

Modelling flavin and substrate substituent effects on the activation barrier and rate of oxygen transfer by *p*-hydroxybenzoate hydroxylase

Lars Ridder^{a,*}, Bruce A. Palfey^b, Jacques Vervoort^a, Ivonne M.C.M. Rietjens^{a,c}

^aLaboratory of Biochemistry, Wageningen University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

^bDepartment of Biological Chemistry, The University of Michigan, Ann Arbor, MI 48109-0606, USA

^cDivision of Toxicology, Wageningen University, Tuinlaan 5, 6703 HE Wageningen, The Netherlands

Received 1 June 2000; accepted 26 June 2000

Edited by Thomas L. James

Abstract The simulation of enzymatic reactions, using computer models, is becoming a powerful tool in the most fundamental challenge in biochemistry: to relate the catalytic activity of enzymes to their structure. In the present study, various computed parameters were correlated with the natural logarithm of experimental rate constants for the hydroxylation of various substrate derivatives catalysed by wild-type *para*-hydroxybenzoate hydroxylase (PHBH) as well as for the hydroxylation of the native substrate (*p*-hydroxybenzoate) by PHBH reconstituted with a series of 8-substituted flavins. The following relative parameters have been calculated and tested: (a) energy barriers from combined quantum mechanical/molecular mechanical (QM/MM) (AM1/CHARMM) reaction pathway calculations, (b) gas-phase reaction enthalpies (AM1) and (c) differences between the HOMO and LUMO energies of the isolated substrate and cofactor molecules (AM1 and B3LYP/6-31+G(*d*)). The gas-phase approaches yielded good correlations, as long as similarly charged species are involved. The QM/MM approach resulted in a good correlation, even including differently charged species. This indicates that the QM/MM model accounts quite well for the solvation effects of the active site surroundings, which vary for differently charged species. The correlations obtained demonstrate quantitative structure activity relationships for an enzyme-catalysed reaction including, for the first time, substitutions on both substrate and cofactor. © 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Substituent effect; Solvation effect; Brønsted correlation

1. Introduction

p-Hydroxybenzoate hydroxylase (PHBH, EC 1.14.13.2) is a flavin dependent monooxygenase catalysing the 3-hydroxylation of 4-hydroxybenzoate. Such hydroxylations are important in the degradation pathway of a wide range of natural abundant compounds (e.g. lignin) as well as industrial aromatic pollutants. The first steps in the reaction cycle involve binding of the substrate, followed by a two electron reduction of the flavin cofactor by NADPH and the incorporation of molecular oxygen to form the C4a-peroxyflavin intermediate.

*Corresponding author. Fax: (31)-317-484801.
E-mail: lars.ridder@p450.bc.wau.nl

Abbreviations: QM/MM, combined quantum mechanics/molecular mechanics; PHBH, *para*-hydroxybenzoate hydroxylase

Upon protonation of the distal oxygen of this C4a-peroxyflavin, the C4a-hydroperoxyflavin is formed which performs an electrophilic attack on the substrate (Fig. 1). The substrate is activated for this electrophilic attack through deprotonation of its C4-hydroxyl moiety [1–3]. The initial products formed are a hydroxy-cyclohexadienone [1,3], which is converted to 3,4-dihydroxybenzoate through keto-enol tautomerism, and a deprotonated hydroxyflavin, which, after protonation and release of a water molecule, converts to oxidised flavin.

The attack of the C4a-hydroperoxyflavin on the substrate is an essential step in the reaction cycle. In previous studies, this step was proposed to be rate limiting on the basis of correlations of the natural logarithm of the rate constant for overall conversion of 4-hydroxybenzoate and four of its fluorinated analogues by PHBH [4] with the calculated (gas-phase) nucleophilic reactivity of these substrates [2], as well as with the calculated activation energies for hydroxylation [5].

In a recent study, the kinetics of PHBH with a series of modified flavin cofactors have been investigated and the results were shown to correlate with Hammett substituent parameters [6]. A drawback of the Hammett constants is that they cannot predict the combined effect of substitutions on both substrate or cofactor. The present study investigates whether several computational approaches to quantify effects of substitutions, used in previous studies on substrate analogues [2,5,7], are able to account for variations in both substrate and cofactor within one theoretical model. A combined quantum mechanical/molecular mechanical (QM/MM) approach [3,5,8], including the electrostatic and steric influence of the protein environment on the activation barrier, is compared to gas-phase approaches in which the isolated reacting species are investigated. The results of the various approaches are compared to experimental results, including experiments with substituted substrates and the above mentioned experiments with substituted flavins [6].

2. Materials and methods

2.1. Stopped-flow experiments

2-Fluoro-4-hydroxybenzoate, 3-fluoro-4-hydroxybenzoate, 2,5-difluoro-4-hydroxybenzoate and 3,5-difluoro-4-hydroxybenzoate were synthesised as described previously [9]. 2,3,5,6-Tetrafluoro-4-hydroxybenzoate was purchased from Aldrich. Rate constants for the hydroxylation of the fluorinated substrates by PHBH were obtained by mixing the anaerobic dithionite-reduced enzyme-substrate complex with buffer containing O₂ in a stopped-flow apparatus [10]. Reactions were in 50 mM potassium phosphate, pH 6.5, 1 mM EDTA, at 4°C. Flavin absorbance or fluorescence changes were analysed as described previously, in order to obtain the rate constant for hydroxylation [10].

2.2. Calculation of activation barriers and reaction energies within the protein environment

Approximate activation barriers and reaction energies for the hydroxylation step with a series of substituted substrates and flavin cofactors (Table 1) were calculated essentially as described previously [3]. Reaction pathway calculations were performed on the basis of the crystal structure of PHBH in complex with substrate, using a QM/MM potential implemented in CHARMM [8]. The dianionic substrate and the C4a-hydroperoxyflavin cofactor were treated with semi-empirical AM1 molecular orbital theory [11,12], whereas the steric and electronic influence of the surrounding enzyme was included by treating all other atoms molecular mechanically [13]. The AM1 parameters for sulphur, provided with CHARMM 24b1, were replaced by the more recently optimised values [14]. All 330 crystal waters were included in the model, using the HBUILD routine of CHARMM to add hydrogen atoms. The 8-hydroxy-FAD and the 8-mercapto-FAD have been indicated to be deprotonated [6] and were treated as such in the QM/MM model.

A reaction coordinate r for the electrophilic attack of the C4a-hydroperoxyflavin on the substrate (Fig. 1) was defined as the difference between Op–Od interatomic distance and the Od–C3 interatomic distance, i.e. $r = d(\text{Op–Od}) - d(\text{Od–C3})$. To simulate the reaction pathway, this reaction coordinate parameter was changed in small steps (0.1 Å) leading to the splitting of the Op–Od bond and the formation of the Od–C3 bond. All atoms within 6 Å distance of any QM atom were optimised for the reactant, product and all intermediate states of the reaction pathway calculated. This definition of the optimised region around the active site is slightly different from the definition used previously [3]. This modification in the protocol was needed to ensure enough flexibility in the model to properly accommodate the different substituents at the 8-position of the flavin cofactor.

2.3. Gas-phase calculations

AM1 calculations were performed on the various reactant and product molecules separately, in gas-phase. The results were compared to the QM/MM reaction pathway calculations to obtain insight into the contributions of the chemical properties of the reactants themselves to the variation in the reaction and activation energies with the different substrates and cofactors within the protein. In order to test the ability of the semi-empirical AM1 method to account for the effect of substituents on the reactivity of the substrate and cofactors, the variations in the calculated AM1 energies of the highest occupied molecular orbital (HOMO) of the various *p*-hydroxybenzoates and the lowest unoccupied molecular orbital (LUMO) of the C4a-hydroperoxyflavin molecules were compared to B3LYP/6-31+G(d)//HF/6-31G(d) results.

In the case of ab initio methods, calculations on anions require basis sets including diffuse functions. Semi-empirical methods, such as AM1, use minimal basis sets which do not include diffuse functions. Nevertheless, AM1 often performs well on anionic systems and

appears to be able to correct for the increased electron repulsion through the parameterisation [11]. The competence of AM1 for even a dianionic system is illustrated by a recent comparison of AM1 with ab initio methods for dianionic oxaloacetate, for which the AM1 results were similar to those from ab initio calculations including diffuse functions and performed even better than ab initio methods without diffuse functions [15].

3. Results

3.1. Comparison of calculated QM/MM energy changes with experimental rate constants

Continuous reaction pathways were obtained for the various substrates and cofactors listed in Table 1. The reaction pathways are comparable with respect to the geometrical changes and the energy profiles [3]. The energy barriers, i.e. the energy differences between the initial energy minimum and the energy maximum of the different profiles, vary between 17.5 and 22.2 kcal/mol (Table 1).

Comparison of these calculated QM/MM energy barriers with the natural logarithm of the experimental rate constants for the hydroxylation step shows a good linear correlation with a coefficient of 0.90 (Fig. 2). In an earlier study, a correlation was found between the calculated QM/MM barrier for hydroxylation and the logarithm of the overall k_{cat} for conversion of the substituted substrates only [5]. The present results show a similar correlation with the rate constants of the individual hydroxylation step. More importantly, the correlation is now extended to include substitutions on the cofactor as well.

3.2. Brønsted relationship between calculated energy barriers and the calculated energy change of the reaction

The approximate transition states were found at slightly different values of the reaction coordinate r , as listed in Table 1. Generally the transition state is found at higher (i.e. less negative) values of r as the energy barrier increases and the reaction energy becomes less negative. This suggests that an increase in activation energy corresponds to a more product-like transition state, in line with the Brønsted [16] and Marcus [17] theories of energy profiles. An excellent linear Brønsted correlation is observed between the calculated QM/MM en-

Table 1

Energy differences between reactant and product complexes $\Delta E_{\text{reaction}}$, and between reactant and transition state complexes ΔE_{act} , of the calculated QM/MM reaction pathways for the various substrates and 8-substituted cofactors

	Substrate	8-Substituted flavin	$\Delta E_{\text{reaction}}$ (kcal/mol)	ΔE_{act} (kcal/mol)	r_{TS} (Å)	$\Delta E_{\text{gas-phase}}$ (kcal/mol)	k_{hydrox} (s^{-1})
1	PHB	-CH ₃	-70.10	17.67	-0.50	-139.97	48
2	2-F-PHB	-CH ₃	-66.88	18.05	-0.50	-134.14	39
3	3-F-PHB	-CH ₃	-65.68	18.57	-0.49	-132.88	39
4	2,5-F ₂ -PHB	-CH ₃	-62.43	18.95	-0.48	-127.17	17.3
5	3,5-F ₂ -PHB	-CH ₃	-61.30	20.71	-0.48	-127.28	2.4
6	2,3,5,6-F ₄ -PHB	-CH ₃	-55.10	22.15	-0.45	-119.48	1.4
7	PHB	-H	-67.43	17.72	-0.50	-140.24	40
8	PHB	-NH ₂	-70.19	17.69	-0.50	-139.24	156
9	PHB	-N(CH ₃) ₂	-70.73	17.84	-0.51	-140.66	34
10	PHB	-SCH ₃	-70.41	18.04	-0.50	-141.24	151
11	PHB	-Cl	-69.18	17.50	-0.49	-143.31	300
12	PHB	-O ⁻	-57.86	21.24	-0.46	-71.619	7.6
13	PHB	-S ⁻	-63.35	19.60	-0.48	-81.050	14.3

Columns 6–8 present the reaction coordinate values corresponding to the approximate transition states r_{TS} , the gas-phase energy differences between the separate reactant and product molecules and the experimental rate constants for the hydroxylation step (k_{hydrox}). Rate constants for 7–13 were taken from [6].

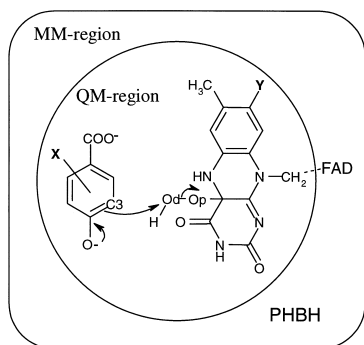


Fig. 1. Schematic representation of the QM/MM model used to simulate the electrophilic attack of C4a-hydroperoxyflavin on dianionic *p*-hydroxybenzoate, for the various substrates and cofactors listed in Table 1.

ergy barriers and the calculated QM/MM energy change from reactants to products for all substrate and cofactor combinations studied (Fig. 3).

3.3. The effect of the active site surroundings on the energy of the reaction

In order to estimate the effect of the protein environment on the energetics of the reaction, AM1 calculations were performed to optimise the separate substrate and cofactor molecules in the gas-phase. Fig. 4 shows that the calculated QM/MM energy differences between the reactant and product states within the active site of PHBH correlate strongly ($r=0.96$) with the energy differences between the separate reactant and product molecules in gas-phase, except for the deprotonated 8-hydroxy-FAD and the 8-mercapto-FAD. With the deprotonated cofactors, the energy differences between the separate reactant and product molecules in gas-phase are less negative than the gas-phase energy differences obtained for the other substrate-cofactor combinations. This may well be explained by the fact that the C4a-hydroxyflavin product molecule is one atomic charge unit more negative than the C4a-hydroperoxyflavin reactant molecule. In the case of the deprotonated 8-hydroxy-FAD and the 8-mercapto-FAD cofactors, the total charge changes from $-1e$ to $-2e$

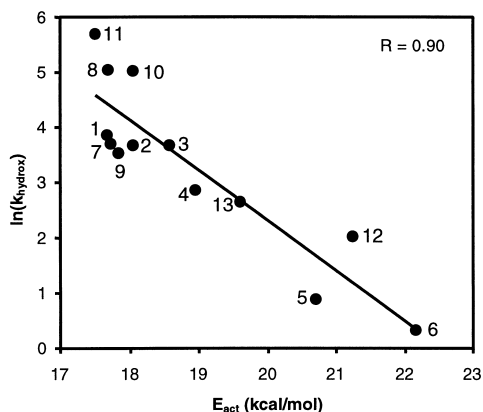


Fig. 2. Linear correlation between the logarithm of the experimental rate constants for the hydroxylation step (k_{hydrox}) and the energy barriers obtained from the QM/MM reaction pathway calculations with the various substrate and cofactors. Labels according to Table 1.

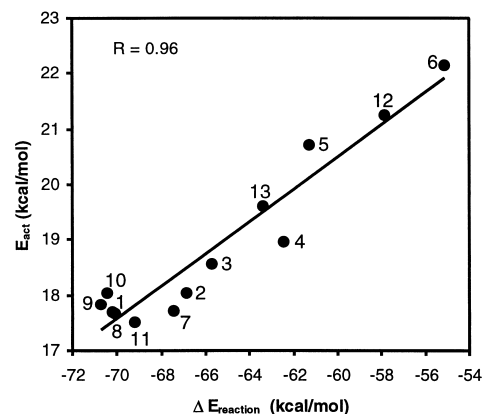


Fig. 3. Linear Bronsted correlation between the energy barriers and the reaction energies of the QM/MM reaction pathway calculations. Labels according to Table 1.

upon the electrophilic reaction with the substrate, instead of changing from 0 to $-1e$ as with the other cofactors. The dianionic 8-hydroxy-C4a-hydroxyflavin and the 8-mercapto-C4a-hydroxyflavin product molecules are relatively unstable due to electronic repulsion.

However, this strong effect present in gas-phase does not appear to be representative for the situation within the active site of PHBH. Fig. 5 shows that the calculated gas-phase energy differences for the reactions with the 8-hydroxy-flavin and the 8-mercapto-flavin cofactors deviate from the Bronsted correlation with the experimental rate constants. Within the actual protein surroundings, the strong effect of the negatively charged substituents on the reaction energy may in fact be reduced by solvation effects, including specific hydrogen bond interactions, polarisation and dielectric screening. The reaction pathway calculations show that this effect of the protein surrounding is accounted for to a quite reasonable extent in the QM/MM calculations. This can be concluded from the fact that the barriers calculated with 8-hydroxy-FAD and the 8-mercapto-FAD fit well in the correlations shown in Figs. 2 and 3. In the QM/MM models, based on the crystal structure, the 8-substituents are surrounded by a couple of water mole-

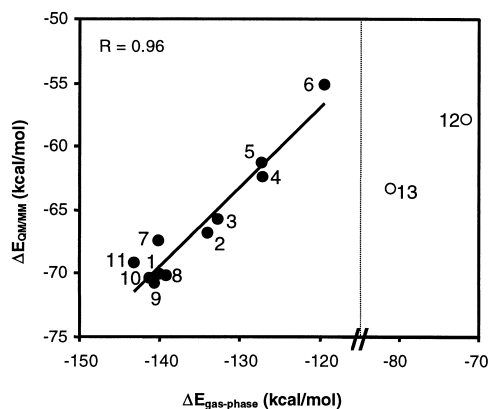


Fig. 4. Linear correlation between the reaction energies of the QM/MM reaction pathways and the AM1 energy difference between the separate reactants and products in gas-phase. Labels according to Table 1. The open circles represent the differently charged hydroxyflavin and mercaptoflavin cofactors which are not included in the linear regression.

cules. In case of 8-hydroxy-FAD and the 8-mercapto-FAD, these water molecules clearly interact with (and stabilise) the negatively charged 8-O⁻ and the 8-S⁻ substituents. This could be one of the features of the active site environment reducing the unfavourable influence of the negative substituents on the reaction energy.

Nevertheless, the correlation in Fig. 4 indicates that, for the similarly charged substrates and cofactors, the variation in the QM/MM calculated energy change of the reaction (and thus the variation in the activation energies (Fig. 3) and the rate constants for hydroxylation (Fig. 2)) is dominated by the chemical properties of the substrates and cofactors themselves and not by specific interactions with the PHBH environment. Thus, the calculated energy difference between separate reactant and products molecules in gas-phase is a good parameter for the enzyme activity with different substrates and cofactors (Fig. 5) as long as they have the same molecular charge.

3.4. Prediction of activation energies and rate constants on the basis of the difference between HOMO and LUMO energies in the reactant complex

In a previous study, it was argued that the electrophilic attack of the C4a-hydroperoxyflavin on the substrate could be described in terms of frontier orbital interactions [3]. This implies that the variation in activation energies should relate to the differences in the HOMO and LUMO energies of the substrate and cofactor, respectively, in accordance with frontier orbital theory [18]. A high HOMO energy indicates a strong nucleophilic reactivity, whereas a low LUMO energy indicates a strong electrophilic reactivity. Fig. 6a,b shows a correlation of the natural logarithm of the rate constants for hydroxylation with the energy difference between the HOMO of the substrate molecules and the LUMO of the C4a-hydroperoxyflavin molecules in gas-phase calculated with the AM1 and B3LYP/6-31+G(d)//HF/6-31G(d) methods. The correlations show that the gas-phase HOMO and LUMO energies are indeed useful parameters to quantify the nucleophilic and electrophilic reactivity of substrate and cofactor, respectively, and to predict variation in activation barriers, as long as similarly charged species are involved. The LUMO energies for the 8-hydroxy and the 8-mercapto cofactor molecules are sig-

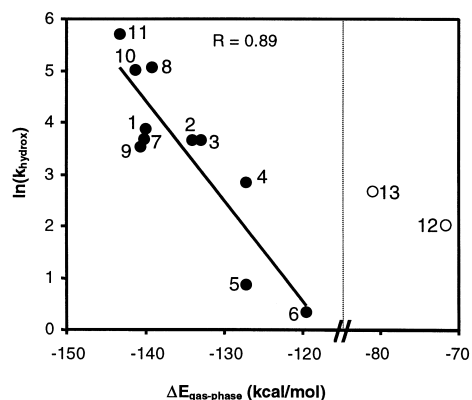


Fig. 5. Linear correlation between experimental rate constants for the hydroxylation step (k_{hydrox}) and the AM1 energy difference between the separate reactants and products in gas-phase. Labels according to Table 1. The open circles represent the differently charged hydroxyflavin and mercaptoflavin cofactors which are not included in the linear regression.

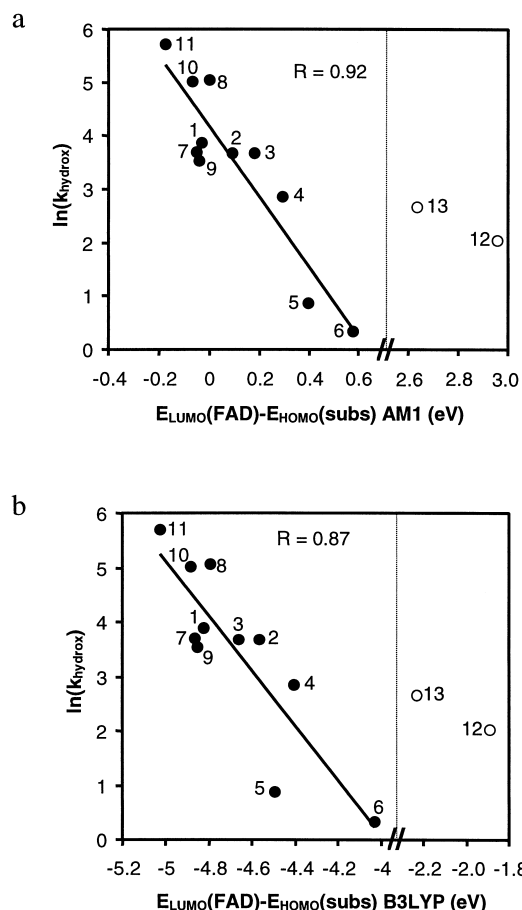


Fig. 6. Linear correlation between the experimental rate constants for the hydroxylation step (k_{hydrox}) and the difference between the HOMO energy of the isolated substrate and the LUMO energy of the isolated cofactor calculated in gas-phase using (a) AM1 and (b) B3LYP/6-31+G(d)//HF/6-31G(d). Labels according to Table 1. The open circles represent the differently charged hydroxyflavin and mercaptoflavin cofactors which are not included in the linear regression.

nificantly higher than the LUMO energies for the other cofactors, indicating that the negatively charged substituents on the flavin ring in these cases make the C4a-hydroperoxyflavin less reactive as electrophile. In analogy to the effects described above, the influence of the negative O⁻ and S⁻ substituents seems to be smaller within the active site compared to the situation in vacuum, probably due to differential solvation effects for differently charged species.

The similarity between the correlations obtained with the AM1 orbital energies and with the more sophisticated B3LYP calculations indicates that AM1 accounts quite accurately for the effect of the substituents on the nucleophilic and electrophilic reactivities of the *p*-hydroxybenzoate substrate and the C4a-hydroperoxyflavin cofactor, respectively.

4. Discussion

Computer simulations of enzymatic reactions are becoming increasingly important in understanding the catalytic activity of enzymes on the basis of their structure. Ultimately, it should be possible to predict a rate constant of a given enzymatic reaction on the basis of a three-dimensional structure of the enzyme and substrate [19]. The present study compares

gas-phase models and QM/MM models of an enzymatic reaction, with respect to their ability to explain the variation in the rate of a specific enzyme-catalysed reaction step, with different cofactors and substrates. For the model enzyme of the present study, PHBH, the structure activity relationships described so far either involve only structural variations in the substrate, e.g. a series of fluorinated *p*-hydroxybenzoate homologues [2,4,5], or include only structural variations in the flavin cofactor [6]. To our knowledge, the present study demonstrates for the first time a quantitative structure activity relationship for an enzyme-catalysed reaction including substitutions on both substrate and cofactor.

It is demonstrated that the reactivities of cofactor and substrate can be predicted on the basis of several parameters. First, the HOMO and LUMO energies calculated for the substrates and cofactors, respectively, both at the AM1 and B3LYP/6-31+G(*d*) levels, are shown to be useful parameters for their reactivity (Fig. 6a,b). Second, the calculated energy differences between the separate reactant and product molecules, using AM1, appear to be a useful parameter as well (Fig. 5). However, these two parameters are valid only for similarly charged species, as illustrated by deviations from the correlations (Figs. 4–6) for the negatively charged 8-hydroxy-flavin and the 8-mercapto-flavin cofactors. The effect of replacement of a neutral group by a charged substituent on molecular properties in a solvent or protein environment often cannot be correctly predicted by parameters derived from gas-phase calculations, due to differential solvation effects on differently charged compounds [20].

A powerful feature of the QM/MM method is that it includes to a certain extent the solvation effects on the reacting species within the active site of the enzyme. It is promising to observe that, without further calibration of the standard parameters involving the interaction of the MM solvent molecules with negatively charged oxygen and sulphur atoms in the QM system, this differential solvation effect is modelled reasonably accurate in the present QM/MM model.

Acknowledgements: This study was especially dedicated to the event of awarding the title 'scientist of the year 1998' at Ann Arbor Michigan University to Prof. Dr Vincent Massey. We thank Dr Adrian J. Mulholland from the University of Bristol for his support. This work was funded by the Netherlands Organisation for Scientific Research (NWO).

References

- [1] Entsch, B., Ballou, D.P. and Massey, V. (1976) *J. Biol. Chem.* 251, 2550–2563.
- [2] Vervoort, J., Rietjens, I.M.C.M., Van Berkel, W.J.H. and Veeger, C. (1992) *Eur. J. Biochem.* 206, 479–484.
- [3] Ridder, L., Mulholland, A.J., Rietjens, I.M.C.M. and Vervoort, J. (1999) *J. Mol. Graph. Mod.* 17, 163–175.
- [4] Husain, M., Entsch, B., Ballou, D.P., Massey, V. and Chapman, J.P. (1980) *J. Biol. Chem.* 255, 4189–4197.
- [5] Ridder, L., Mulholland, A.J., Vervoort, J. and Rietjens, I.M.C.M. (1998) *J. Am. Chem. Soc.* 120, 7641–7642.
- [6] Ortiz-Maldonado, M., Ballou, D.P. and Massey, V. (1999) *Biochemistry* 38, 8124–8137.
- [7] Vervoort, J. and Rietjens, I.M.C.M. (1996) *Biochem. Soc. Trans.* 24, 127–130.
- [8] Field, M.J., Bash, P.A. and Karplus, M. (1990) *J. Comp. Chem.* 11, 700–733.
- [9] Eppink, M.H.M., Boeren, S.A., Vervoort, J. and Van Berkel, W.J.H. (1997) *J. Bacteriol.* 179, 6680–6687.
- [10] Entsch, B., Palfey, B.A., Ballou, D.P. and Massey, V. (1991) *J. Biol. Chem.* 266, 17341–17349.
- [11] Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. and Stewart, J.J.P. (1985) *J. Am. Chem. Soc.* 107, 3902–3909.
- [12] Dewar, M.J.S. and Zoebisch, E.G. (1988) *THEOCHEM* 180, 1–21.
- [13] Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comp. Chem.* 4, 187–217.
- [14] Dewar, M.J.S. and Yuan, Y. (1990) *Inorg. Chem.* 29, 3881–3890.
- [15] Mulholland, A.J. and Richards, W.G. (1998) *J. Mol. Struct. (Theochem)* 429, 13–21.
- [16] Bronsted, J.N. (1928) *Chem. Rev.* 5, 23.
- [17] Marcus, R.A. (1968) *J. Phys. Chem.* 72, 891–899.
- [18] Fleming, I. (1976) John Wiley and Sons, New York.
- [19] Warshel, A. (1992) *Curr. Opin. Struct. Biol.* 2, 230–236.
- [20] Ridder, L., Zuilhof, H., Vervoort, J. and Rietjens, I.M.C.M. (1999) in: *Flavoprotein Protocols* (Chapman, S.K. and Reid, G.A., Eds.), *Methods Mol. Biol.*, pp. 207–227, Vol. 131, Humana Press Inc., NJ.