0.05	11.9 ± 0.4	0.45 ± 0.09	11 ± 8
VOT344	I	-570 ± 3	0.4 ± 0.01	14.1 ± 0.1	0.11 ± 0.02	31 ± 36

	II	-597 ± 5	0.4 ± 0.01	12.6 ± 0.0	0.14 ± 0.03	27 ± 33
	III	-584 ± 5	0.4 ± 0.01	12.6 ± 0.1	0.18 ± 0.02	25 ± 29
	IV	-587 ± 5	0.5 ± 0.01	12.8 ± 0.1	0.22 ± 0.03	29 ± 21
VOD14	I	-618 ± 7	0.4 ± 0.01	-5.8 ± 0.4	0.67 ± 0.07	21 ± 19
	II	-549 ± 11	0.4 ± 0.02	-6.3 ± 0.2	0.58 ± 0.08	12 ± 18
	III	-575 ± 5	0.4 ± 0.02	9.3 ± 0.3	0.79 ± 0.03	16 ± 19
	IV	-499 ± 26	0.5 ± 0.03	9.3 ± 0.0	0.54 ±0.03	11 ± 10
VOD34	I	-610 ± 34	0.3 ± 0.03	6.4 ± 0.3	0.78 ± 0.04	31 ± 16
	II	-529 ± 13	0.4 ± 0.03	6.6 ± 0.1	0.78 ± 0.01	20 ± 22
	III	-518 ± 8	0.3 ± 0.03	9.5 ± 0.1	0.30 ± 0.05	20 ± 20
	IV	-551 ± 16	0.3 ± 0.03	9.5 ± 0.1	0.51 ± 0.04	12 ± 9
VOD44	I	415 ± 4	0.4 ± 0.01	20.1 ± 0.1	0.05 ± 0.02	91 ± 73
	II	435 ± 5	0.3 ± 0.03	19.3 ± 0.5	0.14 ± 0.09	92 ± 66
	III	417 ± 3	0.3 ± 0.02	18.7 ± 0.7	0.30 ± 0.13	82 ± 69
	IV	417 ± 11	0.2 ± 0.03	18.2 ± 0.9	0.18 ± 0.09	93 ± 59
VOS1	I	-317 ± 6	0.3 ± 0.01	9.1 ± 0.1	0.44 ± 0.00	24 ± 11
	II	-423 ± 63	0.4 ± 0.08	8.6 ± 0.5	0.80 ± 0.14	21 ± 19
	III	-392 ± 104	0.4 ± 0.06	8.6 ± 0.2	0.71 ± 0.20	18 ± 13
	IV	-404 ± 82	0.3 ± 0.08	8.1 ± 0.3	0.49 ± 0.17	18 ± 6
VOS3	I	-236 ± 8	0.4 ± 0.01	8.6 ± 0.1	0.27 ± 0.03	33 ± 23
	II	-309 ± 10	0.3 ± 0.03	8.6 ± 0.1	0.32 ± 0.05	31 ± 12
	III	-274 ± 10	0.3 ± 0.01	8.5 ± 0.0	0.24 ± 0.02	37 ± 18
	IV	-295 ± 6	0.3 ± 0.01	8.5 ± 0.0	0.24 ± 0.02	37 ± 16
VOS4	I	334 ±.2	0.6 ± 0.01	13.6 ± 0.1	0.27 ± 0.02	71 ± 63
	II	353 ± 3	0.5 ± 0.01	14.5 ± 0.1	0.14 ± 0.01	76 ± 65
	III	331 ± 2	0.6 ± 0.01	13.0 ± 0.1	0.26 ± 0.02	73 ± 62
	IV	330 ± 2	0.6 ± 0.01	12.9 ± 0.1	0.07 ± 0.01	79 ± 62
VOS4*	I	380 ± 4	0.2 ± 0.02	14.4 ± 0.1	0.14 ± 0.03	83 ± 62

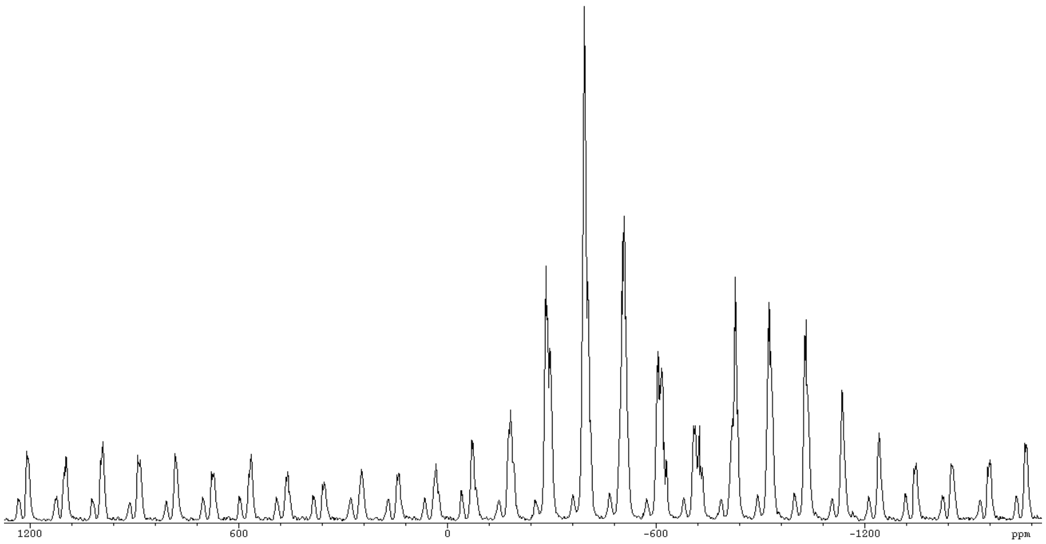
II	439 ± 3	0.3 ± 0.02	19.4 ± 0.1	0.10 ± 0.01	94 ± 66
III	423 ± 3	0.3 ± 0.02	18.9 ± 0.1	0.13 ± 0.01	91 ± 65
IV	419 ± 4	0.1 ± 0.02	18.5 ± 0.1	0.07 ± 0.01	104 ± 51

[a] VOT234 model with no neutrality imposed. (Total system charge of -19e), VOS4* indicates HIS404 being doubly protonated

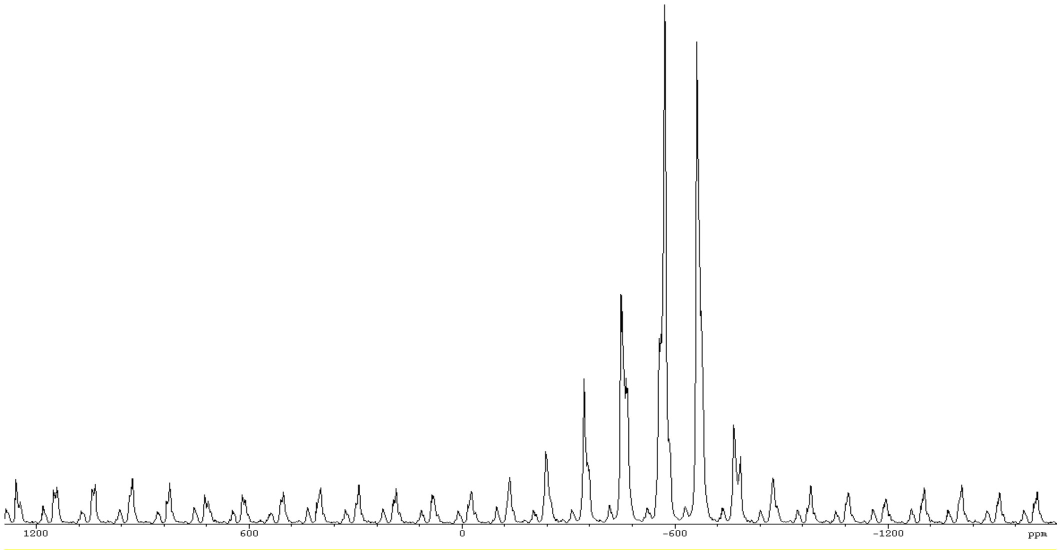
Figure 4.2.5 Simulated 51V MAS spectra for VCPO. The spectra were plotted using the Simpson program[126.] and the NMR and NQR tensor elements taken from experiment (top) and QM/MM computations for models VOT144 (middle) and VOS1 (bottom). As the Euler angles were found to have only very minor effects on the appearance of the spectra, they were set to zero throughout.



a. VCPO-expt.



b.Good candidate:VOT144



c. Poor candidate:VOS4

The quality of the computed NMR parameters can be further gauged by a comparison of the actual solid-state NMR spectrum against those simulated using the DFT-computed NMR parameters. As can be seen in Figure 4.2.5, the overall experimental pattern (top spectrum) is reasonably well reproduced by the "good" candidates (e.g. for VOT144 in the middle spectrum), whereas the other models can show larger discrepancies or even no resemblance whatsoever (e.g. VOS4 in the lower spectrum).

To summarize this section, Table 4.2.7 contains three models that are significantly closer to the experimentally determined parameters than all other candidates. These models are the triply protonated VOT144 and two of the doubly protonated models namely VOD14 and VOD34. Unfortunately, further discrimination between these candidates is not possible at present. Interestingly these models have been identified in the previous QM/MM study by Carlson and co-workers[42] based upon energy. Therefore, by demonstrating that these models can reproduce the spectroscopical observables, we provide compelling evidence that one of them (or a mixture) is indeed a faithful representation of the actual ground state of the enzyme.

Electrostatic and Geometric Effects on the ^{51}V NMR Chemical Shifts

The environment (i.e. everything outside the QM region) is modeled as fixed point charges for the calculation of ^{51}V NMR chemical shifts from the QM/MM optimized geometries. An increase in the size of the QM region led to convergence (generally a downfield shift) in the ^{51}V NMR isotropic chemical shifts (see Figure 4.2.5 and associate discussion). A similar behavior was noted above for the anisotropic ^{51}V NMR chemical shifts, which turned out to be somewhat more sensitive. Despite being a local property, the chemical shifts are thus influenced by the definition of the QM region. We now address the question whether this is due to either electrostatic or geometric differences between the QM/MM partitioning schemes or some combination of both, and discuss the effect of the electrostatic embedding upon the wave function and subsequently the ^{51}V NMR chemical shifts for a few representative cases.

Calculations[127] without surrounding point charges were performed to probe the direct electrostatic effect of the protein environment on the ^{51}V chemical shifts. Model A uses the atomic coordinates of all the QM atoms in the QM/MM optimized VOT134 structure with QM region I. There is a large difference in the parameters that characterize the CSA and EFG tensors as seen by significant deviations (Table 4.2.8) between model A without point charges and model VOT134-I with these point charges (data as shown in Table 4.2.7). However, when an analogous deletion of point charges surrounding VOT134-IV is performed, labeled model B, a reasonable agreement (a low MAPD) with the original model is observed. There are two obvious conclusions; firstly, the electrostatic embedding scheme used within this study does significantly affect the ^{51}V NMR chemical shifts for small QM regions. Secondly, if a large enough QM region is employed then the effects of such an embedding scheme become less influential on the chemical shift of the metal center. Hence, when calculating the ^{51}V NMR chemical shifts within a QM/MM framework, a good electrostatic representation of the MM region seems crucial when using small QM regions.

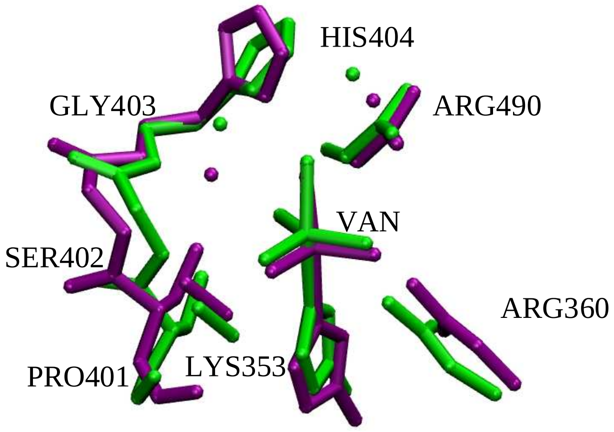


Figure 4.2.6 Overlay of the QM/MM-optimized VOD34 model in the enzyme (green) and the same model QM-optimized in the gas phase (purple, QM region IV employed in both cases, CRYW shown as unlabeled dots).

The next step was to determine the importance of the protein environment on the geometry of the QM region. Starting from the QM/MM optimized VOD34-IV model, the active and inactive MM regions were deleted. The QM region was then re-optimized at the RI-BP86/AE1 level of theory. The starting structure was not stable, i.e. a large structural difference between the original QM/MM optimized geometry and the new QM gas-phase geometry was observed (RMSD 0.79 Å). This is to be expected; as the protein environment is no longer included, the geometry is free to relax without any constraint (see Figure 4.2.5) for an illustration of the geometrical changes). Two separate calculations of CSA and EFG tensors were performed from this gas-phase optimized geometry; model C included only the subset of atoms which are defined in the QM I region and model D included all atoms defined in the QM IV region**.** Table 4.2.8 shows the deviation between the parameters calculated from models C and D and the analogous parameters optimized from the QM/MM optimized geometries. Immediately one can note significant discrepancies. Therefore the inclusion of the MM layer can have an extremely large effect on the ^{51}V NMR chemical shifts, and care needs to be taken when assessing results computed for models in the gas phase, even if these are quite sizable.

The geometric and electrostatic effects were further studied by taking the QM/MM optimized geometry of VOD34-IV and computing the NMR chemical shifts for the atoms defined in the QM region I. This is labeled as model E and the results are given in Table 4.2.8. The low deviation for the parameters computed from model E and the VOD34-IV shows that the small QM region is capable of reproducing the chemical shifts of the larger QM model if the same geometry is used.

The influence of the protein matrix on the ^{51}V chemical shifts is thus largely geometrical in nature, by favoring a particular ligand arrangement about the metal center. The "direct response" of the electronic wave functions in the QM part to the surrounding charge distribution is, in comparison, smaller with an appropriately sized QM region. Similar findings have been reported for solvent effects on $\delta(^{51}V)$ of small vanadates[125] and on other transition metal shifts, for example $\delta(^{57}Fe)$ [128] and $\delta(^{59}Co)$, [129] with somewhat larger "direct solvent" effects noted for $\delta(^{195}Pt)$.[130]

To conclude, a crucial advantage of QM/MM calculations is the ability to effectively optimize the whole system (QM and MM) to realistic geometries, which is a prerequisite for computing reliable chemical shifts. For the actual NMR calculations, it is best to use large QM regions, but realistic chemical shifts can already be obtained with small QM regions provided that the electrostatic effects of the environment are included at least via suitable point charges.

Table 4.2.8 The percent deviation (PD) between the models A-E and the original models (values from Table 4.2.7).The mean absolute percent deviation (MAPD) is also provided for an overall assessment of numerical similarity

Model ID	Brief Description	PD δ_σ	PD η_σ	PD C_Q	PD η_Q	MAPD
A	VOT234-I : EFG and CSA tensors calculated without point charges	-12	+45	+36	-55	37
B	VOT234-IV : EFG and CSA tensors calculated without point charges	-2	+14	-1	-3	5
C	VOD34-I : Optimization (at QM IV), EFG and CSA tensors calculated without point charges	+21	-4	-81	+32	35
D	VOD34-IV : Optimization (at QM IV), EFG and CSA tensors calculated without point charges	+18	+45	-43	+11	30
E	QM/MM optimized geometry of VOT34 IV with EFG and CSA tensors calculated including only atoms from QM region I plus all point charges	-4	+8	-2	+4	5

Conclusions

^{51}V NMR parameters of VCPO, for a large number of QM/MM optimized models, are computed, calling special attention to different possible protonation states and positional isomers. In terms of non-hydrogen atomic coordinates all of the many models considered show reasonable agreement with the initial X-ray structure. This makes exclusion of candidates based solely on geometric criteria (RMSD) rather difficult. The bond lengths for atoms coordinated to vanadium across individual models, as well as the computed ^{51}V NMR parameters, turned out to be rather insensitive to the conformational flexibility of VCPO. The latter was sampled by classical MD simulations with subsequent QM/MM optimizations starting from a number of snapshots. The isotropic ^{51}V chemical shifts were found to be a poor indicator of protonation states due to small variation amongst models.

Investigation of the EFG and CSA tensors has identified three models that agree with the experimental derived values from the ^{51}V NMR spectrum. These models, VOT144, VOD14, and VOD34, all have similar calculated isotropic and anisotropic chemical shifts, as well as similar geometries (only 0.1 Å difference in RMSD across the three models). Therefore these three models are equally consistent with both the X-ray and the NMR experimental. It is encouraging that by modeling the protein in a rational and systematic way the computed ^{51}V NMR anisotropic chemical shifts agree with the recently reported experimental ^{51}V NMR spectrum of VCPO. It is reassuring that the conclusions from the present spectroscopic QM/MM work are consistent with those from previous QM/MM studies,[42,43] which identified the same models as being good candidates solely on the basis of energetic considerations.

4.3 ^{51}V NMR Chemical shifts calculated from QM/MM models of Vanadium Bromoperoxidase – Native Form

QM/MM setup

The vanadate moiety has four crystallographically resolved oxygen atoms (one axial and three equatorial) as shown in Figure 1.2.1. Consistent with the previous study on VCPO, triply, doubly and singly protonated models are considered here along with a model that is fully deprotonated, which are hereafter referred to as VBPO-T, VBPO-D, VBPO-S and VBPO-Z respectively. The oxygen labels from the X-ray study (1QI9.pdb) are used throughout this work. Protonation at different vanadate oxygen atoms generates positional isomers. Those that are studied presently are listed in Table 4.3.1 which also introduces the chosen notation. The most notable difference between the immediate vicinities of the active site of these proteins is the non-polar PHE397 in VCPO being a more polar in HIS411 VBPO, which may interact via hydrogen bonding interactions with the vanadate moiety. In the previous chapter on VCPO, the models which included protonated O2 were not intensively studied because the closely located PHE397 will act as H-bond acceptor. The corresponding HIS411 in VBPO is such a potential H-bond parter, so that additional models protonated at O2 were considered here.

Table 4.3.1. Selected models labeled according to the protonation state (VBPO-T, VBPO-D and VBPO-S) and the oxygen atoms that are protonated (1,2,3,4).

Triply protonated	Doubly protonated	Singly protonated
VBPO-T144	VBPO-D14	VBPO-S1
VBPO-T344	VBPO-D24	VBPO-S2
VBPO-T234	VBPO-D34	VBPO-S3
Unprotonated	VBPO-D12	VBPO-S4
VBPO-Z	VBPO-D13	

The models are selected based on results of the previous work on VCPO, as there is high sequence homology in the active site and very similar X-ray refined coordinates. Preference has been given to models protonated at O4, as suggested by the rather long V-O_4 bond in the solid state (see below). Three QM regions were investigated for each of the selected models in Table 4.3.1 and Figure 4.3.1 displays the residues, which are included in each QM region, see Table 4.3.2. Note that QM region III herein (i.e. the largest one for VBPO) is analogous to region IV from the previous chapter on VCPO.[131] HIS411 has been included as this represents the major difference in the sequence of the active site of VBPO.

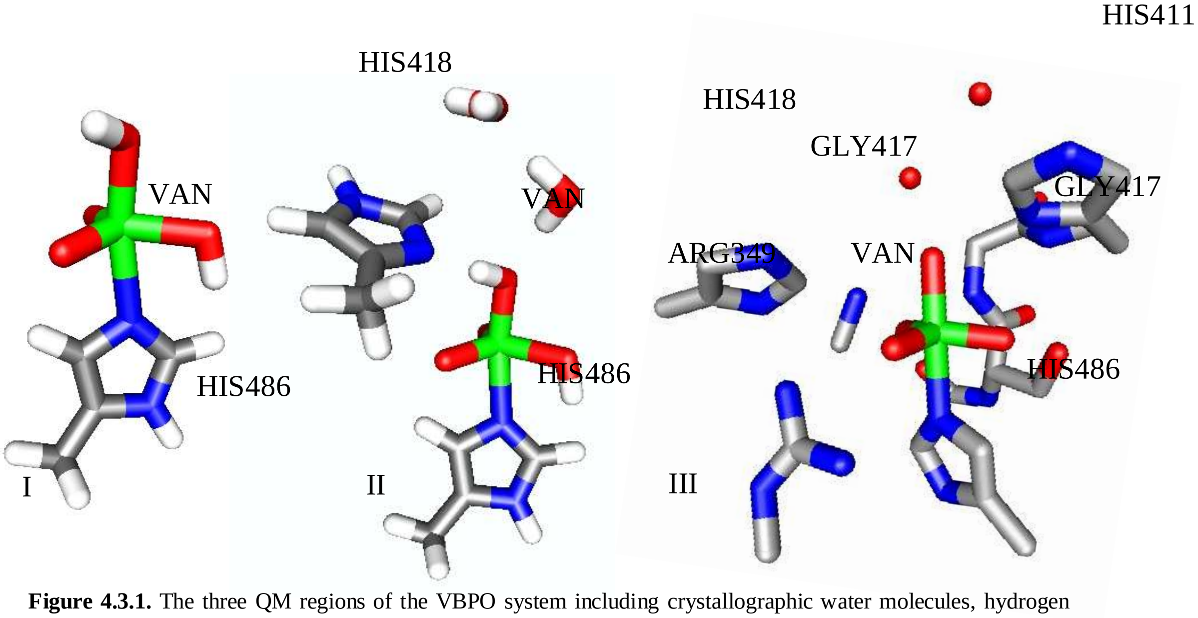


Figure 4.3.1. The three QM regions of the VBPO system including crystallographic water molecules, hydrogen atoms are omitted for clarity for QM region III.

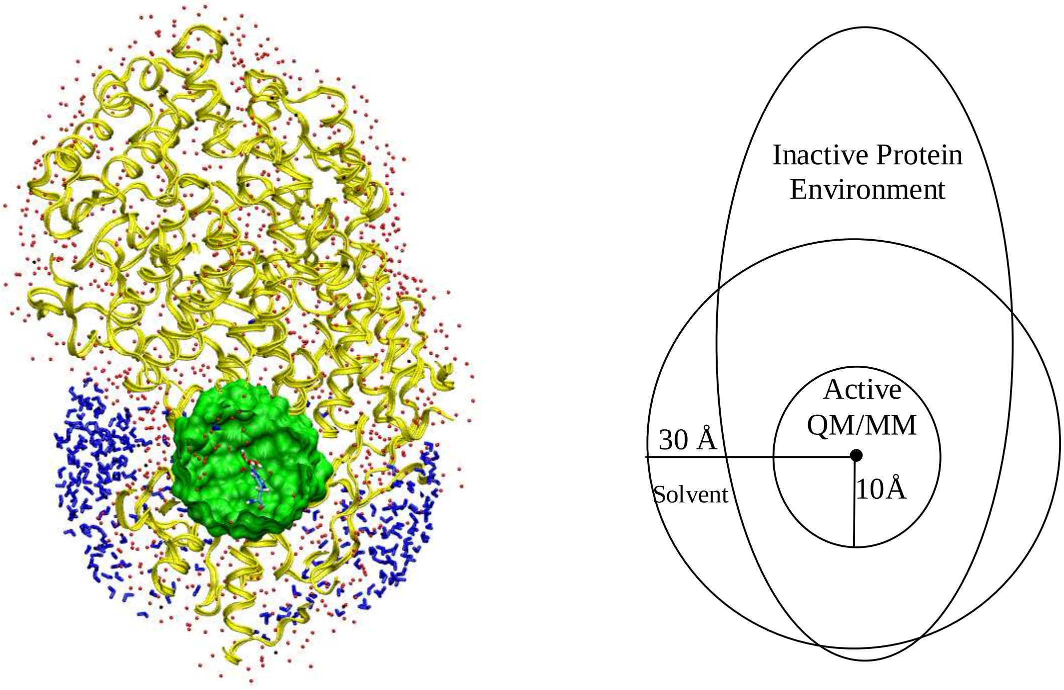
Table 4.3.2. The residues, which are, included in each of the QM regions for VBPO.

QM region (Number of atoms)	Residues included
I (19 atoms)	Vanadate moiety (VAN), HIS486
II(37 atoms)	I + HIS418, CRYW772, CRYW774
III(97 atoms)	II + LYS341, ARG349, GLY417, HIS411,SER416,PRO415

The residues within 10 Å of the vanadium cofactor were included in the active region (see Figure 4.3.2). The residues in the active region (~1000 atoms) are shown in Table 4.3.3.

Table 4.3.3 Example of the residues included in the active region for VOD24(1194 atoms, residue numbering according to that in 1QI9.pdb)

PHE53	LYS55	VAL102	ASN103	PRO104	THR105	ALA106	ILE122
VAL273	THR274	PHE275	THR276	ASP277	ASN278	ILE279	ASN280
THR281	GLU282	GLU330	GLN333	ARG334	SER335	SER336	TRP337
TYR338	GLN339	LYS340	TRP341	PHE346	ARG348	PRO349	GLU350
GLN401	ILE403	GLU405	GLY406	SER407	PRO408	THR409	HSE410
PRO411	SER412	TYR413	PRO414	SER415	GLY416	HSE417	ALA418
THR419	GLN420	ASN421	GLY422	PHE424	ASN474	VAL475	ALA476
PHE477	GLY478	ARG479	GLN480	MET481	LEU482	GLY483	ILE484
HSD485	TYR486	PHE488	ASP489	GLY490	GLY493	LEU496	ARG298
PHE314	VAL315	ILE319	SER320				



a.

b.

Figure 4.3.2 **a.** A cross section through the solvated protein: the region in blue is solvent (30 Å sphere centered on vanadium), the fixed protein environment during the QM/MM optimization (residues > 10 Å from vanadate) is displayed as a yellow ribbon. The green sphere centered on the vanadate moiety is the active region subjected to the QM/MM optimization (including all residues < 10 Å from vanadate). The vanadate moiety and the bound HIS486 are shown as a visual guide. Crystallographically resolved waters (oxygen position) are displayed as red spheres.

b. Schematic system partitioning with the filled black circle (roughly in the centre) representing the vanadate moiety.

QM/MM results are available for each of the selected models (Table 4.3.1) with three different QM regions I – III of increasing size (Figure 4.3.1 and Table 4.3.2). The results for the largest QM region III should be the most reliable and will therefore be discussed preferentially, while those for the smaller QM regions I –II will be presented primarily in the context of establishing their convergence (or sensitivity) with regard to the size of the QM region. In the majority of cases, we have generated data for 6 individual snapshots (see Computational details). As before, these will not be discussed individually, but only in terms of their mean values and standard deviations.

Geometries

The X-ray structure of the native form of VBPO solved by Weyand et al.[49] has the equatorial V-O^1 and V-O^2 bond lengths almost identical (within ~0.01 Å) with the V-O^3 being comparatively shorter (~0.08 Å), see top of Table 4.3.4. The axial VO distance is reported to be longer than the equatorial ones by around 0.17-0.25 Å. Both V-O^3 and V-O^4 distances are noticeably shorter in VBPO than VCPO, otherwise there is little variation in the bond lengths between both these enzyemes. This has led to suggestions of one axial and one equatorial hydroxyl ligand, in agreement with the proposed proton arrangement of VCPO enzyme. However due to the current resolution limitations of X-ray crystallography, and the unresolved protonation state of VCPO, a wider range of possible protonation states were considered. The average bond lengths and standard deviations for the atoms coordinated to vanadium are given in Table 4.3.4 for QM regions I to III.

In general there appears to be excellent geometric convergence (for individual models) across the six snapshots after QM/MM optimization. The standard deviations in Table 4.3.4 (values in parentheses) reveals that the geometrical parameters for each model and QM region are generally very similar within the 6 selected snapshots. This finding indicates that in each case, there is a predominant population of minima that are closely related as far as the geometry of the active site is concerned, and that there are only small perturbations in the latter due to the protein environment.

The mean absolute deviation (MAD) against the analogous models for the QM/MM optimized models of VCPO and against the X-ray-derived bond lengths are given in columns 8 and 9 of Table 4.3.4, respectively, along with identifying the bond which represents the maximum absolute deviation (Δ_{Max} where $\Delta = r_{\mathrm{VCPO}} - r_{\mathrm{VBPO}}$) between VBPO and VCPO. The MAD values are extremely low and there appears to be no particular model presenting a strong difference between VBPO and VCPO. The Δ_{Max} are also consistently low and show a small deviation across the different models, and no particular bond appears to persistently dominate the Δ_{Max} values. The agreement between the QM/MM models of VBPO and VCPO appears to be even better than that between the VBPO and VCPO X-ray structures.

Overall the results for each of the models are almost completely insensitive toward extension of the QM regions from I to III. Apparently, the effects of the local environment on the vanadate geometry are reasonably well described already by the smallest QM region and the electrostatic embedding scheme. The only notable exception is the case where a water molecule is placed at the axial position (VBPO-T144) the V-O^{4} bond length contracts by 0.13 Å and there is a concurrent elongation (albeit less dramatic) of the V-N HIS^{486} of 0.06 Å, is observed upon going from QM region I to III. In the VBPO case, no proton transfer is observed from axial water molecule (as seen in VCPO for the VOT134 model previously), forming a persistent hydrogen bond to HIS418. The axial water ligand in the VBPO-T144 model undergoes a notable change in geometry when increasing the size of the QM region, this was also observed in the analogous VCPO models. The axial water ligand is less tightly bound to the vanadium metal centre, thus making it more susceptible to the specific hydrogen-bonding interactions and the way these are treated computationally.

Optimized V-O bond distances invariably adopt values that are typical for the particular ligand involved, around ca. 1.7 Å, 1.8-2.0 Å, and 2.1 Å for oxo, hydroxy, and water ligands, respectively. The V-N bond length involving HIS486 is somewhat more variable across models, ranging from ca. 2.1 – 2.3 Å. Closest V-N contacts are observed when the other axial ligand trans to HIS486 is a water molecule (VBPO-T144) longer values are found when this is just an oxo atom (as in VBPO-S1 and VBPO-S3). To summarize, essentially the same observations were noted for VCPO and VBPO, and there is no significant geometric differences between the vanadate moiety in corresponding QM/MM models for both enzymes, see Figure 4.3.3.

Table 4.3.4 Bond lengths involving atoms coordinated to vanadium for QM/MM regions I and III. Mean bond lengths and standard deviations are reported in Å. X-ray refined experimental values[39,49] are shown for comparison.

		$\text{V-N}^{\text{HIS486}}$	V-O^1	V-O^2	V-O^3	V-O^4	MAD (Δ_{Max}) VBPO-VCPO	MAD (Δ_{Max}) QM/MM-Expt
VCPO[39]	X-ray	2.08	1.60	1.61	1.62	1.88	-	-
VBPO[49]	X-ray	2.11	1.59	1.60	1.52	1.77	0.06(0.12)	-
VBPO-T344	I	2.09(0.00)	1.65(0.00)	1.65(0.00)	1.85(0.00)	2.13(0.02)	0.02(0.02)[a,d,e]	0.16(0.36)[e]
	II	2.16(0.02)	1.65(0.00)	1.65(0.00)	1.85(0.01)	2.00(0.03)	0.03(0.03)[a,e]	0.14(0.33)[d]
	III	2.15(0.00)	1.66(0.00)	1.65(0.00)	1.85(0.00)	2.00(0.04)	0.03(0.07)[d]	0.14(0.23)[e]
VBPO-T144	I	2.11(0.01)	1.85(0.00)	1.65(0.00)	1.65(0.01)	2.13 (0.02)	0.03(0.04)[a]	0.16(0.36)[e]
	II	2.16(0.01)	1.88(0.02)	1.65(0.00)	1.66(0.01)	2.00(0.08)	0.03(0.07)[a]	0.15(0.29)[b]
	III	2.20(0.01)	1.90(0.00)	1.67(0.00)	1.64(0.01)	1.92(0.08)	0.07(0.23)[e]	0.15(0.31)[e]
VBPO-T234	I	2.19(0.03)	1.63(0.00)	1.86(0.01)	1.81 (0.01)	1.81 (0.01)	0.02(0.05)[a]	0.14(0.29)[d]
	II	2.32(0.07)	1.66(0.02)	1.88(0.01)	1.86(0.03)	1.68(0.05)	0.02(0.06)[a]	0.20(0.34)[d]
	III	2.31(0.39)	1.66(0.02)	1.90(0.04)	1.84(0.04)	1.69(0.06)	0.02(0.05)[d]	0.19(0.32)[d]
VBPO-T234*	I	2.19(0.02)	1.62(0.00)	1.82(0.00)	1.85(0.01)	1.84(0.00)	0.04(0.08)[c]	0.15(0.33)[d]
	II	2.38(0.04)	1.67(0.00)	1.86(0.00)	1.89(0.00)	1.67(0.00)	0.04(0.06)[c]	0.22(0.37)[d]

	III	2.39(0.03)	1.66(0.00)	1.87(0.00)	1.87(0.03)	1.52(0.28)	0.06(0.15)[e]	0.24(0.35)[d]
VBPO-D14	I	2.21(0.03)	1.91(0.01)	1.68(0.02)	1.65(0.00)	1.89(0.00)	0.01(0.02)[c]	0.15(0.32)[b]
	II	2.23(0.02)	1.91(0.01)	1.67(0.00)	1.65(0.00)	1.88(0.01	0.01(0.02)[a]	0.15(0.32)[b]
	III	2.21(0.01)	1.92(0.01)	1.67(0.00)	1.66(0.00)	1.87(0.01)	0.02(0.03)[a,d]	0.15(0.33)[b]
VBPO-D34	I	2.22(0.01)	1.67(0.01)	1.66(0.00)	1.89(0.01)	1.87(0.01)	0.01(0.03)[e]	0.14(0.37)[d]
	II	2.24(0.01)	1.68(0.01)	1.66(0.00)	1.90(0.00)	1.86(0.02)	0.03(0.05)[e]	0.15(0.38)[d]
	III	2.23(0.01)	1.67(0.01)	1.68(0.00)	1.89(0.01)	1.86(0.02)	0.03(0.05)[e]	0.15(0.38)[d]
VBPO-D34[*]	I	2.20(0.02)	1.67(0.00)	1.67(0.00)	1.90(0.01)	1.88(0.00)	0.01(0.02)[e]	0.15(0.38)[d]
	II	2.22(0.03)	1.67(0.00)	1.67(0.00)	1.91(0.01)	1.87(0.00)	0.01(0.04)[e]	0.15(0.39)[d]
	III	2.21(0.03)	1.67(0.00)	1.68(0.01)	1.90(0.01)	1.87(0.04)	0.02(0.04)[e]	0.15(0.38)[d]
VBPO-D34[*+]	I	2.2(0.00)	1.67(0.00)	1.69(0.00)	1.88(0.00)	1.87(0.00)	0.02(0.03)[e]	0.14(0.36)[d]
	II	2.24(0.05)	1.68(0.01)	1.67(0.02)	1.89(0.01)	1.83(0.02)	0.03(0.08)[e]	0.14(0.37)[d]
	III	2.25(0.07)	1.67(0.01)	1.72(0.01)	1.88(0.01)	1.83(0.04)	0.05(0.08)[e]	0.15(0.36)[d]
VBPO-D12	I	2.31(0.02)	1.90(0.00)	1.66(0.00)	1.91(0.00)	1.64(0.00)	-	0.22(0.39)[d]
	II	2.33(0.02)	1.89(0.00)	1.65(0.00)	1.91(0.00)	1.64(0.00)	-	0.22(0.39)[d]
	III	2.29(0.03)	1.92(0.00)	1.89(0.00)	1.66(0.00)	1.64 (0.00)	-	0.21(0.33)[b]
VBPO-D24	I	2.25(0.05)	1.67(0.01)	1.91(0.00)	1.66(0.01)	1.87(0.03)	-	0.15(0.31)[c]
	II	2.29(0.00)	1.68(0.00)	1.91(0.01)	1.66(0.00)	1.82(0.01)	-	0.15(0.31)[c]
	III	2.27(0.03)	1.68(0.00)	1.93(0.00)	1.66(0.00)	1.82(0.00)	-	0.15(0.33)[c]

VBPO-S4	I	2.38(0.01)	1.72(0.00)	1.70(0.00)	1.70(0.00)	1.93(0.01)	0.01(0.01)	0.17(0.27)[a]
	II	2.36(0.02)	1.72(0.00)	1.69(0.00)	1.70(0.00)	1.96(0.01)	0.01(0.01)	0.17(0.25)[a]
	III	2.37(0.01)	1.72(0.00)	1.72(0.00)	1.70(0.00)	1.93(0.00)	0.01(0.05)[a]	0.17(0.26)[a]
VBPO-S4[*]	I	2.4(0.04)	1.72(0.00)	1.73(0.04)	1.70(0.00)	1.92(0.01)	0.01(0.03)[c]	0.18(0.29)[a]
	II	2.36(0.01)	1.72(0.00)	1.70(0.00)	1.76(0.12)	1.96(0.01)	0.01(0.05)[d]	0.18(0.25)[a]
	III	2.37(0.02)	1.71(0.00)	1.74(0.00)	1.70(0.03)	1.93(0.01)	0.02(0.05)[a]	0.17(0.26)[a]
VBPO-S4[*+]	I	2.32(0.02)	1.72(0.01)	1.73(0.00)	1.69(0.00)	1.91(0.02)	0.03(0.07)[a]	0.16(0.21)[a]
	II	2.32(0.03)	1.71(0.01)	1.72(0.00)	1.69(0.00)	1.94(0.02)	0.02(0.05)[a]	0.16(0.21)[a]
	III	2.32(0.02)	1.68(0.04)	1.81(0.07)	1.68(0.01)	1.89(0.04)	0.04(0.09)[c]	0.16(0.21)[a,b]
VBPO-S3	I	2.69(0.08)	1.72(0.00)	1.69(0.00)	1.91(0.00)	1.68(0.00)	0.04(0.15)[a]	0.26(0.58)[a]
	II	2.65(0.07)	1.72(0.00)	1.69(0.00)	1.91(0.00)	1.69(0.00)	0.07(0.24)[a]	0.25(0.54)[a]
	III	2.64(0.08)	1.72(0.00)	1.71(0.00)	1.90(0.00)	1.68(0.00)	0.08(0.22)[a]	0.25(0.53)[a]
VBPO-S3[*]	I	2.72(0.05)	1.71(0.00)	1.69(0.00)	1.91(0.01)	1.69(0.01)	0.05(0.18)[a]	0.26(0.61)[a]
	II	2.72(0.05)	1.71(0.01)	1.69(0.01)	1.91(0.01)	1.69(0.01)	0.09(0.31)[a]	0.26(0.61)[a]
	III	2.66(0.04)	1.71(0.01)	1.71(0.00)	1.90(0.02)	1.69(0.01)	0.08(0.24)[a]	0.22(0.55)[a]
VBPO-S2	I	2.44(0,07)	1.71(0.01	1.96(0.01)	1.9(0.00)	1.69(0.01)	-	0.25(0.38)[a]
	II	2.53(0.05)	1.69(0.00)	1.93(0.00)	1.69(0.00)	1.72(0.00)	-	0.21(0.42)[a]
	III	2.42(0.04)	1.71(0.01)	1.99(0.01)	1.69(0.00)	1.71(0.01)	-	0.21(0.39)[c]
VBPO-S1	I	2.48(0.04)	1.94(0.01)	1.70(0.00)	1.70(0.00)	1.70(0.00)	0.08(0.33)[a]	0.21(0.37)[a]

	II	2.52(0.09)	1.93(0.01)	1.69(0.00)	1.70(0.00)	1.71(0.01)	0.07(0.08)[e]	0.22(0.41)[a]
	III	2.31	1.98	1.72	1.70	1.69	0.03(0.05)[c]	0.19(0.39)[b]
VBPO-Z	III	2.87	1.75	1.75	1.71	1.72	-	0.26(0.75)[a]

* denotes a flipped HIS411 conformation. + denotes a doubly protonated HIS411. [a] Δ_{Max} observed for V-N^{HIS}. [b] Δ_{Max} observed for V-O^1. [c] Δ_{Max} observed for V-O^2. [d] Δ_{Max} observed for V-O^3. [e] Δ_{Max} observed for V-O^4.

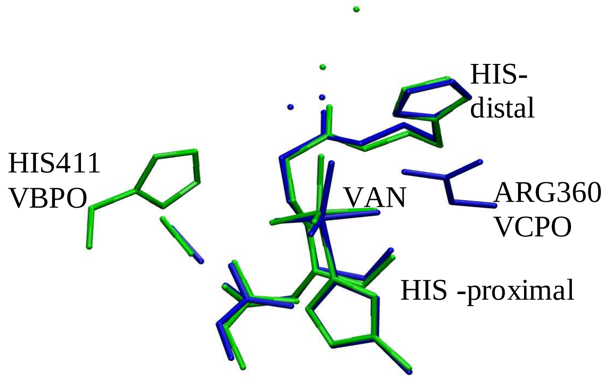


Figure 4.3.3 An overlay of the QM parts of VBPO(green) and VCPO(blue), for model D34, omitting hydrogen for clarity.

Considering the main geometric difference between the VBPO and VCPO X-ray structures in the active site, Zameplla et al. suggested that the HIS vs. PHE differences between VBPO and VCPO might be due to HIS411 tuning the hydrogen bond strength of LYS341.[] A set of new calculations containing the HIS411 in the 'flip' conformation, that is rotating the aromatic group of the HIS residue so that the N^{ε} is facing away from the LYS 341, was performed (see Figure 4.3.4). This procedure was motivated by the fact that the current resolution of protein X-ray crystallography makes a clear distinction between N and C atoms in the five-membered rings of HIS residues difficult, and that, hence, two orientations are possible a priori. The geometric difference between the flip conformation of a given protonation state is not particularly pronounced.

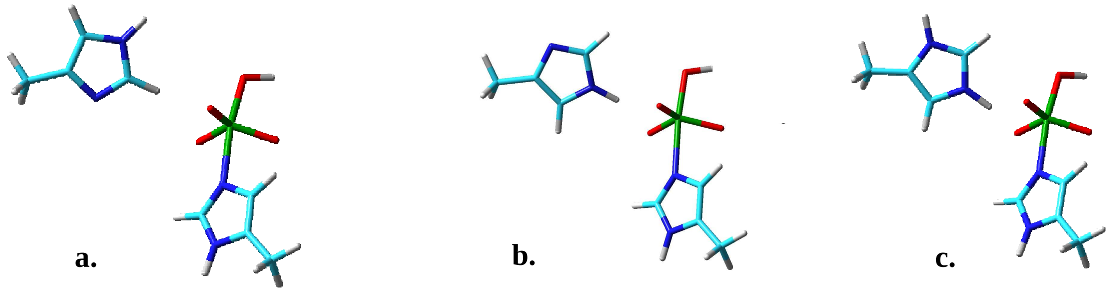


Figure 4.3.4 Vanadate moiety with the HIS411 a. 'normal', b. flipped and c. doubly protonated.

As the observed high shielding of the ^{51}V resonance in VBPO (see below) could be indicative of a very high negative charge, exploration was done on a fully unprotonated, triply charged model (VBPO-Z). During QM/MM optimization of the first snapshot for this model, the HIS486 residue essentially detached from the vanadate moiety (attaining a V···N distance of nearly 2.9 Å, see Table 4.3.4). This model is therefore in severe discord with the structure in the solid and was no longer considered.

Hydrogen Bonding

The analysis of the H-bond patterns is analogous to those in the VCPO variants discussed in the preceding chapters. The models generally show rather small standard deviations (see Table 4.3.5) across the six snapshots (values in parentheses), which indicates that the individual models of VBPO, as those of VCPO, have a dominant hydrogen-bonding network and not a multitude of different networks. As expected, the hydrogen bonds formed with crystallographically resolved water molecules exhibit much larger variations than those involving the amino acid residues due to the inherent mobility of water.

The O^1-ARG349 hydrogen bond is found to be around 1.9-2 Å for all models, except for the shorter VBPO-S1 due to the protonated O^1, and the VBPO-Z model which is also significantly shorter than the other models in Table 4.3.5. This hydrogen bond is stable across the six snapshots (cf. the low standard deviation). The O^2-LYS341 is generally shorter than the O^2-GLY417 distances, in the VCPO models these two hydrogen bonds tended to be in greater agreement with one another. The O^3-SER415 follows the same basic pattern with stability across snapshots and variation between models. The H^4 atom (when present) is hydrogen bonded to HIS418 and generally has the shortest distance of all possible hydrogen bonds considered around the vanadate moiety. In summary, the hydrogen bonding around the vanadate moiety is generally similar for different snapshots of a given model, but there is some variation between the models. A rather similar hydrogen-bonding network was noted for the VCPO system as shown by the low MAD in Table 4.3.5. The only hydrogen bond, which shows an appreciable difference, is the H4 to HIS418 that tends to be longer for the VBPO than the VCPO models. Whether this can be partly responsible for experimentally observed differences in the properties of both enzymes is unclear at present.

Table 4.3.5 Possible hydrogen bonds in QM/MM calculations using QM region III. Mean bond lengths and standard deviations are reported in angstroms. CHARMM atom types are given in brackets and superscripted residue labels are consistent with the X-ray 1QI9.pdb file.

	O^1 - H [HH11]^{ARG349}	O^2 - H [HZ]^{LYS341}	O^2 - H [HN]^{GLY417}	O^3 - H [HG]^{SER416}	H^4 - N [ND1]^{HIS418}	MAD(Δ_{Max})
VBPO-T344	1.99 (0.05)	1.78 (0.04)	1.87 (0.01)	1.85 (0.03)	1.50 (0.10)	0.08 (0.18)[c]
VBPO-T144	2.08 (0.21)	1.82 (0.06)	1.78 (0.03)	1.72 (0.04)	1.08 (0.00)	0.48 (1.15)[a]
VBPO-D234	2.02 (0.20)	1.73 (0.08)	2.01 (0.18)	1.91 (0.05)	1.56 (0.93)	0.17 (0.52)[c]
VBPO-D234*	2.09 (0.04)	4.51 (0.28)	1.85 (0.02)	2.08 (0.02)	1.21 (0.27)	0.61 (2.72)[a]
VBPO-D14	1.86 (0.05)	1.79 (0.03)	1.76 (0.04)	1.74 (0.03)	1.80 (0.12)	0.13 (0.46)[b]
VBPO-D34	1.95 (0.08)	1.70 (0.08)	1.82 (0.02)	1.93 (0.03)	1.86 (0.10)	0.06 (0.09)[a]
VBPO-D34*	1.87 (0.04)	1.69(0.02)	1.84(0.04)	1.96 (0.06)	1.84 (0.04)	0.07(0.09)[c]
VBPO-D34*+	1.93 (0.05)	1.69 (0.02)	1.97 (0.04)	1.86 (0.17)	1.64 (0.15)	0.09 (0.29)[c]
VBPO-D24	1.88 (0.02)	1.71 (0.02)	1.87(0.00)	1.83(0.04)	1.63(0.05)	-
VBPO-D12	1.89 (0.03)	2.99 (0.02)	1.82(0.01)	1.83 (0.00)	-	-
VBPO-S4	1.88 (0.02)	1.62 (0.01)	1.62 (0.01)	1.68 (0.01)	1.93 (0.05)	0.05 (0.12)[a]
VBPO-S4*	1.87 (0.02)	1.61 (0.01)	1.83 (0.01)	1.75 (0.01)	1.84 (0.03)	0.04 (0.09)[a]
VBPO-S4*+	1.95 (0.14)	1.63 (0.04)	1.86 (0.04)	1.76 (0.06)	1.83 (0.05)	0.06 (0.12)[a]
VBPO-S3	1.81 (0.00)	1.58 (0.02)	1.76 (0.01)	2.05 (0.01)	-	0.29 (0.38)[b]
VBPO-S3*	1.83 (0.05)	1.60 (0.03)	1.87 (0.08)	1.99 (0.16)	-	0.24 (1.67)[b]
VBPO-S2	1.72 (0.04)	1.73 (0.07)	1.87 (0.02)	1.76 (0.01)	-	-
VBPO-S1	1.74	1.57	1.64	1.75	-	0.40 (0.77)[b]

* denotes a flipped HIS411 conformation. [+] denotes a doubly protonated HIS411

[a.] Δ_{Max} observed for O^2 - [HZ]^{LYS341}

[b.] Δ_{Max} observed for O^3 - [HG]^{SER416}

[c.] Δ_{Max} observed for O^4- N^{HIS418}

[d.] Δ_{Max} observed for O^1- [HH11]^{ARG349}

Isotropic Chemical Shifts

The computed isotropic ^{51}V chemical shifts (δ_{iso}) are displayed in Table 4.3.6. A cursory glance at Table 4.3.6 shows that the δ_{iso} values of all models are rather stable across the QM regions I-III, where variations typically amount to less than 25 ppm. Fluctuations across QM regions I–III tend to be somewhat smaller than observed for VCPO. This is primarily because no proton transfer to HIS 418 occurred, which had been observed for VCPO to the corresponding, proximal HIS404 residue.[131]

The isotropic NMR shifts in VBPO range from around -550 to -690 pmm (VCPO range from -550 ppm to -620 ppm), without showing a clear dependence on the protonation state. This range is bracketed by VBPO-S4* being the most de-shielded and VBPOS1 is the most shielded of the models considered. Although the latter seemingly is in best agreement with the experimental solid-state value, the substantial shielding of the ^{51}V resonance with respect to that in solid VCPO is not very well reproduced. Furthermore, all the models in Table 4.3.6 are less shielded than the experimental value. This is completely opposite to prior experience, where previous results at the same or comparable QM levels for other vanadate complexes displayed isotropic values that were generally too strongly shielded by roughly 100ppm.[47,125] Thus, no QM/MM optimized model can be proposed as a solid candidate based on the agreement between computed and experimental isotropic shifts.

Table 4.3.6 ^{51}V isotropic chemical shifts (ppm) averaged over 6 snapshots, together with the corresponding standard deviations. For each of the QM/MM models considered, results are given for QM regions I – III. The MAD and Δ_{Max} relative to the VCPO are also reported.

	I	**II**	**III**	**MAD (Δ_{Max})**	**Δδ VBPO(III) rel. to VCPO(IV)**
VBPO-T344	-616 ± 7	-645 ± 9	-613 ± 7	14 (20)$^{\mathrm{II}}$	-11
VBPO-T144	-593 ± 9	-623 ± 9	-592 ± 13	14 (20)$^{\mathrm{II}}$	10
VBPO-D234	-573 ± 24	-640 ± 12	-622 ± 10	26 (39)$^{\mathrm{I}}$	-39
VBPO-D234*	-549 ± 7	-603 ± 4	-603 ± 6	24 (36)$^{\mathrm{I}}$	-20
VBPO-D14	-633 ± 21	-641 ± 15	-624 ± 14	23 (28)$^{\mathrm{III}}$	-44
VBPO-D24	-624 ± 7	-642 ± 12	-622 ± 7	8 (12)$^{\mathrm{II}}$	-
VBPO-D34	-630 ± 2	-635 ± 3	-605 ± 4	12 (18)$^{\mathrm{III}}$	2
VBPO-D34*	-633 ± 5	-636 ± 6	-607 ± 8	12 (18)$^{\mathrm{III}}$	0
VBPO-D34^{*+}	-617 ± 2	-623 ± 2	-591 ± 2	13 (19)$^{\mathrm{III}}$	16
VBPO-34§	-567 ± 4	-568 ± 3	-573 ± 10	2 (4)$^{\mathrm{III}}$	34
VBPO-34*§	-567 ± 5	-569 ± 5	-587 ± 14	8 (13)$^{\mathrm{III}}$	20
VBPO-S4	-562 ± 11	-576 ± 7	-561 ± 7	6 (10)$^{\mathrm{II}}$	-8
VBPO-S4*	-566 ± 5	-578 ± 5	-546 ± 4	12 (17)$^{\mathrm{III}}$	7
VBPO-S4^{*+}	-576 ± 7	-587 ± 1	-554 ± 10	11 (17)$^{\mathrm{II}}$	-1
VBPO-S3	-686 ± 10	-683 ± 10	-655 ± 10	12(18)$^{\mathrm{III}}$	-45
VBPO-S3*	-687 ± 2	-682 ± 2	-655 ± 4	13 (20)$^{\mathrm{II}}$	-45
VBPO-S2	-645 ± 4	-628 ± 3	625 ± 6	8(12)$^{\mathrm{I}}$	-
VBPO-S1	-681	-687	-670	6 (9)$^{\mathrm{III}}$	-55
Experiment^{a}			-687		-179

^{a}From references [46,51]. * denotes a flipped HIS411 conformation. $^{+}$ denotes a doubly protonated HIS411§ CSA tensors calculated without point charges.

The role of the HIS411 residue, as compared to PHE397 in VCPO, was studied in a number of ways. Firstly the difference between having HIS411 in the QM region or the MM active region has little effect on the chemical shifts (compare entries I and III in Table 4.3.6). Secondly, flipping the HIS into the two rotamers also has little effect on the chemical shifts (compare starred and unstarred results in Table 4.3.6). Therefore the electrostatic forces either using the charge distribution of the QM region or modeled as point charges is not significant enough to alter the chemical shift of vanadium by more than 26ppm. Thirdly, site directed mutagenesis experiments on VCPO converted PHE to a HIS, this did change the activity of the enzyme. Computational we trialed the reverse, that is we mutated the HIS411 into a PHE in VBPO, resulting in a $\Delta\delta_{iso}$=20ppm downfield shift. This deviation in chemical shifts was very small and therefore this cannot be used to rationalize the unusually high chemical shielding of V chemical shift in VBPO using steric arguments. In short, the inclusion of the HIS411 has little direct effect on chemical shifts in either conformation or either region and therefore cannot be proposed as the basis for the unusually high chemical shielding of ^{51}V in VBPO.

The QM/MM regions are centered on one of the active sties (VO4, notation as in the pdb file), the other active site (VO5) is too far beyond the radial cuttoff to be considered in the active QM/MM regions, and therefore remains fixed. To ensure the choice of which active site to study had no bearing on the results, a series of calculations were performed where QM/MM active site was centered on VO5, and subjected to QM/MM minimization. The results do not significantly differ between the VO4 and VO5 models and the values are therefore not included in Table 4.3.5 and Table 4.3.6.

In summary, and in keeping with our results for VCPO, the isotropic chemical shifts are found to be a poor discriminator of protonation state of VBPO, due to the small variation resulting from changes in the proximal proton environment. Furthermore, all of the models are noticeably deshielded when compared to experiment.

Anisotropic ^{51}V NMR Chemical Shifts: The CSA and EFG Tensors

The computed parameters characterizing the CSA and EFG tensors are collated in Table 4.3.7. The mean absolute percent deviation (MAPD) to those computed for the corresponding VCPO with the associated standard deviation (SD) is also included. These deviations provide a numerical feel for the similarity to the VCPO QM/MM models.

In general, the computed CSA and EFG values are somewhat more sensitive to the size of the QM region than the isotropic chemical shifts above. As expected, changes between QM region I and III are not particularly pronounced, in agreement with the previous VCPO conclusions, which were taken to indicate good convergence.

The difference between the flipped conformation of a given protonation state is larger than the difference between model I and model III for the same protonation state. The flip conformation is simply a computational tool to investigate the effect of such a residue.

A critical assessment of the models would require precise knowledge of the experimental CSA and EFG tensor elements. However, only a preliminary spectrum is published so far (as noted above), with only the isotropic chemical shift reported.[51] Therefore confident assignment of the protonation state is not possible at present, but may become possible as better experimental spectra and reliable refinements of the tensorial quantities become available. The computed values were presented in Table 4.3.7 to illustrate the differences amongst these members of the VHPO family and as predictions for future reference.

The previous section showed that the bromination of SER402 had no notable effect on the ^{51}V isotropic chemical shift of VBPO. Table 4.3.7 shows that the parameters that characterize the solid-state ^{51}V NMR are severely effected by SER402 becoming brominated in VBPO. This finding does suggest that the experimental method of solid-state ^{51}V NMR could be beneficial in addressing this question.

Inspection of the MAPD values in Table 4.3.7 reveals noticeable differences between the parameters that characterize the EFG and CSA tensor between individual models of VBPO and VCPO. The asymmetry parameter of the EFG generally shows the greatest deviation. Therefore the solid-state ^{51}V NMR spectrum of VBPO should be rather different from the previously published one of VCPO, which is in qualitative agreement with the preliminary experimental results for VBPO.[51] This is in spite rather similar geometries, hydrogen bonding networks and isotropic chemical shifts of QM/MM optimized models. It appears that experimentally observed differences between VBPO and VCPO is at least partially related to the PHE vs. HIS difference between VBPO and VCPO, and the longer-range electrostatic effects of more distant structural differences may also contribute to some degree.

Table 4.3.7 Reduced anisotropy δ_σ, asymmetry η_σ, nuclear quadrupole coupling constant C_Q(MHz), and asymmetry parameter η_Q with the associated standard deviations for six snapshots. The mean absolute percent deviation (MAPD) compared to the respective QM/MM VCPO models[131] is also tabulated. Results are given for QM regions I to III.

		δ_σ	η_σ	C_Q	η_Q	MAPD(Δ_{Max})
VBPO-T344	I	-536 ± 6	0.5 ± 0.02	15 ± 0.0	0.14 ± 0.02	16 (27)[g]
	II	-549±12	0.5±0.02	11±1.0	0.31±0.13	42(121)[g]
	III	-550 ± 5	0.5 ± 0.02	12 ± 1.4	0.28 ± 0.06	10 (27)[g]
VBPO-T144	I	-646±20	0.4±0.02	16±7.0	0.57±0.27	21(30)[g]
	II	-620±47	0.5±0.03	8±9.9	0.77±0.37	19(35)[f]
	III	-511 ± 26	0.5 ± 0.03	10 ±3.90	0.67 ±0.33	26 (48)[g]
VBPO-T234	I	-460±76	0.3±0.18	5±0.2	0.60±0.38	21(50)[e]
	II	-379±78	0.4±0.20	7±1.4	0.54±0.40	38(116)[g]
	III	-323 ± 117	0.6 ± 0.29	7 ±1.10	0.57 ±0.30	26 (58)[g]
VBPO-T234*	I	-464±13	0.4±0.03	-6.5±0.19	0.34±0.05	35(100)[e]
	II	-317±28	0.5±0.08	-5.9±0.34	0.58±0.07	44(132)[g]
	III	-303 ± 26	0.5 ± 0.04	-5.6 ±3.2	0.39 ±0.10	14 (35)[f]
VBPO-D14	I	-568 ± 7	0.4 ± 0.01	9 ± 0.1	0.58 ± 0.11	13 (26)[f]

	II	-550±24	0.4±0.03	9±0.6	0.67±0.20	15(43)[f]
	III	-598 ± 51	0.7 ± 0.16	8 ± 0.8	0.58 ± 0.24	20 (40)[e]
VBPO-D34	I	-531 ± 34	0.4 ± 0.32	9 ± 0.6	0.18 ± 0.09	26 (51)[g]
	II	-552±12	0.4±0.02	9±0.7	0.16±0.11	14(36)[f]
	III	-532 ± 16	0.4 ± 0.02	9 ± 0.6	0.27 ± 0.14	22 (47)[g]
VBPO-D34*	I	-515±16	0.4±0.03	9±0.21	0.3±0.04	38(61)[g]
	II	-508±22	0.4±0.01	9±0.32	0.3±0.04	25(61)[g]
	III	-519 ± 41	0.4 ± 0.01	9 ± 0.35	0.4 ± 0.07	16 (33)[e]
VBPO-D34^{*+}	I	-480±8	0.5±0.03	8±0.08	0.5±0.04	37(66)[e]
	II	-426±25	0.5±0.00	9±0.25	0.4±0.07	32(49)[g]
	III	-449 ± 41	0.6 ± 0.03	7 ± 0.21	0.7 ± 0.05	45 (100)[e]
VBPO-D34§	I	-653±11	0.3±0.00	11±0.48	0.3±0.06	-
	II	-631±9	0.3±0.00	11±0.59	0.3±0.08	-
	III	-652 ± 11	0.3 ± 0.00	11 ± 0.5	0.28 ± 0.06	24 (64)[g]
VBPO-D34*§	I	-642±7	0.3±0.01	11±0.17	0.3±0.03	-
	II	-661±20	0.3±0.01	11±0.34	0.3±0.02	-
	III	-409 ± 34	0.5 ± 0.06	10 ±0.3	0.33 ± 0.06	40 (67)[e]
VBPO-D12	III	320 ± 27	0.2 ± 0.03	-8 ± 0.0	0.24 ± 0.06	-
VBPO-D24	I	-311±5	0.6± 0.01	8±0.07	0.3±0.03	-
	II	-305±2	0.6±0.01	8±0.17	0.3±0.03	-
	III	-425 ± 9	0.3 ± 0.02	8 ± 0.2	0.56 ± 0.05	-
VBPO-S4	I	362 ± 7	0.2 ±0.08	13 ±0.4	0.20 ± 0.06	58 (100)[e]
	II	379±8	0.2±0.24	15±0.0	0.22±0.06	32(57)[g]
	III	359 ± 11	0.1 ±0.05	14 ±0.7	0.27 ± 0.05	88 (285)[g]
VBPO-S4*	I	358±19	0.3±0.09	15±0.23	0.32±0.04	43(128)[g]
	II	379±5	0.4±0.03	15±0.38	0.26±0.01	39(128)[g]
	III	353 ± 5	0.3 ± 0.022	14 ± 0.28	0.3 ± 0.02	99 (328)[g]
VBPO-S4^{MC2}	I	366	0.26	14	0.32	23(56)[e]
	II	390	0.25	15	0.19	25(35)[g]
	III	363	0.1	14	0.18	57 (157)[g]

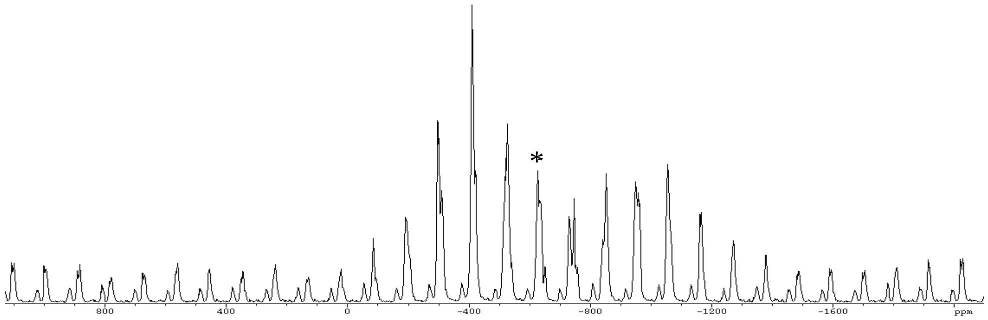
VBPO-S3	I	-121±18	0.6±0.07	6±0.3	0.51±0.12	54(88)[g]
	II	-133±15	0.7±0.04	6±0.3	0.55±0.12	73(113)[e]
	III	-143 ± 17	0.9 ± 0.03	5.0 ± 0.30	0.49 ± 0.13	107 (200)[f]
VBPO-S3*	I	-515±16	0.4±0.03	9±0.21	0.3±0.04	33(118)[h]
	II	-508±22	0.4±0.01	9±0.32	0.3±0.04	27(64)[h]
	III	-151 ± 23	1.0 ± 0.21	5.4 ± 0.61	0.70 ± 0.26	114 (233)[f]
VBPO-S2	I	-180±9	0.8±0.03	7±0.18	0.4±0.00	-
	II	-207±8	0.9±0.01	7±0.21	0.6±0.05	-
	III	242± 27	0.2 ± 0.02	7.8 ± 0.35	0.37 ± 0.16	-
VBPO-S1	I	-258	0.32	9	0.40	25 (75)[g]
	II	-241	0.4	7	0.7	19(43)[h]
	III	-270	0.43	7	0.32	39 (47)[h]

* denotes a flipped HIS411 conformation. $^+$denotes a doubly protonated HIS411. §CSA tensors calculated without point charges. ^{MC2} indicates the active site of the other half of the homodimer was investigated.

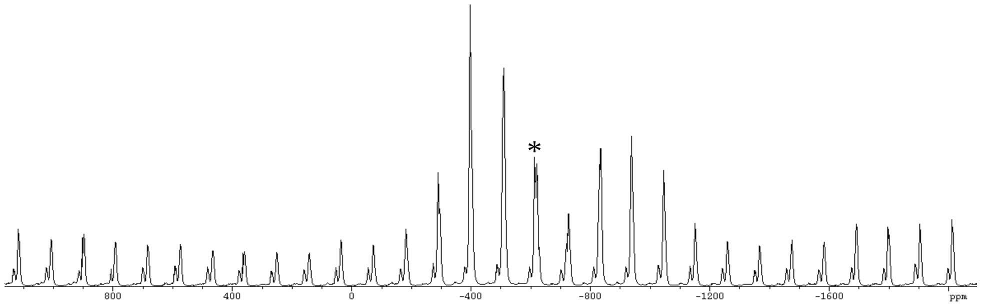
[e] Δ_{Max} observed for η_σ
[f] Δ_{Max} observed for C_Q
[g] Δ_{Max} observed for η_Q

Simulated MAS NMR spectra of selected models are provided in Figure 4.3.5. The different protonation states are clearly producing grossly different spectral features, e.g. VOD34 vs. VOS4. Qualitatively, the triply and doubly protonated models with their large absolute values of δ_σ seem to fit somewhat better to the preliminary spectrum of VBPO than the singly protonated models. More definite conclusions, however, would have to await more precise experimental data. Eventually, this may provide an indication as to the assignment of protonation state, without explicit recourse to energetic arguments from theoretical calculations.

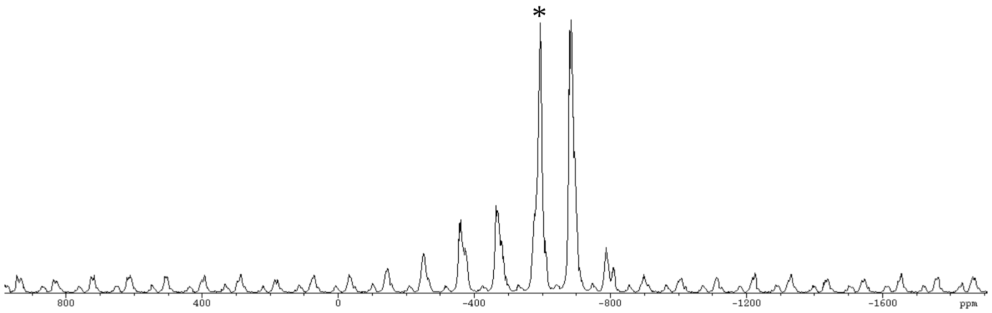
Figure 4.3.5 Theoretical spectra simulated using SIMPSON[126] with parameters taken from Table 4.3.7 for QM region III of selected models (star: isotropic shift). * denotes a flipped HIS411 conformation. + denotes a doubly protonated HIS411§ CSA tensors calculated without point charges.



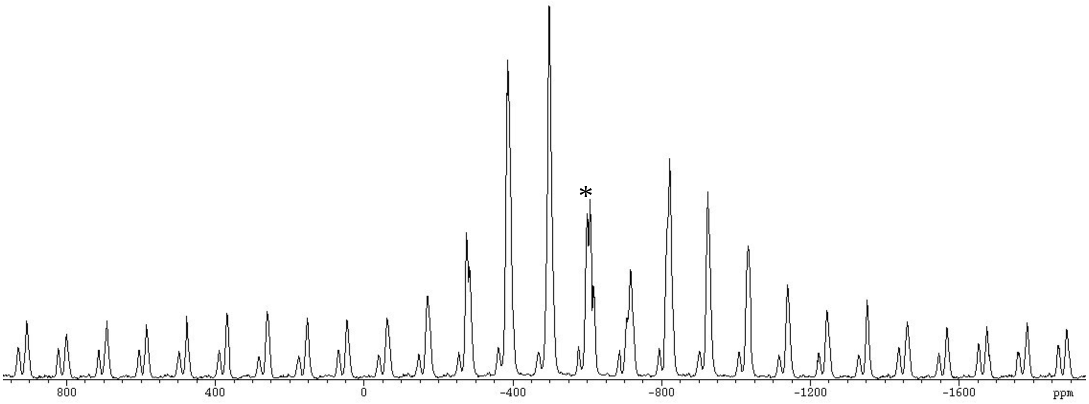
a. VBPO-T344



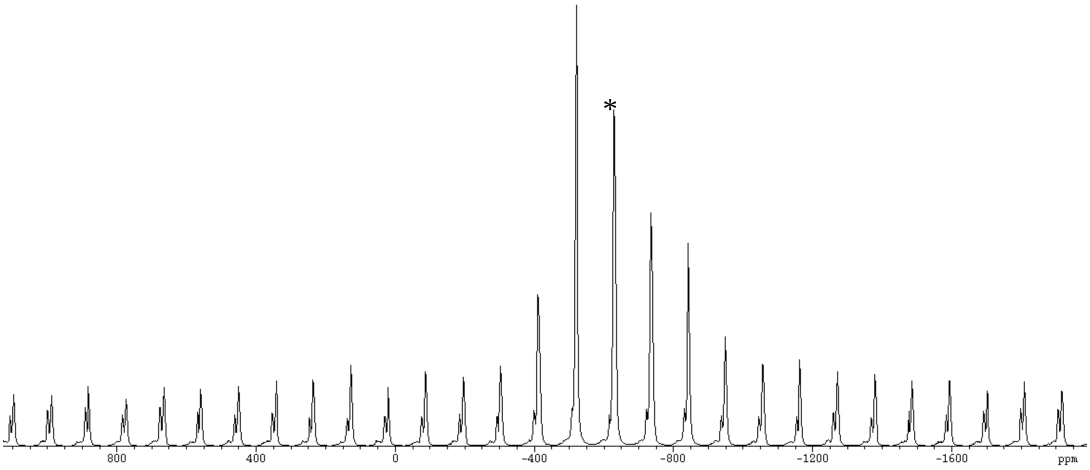
b. VBPO-D34



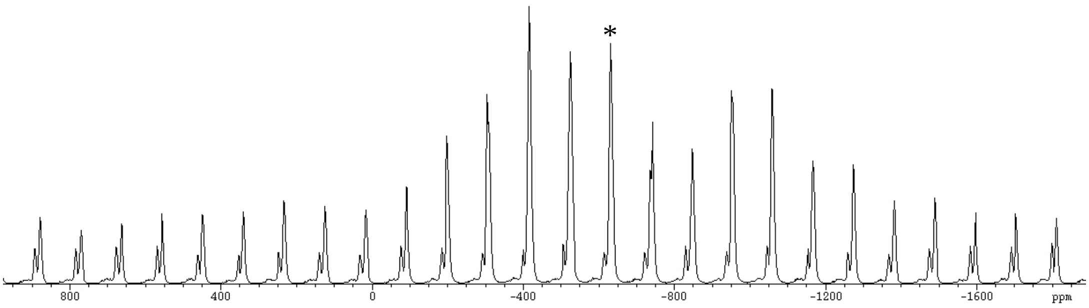
c. VBPO-S4



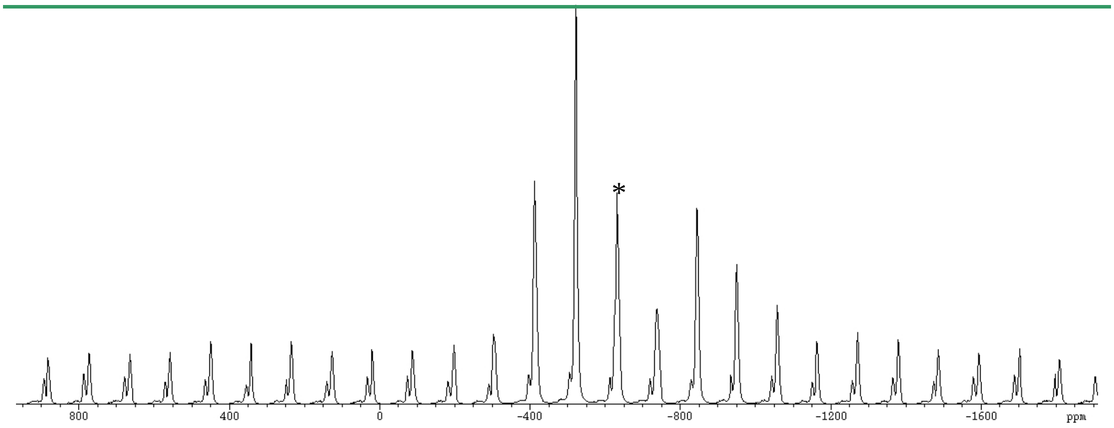
d. VBPO-T144



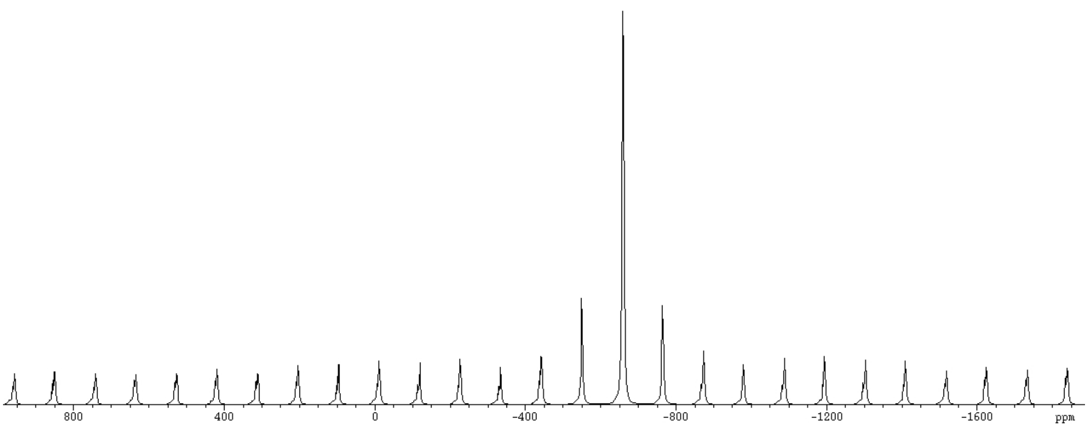
e. VBPO-T234



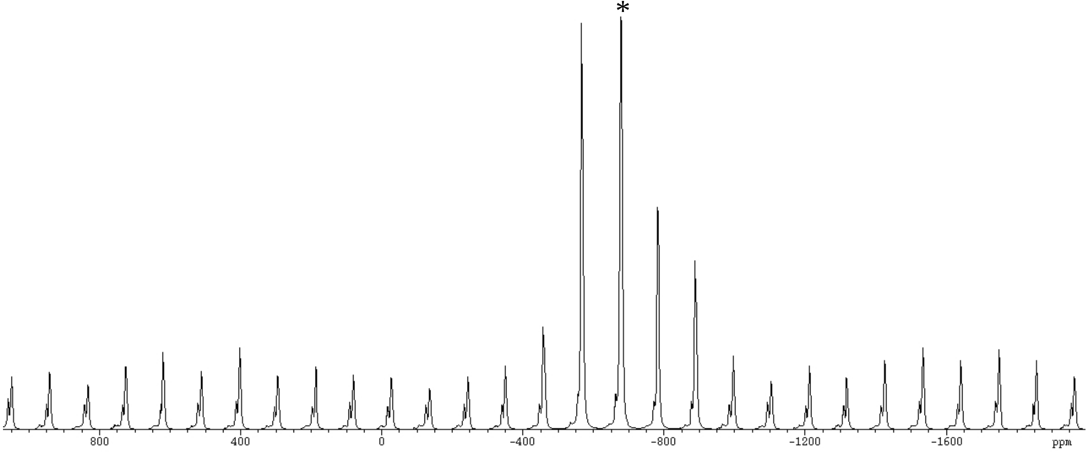
f. VBPO-D14



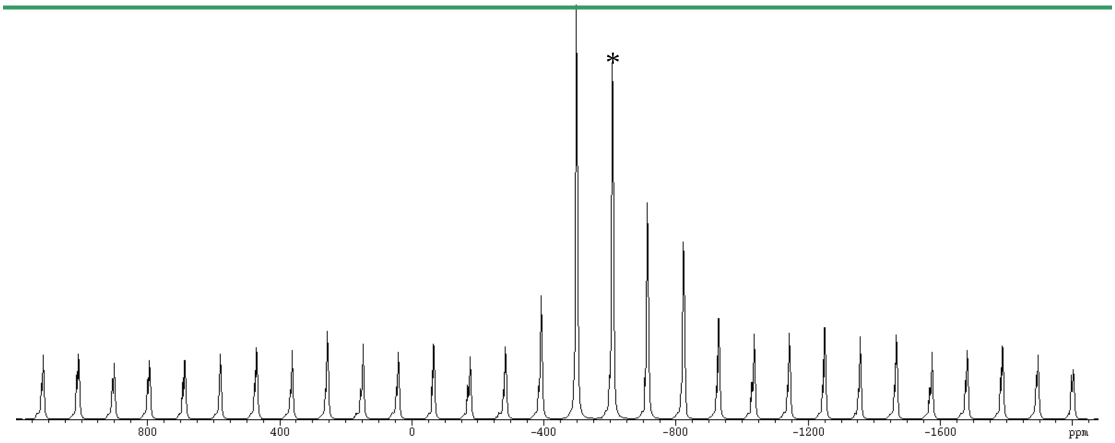
g. VBPO-D24



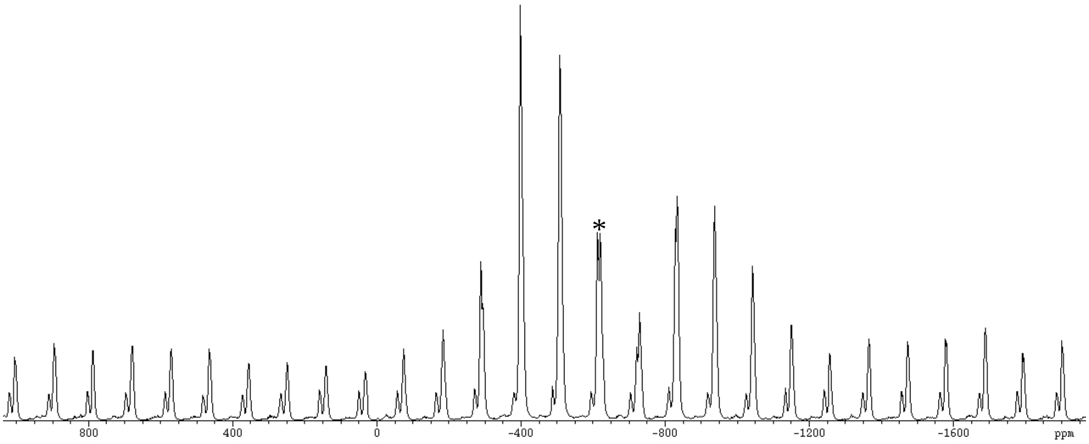
h. VBPO-S3



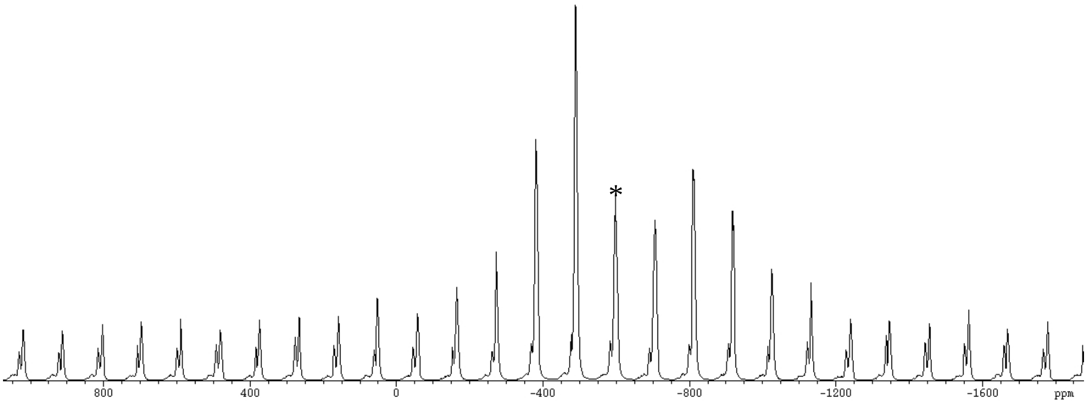
i. VBPO-S1



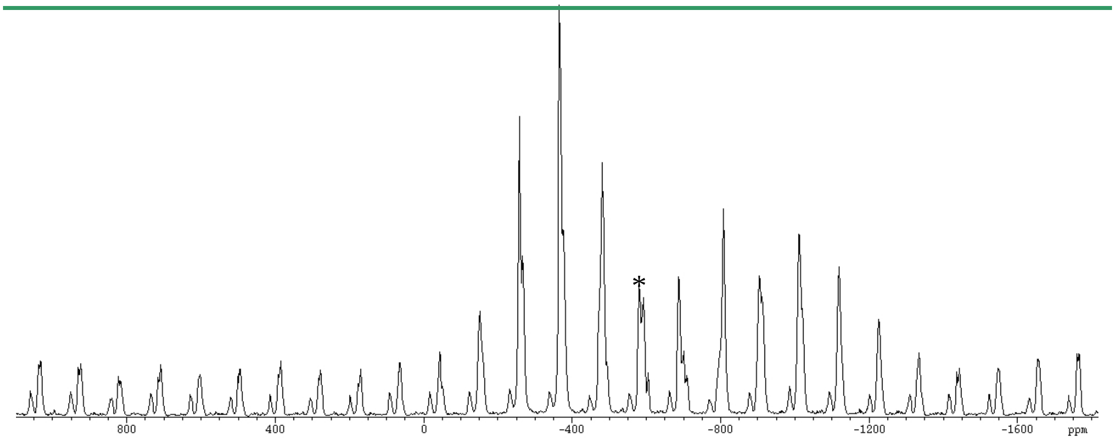
j. VBPO-T234*



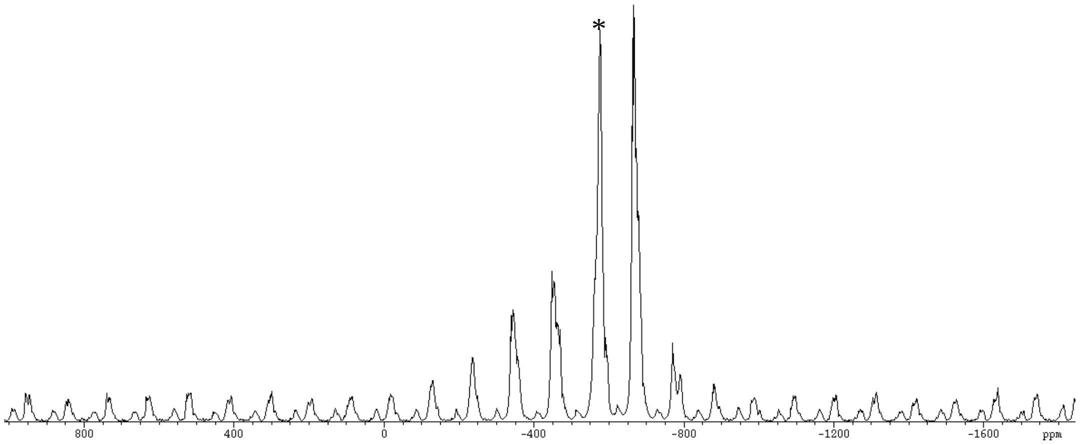
k. VBPO-D34*



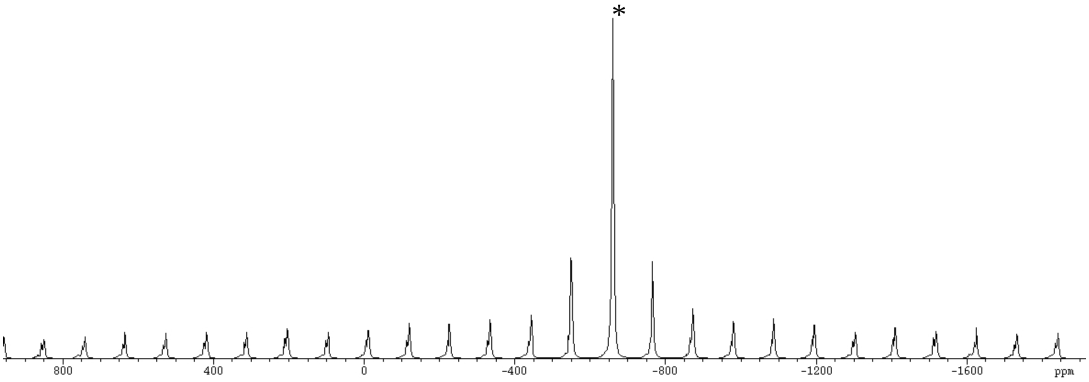
l. VBPO-D34*+



m. VBPO-D34 §



n. VBPO-S4*



o. VBPO-S3*

Conclusions

^{51}V NMR parameters of VBPO were computed for a large number of QM/MM optimized models, calling special attention to different possible protonation states and positional isomers. VBPO shows a small degree of conformational flexibility in terms of bond lengths for atoms coordinated to vanadium across individual models, as well as the computed ^{51}V NMR parameters, evidenced by the low standard deviation amongst models sampled along a classical MD simulation with subsequent QM/MM optimizations. In terms of non-hydrogen atomic coordinates all of the many models considered show good agreement with those of VCPO.

As had been the case with VCPO, the computed isotropic ^{51}V chemical shifts of VBPO show rather little variation between the models, making structural assignments based on this quantity difficult. The notable upfield shift of the isotropic ^{51}V NMR resonance that is observed on going from solid VCPO to solid VBPO is only partly recovered in the computations, and only for some of the models. No structural candidate could be identified that could account for the exceptional shielding of the VBPO ^{51}V NMR signal on going from the solid state into solution.

As there is a notable difference between the EFG and CSA tensors computed from models of VBPO and VCPO, and furthermore, the latter showed good agreement with the experimental derived values from the ^{51}V NMR spectrum, it is likely that the anisotropic chemical shifts reported here would be of use to future studies of VBPO using solid-state MAS NMR spectroscopy.

4.4 ^{51}V NMR chemical shifts calculated from QM/MM models of Vanadium haloperoxidase - peroxo forms

QM/MM setup

Starting from the X-ray structures[39], a number of possible protonation states of the peroxovanadate cofactor were modeled using the same QM/MM optimization protocol as for the native forms discussed in the preceding chapters. Our previous QM/MM studies[132] showed that two vanadium active sites of native VBPO are similar, and therefore only one such 'homology type' model was studied, focusing only on one of the metal centers. The large degree of homology, particularly within the active site, provides validity to this approach. In order to provide a convenient comparison to the pure DFT work by Pecoraro et al.[55] a range of models was considered, see Table 4.4.1.

Table 4.4.1 Selected models labeled according to the protonation state.

Unprotonated	Singly protonated	Doubly protonated
p-VHPO-Z	*p*-VHPO-S1	*p*-VHPO-D13
	p-VHPO-S2	*p*-VHPO-D12†
	p-VHPO-S3	*p*-VHPO-D14†
	p-VHPO-S4	

† Only investigated for *p*-VCPO.

Results for QM regions I (small) to III (large) are herein presented for each of the selected models in Table 4.4.1 and Figure 4.4.1 shows the segments of the residues and the crystallographic water molecules that are included in the QM I and III regions. Note that QM region III corresponds to QM region IV from our first study on native VCPO.[39]

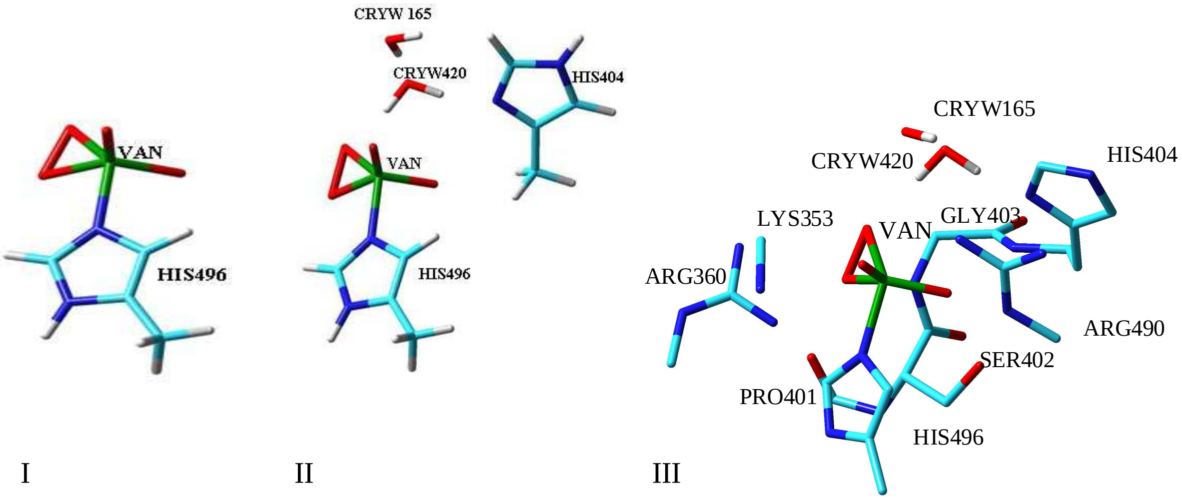


Figure 4.4.1 QM regions I to III for peroxo-form of haloperoxidases, deprotonated vanadate moiety shown as an example.

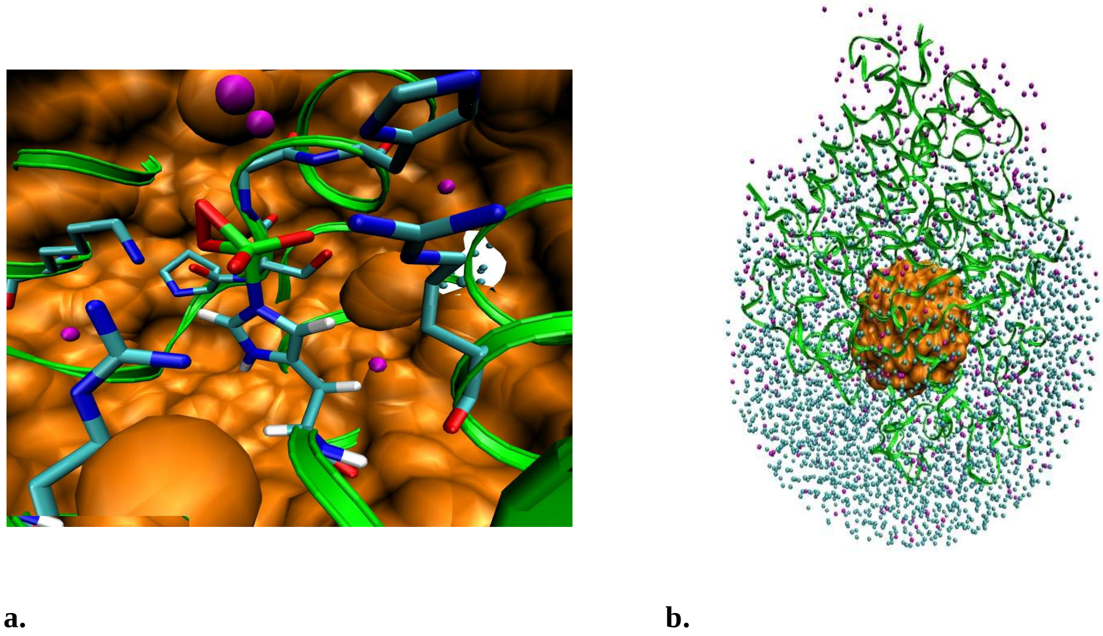


Figure 4.4.2 a. Vanadate co-factor surrounded by residues in the first shell of the active regions that are included in the larger QM III region. b. Solvated *p*-VCPO protein,. The region in blue is solvent (30 Å sphere centered on vanadium), and the fixed protein environment during the QM/MM optimization (residues > 10 Å from vanadate) is displayed as a green ribbon. The orange sphere centered on the vanadate moiety is the active region subjected to the QM/MM optimization (including all residues <10 Å from the vanadate). Crystallographically resolved waters (oxygen positions) are displayed as purple spheres.

Results and discussion

Geometrical Details

Because the constituent $V(O_2)O_2$ unit bears one overall negative charge and because neutral vanadates or peroxovanadates are usually not acting as strong Brønsted bases, an unprotonated (*p*-VHPO-Z) or a singly protonated state would seem as the most probable. Consistent with this expectation, most of the doubly protonated models that were considered turned out to be unstable with respect to maintaining the peroxo-coordinated geometry. For instance, QM/MM optimization of *p*-VCPO-D12 and *p*-VCPO-D14 resulted in partial detachment of the hydroperoxy ligand, which rearranged from the side-on η^2 coordination to an essentially end-on η^1 mode. Because of this clear disagreement with experiment, these models were not explored further. The only stable doubly protonated species was *p*-VHPO-D13. Salient distances of all final *p*-VCPO models and those of the corresponding *p*-VBPO homologues are summarized in Table 4.4.2 and Table 4.4.3.

The bond lengths of the peroxovanadate cofactor fall within the expected range, ca. 1.6-1.7 Å for vanadium oxo bonds, 1.8-1.9 Å for vanadium hydroxy and peroxo bonds, and 2.0- 2.1 Å for the V-N bond to HIS404. As the standard deviations across the six snapshots for the bond lengths of the peroxovanadate cofactor are rather small, all the models are generally very similar within the six selected snapshots .The geometric convergence for the individual models is comparable to our prior QM/MM optimizations of the native VHPO forms,[131] indicating that the active site is rather rigid, and that the use of a small number of snapshots seems valid. Furthermore, a good degree of similarity is observed between the bond lengths of the peroxovanadate, across many of the different protonated models of *p*-VCPO and *p*-VBPO, see Table 4.4.2 and Table 4.4.3, respectively.

Table 4.4.2 Bond lengths involving vanadium cofactor for QM/MM regions I to III of *p*-VCPO.[a]

		V-N$^{\text{HIS 496}}$	V-O^{1}	V-O^{2}	V-O^{3}	V-O^{4}	O^{2}-O^{4}	V-O$^{\text{wat}}$ [b]	MAD(Δ_{MAX}) [c]	MAD(Δ_{MAX}) [c]
X-ray *p*-VCPO		2.19	1.93	1.89	1.60	1.86	1.47	4.39	(including V-O$^{\text{wat}}$)	(without V-O$^{\text{wat}}$)
p-VCPO-Z	I	2.06(0.01)	1.68(0.00)	1.89(0.00)	1.65(0.00)	1.87(0.01)	1.47(0.00)	4.02(0.57) [d]	0.12(0.37)[h]	0.07(0.25) [f]
	II	2.07(0.01)	1.66(0.01)	1.89(0.01)	1.67(0.01)	1.87(0.01)	1.47(0.00)	3.86(0.26)		
	III	2.06(0.01)	1.67(0.00)	1.93(0.01)	1.66(0.00)	1.87(0.00)	1.47(0.00)	3.90(0.22)	0.14(0.49)[h]	0.08(0.26) [f]
p-VCPO-S1	I	2.03(0.01)	1.83(0.00)	1.87(0.00)	1.61(0.00)	1.84(0.00)	1.45(0.00)	2.64(0.01) [d]	0.30(1.75)[h]	0.06(0.16) [e]
	II	2.14(0.00)	1.88(0.01)	1.86(0.00)	1.62(0.00)	1.85(0.01)	1.45(0.01)	2.12(0.01)		
	III	2.13(0.00)	1.88(0.00)	1.88(0.00)	1.62(0.00)	1.86(0.00)	1.45(0.00)	2.13(0.01)	0.35(2.26)[h]	0.03(0.06) [e]
p-VCPO-S2	I	2.06(0.02)	1.65(0.01)	2.10(0.03)	1.63(0.00)	1.90(0.01)	1.47(0.01)	3.11(0.60) [d]	0.21(1.28)[h]	0.12(0.28) [f]
	II	2.18(0.01)	1.67(0.01)	2.15(0.05)	1.64(0.01)	1.91(0.01)	1.47(0.01)	2.22(0.03)		0.11(0.27) [f]
	III	2.13(0.04)	1.66(0.00)	2.11(0.06)	1.65(0.01)	1.89(0.02)	1.47(0.01)	2.51(0.31)	0.36(1.88)[h]	
p-VCPO-S3	I	2.04(0.02)	1.63(0.00)	1.85(0.00)	1.81(0.02)	1.85(0.01)	1.45(0.00)	2.89(0.73) [d]	0.32(1.50)[h]	0.12(0.21) [g]
	II	2.13(0.03)	1.65(0.01)	1.86(0.00)	1.86(0.02)	1.85(0.01)	1.45(0.00)	2.14(0.07)		
	III	2.13(0.04)	1.64(0.00)	1.87(0.01)	1.86(0.02)	1.87(0.01)	1.45(0.00)	2.15(0.10)	0.41(2.24)[h]	0.11(0.26) [g]
p-VCPO-S4	I	2.04(0.01)	1.64(0.00)	1.94(0.00)	1.63(0.01)	2.05(0.00)	1.47(0.00)	3.15(0.36) [d]	0.28(1.24)[h]	0.12(0.29) [f]
	II	2.07(0.01)	1.66(0.00)	1.9(0.01)	1.89(0.01)	1.89(0.01)*	1.47(0.00)	3.51(0.13)		
	III	2.06(0.01)	1.66(0.00)	1.94(0.00)	1.66(0.00)	1.88(0.00)*	1.47(0.00)	3.68(0.01)	0.18(0.71)[h]	0.09(0.27) [f]

p-VCPO-D12	I	2.01(0.01)	1.75(0.09)	2.37(0.04)	1.59(0.00)	1.84(0.02)	1.46(0.00)	2.65(0.97) [f]		
p-VCPO-D13	I	2.05(0.01)	1.78(0.00)	1.82(0.02)	1.77(0.00)	1.84(0.02)	1.42(0.01)	2.30(0.01) [d]	0.38(2.09)[h]	0.10(0.17) [g]
	II	2.09(0.01)	1.81(0.01)	1.82(0.02)	1.80(0.02)	1.84(0.00)	1.43(0.00)	2.09(0.01)		
	III	2.08(0.01)	1.83(0.00)	1.82(0.00)	1.77(0.01)	1.83(0.00)	1.43(0.00)	2.11(0.01)	0.40(2.28)[h]	0.09(0.17) [g]
p-VCPO-D14	I	2.04(0.01)	1.79(0.00)	1.88(0.00)	1.59(0.00)	2.25(0.05)	1.46(0.01)	2.21(0.02) [f]		

---

[a] Mean bond lengths and standard deviations are reported in Å, X-ray refined experimental values of *p*-VCPO are shown for comparison. [b] O^{wat} is in the MM region for the small QM I models. [c] MAD is the mean absolute deviation ($\overline{|\Delta|}$, where $\Delta = r_{vcpo} - r_{exp}$) and Δ_{MAX} is the maximum deviation relative to the X-ray distances of 1IDU.pdb (with and without V-O^{wat}, see text). [d] Water in MM region. [e] Δ_{MAX} observed for V-N^{His}. [f] Δ_{MAX} observed for V-O^1. [g] Δ_{MAX} observed for V-O^3. [*] Proton transfers to HIS404. [h] Δ_{MAX} observed for V-O^{wat}.

Table 4.4.3 Bond lengths involving vanadium cofactor for QM/MM regions I, II and III for *p*-VBPO.[§]

		V-N$^{HIS\ 496}$	V-O^{1}	V-O^{2}	V-O^{3}	V-O^{4}	O^{2}-O^{4}	V-O^{wat‡}	MAD(Δ_{MAX})†
X-ray *p*-VCPO		2.19	1.93	1.89	1.60	1.86	1.47	4.39	
p-VBPO-Z	I	2.07(0.01)	1.68(0.00)	1.90(0.01)	1.65(0.01)	1.86(0.00)	1.47(0.01)	4.30(0.53)^{f}	0.07(0.25)^{b}
	II	2.08(0.01)	1.68(0.00)	1.89(0.01)	1.65(0.01)	1.87(0.01)	1.47(0.00)	4.09(0.55)	0.07(0.25)^{b}
	III	2.07(0.01)	1.67(0.00)	1.93(0.01)	1.65(0.00)	1.87(0.00)	1.47(0.00)	4.10(0.50)	0.08(0.26)^{b}
p-VBPO-Z^{+}	III	2.09	1.68	1.94	1.64	1.88	1.48	3.32	0.08(0.25)^{b}
p-VBPO-S1	I	2.04(0.01)	1.83(0.01)	1.88(0.01)	1.60(0.00)	1.83(0.01)	1.45(0.00)	3.05(0.25)^{f}	0.05(0.15)^{a}
	II	2.14(0.09)	1.91(0.03)	1.88(0.01)	1.61(0.01)	1.86(0.02)	1.45(0.00)	2.23(0.31)	0.02(0.05)^{a}
	III	2.20(0.04)	1.94(0.02)	1.90(0.00)	1.61(0.00)	1.88(0.01)	1.45(0.00)	2.03(0.10)	0.01(0.02)^{e}
p-VBPO-S1^{+}	III	2.17	1.94	1.90	1.63	1.88	1.45	1.96	0.02(0.03)^{d}
p-VBPO-S2	I	2.03(0.01)	1.65(0.00)	2.24(0.08)	1.62(0.00)	1.88(0.02)	1.47(0.00)	3.03(0.13)^{f}	0.14(0.35)^{c}
	III	2.05(0.02)	1.65(0.01)	2.12(0.03)	1.63(0.01)	1.88(0.00)	1.48(0.01)	3.05(0.11)	0.12(0.28)^{b}
p-VBPO-S3	I	2.05(0.02)	1.62(0.01)	1.87(0.01)	1.79(0.01)	1.82(0.01)	1.45(0.00)	3.13(0.29)^{f}	0.12(0.31)^{b}
	II	2.11(0.05)	1.67(0.01)	1.90(0.01)	1.70(0.06)	1.85(0.01)	1.46(0.01)	3.37(0.53)	0.08(0.26)^{b}
	III	2.06(0.06)	1.64(0.01)	1.89(0.03)	1.76(0.04)	1.84(0.01)	1.46(0.00)	3.49(0.17)	0.10(0.29)^{b}
p-VBPO-S4	I	2.07(0.01)	1.65(0.00)	1.95(0.01)	1.62(0.00)	2.05(0.01)	1.47(0.00)	3.10(0.20)^{f}	0.11(0.28)^{b}
	II	2.08(0.01)	1.69(0.01)	1.91(0.01)	1.64(0.02)	1.87(0.01)*	1.47(0.00)	3.66(0.08)	0.07(0.24)^{b}
	III	2.08(0.01)	1.68(0.02)	1.95(0.01)	1.65(0.02)	1.87(0.02)*	1.47(0.00)	2.31(0.51)	0.08(0.25)^{b}
p-VBPO-D13	I	2.03(0.01)	1.77(0.01)	1.83(0.01)	1.76(0.01)	1.82(0.01)	1.42(0.00)	2.52(0.39)^{f}	0.11(0.16)^{b,d}
	II	2.17(0.03)	1.86(0.01)	1.83(0.00)	1.80(0.00)	1.84(0.01)	1.43(0.00)	1.89(0.02)	0.07(0.07)^{b}
	III	2.19(0.01)	1.88(0.02)	1.84(0.00)	1.77(0.01)	1.86(0.00)	1.43(0.00)	1.89(0.01)	0.05(0.17)^{d}

[§] Mean bond lengths and standard deviations are reported in Å, X-ray-refined experimental values of *p*-VCPO are shown for comparison. [+] denotes a flipped HIS411.[†] MAD is the mean absolute deviation and Δ_{MAX} is the maximum deviation relative to the X-ray distances of 1IDU.pdb. [a] Δ_{MAX} observed for V-N^{His}. [b] Δ_{MAX} observed for V-O^{1}. [c] Δ_{MAX} observed for V-O^{2}. [d] Δ_{MAX} observed for V-O^{3}. [e] Δ_{MAX} observed for V-O^{4}, [f]Water in MM region. [+] denotes models with a flipped HIS411 conformation. [‡]O^{wat} is in the MM region for the small QM I models. [*]Proton transfers to HIS404.

The X-ray structure of *p*-VCPO with a resolution of 2.24Å determined by Messerschmidt et al.[39] shows the terminal oxo ligands O^{3} and O^{1} at distances of 1.60 Å and 1.93Å , respectively, suggesting that these atoms are unprotonated and protonated, respectively. The QM/MM-optimized bond distances from vanadium to the O^{3} and O^{1} ligands for the *p*-VCPO-S1 model fit best to the experimentally refined values, however the experimental uncertainty is too large to provide definitive conclusions. All other models that are not protonated at O^{1} have the V-O^{1} bond length elongated by up to 0.27Å relative to the X-ray structure.

The rather long refined V-O^1 distance might also nurture the speculation that this O atom could be doubly protonated, i.e. be present as a water ligand. In fact, the resulting ligand environment about the metal would be reminiscent of $VO(O_2)(H_2O)_n^+$, the formulation of peroxovanadate at low pH. Because the nearby Arg360 residue is probably protonated and a good H-bond donor, this situation is unlikely so that we did not set up and prepare any *p*-VHPO-D11 models from the onset. In order to test this expectation, we constructed such a model from one of the *p*-VCPO-S1 minima (optimized with QM region III) by manually moving the proton from Arg360 to O^1. During optimization of this *p*-VHPO-D11 structure with a neutral Arg, the proton jumped back to the arginine, affording the regular *p*-VCPO-S1 model.

In general, most of the vanadium-ligand bond distances increase between ~0.01 to ~0.05 Å when going from the smaller QM region I to the larger QM III region. In contrast, the V-O^4 bond in the singly protonated model *p*-VCPO-S4 shortens by about 0.17 Å when going from the smaller to the larger QM III region. This is attributed to a proton transfer from the O^4 ligand to a crystal water, and a subsequent proton transfer to the neighboring HIS404, which occurred spontaneously during optimization. This creates an active site which is structurally similar to the originally unprotonated model, *p*-VCPO-Z, except for an additionally charged histidine group (HIS404). This occurs for both the *p*-VCPO and *p*-VBPO active sites, see Table 4.4.2 and Table 4.4.3, respectively. Likewise, for *p*-VBPO-S3 the proton is transferred from the O^3 ligand to HIS404 during QM/MM optimisation with QM region II, but the same proton hops back to O^3 during optimization with QM region III (no such event occurred with the corresponding *p*-VCPO models).

The distance between the vanadium and N^{HIS496} is typically shorter in the QM/MM optimized models than in the X-ray structure, by up to 0.13Å for *p*-VCPO-Z and *p*-VCPO-S4; smaller deviations are typically seen when a crystal water molecule is bound to the peroxovanadate center, as in the other singly protonated models. The computed V-N^{HIS496} distances vary between 2.06Å up to 2.19Å for the *p*-VCPO models, while they range from 2.3Å to 2.6Å in the native VCPO forms, suggesting that peroxovanadate binds more strongly than vanadate to HIS496. This trend has already been observed by Renirie et al.,[58] and aqueous peroxovanadate has been found to bind imidazole more strongly than vanadate.[133]

Typically the QM/MM-optimized bond distances of *p*-VCPO and *p*-VBPO, with the same protonation state of the peroxovanadate cofactor, show little variation between these two proteins, which adds support to the validity of homology modeling in the case of *p*-VBPO. The *p*-VCPO-Z and the related *p*-VCPO-S4 models (essentially *p*-VCPO-Z with protonated HIS) contain an empty cavity above the peroxovanadate moiety, in agreement with the X-ray structure. The remaining models contain a water molecule in this position, with a computed distance short enough to be considered bound to the metal (V-O distances between 2.1 and 2.5 Å), in apparent disagreement with experiment, where this site remains empty.

It is difficult at this point to judge the significance of the seemingly empty coordination site in the solid-state structure. If this site is truly vacant, *p*-VCPO-Z and *p*-VCPO-S4 would fit much better to the refined coordinates than all other models. For these two models, the mean absolute deviation (MAD) between all optimized and experimental bond lengths in Table 4.4.2 amounts to 0.14 and 0.18 Å, respectively (QM region III), whereas those of the other candidates approach or exceed 0.4 Å (see last but one column in Table 4.4.2). If, on the other hand, this site was occupied by a very mobile or labile water ligand, the latter might well escape detection in the X-ray analysis, given the low resolution presently achievable. In that case, i.e. when the V-O^{wat} distance is disregarded in the analysis, the resulting MAD is lowest for the *p*-VCPO-S1 model (0.03Å, see last column in Table 4.4.2) and shows little variation across the remaining models, typically around 0.1Å. We refrain from making structural assignments based on these data and note that essentially all models presented in Table 4.4.2 could be viable candidates, with a slight preference for *p*-VCPO-Z and *p*-VCPO-S1 (or, perhaps, an equilibrium mixture between both).

In this context the intrinsic water binding energy of the pristine $V(O_2)O(OH)(H_2O)(Im)$ fragment (Im = methyl-imidazole) common to all *p*-VCPO-S models is of interest. Starting from the corresponding coordinates of the *p*-VCPO-S1 minimum obtained with QM region II, this complex was optimized in the gas phase. In the resulting minimum, the V-O^{wat} distance is 2.35 Å. At the RI-BP86/AE1 level, water dissociation affording $V(O_2)O(OH)(Im)$ is computed to be endothermic by 11.5 kcal/mol (6.0 kcal/mol when corrected for basis-set superposition error). The water is thus indicated to be weakly bound. An attempt to optimize a non-protonated hydrated complex, i.e. $[V(O_2)O_2(H_2O)(Im)]^-$, resulted in detachment of the water ligand from the metal, in agreement with the results from the full QM/MM computations.

Hydrogen bonding

In all of the above-considered models of *p*-VCPO the HIS404 residue (which is included in QM region III) is no longer hydrogen bonded to the peroxovanadate cofactor, see Figure 1.2.2 This residue acts as a hydrogen bond acceptor from the nearest solvent water molecule, which in turn donates a hydrogen bond to the peroxovanadate cofactor. The O^1 ligand of the peroxovandate is hydrogen bonded to ARG360 and ARG490. The singly protonated *p*-VCPO-S1 model no longer forms a hydrogen bond with ARG490 , as compared to its native counterpart, and this is also observed for the *p*-VCPO-D13 doubly protonated models. This causes a gross structural deviation relative to the other models, as seen in overlay of all the models in Figure 4.4.3, and the RMSD is much larger for *p*-VCPO-D13 when compared to other protonated and unprotonated models.

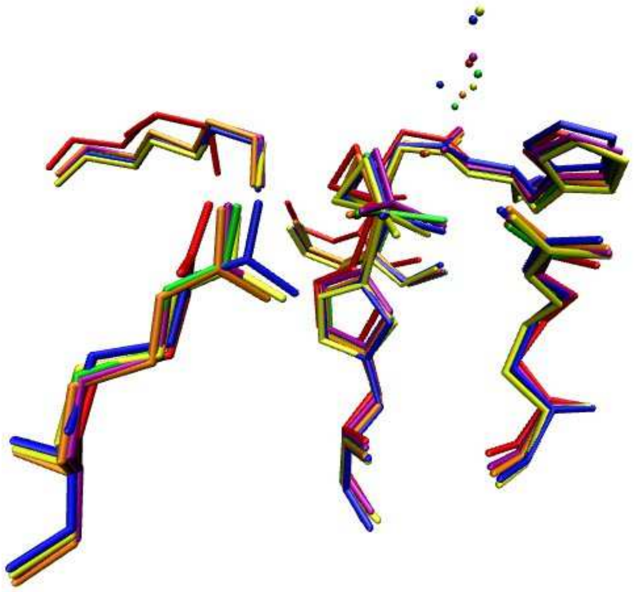


Figure 4.4.3 Overlay of the heavy atoms in the active site of *p*-VCPO for X-ray (Yellow),VOD13 (Red), VOS4 (Green), VOS3 (Orange), VO4 (Blue), VOS1 (Magenta) for the QM region III.

Table 4.4.4 Selected hydrogen bonds in QM/MM calculations using the QM III region for *p*-VCPO. Mean bond distances and standard deviations are reported in Å. CHARMM atom types are given in brackets and superscripted residue labels are consistent with the X-ray pdb file.

		O^1-H, [HH11]^{ARG360}	O^1-H, [HH21]^{ARG360}	O^1-H, [HH22]^{ARG490}	O^2-H, [HZ1]^{LYS353}	O^2-H, [HN]^{GLY403}	O^3-H, [HE]^{ARG490}	O^3-H, [HN]^{SER402}	O4-H^{wat}
p-VCPO-Z	III	1.91(0.06)	2.21(0.10)	2.15(0.12)	1.69(0.03)	1.80(0.02)	1.87(0.02)	2.07(0.10)	1.94(0.04)
p-VCPO-S1	III	1.94(0.00)	2.26(0.03)	--	1.79(0.01)	1.82(0.01)	2.07(0.01)	--	--
p-VCPO-S3	III	2.39(0.33)	2.04(0.13)	2.00(0.08)	1.76(0.03)	1.90(0.05)	2.00(0.10)	--	--
p-VCPO-S4	III	1.96(0.04)	2.00(0.01)	1.91(0.01)	1.74(0.01)	1.76(0.01)	1.96(0.02)	--	1.84(0.01)
p-VCPO-D13	III	--	--	--	--	2.13(0.05)	--	1.72(0.02)	--

Table 4.4.5 Selected hydrogen bonds in QM/MM calculations using the QM III region for *p*-VBPO. Mean bond distances and standard deviations are reported in Å. CHARMM atom types are given in brackets and superscripted residue labels are consistent with the X-ray pdb file. ARG 417 is not included in the QM region III of *p*-VBPO, but it is included in the *p*-VCPO QM region III. [+] Model with flipped HIS411

		O^1-H, [HH11]^{ARG380}	O^1-H, [HH21]^{ARG380}	O^2-H, [HZ1]^{LYS343}	O^2-H, [HN]^{GLY417}	O^3-H, [HN]^{SER416}	O4-H^{wat}
p-VBPO-Z	III	2.02(0.02)	2.15(0.11)	1.75(0.03)	1.84(0.01)	1.98(0.02)	1.96(0.14)
p-VBPO-Z[+]	III	1.91	2.08	1.69	1.81	--	--
p-VBPO-S1	III	>2.5	>3	1.88(0.01)	1.85(0.01)	--	--
p-VBPO-S2	III	1.85(0.08)	2.28(0.51)	1.83(0.03)	--	--	--
p-VBPO-S3	III	2.06(0.06)	2.35(0.52)	1.86(0.15)	2.07(0.14)	1.80(0.11)	--
p-VBPO-S4	III	1.88(0.04)	2.03(0.06)	1.75(0.03)	1.81(0.01)	--	--
p-VBPO-D13	III	--	--	2.11(0.07)	1.58(0.02)	1.72(0.02)	--

Table 4.4.6 Root-mean-square deviations (RMSD, in Å) for all heavy atoms included in the QM region I and III, relative to the X-ray structure.

	I	III
p-VCPO-Z	0.0 ± 0.01	0.2 ± 0.03
p-VCPO-S1	0.0 ± 0.00	0.3 ± 0.01
p-VCPO-S3	0.1 ± 0.01	0.3 ± 0.02
p-VCPO-S4[†]	0.1 ± 0.01	0.3 ± 0.00
p-VCPO-D13	0.1 ± 0.01	0.5 ± 0.06

[†] *p*-VCPO-S4 deprotonates to resemble the *p*-VCPO-Z model, with a charged histidine.

The LYS353 residue in *p*-VCPO, like the corresponding LYS341 in *p*-VBPO, forms a strong hydrogen bond to the O^2 peroxo oxygen ligand see Table 4.4.4 and Table 4.4.5. This hydrogen bond is generally believed to induce the bond polarization necessary for the heterolytic cleavage of the side-on-bound peroxide.[55] The major difference in the active sites of *p*-VCPO and the homology-modeled *p*-VBPO is a HIS411 in the position corresponding to a PHE397 in the *p*-VCPO. Since the HIS411 is within hydrogen bonding distance of LYS353, it has been suggested[49] that this residue in VBPO may form significant hydrogen bonds, thereby indirectly altering the reactivity of the enzyme. Previous site-directed mutagenesis studies have confirmed that these residues are crucial for activity.[64,52] This HIS411 may affect the reactivity, by either protonating or deprotonating the LYS353, thereby indirectly affecting the oxidative strength of these enzymes and their affinity for the halide. To understand the possible roles of the HIS411, a set of calculations were performed with HIS411 in a flipped conformation, that is, rotating the aromatic group of HIS residue so that the N^{ε} atom is facing away from the LYS341. Keeping in mind that the resolution of the X-ray structure makes it difficult to distinguish clearly between N and C atoms in the five-membered rings of histidine residues, special attention is called to the two possible orientations. Such a rotated HIS411 residue is stable only for two models, *p*-VBPO-Z and *p*-VBPO-S1, see Table 4.4.5, and for the rest of the models it flipped back to the original position during the QM/MM optimization. The geometric difference between the flipped conformations of a given protonation state is not particularly pronounced.

To summarize this part, the non- and singly protonated models *p*-VCPO-Z and *p*-VCPO-S1 appear to be the best candidates for the protonation state of *p*-VCPO. It is difficult to exclude all the other models, however, and many of these may be further regarded as possible candidates. A comparable degree of similarity was also observed in our earlier studies on the native-forms of VCPO and VBPO. The lack of *p*-VBPO X-ray data prevents a similar type of analysis for this enzyme at present.

Solvation

As the apical oxygen (O^4) in the native-form of VCPO is released and the peroxide binds side-on in the equatorial plane to the vanadium, an empty coordination site is generated and is able to accept the chloride ion during the catalytic cycle. In the crystallographic structure of *p*-VCPO, the nearest water molecules are about 4.4 Å and 5.7 Å away from the metal center, leaving a visibly large empty space at the latter. During the MD equilibration of *p*-VCPO-Z, these two water molecules diffuse closer to the vanadium, occupying this void that exists in the experimentally determined structure. The solvent molecules, Cryw420 and Cryw165, oscillate between ~3 and ~5.5 Å from the vanadium, due to an interchange of their positions, which is shown in Figure 4.4.4a. Whether this void can accommodate any additional water molecules that are as yet unidentified crystallographically, is an important question that may have mechanistic consequences.[134]

To probe if such an unidentified water molecule could have been missed by our standard solvation/equilibration protocol, three additional water molecules were placed manually into this cavity of *p*-VCPO-Z and a series of classical molecular dynamics simulations were performed. The additional solvent molecules were found to interchange with one of the two crystal water molecules within the active site, subsequently undergoing the same oscillation of ~3 or ~5.5Å from the vanadium throughout the MD simulation, see Figure 4.4.4b. Such frequent exchange processes are an indication of low barriers, presumably due to weak interactions. In general, there appears to be only sufficient room for two water molecules close to the metal, either one of the two crystallographically observed, or one of the added waters. The overall affect of the additional water is only marginal on the isotopic chemical shifts, typically less than 10ppm, as discussed in further detail below.

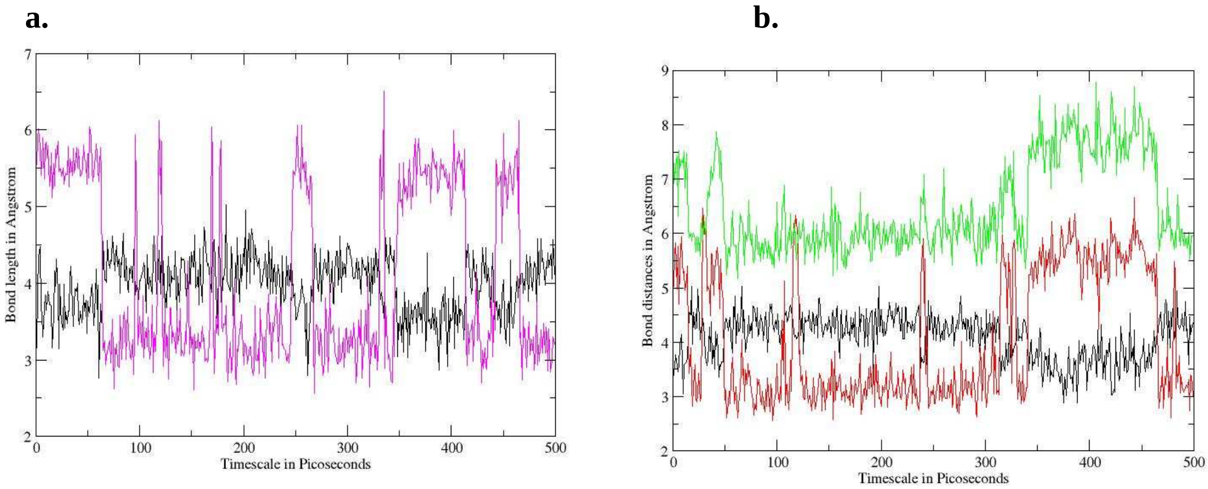
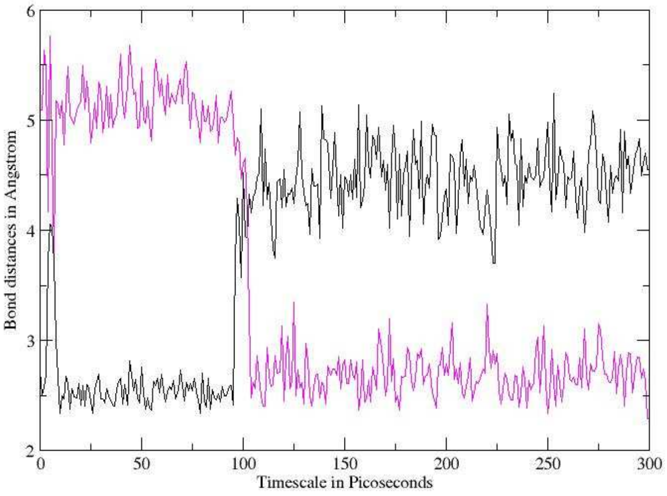


Figure 4.4.4 Plot of the distances between vanadium and the oxygen atoms of the nearest solvent molecules during MD initial equilibration. a. The crystal water molecule Cryw420 (black) and Cryw165 (pink) b. The crystal water molecule Cryw420 (black)and the additional water molecules SOLV$_z$(red) and SOLV$_y$(green), that were manually added during the solvation phase.

Figure 4.4.5 Plot of the distances between vanadium and the O atoms of the nearest solvent molecules during initial MD equilibration for 300ps. The crystal water molecules Cryw420 (black) and Cryw165 (pink).



Isotropic chemical shifts

An investigation is made whether ^{51}V chemical shifts may act as a useful probe for the assignment of protonation states of the peroxovanadate active site in vanadium haloperoxidase enzymes. The ^{51}V chemical shifts (δ_{iso}) of the vanadium nuclei in the different QM regions are given in Table 4.4.7 (computed with inclusion of the MM point charges).

Table 4.4.7 ^{51}V Isotropic chemical shifts (ppm) averaged over six snapshots for the two regions, together with the corresponding standard deviations. $\overline{\Delta}_{\delta}$ and $\overline{|\Delta|}_{\delta}$ are the mean signed and mean absolute deviations between the VCPO and VBPO models.

Model ID	QM region I		QM region III	
	p-VCPO	***p*-VBPO**	*p*-VCPO	***p*-VBPO**
Z	-704 ± 7	**-708 ± 6**	-696 ± 8	**-693 ± 15**
S1	-611 ± 3	**-668 ± 13**	-683 ± 20	**-757 ± 21**
S2	-741 ± 13	**-694 ± 64**	-700 ± 30	**-735 ± 30**
S3	-599 ± 23	**-612 ± 17**	-658 ± 35	**-617** ±34
S4	-751 ± 7	**-755 ± 9**	-739 ± 6	**-703 ± 16**
D13	-227 ± 10	**-327 ± 5**	-379 ± 28	**-428 ± 9**
$\overline{\Delta}_{\delta}$ ($\overline{|\Delta|}_{\delta}$)	58(58)		-9 (41)	

The isotropic chemical shifts across the QM regions I-III are fairly stable, where variations typically amount to less than 30ppm for most of the models (*p*-VHPO-Z, *p*-VCPO-S2,S4 and *p*-VBPO-S2,S3). The other models show larger fluctuations, with more significant shielding occurring on going from QM region I to III, up to ca. -90ppm. This shielding may be partially attributed to the binding of the apical water molecule as it gets included in the larger QM region (see discussion above). In the following, focus is set on the QM III results. For the isotropic chemical shifts there are some notable differences between *p*-VCPO and *p*-VBPO models, cf. the $\overline{|\Delta|}_{\delta}$ value of ca. 40 ppm in Table 4.4.7. The isotropic chemical shifts may be loosely collected together into two groups, based on their locations. Firstly, a downfield region from around ca. -400ppm has the doubly protonated models located within it. Secondly, an upfield region from -620 to -760ppm contains all singly and non-protonated models, with the axially protonated models representing the more upfield of these shifts.

Table 4.4.8 ^{51}V Isotropic chemical shifts (ppm) averaged over six snapshots, together with the corresponding standard deviations for *p*-VCPO, where $\Delta\delta_1$ is the difference between *p*-VCPO(III) and *n*-VOT144(IV), $\Delta\delta_2$ is difference between *p*-VCPO(III) relative to *n*-VOD14(IV) and $\Delta\delta_3$ is the difference between *p*-VCPO(III) and *n*-VOD34(IV).

	I	II	III	$\Delta\delta_1$	$\Delta\delta_2$	$\Delta\delta_3$	$\overline{\Delta}\delta$
p-VCPO-Z	-704 ± 7	-699 ± 9	-696 ± 8	-94	-116	-89	-100
p-VCPO-Z*	-650 ± 5	-646 ± 9	-669 ± 9				
p-VCPO-Z**	--	-704	-706	-104	-126	-99	-110
p-VCPO-S1	-611 ± 3	-713 ± 4	-683 ± 20	-81	-103	-76	-87
p-VCPO-S1*	-490 ± 3	-612 ± 3	-647 ± 3				
p-VCPO-S2	-741 ± 13	-697 ± 29	-700 ± 30	-98	-120	-93	-104
p-VCPO-S3	-599 ± 23	-675 ± 35	-658 ± 35	-56	-78	-51	-62
p-VCPO-S4	-751 ± 7	-729 ± 12	-739 ± 6	-137	-159	-132	-143
p-VCPO-S4[b]	-681 ± 10	-671 ± 15	-720 ± 7				
p-VCPO-D13	-227 ± 10	-314 ± 68	-379 ± 28	223	201	228	217

*Denotes calculation of isotropic chemical shifts without point charges.
**Denotes model with three additional solvent water molecules along with two crystal water molecules near the vanadium cofactor.

Table 4.4.9 ^{51}V Isotropic chemical shifts (ppm) averaged over six snapshots, together with the corresponding standard deviations[a] for *p*-VBPO, $\Delta\delta_1$ is the difference between *p*-VBPO(III) and *n*-VBPO-D34(III), $\Delta\delta_2$ is the difference between *p*-VBPO(III) and *n*-VBPO-T344(III).

	I	II	III	$\Delta\delta_1$	$\Delta\delta_2$	$\overline{\Delta}\delta$
p-VBPO-Z	-708 ± 6	-701 ± 11	-693 ± 15	-88	-80	-84
p-VBPO-Z*	-649 ± 5	-650 ± 17	-666 ± 13			
p-VBPO-Z[+]	--	--	-695			
p-VBPO-S1	-668 ± 13	-717 ± 49	-757 ± 21	-152	-144	-148
p-VBPO-S1*	-583 ± 11	-662 ± 50	-742 ± 14			
p-VBPO-S1[+]	--	--	-733			
p-VBPO-S2	-694 ± 64	--	-735 ± 30	-130	-122	-126
p-VBPO-S3	-612 ± 17	-669±20	-617 ± 34	-12	-4	-8
p-VBPO-S4	-755 ± 9	-695 ± 15	-703 ± 16	-98	-90	-94
p-VBPO-S4*	-676 ± 8	-649 ± 15	-685 ± 9			
p-VBPO-D13	-327 ± 5	-408 ± 21	-428± 9	177	185	181

*Denotes calculation of isotropic chemical shifts without point charges.
[+] Model with flipped HIS411.

Comparison of these isotropic shifts with experiment is hampered by the fact that only a single value is known, namely that of *p*-VBPO in solution at δ = -1135 ppm.[45] None of the computed values comes even close to this exceptionally shielded region. The same had been noted for the native VBPO, where the solution δ(^{51}V) value of -931 ppm reported in the same study was reproduced neither computationally[132] nor in a preliminary solid-state NMR study of the same enzyme.[51] What is clear, however, is that transformation of the native into the peroxo form entails a substantial shielding of the ^{51}V resonance, by ca. -200 ppm according to the VBPO solution study. Assessment of the corresponding trend for the QM/MM data depends noticeably on the particular model(s) used for the native forms. For instance, the three VCPO models that have reproduced the solid-state ^{51}V NMR-tensor elements reasonably well, VCPO-D14, -D34, and -T144, showed isotropic δ(^{51}V) values of -580, -607, and -602 ppm, respectively (largest QM region).[131] Except for the diprotonated *p*-VCPO-D13, all other *p*-VCPO models show clear upfield shifts from these numbers. For the *p*-VCPO-Z and *p*-VCPO-S1 variants, which emerged as particularly promising from the structural parameters discussed above, this computed upfield shift is between ca. -80 ppm and -120 ppm. Comparing the same VBPO and *p*-VBPO models models with each other135 affords computed upfield shifts around ca. -70 to -160 ppm. These (and the other singly protonated) models thus reproduce the observed trend qualitatively reasonably well. The quantitative underestimation of this effect may be rooted in shortcomings of the particular exchange-correlation functionals employed as these have shown problems to accurately reproduce similar trends between simple aqueous vanadate and peroxovanadate complexes.[136] In any event the qualitative agreement with experimental trend for most of the QM/MM models is encouraging. One model for each enzyme fails to reproduce this trend, namely the doubly protonated *p*-VHPO-D13, for which substantial downfield shifts are computed (Table 4.4.7). This candidate can thus be safely excluded.

Anisotropic chemical shifts

The four parameters that characterize a solid-state ^{51}V spectrum are are collected in Table 4.4.10.

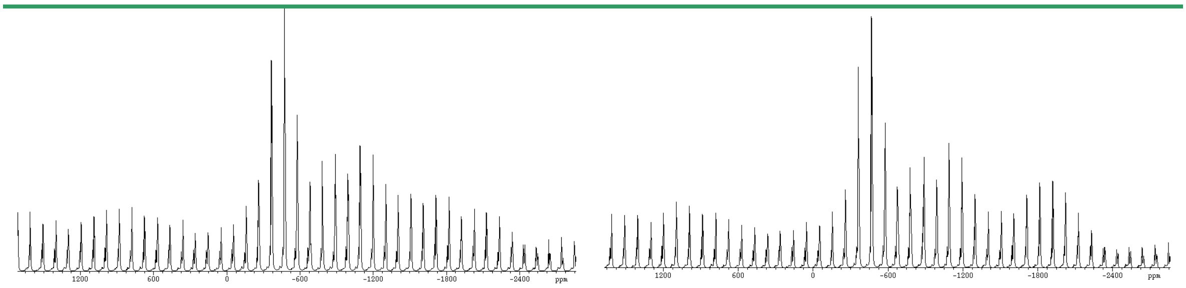
Table 4.4.10 Reduced anisotropy δ_σ, asymmetry $\eta_{\sigma,}$ nuclear quadrupole coupling constant C_Q (MHz) and asymmetry parameter η_Q with the associated standard deviations for six snapshots. MAPD is the mean absolute percent deviation across the four parameters relative to the average values for the S1 model (QM region III) of the respective haloperoxidases.

*Denotes calculation of isotropic chemical shifts without point charges.
**Denotes model with three additional solvent water molecules along with two crystal water molecules near the vanadium cofactor. +Model with flipped HIS411.

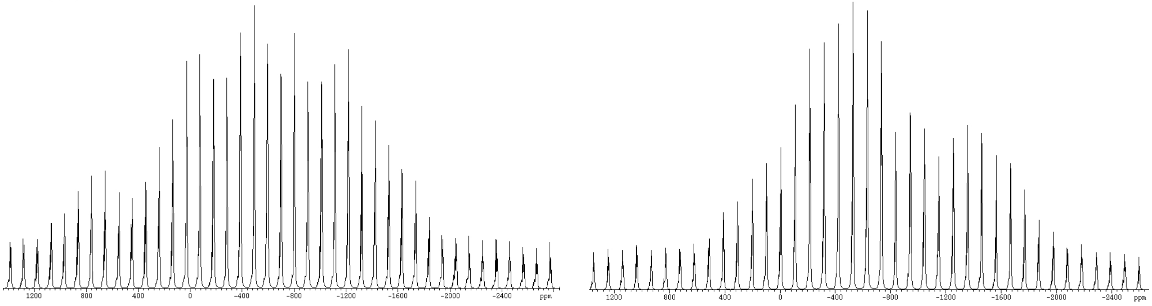
Model ID	*p*-VCPO ***p*-VBPO**	δ_σ	η_σ	C_Q	η_Q	MAPD
Z	I	-685 ± 9	0.4 ± 0.01	9 ± 0.35	0.30 ± 0.14	
	I	**-729 ± 7**	**0.3 ± 0.01**	**8 ± 0.12**	**0.25 ± 0.05**	
	II	-738 ± 6	0.3 ± 0.05	9 ± 0.23	0.32 ± 0.13	
	II	**-733 ± 8**	**0.3 ± 0.03**	**9 ± 0.42**	**0.22 ± 0.09**	
	III	-699 ± 6	0.3 ± 0.03	8 ± 0.36	0.28 ± 0.22	44
	III	**-695 ± 11**	**0.3 ± 0.06**	**8 ± 0.24**	**0.18 ± 0.06**	**52**
Z*	I	-684 ± 9	0.4 ± 0.01	11 ± 0.27	0.32 ± 0.13	
	I	**-686 ± 4**	**0.4 ± 0.01**	**11 ± 0.15**	**0.11 ± 0.04**	
	II	-698 ± 6	0.3 ± 0.05	11 ± 0.24	0.34 ± 0.17	
	II	**-728 ± 10**	**0.3 ± 0.02**	**11 ± 0.41**	**0.09 ± 0.03**	
	III	-677 ± 7	0.3 ± 0.03	9 ± 0.37	0.27 ± 0.15	--
	III	**-780 ± 12**	**0.1 ± 0.04**	**9 ± 0.23**	**0.2 ± 0.12**	**--**
p-VCPO-Z**	II	-730	0.3	10	0.29	
	III	-668	0.3	9	0.52	
p-VBPO-Z$^+$	III	-684	0.3	8	0.18	
S1	I	-809 ± 4	0.6 ± 0.01	-8 ± 0.06	0.84 ± 0.03	
	I	**-750 ± 12**	**0.6 ± 0.04**	**-8 ± 0.02**	**0.34 ± 0.04**	
	II	-951 ± 4	0.8 ± 0.01	6 ± 0.23	0.77 ± 0.07	
	II	**-965 ± 80**	**0.7 ± 0.17**	**-6 ± 1.36**	**0.67 ± 0.08**	
	III	-923 ± 6	0.8 ± 0.04	6 ± 0.30	0.65 ± 0.06	--
	III	**-1014 ± 5**	**0.5 ± 0.03**	**5 ± 0.45**	**0.86 ± 0.10**	**--**
S1*	I	-727 ± 2	0.9 ± 0.01	9 ± 0.11	0.95 ± 0.03	
	I	**-644 ± 4**	**0.9 ± 0.01**	**-9 ± 0.03**	**0.3 ± 0.08**	
	II	904 ± 4	0.9 ± 0.00	8 ± 0.20	0.49 ± 0.04	
	II	**-923 ± 18**	**0.9 ± 0.05**	**-7 ± 0.33**	**0.20 ± 0.15**	
	III	-913 ± 5	0.9 ± 0.00	7 ± 0.17	0.66 ± 0.05	--

	III	**-1032 ± 10**	**0.5 ± 0.04**	**-5 ± 0.76**	**0.72 ± 0.09**	**--**
p-VBPO-						
S1$^+$	III	-1052	0.4	7	0.6	--
S2	I	-668 ± 40	0.23 ± 0.03	9 ± 0.26	0.82 ± 0.10	
	I	**-501 ± 108**	**0.60 ± 0.26**	**9 ± 1.29**	**0.93 ± 0.02**	
	II	-837±8	0.2±0.02	14±0.7	0.51±0.07	
	II	**-466±114**	**0.5±0.32**	**9±0.12**	**0.10±0.03**	
	III	-739 ± 93	0.19 ± 0.08	12 ± 2.34	0.44 ± 0.13	57
	III	-618 ± 59	0.20 ± 0.10	8 ± 0.20	0.80 ± 0.13	42
S3	I	-800 ± 38	0.7 ± 0.02	4 ± 0.11	0.47 ± 0.35	
	I	**-708 ± 29**	**0.7 ± 0.03**	**7 ± 0.59**	**0.74 ± 0.17**	
	II	-918 ± 34	0.9 ± 0.03	6 ± 1.52	0.57 ± 0.25	
	II	**-711 ± 89**	**0.6 ± 0.19**	**9 ± 1.40**	**0.58 ± 0.28**	
	III	-901 ± 35	0.8 ± 0.18	7 ± 1.57	0.52 ± 0.34	10
	III	**-716 ± 74**	**0.7 ± 0.19**	**8 ± 1.63**	**0.67 ± 0.19**	**37**
S4	I	-800± 23	0.9 ± 0.01	13 ± 0.42	0.64 ± 0.05	
	I	**-772 ± 8**	**0.1 ± 0.02**	**13 ± 0.41**	**0.32 ± 0.03**	
	II	-725 ± 12	0.3 ± 0.01	8 ± 0.82	0.27 ± 0.03	
	II	**-692 ± 5**	**0.4 ± 0.07**	**8 ± 0.14**	**0.48 ± 0.02**	
	III	-672 ± 6	0.21± 0.00	7 ± 0.04	0.18 ± 0.03	47
	III	**-674 ± 17**	**0.3 ± 0.06**	**7 ± 0.58**	**0.24 ± 0.08**	**46**
S*	I	-834± 19	0.1 ± 0.00	15 ± 0.42	0.61 ± 0.04	
	I	**-815± 3**	**0.1 ± 0.01**	**15 ± 0.42**	**0.35 ± 0.03**	
	II	-695 ± 14	0.3 ± 0.02	10 ± 0.71	0.35 ± 0.02	
	II	**-688 ± 10**	**0.4 ± 0.07**	**10 ± 0.12**	**0.44 ± 0.10**	
	III	-636 ± 7	0.3 ± 0.01	7 ± 0.05	0.21 ± 0.00	--
	III	**-756 ±17**	**0.1 ± 0.03**	**9 ± 0.48**	**0.15 ± 0.13**	**--**
D13	I	1070 ± 10	0.7 ± 0.03	10 ± 0.57	0.62 ± 0.11	
	I	**981 ± 19**	**0.9 ± 0.04**	**9 ± 0.29**	**0.45 ± 0.06**	
	II	1157 ± 60	0.7 ± 0.06	5 ± 0.64	0.47 ± 0.06	
	II	**-1040 ± 18**	**0.9 ± 0.03**	**10 ± 0.90**	**0.17 ± 0.01**	
	III	1107 ± 34	0.7 ± 0.02	5 ± 0.24	0.79 ± 0.27	68
	III	**-1028 ± 18**	**0.9 ± 0.04**	**10 ± 0.93**	**0.04 ± 0.02**	**69**

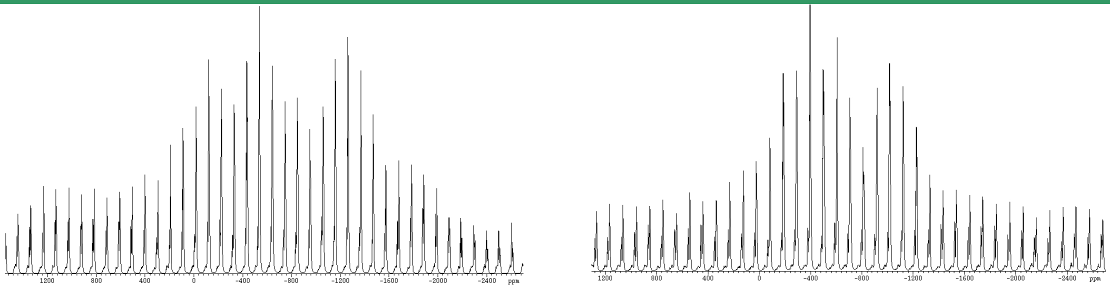
In general, small variations are observed in the calculated CSA and EFG values across the six snapshots. There are variations in the reduced anisotropy(δ_σ) when increasing the size of the QM region for all models, and these variations are heavily dependent upon the distance of the crystal water to the vanadium metal center. They are particularly pronounced for both the *p*-VCPO-S4 and *p*-VBPO-S4 models, which can be traced back to the proton transfer concomitant with the increase of the QM region. For the *p*-VBPO-S3 model, where a similar intermittent proton transfer was observed (see section on Geometric Details above), the variation in the reduced anisotropy is less pronounced.



a. *p*-VCPO-Z **b.** *p*-VBPO-Z



c. *p*-VCPO-S1 **d.** *p*-VBPO-S1



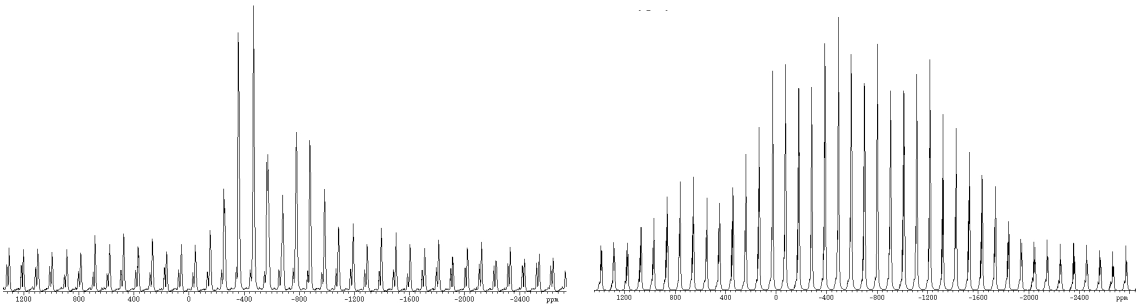
e. *p*-VCPO-S3 **f.** *p*-VBPO-S3

Figure 4.4.6 Simulated spectra of a. *p*-VCPO-Z, b. *p*-VBPO-Z, c. *p*-VCPO-S1, d. *p*-VBPO-S1 e. *p*-VCPO-S3, f. *p*-VBPO-S3, to visualize the difference between the unprotonated and singly protonated candidates in both *p*-VCPO and *p*-VBPO.

Considering the MAPD in Table 4.4.10, which are given relative to the *p*-VHPO-S1 models, it is apparent that there are substantial differences between the various candidates. The *p*-VHPO-S1 models are simply chosen as a reference since they are strong candidates. These differences can be even better appreciated in the spectra that are simulated using the QM/MM data (Figure 4.4.6). These results thus provide a promising avenue of discriminating between the likely candidates, once the corresponding experimental data becomes available.

In particular, the breadth of the experimental spectrum should be a good indicator for the protonation state, i.e. non- vs. singly protonated (compare, e.g. Figure 4.4.6a with 4.4.6c/e). It remains unclear at present whether the differences within the singly protonated models will be sufficient for specific assignments of the protonation site. According to the calculations, the differences should be more prominent for *p*-VPBO, which therefore appears to a better experimental target than *p*-VCPO for solid-state MAS ^{51}V NMR spectroscopy.

When the same *p*-VCPO and *p*-VBPO models are compared with each other, the variations in the tensorial properties are much less pronounced than when going from one protonation model to another. To appreciate the difference between a VHPO protein in the native or in the peroxo-form, illustrative examples of simulated ^{51}V NMR spectra for VCPO are given in Figure 4.4.7a and 4.4.7b.



a. *n*-VCPO-D14 **b.** *p*-VCPO-S1

Figure 4.4.7 Simulated spectra for one of the best candidates of the native-form of VCPO, a. VCPO-D14, compared to one of the best candidates for the peroxo-form, b. *p*-VCPO-S1.

As in our previous studies on the native enzymes, we assessed the direct effect of the surrounding peptide regions on the ^{51}V NMR and EFG tensors by simply deleting the point charges from the MM part in the property computation. The corresponding results, denoted by an asterisk, show similar trends as the tensor values for the full models given in Table 4.4.7. It was also checked for selected models that the inclusion of additional water molecules in the apical cavity discussed above (data denoted by a double asterisk), or the possible ring-flip of the extra HIS411 in *p*-VBPO (denoted by a dagger) , have only minor effects on the computed NMR and EFG tensors of the metal center.

TD-DFT

TD-DFT was employed to calculate electronic excitation energies from the QM/MM optimized models of both the native and peroxo-forms of VCPO and VBPO. Despite the known deficiencies of TD-DFT for long-range charge-transfer states,[137] this method has proven to be very promising for many transition-metal systems[138] and has recently been used to assess the protonation state of native VCPO.[139] Results for the two strongest, low-lying excitations for the *p*-VCPO and *p*-VBPO models are presented in Table 4.4.11. The last column of Table 4.4.11 shows the difference between the calculated excitation energy for the most intense of these two bands (QM I region) and the experimentally observed absorption band at 384nm reported by Renirie et al.[58] We list only results for the smaller QM region I because those for the larger QM region III were plagued by spurious long-range charge transfer states (e.g. from a lone pair on a distant peptide N atom to an empty metal-centered MO). In a series of test calculations on the p-VCPO-Z model, only small differences between the snapshots were observed, and therefore only results for a single snapshot are presented in Table 4.4.11.

In most cases there are two close-lying absorptions with similar intensities, which would probably not be resolved, as the UV-vis absorption bands reported for *p*-VCPO[58] or *p*-VBPO[140] are very broad. According to visual inspection of the MOs involved (QM I region) the excitations occur predominantly from the occupied π-orbitals on the imidazole ring of HIS496, or an occupied π* orbital on the peroxo moiety, into anti-bonding orbitals with a large d-character on vanadium.

Table 4.4.11 The transition energy λ in nm, the oscillator strength (f), and the coefficients (C_{ij}) of the CI expansion from TD-DFT calculations of the peroxo-form of *p*-VCPO and *p*-VBPO (QM region I). Δ denotes the difference between the calculated absorption maximum and the experimental absorption maximum at 384nm.

Model[a]	λ_{nm}(f)	Dominant Contribution	C_{ij}	Δ(λ-384)
p-VCPO-Z	336(0.01)	H → L	0.532	-48
	292(0.01)	H-1 → L	0.608	
_p_-VBPO-Z	**335(0.009)**	**H → L+1**	**0.457**	**-49**
	297(0.007)	**H → L**	**0.615**	
p-VCPO-S1	320(0.03)	H-2 → L	0.567	-8
	376(0.03)	H → L+1	0.693	
_p_-VBPO-S1	**469(0.006)**	**H → L**	**0.490**	**43**
	427(0.006)	**H → L**	**0.495**	
p-VCPO-S2	331(0.01)	H → L+2	0.661	-53
	336(0.01)	H-1 → L+1	0.461	
_P_-VBPO-S2	**329(0.008)**	**H → L+1**	**0.697**	**-55**
	292(0.004)	**H-2 → L**	**0.573**	
p-VCPO-S3	453(0.01)	H → L	0.669	69
	407(0.01)	H-1 → L	0.601	
_p_-VBPO-S3	**457(0.006)**	**H → L**	**0.653**	**73**
	342(0.005)	**H-1 → L+1**	**0.587**	
p-VCPO-D13	773(0.01)	H → L	0.687	389
	396(0.01)	H-1 → L	0.671	
_p_-VBPO-D13	**739(0.01)**	**H → L**	**0.678**	**355**
	437(0.06)	**H-1 → L+1**	**0.582**	

[a] No results are given for *p*-VHPO-S4 models, because they are not stable as such when the size of the QM region is increased.

Table 4.4.12 The transition energy λ in nm, the oscillator strength (f), and the coefficients (C_{ij}) of the CI expansion from TD-DFT calculations.

	λ_{nm}(f)	Dominant contribution	C_{ij}
$V(=O)_2(OH)_2$	268(0.008)	H-1 → L	0.598

I

	λ_{nm}(f)	Dominant contribution	C_{ij}
p-VCPO-S1*	337(0.02)	H-2 → L	0.558
	310(0.01)	H-1 → L	0.627
n-VCPO-S1	269(0.02)	H-1 → L+1	0.557
***n*-VBPO-S1**	**279(0.02)**	**H-1 → L+2**	**0.439**
n-VCPO-S3	281(0.02)	H-1 → L+1	0.402
***n*-VBPO-S3**	**274(0.02)**	**H-2 → L**	**0.392**
n-VCPO-S4	273(0.02)	H-3 → L	0.368
***n*-VBPO-S4**	**278(0.01)**	**H → L**	**0.381**
n-VCPO-T234	431(0.009)	H → L+1	0.699
***n*-VBPO-T234**	**340(0.01)**	**H-2 → L**	**0.552**

*Denotes point-charges are not present.

Most models agree with experiment within ca. 50 nm (or ca. 0.4 eV - 0.5 eV in excitation energies), a rather modest accuracy for TD-DFT, and it is difficult to assign or exclude one particular model based on these data. Because only the small QM region could be employed in the TDDFT computations, the effect of the surrounding MM charges is noticeable. For the p-VCPO-S1 model, for instance, the λ_{max} value decreases from 376 nm (Table 4.4.11) to 337 nm (see first row in the above Table 4.4.12) upon deletion of these point charges.

In the resting state of the native VCPO enzyme, the UV-vis band appears at 305 nm - 315nm, depending on the pH value.[58] TD-DFT computations for selected native VCPO and VBPO models furnished strong absorptions in the observed region, again without clear distinctions for specific protonation states (See Table 4.4.12). The computed changes between the native and peroxo forms are more significant than differences between the two proteins at a particular stage of the catalytic cycle, as seen above for anisotropic chemical shifts. Interestingly the difference between VCPO and VBPO appears more pronounced in the peroxo-form. For the parent vanadate, $H_2VO_4^-$, a band at ca. 268 nm is computed,[141] i.e. distinctly blue-shifted relative to the vanadate cofactors in the enzymes and in good accord with experiment (λ = 270 nm).[141]

Conclusions

We have used a QM/MM scheme to compute structures, ^{51}V NMR tensors, and UV-Vis excitation energies of two vanadium-dependent haloperoxidases in their peroxo forms, namely *p*-VCPO and *p*-VBPO. The latter has been modeled for the first time, assuming homology of the active site with that of *p*-VCPO, for which the structure is known in the solid. Special attention was called to the dependence of the properties on the protonation state of the peroxovanadate moiety and the ensuing hydrogen-bond network. The ligand environment about the metal as observed in solid *p*-VCPO is best preserved in the unprotonated models, whereas most of the singly protonated variants bind a nearby water molecule, which is present as free crystal water in the solid. Judging from the agreement of optimized bond lengths and the isotropic ^{51}V chemical shifts with experiment, the non- or singly protonated models are the most likely structural candidates for these enzymes. Models that are doubly protonated (whether at the terminal oxo or the peroxo atoms) can be excluded, as they are either unstable during optimization, or afford isotropic chemical shifts that fail to reproduce the experimentally observed increase in shielding on going from native to peroxo forms.

For a further discrimination between non- and singly protonated models (and the specific site of protonation in the latter), isotropic chemical shifts and TDDFT-derived excitation energies are not sufficient. Theoretical chemical-shift and EFG tensors, on the other hand, predict substantial intrinsic differences between the various protonation models, in particular between non- and singly protonated ones. When an experimental solid-state ^{51}V MAS NMR of *p*-VCPO or *p*-VBPO will be reported eventually, a comparison to the simulated spectra reported herein should enable the discrimination between the possible candidates and furnish detailed insights into the structure of the active site.

The VCPO and VBPO proteins have been experimentally shown to have a different oxidation profile. Whether this is a result of a single step in the catalytic cycle or whether this is a fundamental difference throughout the cycle is a question that may be addressed systematically with QM/MM calculations. This study takes the first steps toward answering this question, as it was shown that the secondary environment of the proteins, i.e. outside the first hydrogen bonding sphere, effectively tunes the properties of the active site via indirect hydrogen bonds and electrostatic interactions. The subtle differences between the active sites of both native and peroxo forms of VCPO and VBPO are best probed by the anisotropic ^{51}V NMR properties, i.e. CSA and EFG tensors.

In summary, we have successfully analyzed a challenging reactive intermediate by using a combination of theoretical methods for a comprehensive comparison to a number of previously published experimental reports. The multidisciplinary nature of this comparison serves to gauge the strengths of the various theoretical approaches and to increase the confidence in the results. New interpretations of available experimental data, and predictions where such data is currently lacking, open the way for further studies on these intriguing systems.

5 Summary

This thesis has contributed to understanding of vanadium containing complexes via a study of the chemical shifts computed using Density Functional Theory. A series of biologically important Vanadium containing complexes in gaseous and condensed phases were investigated using static and dynamic calculations, and the results have been critically compared with the experimental data.

DFT has been used to determine the innocence of oxidized amavadin, a derivative of an intriguing natural product with an unusual structure. ^{51}V NMR chemical shifts have been computed for non-oxo vanadium(V) complexes related to oxidized amavadin, $[\Delta\text{-V}^{V}\{(S,S)\text{-hidpa}\}_2]^-$, ($H_3$hidpa = 2,2′-hydroxyiminodipropionic acid). According to model calculations, the unusual deshielding of the ^{51}V resonance is due to a combination of conventional substituent effects (e.g. oxo- vs. dihydroxo-, or alkoxy- vs. carboxylato-ligands), rather than to a non-innocent nature of the hidpa ligand. For selected diastereomeric vanadium hidpa complexes, Born-Oppenheimer molecular dynamics simulations have been carried out in order to rationalize the observed differentiation of ^{51}V NMR chemical shifts.

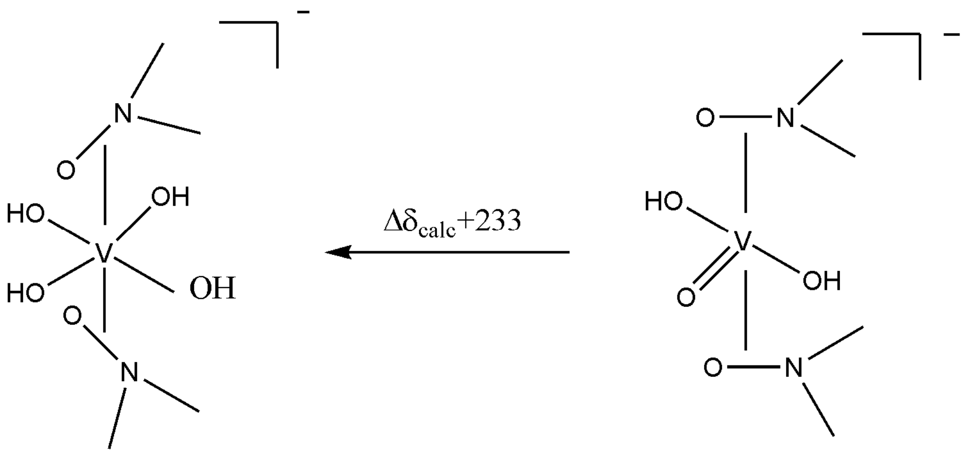
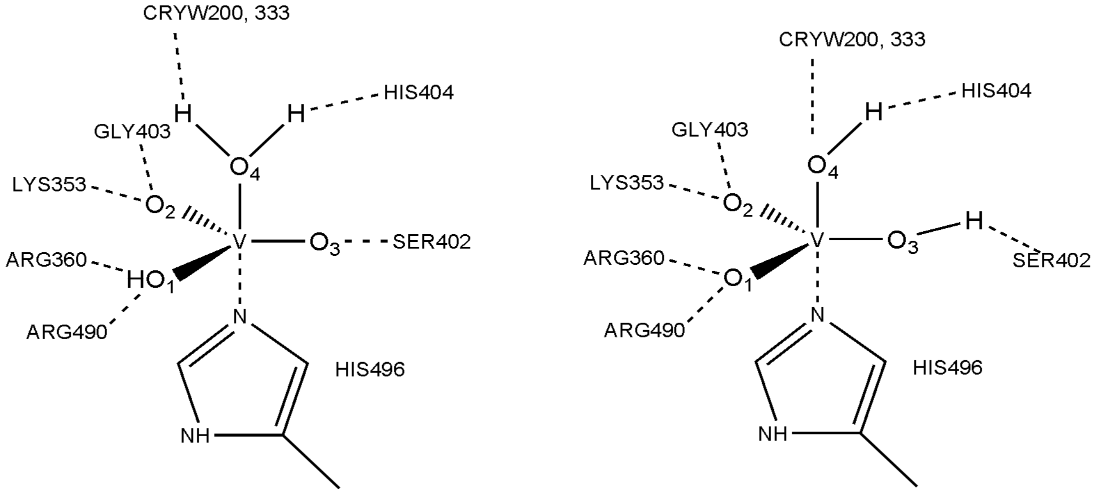


Figure 5.1 Going from an oxovanadium complex to a nonoxo species modelling the unusual ligand environment in oxidized amanadin is responsible for a large part of the low shielding of the ^{51}V nucleus in the latter.

In the main part of this thesis, a protocol for computation of ^{51}V NMR chemical shifts of large molecules has been developed, validated, and applied to elucidate structural details of vanadium-containing enzymes. This protocol is based on a popular hybrid QM/MM model and has been extensively tested for computations of the ^{51}V NMR tensor parameters of an entire enzyme, vanadium-dependent chloroperoxidase (VCPO). Special attention is called to the protonation state and protonation sites of the vanadates cofactor, important information that has not yet been obtained experimentally. An extensive number of protonation states for the vanadium cofactor (active site of the protein) and a number of probable positional isomers for each of the protonation states have been considered. Isotropic chemical shifts are found to be a poor indicator of the protonation state, however, anisotropic chemical shifts and the nuclear quadrupole tensors appear to be sensitive to changes in the proton environment of the vanadium nuclei. It turns out that only a small number out of many possible candidates with different protonation states and H-bond networks can be reconciled with the experimental solid-state MAS ^{51}V NMR data. The computations show that the effect of the peptide environment on the metal shift in the active site is indirect: the peptide backbone acts as scaffold dictating the precise location of the ligands about the metal, which is more important for the tensor elements than the direct response of the wave function to the charge distribution of the environment.



VOT144

VOD34

Figure 5.2 Schematic sketches of VCPO candidates that can reproduce the experimental NMR data)

The observed unusual ^{51}V chemical shift in the related vanadium-containing bromoperoxidase (VBPO) is only partly reproduced using that QM/MM based computational approach. The primary difference between these related structures, the presence of a histidine in VBPO whereas a phenylalanine is located at that position in VCPO, is studied via analysis of the respective theoretical ^{51}V NMR parameters. Theoretical ^{51}V chemical shift and EFG tensors for VBPO and VCPO show only minor differences in the isotropic $\delta(^{51}V)$ value, but somewhat more pronounced changes in the tensor values of both enzymes. Once these tensor values can be accurately refined experimentally for VBPO, selection of the best model(s) should be possible, which would furnish insights into structural details (protonation state, H-bond network) of this enzyme.

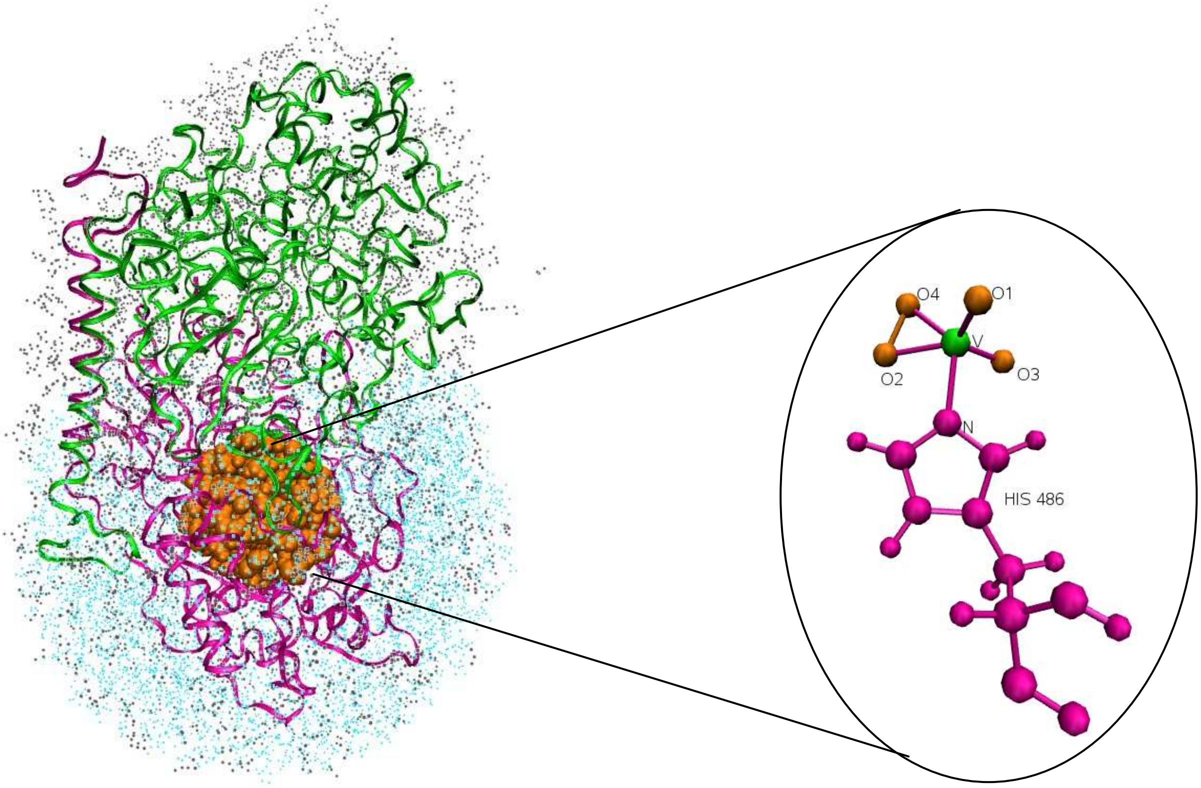


Figure 5.3 Structure predicted for peroxo-VBPO via homology modeling. The predicted V NMR and EFG tensors should allow to assign protonation state (and site) once experimental data become available.

Finally, theoretical CSA and EFG tensors have been computed from QM/MM models of VCPO and VBPO in their peroxo-forms. As the structure of the peroxo form of VBPO has not been determined experimentally or studied computationally so far, the homology model presented in this thesis constitutes the first explicit structural proposal for this important intermediate. Based on the trends in isotropic shifts on going from the native to the peroxo-forms and on the requirement for a stable side-on peroxo unit, the possibility for doubly protonated models can be excluded. The computed NMR and EFG tensors combined with TD-DFT results for UV/Vis bands highlight intrinsic differences that may provide means to elucidate structural differences between the remaining non- and singly protonated models. When an experimental solid-state ^{51}V MAS NMR of VBPO in the peroxo form will be reported eventually, a comparison to the simulated spectra reported herein should enable the discrimination between the possible candidates.

In summary, this thesis has helped to expand the limits of NMR computations significantly. Using an appropriate hybrid QM/MM methodology, it is now possible to compute transition-metal NMR properties for entire metalloenzymes. For a vanadium containing haloperoxidase, this approach turned out to be a valuable complement to experimental ^{51}V solid state NMR spectroscopy and X-ray crystallography, because a combination of all three methods can now be used to obtain refined structural information concerning the H-bond network in the active site. The multidisciplinary nature of this approach only serves to increases the confidence in the results, and holds great promise for further applications in the bioinorganic chemistry of transition metals. The calculations were not only used for interpretation and assignment of available experimental data, but also to make predictions where the experimental data is currently lacking, thus leading the way for further experimental studies.

6 Outlook

One area where the present approach can be immediately applied is that of VCPO mutants, which are being prepared and studied experimentally in order to assess the effects of individual residues on structure and catalytic activity. QM/MM computations of the type presented in this thesis can be expected to furnish detailed insights into the way how specific residues can fine-tune the H-bond pattern in the active site. Eventually it would be desirable to perform not just geometry optimizations, but also MD simulations at a suitable QM/MM level, in order to improve the sampling for the NMR properties.

Publications

Articles resulting form this thesis:

1. The Presumption of Innocence? A DFT-Directed Verdict on Oxidized Amavadin and Vanadium Catecholate Complexes. K. R. Geethalakshmi, Mark P. Waller, and Michael Bühl, *Inorg. Chem.* **2007,** *46*, 11297−11307.

2. ^{51}V NMR Chemical Shifts Calculated from QM/MM Models of Vanadium Chloroperoxidase. Mark P. Waller, Michael Bühl, K. R. Geethalakshmi, Dongqi Wang, and Walter Thiel, *Chem. Eur. J.* **2007**, *13*, 4723 – 4732

3. ^{51}V NMR chemical shifts from QM/MM models of Vanadium Bromoperoxidase. Waller, M. P.; Geethalakshmi K. R.; Buhl, M. *J. Phys.Chem.B.* **2008**, *112*, 5813-5823.

4. ^{51}V NMR chemical shifts calculated from QM/MM models of Peroxo-forms of Vanadium Haloperoxidases. K. R. Geethalakshmi, Mark P. Waller, Walter Thiel and Michael Bühl. *J. Phys. Chem. B*. Publication date(web) March 4th , **2009**.

References

[1] See for instance (a) Rehder, D.; in: Pregosin, P. S. (Ed.) *Transition Metal Nuclear Magnetic Resonance*, Elsevier, Amsterdam, **1991**, pp. 1-58; (b) Rehder, D.; Polenova, T.; Bühl, M. *Ann. Rep. NMR. Spectrosc*., **2007**, 62, 49-114.

[2] Rehder, D.,*Angew. Chem. Int. Ed. Engl.* **1991**, 30, 148-167.

[3] Rehder, D. *Biometals,* **1992**, 5, 3-12.

[4] Wilson, P. J.; Amos, R. D.; Handy, N. C. *Phys. Chem. Chem. Phys.,* **2000**, 2,187.

[5] Car, R.; Parrinello, M. *Phys Rev Lett* **1985**, 55, 2471.

[6] Senn, H. M.; Thiel, W. *Topics in current chemistry,* **2007**, 268, 173-290.

[7] Cornman, C. R.; Colopas, G. J.; Hoeschele, J. D.; Kampf, J.; Pecoraro, V. L. *J. Am. Chem. Soc.*, **1992**, *114*, 9925-9933.

[8] Bayer, E.; Kneifel, H. *Z. Naturforsch*., **1972** *,27B*, 20.

[9] Nawi, M. A.; Reichel, T.L. *Inorg. Chim. Acta.*, **1987**, *136*, 33.

[10] Fraústo da silva, J. J. R. *Chem. Speciation Bioavailability.*, **1989**, *1,* 139.

[11] Kneifel, H.; Bayer, E. *Angew. Chem.* **1973**, *85*,542-543.

[12] Krauss, P.; Bayer, E.; Kneifel, H. *Z. Naturforsch. B: Anorg. Chem. Org. Chem.*, **1984**, *39b*, 829-832.

[13] Bayer, E.; Koch, E.; Anderegg, G.. *Angew. Chem.* **1987**, *99*, 570-572.

[14] Bayer, E.; Koch, E.; Anderegg, G. *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 545-546.

[15] Carrondo, M. A. A. F. de C. T.; Duarte, M. T. L.S.; Pessoa, J .C.; Silva, J. A. L.; Fraústo da Silva, J. J. R.; Vaz, M. C. T. A.; Vilas-Boas, L. F. *J. Chem. Soc. Chem. Commun.* **1988**, 1158-1159.

[16] Fraústo da Silva, J. J. R.; Guedas Da Silva, M.F.C.; Da Silva, J. A. L. *Mol. Electrochem. Inorg. Bioinorg Organomet. Compds.* **1993**, 411-415.

[17] Remenyi, C.; Munzarova, M. L.; Kaupp, M. *J. Phys. Chem. B* **2005**, 109, 4227-4233.

[18] Butler, A.; Clague, M. J.; Meister, G. *Chem. Rev.* **1994**, *94*, 625-638.

[19] Nuber, B.; Wriss, J. *Acta. Crystallogr.* **1981**, *B37*, 947-948.

[20] Quilitzsch, U.; Wieghardt, K. *Z. Naturforsch.* **1979**, *34b*, 640-641.

[21] Meicheng, S.; Jing, L.; Zuohua, P.; Huadong, Z.; Youqi, T. *HuaXue XueBao*, **1990**, *11*, 280-285.

[22] Armstrong, E. M.; Beddoes, R. L.; Calviou, L.J; Charnock, J.M; Collison, D.; Ertok, S.N; Naismith, J. H.; Garner, C. D. *J. Am. Chem. Soc.*, **1993**, *115*, 807.

[23] Smith, P. D.; Berry, R.E.; Harben, S. M.; Beddoes, R. L.; Helliwell, M.; Collison, D.; Garner, C.D. *J. Chem. Soc., Dalton Trans.*, **1997**, 4509. See Reference 8 within.

[24] Lenhardt, J.; Baruah, B.; Crans, D.C.; Johnson, M. D. *Chem.Commun.*, **2006**, 4641.

[25] Paul, P.C.; Angus-Dunne, S.J.; Batchelor, R.J.; Einstein, F.W.B.; Tracey, A.S. *Can. J. Chem.*, **1997**, *75*, 183.

[26] (a) Keramidas, A.D.; Miller, S.M.; Anderson, O.P.; Crans, D.C. *J. Am. Chem. Soc.*, **1997**, *119*, 8901. (b) Tracey, A.S.; Jaswal, J.S. *J. Am. Chem. Soc.*, **1992**, *114*, 3835.

[27] Van Schijndel, J. W. P. M.; Barnett, P.; Roelse, J.; Vollenbroek, E. G. M.; Wever, R. *Eur. J. Biochem*. **1994**, *225*, 151-157.

[28] Martinez, J. S.; Carrol, G.L. ; Tschirret-Guth, R.A.; Altenhoff, G.; Little,R.D.; Butler, A. *J. Am. Chem. Soc.* **2001**, *123*, 3289-3294.

[29] Smith, T. S.; Pecoraro, V. L. *Inorg. Chem.* **2002**, *41*, 6754-6760.

[30] Butler, A. *Coord. Chem. Rev*. **1999**, *187*, 17-35.

[31] Everett, R.; Butler, A. *Inorg. Chem.* **1989**, *28*, 393-395.

[32] a) Vollenbroek, E. G. M.; Simons, L.H.; Van Schijndel, J. W. P. M.; Barnett, P.; Blazer, M.; Decker, H.; Vanderlinden, C.; Wever, R.; *Biochem. Soc. Trans*. **1995**, *23*, 267-271. b) Butler, A.; *Science*, **1998**, *281*, 207-210.

[33] Ligtenbarg, A. G. J. ; Hage, R. ; Feringa, B. L. *Coord. Chem. Rev.* **2003,** *237*, 89-101.

[34] Thompson, K. H.; McNeill, J. H.; Orvig, C.; *Chem. Rev*. **1999**, *99*, 2561-2571.

[35] R. Wever; H.L Dekkar; J.W. van Schijindel; E.G.M. Vollenbroek, WO patent 9527009, 1995

[36] D. C. Crans, J. J. Smee, E. Gaidamauskas, L. Yang, *Chem. Rev*. **2004**, *104*, 849-902.

[37] E. deBoer, K. Boon, R. Wever, *Biochemistry,* **1988**, *27*, 1629-1635.

[38] G. J. Colpas, B. J. Hamstra, J. W. Kampf, V. L. Pecoraro, *J. Am. Chem. Soc.* **1996,** *118*, 3469-3478.

[39] a) A. Messerschmidt, R. Wever, *Proc. Natl. Acad. Sci*. **1996**, *93*, 392-396; b) A. Messerschmidt, L. Prade, R. Wever, *Biol Chem*. **1997**, *378*, 309-315.

[40] G. Zampella, J. Y. Kravitz, C. E. Webster, P. Fantucci, M. B. Hall, H. A. Carlson, V. L. Pecoraro, L. De Gioia, *Inorg. Chem.* **2004**, *43*, 4127-4136.

[41] M. Bangesh, W. Plass, *J. Mol. Struct. (Theochem*) **2005**, *725*, 163-175.

[42] J. Y. Kravitz, V. L. Pecoraro, H. A. Carlson, *J. Chem. Theory Comput.* **2005,** *1*, 1265-1274.

[43] S. Raugei, P. Carloni, *J. Phys. Chem. B* **2006**, *110*, 3747-3758.

[44] A. Butler, H. Eckert, *J. Am. Chem. Soc.* **1989**, *111*, 2802-2809.

[45] D. Rehder, M. Casny, R. Grosse, *Magn. Reson. Chem*. **2004**, *42*, 745-749.

[46] Pooransingh-Margolis, N.; Renirie, R.; Hasan, Z.; Wever, R.; Vega, A. J.; Polenova, T. *J. Am. Chem. Soc.* **2006**, *128*, 5190-5208.

[47] Bühl, M.; Schurhammer, R.; Imhof, P. *J. Am. Chem. Soc*. **2004**, *126*, 3310-3320.

[48] For selected examples see: a) Cui, Q.; Karplus, M.; *J. Phys. Chem. B* **2000**, *104*, 3721-3743. b) Sebastiani, D. ; Röthlisberger, U.; *J. Phys. Chem. B* **2004**, *108*, 2807-2815. c) Moon, S.; Christiansen, P. A.; DiLabio, G. A.; *J. Chem. Phys.* **2004**, *120*, 9080-9086. d) Wang, B.; Merz, K. M.; *J. Chem. Theory Comput.* **2006**, *2*, 209-215. e) Bagno, A.; Amico, F. D'; Saielli, G. *J. Phys. Chem. B* **2006**, *110*, 23004-23006.

[49] Weyand, M.; Hecht, H. -J.; Kiess, M.; Liaud, M. F.; Vilter, H.; Schomburg, D. *J. MOL. BIOL.* **1999**, *293*, 595.

[50] Littlechild, J.; Garcia-Rodriguez, E.; Dalby, A.; Isupov, M. *J. Mol. Recognit.* **2002**, *15*, 291-296

[51] Pooransingh-Margolis, N. **2005**, PhD Thesis, University of Delaware

[52] Tanaka, N.; Hasan, Z.; Wever, R. *Inorg. Chim. Acta,* **2003**, 356, 288-296.

[53] Dau, H.; Dittmer, J.; Epple, M.; Hanss, J.; Kiss, E.; Rehder, D.; Schulzke, C.; Vilter, H. FEBS Letters 1999, 457, 237-240.

[54] Feiters, M. C.; Leblanc, C.; Küpper, F. C.; Meyer-Klaucke, W.; Michel, G.; Potin, P. *J. Am. Chem. Soc.*, **2005**, *127*, 15340-15341.

[55] Zampella, G.; Fantucci, P.; Pecoraio, V. L.; Gioia, L. D. *J. Am. Chem. Soc.* **2005**, *127*, 953-960.

[56] Crans, D. C.; Smee, J. J.; Graidamauskas, E.; Yang, L. *Chem. Rev.* **2004,** *104*, 849-902.

[57] Thompson, K. H.; McNeill, J. H.; Orvig, C. *Chem. Rev.* **1999,** *99*, 2561-2571.

[58] Renirie, R.; Hemrika, W.; Wever, R. *J. Biol. Chem.* **2000**, *275*, 11650.

[59] De Boer, E.; Tromp, M. G. M.; Plat, H.; Krenn, G. E.; Wever, R. *Biochim, Biophys. Acta.* **1986**, *872*, 104-115.

[60] Colpas, G. J.; Hamstra, B. J.; Kampf, J. W.; Pecoraro, V. L. *J. Am. Chem. Soc.* **1996**, *118*, 3469-3478.

[61] Zampella, G.; Kravitz, J. Y.; Webster C. E.; Fantucci, P.; Hall, M. B.; Carlson H. A.; Pecoraro, V.L.; De Gioia, L. *Inorg. Chem.* **2004**, *43*, 4127-4136.

[62] Hamstra, B. J.; Colpas, G. J.; Pecoraro, V. L. *Inorg. Chem.* **1998**, *37*, 949-955.

[63] Colpas, G. J.; Hamstra, B.J.; Kampf, J.W.; Pecorraro, V.L. *J. Am. Chem. Soc.* **1996**, *118*, 3469-3478.

[64] Hemrika, W.; Renirie, R.; Macedo-Riberio, S.; Messerschmidt, A.; Wever, R. *J. Biol. Chem.* **1999**, *274*, 23820-23827.

[65] Amara, P.; Field, M. J. *In Computational Molecular Biology*; Leszczynski, J. , Ed.; Elsevier: Amsterdam, **1999**,

[66] McCammon, J; Gelin, J. B.; Karplus, M.. Nature, 1977, 267: 585–590

[67] Phineus R. L. Markwick, Malliavin, T.; Nilges, M. *PLoS Computational Biology*, **2008**, 4, 1-7.

[68] Thar, J.; Reckein, W.; Kirchner., B. *Topics in Current Chemistry*, **2007**, 268, 133-171.

[69] Frank Jensen, *Introduction to computational chemistry* Wiley, **2007**, 451.

[70] Hohenberg, P.; Kohn, W., *Phys. Rev*. 136,864B (1964)

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

[72] a) Martin, J. M. L.; Bauschlicher, Jr. C. W.; Ricca, A., *Comput. Phys. Commun.* 133, 189 (**2001**) b) Pe´ rez-Jorda´, J. M.; A. Becke, San-Fabia´ n, E., *J. Chem. Phys.* 100, 6520 (**1994**). c) Pederson, M. R.; Jackson, K. A., *Phys. Rev. B* 41, 7453 (**1990**) d) Delley, B. ; in *Modern Density Functional Theory: A Tool for Chemistry,* edited by P. Politzer and J. M. Seminario Elsevier, New York (1994) e) Krack, M.; Köster, A. M., *J. Chem. Phys.* 108, 3226 (**1998**)

f) Challacombe, M., *J. Chem. Phys*. 113, 10037, **(2000)**

[73] Gross, E. K. U., W. Kohn, *Phys. Rev. Lett.*, **1985**, 55, 2850 (1985).

[74] Dreuw, A., Weisman, J. L., Head-Gordon M. *J. Chem. Phys*. **2003**, 119, 2943-2346

[75] a) Hehre, W. J.; Stewart, R.F.; Pople, J.A., *J.Chem. Phys*. 51, 2657 (1969) b) Hehre, W. J. ; Stewart, R.F. ; Pople, J.A., *J.Chem. Phys*. 52, 2769 (1970) c) Gordon, M.S.; Bjorke, M.D.; Marsh, M.J.; Korth, M.S., *J. Am. Chem. Soc.* 100, 2670 (1978)

[76] Dunning, Jr. T. H., *J. Chem. Phys.* 53, 2823 (1970)

[77] Hehre, W. J.; Ditchfield, R. ; Pople, J.A., *J. Chem. Phys.* 56, 2257 (1972)

[78] a) Hariharan, P.C.; Pople, J.A., *Theoret. Chimica. Acta*, 213 (1973) b) Francl, M.M.; Petro, W.J.; Hehre, W.J.; Binkley, J.S.; Gordon, M.S.; DeFress, D.J.; Pople, J.A., *J. Chem. Phys.* 77, 3654 (1982) c) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* 1984, 809, 3265.

[79] Bakowies,D.; Thiel, W. *J. Phys. Chem.* **1996**, 100, 10580–10594.

[80] Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, 7, 718–730.

[81] Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, 11, 700–733.

[82] Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **1999**, 20,1468–1494.

[83] Dick, B. G.; Overhauser, A. W. *Phys. Rev*. **1958**, 112, 90.

[84] Thole, B. T. *Chem. Phys.* **1981**, 59 341–350

[85] Bühl, M. *Ann. Rep. NMR Spectrosc*. **2008**, *64*, 77-126

[86] Ditchfield, R. Mol. Phys. **1974**, 27, 789.

[87] Wolinski, K.; Hinton, H.F.; Paulay, P. *J. Am. Chem. Soc.* **1990,** 112, 8251.

[88] Womacott, T.M.; Womacott, R. J. Introductory Statistics (fifth ed.), **1990,** Wiley, New York 1990.

[89] See, for example: (a) Grant, D. M.; Harris, R. K., Eds. Encyclopedia of nuclear magnetic resonance; Wiley: New York, 1996. (b) Wüthrich, K. NMR in structural biology; World Scientific: Singapore, 1995. (c) Brooks, C. L., III; Karplus, M.; Pettitt, B. M. Proteins: A Theoretical Perspective of Dynamics, Structure, & Thermodynamics; Adv. Chem. Phys. LXXI; John Wiley & Sons: New York, 1988.

[90] See, for example: (a) Ösapay, K.; Case, D. A. J. Biomol. NMR 1994, 4, 215. (b) Case, D. A.; Dyson, H. J.; Wright, P. E. Methods Enzymol. 1994, 239, 392. (c) Case, D. A. In Encyclopedia of Computational Chemistry; John Wiley: New York, 1998. (d) Sitkoff, D.; Case, D. A. Prog. NMR Spectrosc. 1998, 32, 165. (e) deDios, A. C.; Pearson, J. G.; Oldfield, E. Science 1993, 260, 1491. (f) Le, H.; Oldfield, E. J. Biomol. NMR 1994, 4, 341. (g) Halvin, R. H.; Le, H.; Laws, D. D.; deDios, A. C.; Oldfield, E. J. Am. Chem. Soc. 1997, 119, 11951. (h) Wishart, D. S.; Sykes, B. D.; Richards, F. M. J. Mol. Biol. 1991, 222, 311.

[91] See, for example: Parr, R. G.; Yang, W. Density-Functional Theory of Atoms and Molecules; Oxford University Press: New York, 1989.

[92] Sapay, K.; Theriault, Y.; Wright, P. E.; Case, D. A. J. Mol. Biol. 1994, 244, 183.

[93] Giessner-Prettre, C.; Pullman, B. Q. Rev. Biophys. 1987, 20, 113.

[94] (a) Ösapay, K.; Case, D. A. J. Am. Chem. Soc. 1991, 113, 9436. (b) Wijmenga, S. S.; Kruithof, M.; Hilbers, C. W. J. Biol. NMR 1997, 10, 337. (c) Williamson, M. P.; Asakura, T. J. Magn. Reson. B 1993, 101, 63.

[95] See, for example: Case, D. A. Curr. Opin. Struct. Biol. 1998, 8, 624.

[96] Mulliken, R. S., *J. Chem. Phys.*, **1995**, 23, 1833.

[97] Reed, A. E.; Weinstock, R. B.; Weinhold, F. J. *Chem. Phys.*, **1985,** 83, 735

[98] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M.

Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian 03, Gaussian, Inc., Pittsburgh PA, 2003.

[99] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100.

[100] Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822-8824; Perdew, J. P. *Phys. Rev. B* **1986**,*34*, 7406.

[101] (a) Wachters, A. J. H. *J. Chem. Phys.* **1970**, *52*, 1033-1036. (b) Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377-4384

[102] (a) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257-2261. (b) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta.* **1973**, *28*, 213-222. (c) Binning Jr., R. C.; Curtiss, L. A. *J. Comput. Chem.* **1990**, *11*, 1206.

[103] See for instance Koch, W.; Holthausen, M. C., *A Chemist's Guide to Density Functional Theory,* Wiley-VCH, Weinheim, 2000, and the extensive bibliography therein.

[104] Ahlrichs, R.; Bär, M.; Häser, M.; Korn, H.; Kölmel, M. *Chem. Phys. Lett.*, 1989, **154**, 165.

[105] (a) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R., *Chem. Phys. Lett.*, 1995, **240**, 283. (b) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.,* 1997, **97**, 119.

[106] http://www.rcsb.org/pdb/home/home.do

[107] Li, H.; Robertson, A. D.; and Jensen, J. H. *Proteins* **2005**, *61*, 704-721.

[108] Accelrys Inc., InsightII v. 2000. San Diego: Accelrys Inc., 2004.

[109] Klähn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005,** *109*, 15645-1650.

[110] Sherwood, P.; de Vries, A. H.; Guest, M. F., Schreckenbach, G. ; Catlow, C. R. A.; French, S. A. ; Sokol, A. A.; Bromley, S. T. ; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C. ; Casci, J.; Watson, M.; King, Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C; *J. Mol. Struct.* (*Theochem*) **2003**, *632*, 1-28.

[111] Smith, W.; Forester, T. ; *J. Molec. Graphics,* **1996**, *14*, 136.

[112] a) MacKerell, A. D.; Bashford, Jr., D.; Bellott, M.; Dunbrack, R. L.; Evanseck, Jr., J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D. Kuchnir, L. ; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, III, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J. ; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M .*J. Phys. Chem. B* **1998**, *102,* 3586-3616; b) MacKerell, A. D.; Feig, Jr., M.; Brooks III, C. L. *J. Comp. Chem.* **2004**, *25*, 1400-1415.

[113] Corresponding to model B in: D. Bakowies, W. Thiel, *J. Phys. Chem.* **1996**, *100*, 10580-10594.

[114] Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785-789.

[115] T. Clark, J. Chandrasekhar, G. Spitznagel, P. Schleyer, *J. Comput. Chem.* **1983**, *4*, 294-301.

[116] a) U. Haeberlen, In Advances in Magnetic Resonance; Suppl. 1; J. S. Waugh, Ed.; Academic Press: New York, **1976**. b) M. Mehring, Principles of High Resolution NMR in Solids, 2nd. ed.; Springer Verlag: Berlin, **1983**. c) H. W. Spiess, In NMR Basic Principles and Progress; P. Diehl, E. Fluck, R. Kosfeld, Eds.; Springer Verlag, Berlin, **1978**, Vol. 15.

[117] as implemented in G03: (a) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 404-417; (b) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys*. **2002**, *117*, 43-54. (c) Cossi, M.; Crescenzi, O. *J. Chem. Phys.* **2003**, *19*, 8863-8872.

[118] Bühl, M.; Kabrede, H. *J. Chem. Theory Comput*. **2006**, *2*, 1282-1290

[119] The stationary point for **6** turned out be a transition state, we discuss it here because it has the same orientations of the OH groups as in **2a.** The true minimum with an additional intramolecular H bridge is 0.65 kcal/mol lower in energy and it has δ =-630.

[120] Geethalakshmi K.R., Waller M.P., Bühl, M. *Inorganic Chem.,* **2007,** 46, 11297.

[121] Koch, E.; Kneifel, H.; Bayer, E. *Z. Naturforsch*, **1986**, *41b*, 359-362.

[122] (a) Grigoleit, S.; Bühl, M. *Chem. Eur. J.*, **2004**, *10,* 5541-5552. (b) Bühl, M.; Grigoleit, S. *Organometallics* **2005**, *24*, 1516-1527.

[123] In each case, the mean values over the entire trajectories correspond closely to the mean values over the snapshots selected for the NMR computations.

[124] a) G. Schreckenbach, T. Ziegler, *Int. J. Quantum. Chem.* **1997**, *61*, 899-918; b) R. Bouten, E. J. Baerends, E. van Lenthe, L. Visscher, G. Schreckenbach, T. Ziegler, *J. Phys. Chem. A* **2000,** *104*, 5600-5611.

[125] M. Bühl, M. Parrinello, *Chem. Eur. J.* **2001**, *7*, 4487-4494.

[126] Bak, M.; Rasmussen, J.T.; Nielsen, N.C.J.; *Magnetic. Reson.* **2000**, 147, 296-330.

[127] Study done by Mark P. Waller.

[128] M. Bühl, F. T. Mauschick, F. Terstegen, B. Wrackmeyer, *Angew. Chem. Int. Ed.* **2002**, *41*, 2312-2315.

[129] M. Bühl, S. Grigoleit, H. Kabrede, F. T. Mauschick, *Chem. Eur. J.* **2006**, *12*, 477-488.

[130] M. Sterzel, J. Autschbach, *Inorg. Chem.* **2006**, *45*, 3316-3324.

[131] Waller, M. P.; Bühl, M.; Geethalakshmi K.R.; Wang, D.; Thiel, W. *Chem. Eur. J.* **2007**, 13, 4723-4732.

[132] Waller, M. P.; Geethalakshmi K. R.; Bühl, M. *J. Phys.Chem.B.* **2008**, *112*, 5813-5823

[133] Andersson, I.; Angus-Dunne, S.; Howarth, O.; Pettersson, L. *J. Inorg. Biochem.* **2000**, *80*, 51-58.

[134] For instance, the presence of a single water molecule in the active site can affect activation barriers significantly: Altun, A.; Guallar, V.; Friesner, R. A.; Shaik, S.; Thiel, W. *J. Am. Chem. Soc.* **2006**, 128, 3924-3925.

[135] I.e. VCPO-D14, -D34, and -T144 with $\delta(^{51}V)$ = -624, -605, and -592 ppm, respectively, for the native form, and *p*-VBPO-Z and *p*-VBPO-S1 for the peroxo form.

[136] a) Bühl, M.; Mauschick, F. T.; Schurhammer, R., in: *High Performance Computing in Science and Engineering, Munich 2002,* b) Wagner, S.; Hanke, W.; Bode, A.; Durst F. (Eds.), Springer Verlag, Berlin, 2003, pp.189-199.

[137] See e.g. Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126,* 4007

[138] For some recent examples see: a) Adams, C. J.; Fey, N.; Harrison, Z. A.; Sazanovich, I. V.; Towrie, M.; Weinstein. J. A. *Inorg. Chem.* **2008,** *47,* 8242–8257. b) Lage, M. L.; Fernandez, I.; Mancheno, M. J.; Sierra, M. A.; *Inorg. Chem.* **2008**, *47*, 5253–5258.

[139] Zhang, Y.; Gascon, J. A. *J. Inorg. Biochem.* **2008**, 102, 1684-1690

[140] Tromp, M. G. M.; Olafsson, G.; Krenn, B. E.; Wever, R. *Biochem. Biophys. Acta.* **1990**, *1040*, 192-198.

[141] For consistency with the QM/MM computations for the enzymes, an analogous setup has been employed for this system, namely MD equilibration of a hydrated cluster (vanadate and 18,000 water molecules), followed by BP86/AE1 / TIP3P optimization for six snapshots and averaging of the TD-B3LYP/AE1(+) data computed for these minima (including the MM charges - if these are omitted, the computed λ_{max} value increases by ca. 10 nm),