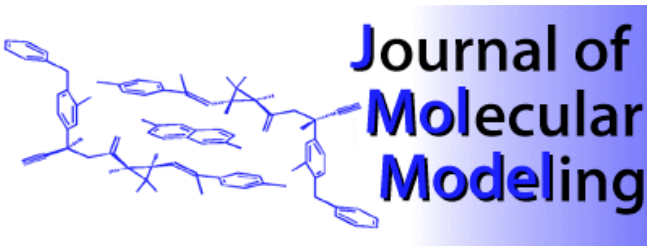
J. Mol. Model. **1998**, *4*, 395 – 404



FULL PAPER

Molecular Modelling Studies on the Catalytic Mechanism of *Candida Rugosa Lipase*

Peter Monecke, Rudolf Friedemann, Stefan Naumann, and René Csuk

Institute of Organic Chemistry, Martin-Luther-University Halle-Wittenberg, Kurt-Mothes-Strasse 2, D-06120 Halle (Saale), Germany. Tel: +49-345 5525668; Fax: +49-345 5527608. E-mail: friedemann@chemie.uni-halle.de

Received: 25 September 1998 / Accepted: 12 November 1998 / Published: 11 December 1998

Abstract Quantum chemical and molecular dynamics investigations have been performed on model systems for *Candida rugosa* lipase (CRL) to study mechanistic and conformational features of the catalytic hydrolysis. Based on X-ray data, a simplified model of the CRL substrate complex was created for the PM3 and *ab initio* calculations, including the amino acid residues both of the catalytic triad and the oxyanion hole.

The energetic and structural properties of significant species along the pathway of the hydrolysis of the model substrate acetic acid methyl ester have been calculated. By modifications of the residues of the oxyanion hole as well as the catalytic triad, the influence of these parts of the active site on the pathway of the reaction was analysed in more detail.

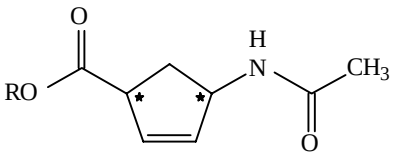
Moreover, molecular dynamics simulations have been carried out on CRL adducts with (±)-cis-4-acetamido-cyclopent-2-ene-1-carboxylic esters with different lengths of their alkyl chain and their absolute configuration as substrates. For the MD simulations using the AMBER program, all amino acid residues and water molecules with a cut-off radius less than 1500 pm from the substrate were taken into account. From the analysis of the trajectories and histograms for significant hydrogen bonds in the active site of the enzyme adducts, some hints were obtained for the enantiodifferentiation and the chain dependence of the esters in catalytic hydrolysis by CRL.

Keywords Computer simulations, Catalytic mechanism, Ester hydrolysis, Protein modelling, Molecular orbital calculations

Introduction

In the last few years lipases have acquired great importance in asymmetric synthesis, especially of enantiomerically pure building blocks for the creation of pharmaceuticals [1-3].

correspondence to: R. Friedemann



R =		
	Me	**1a**
	Et	**1b**
	1-Pr	**1c**
	1-Bu	**1d**
	1-Hex	**1e**
	1-Oct	**1f**

Figure 1 *Acronyms of the considered cis-4-acetamido-cyclopent-2-ene carboxylic esters*

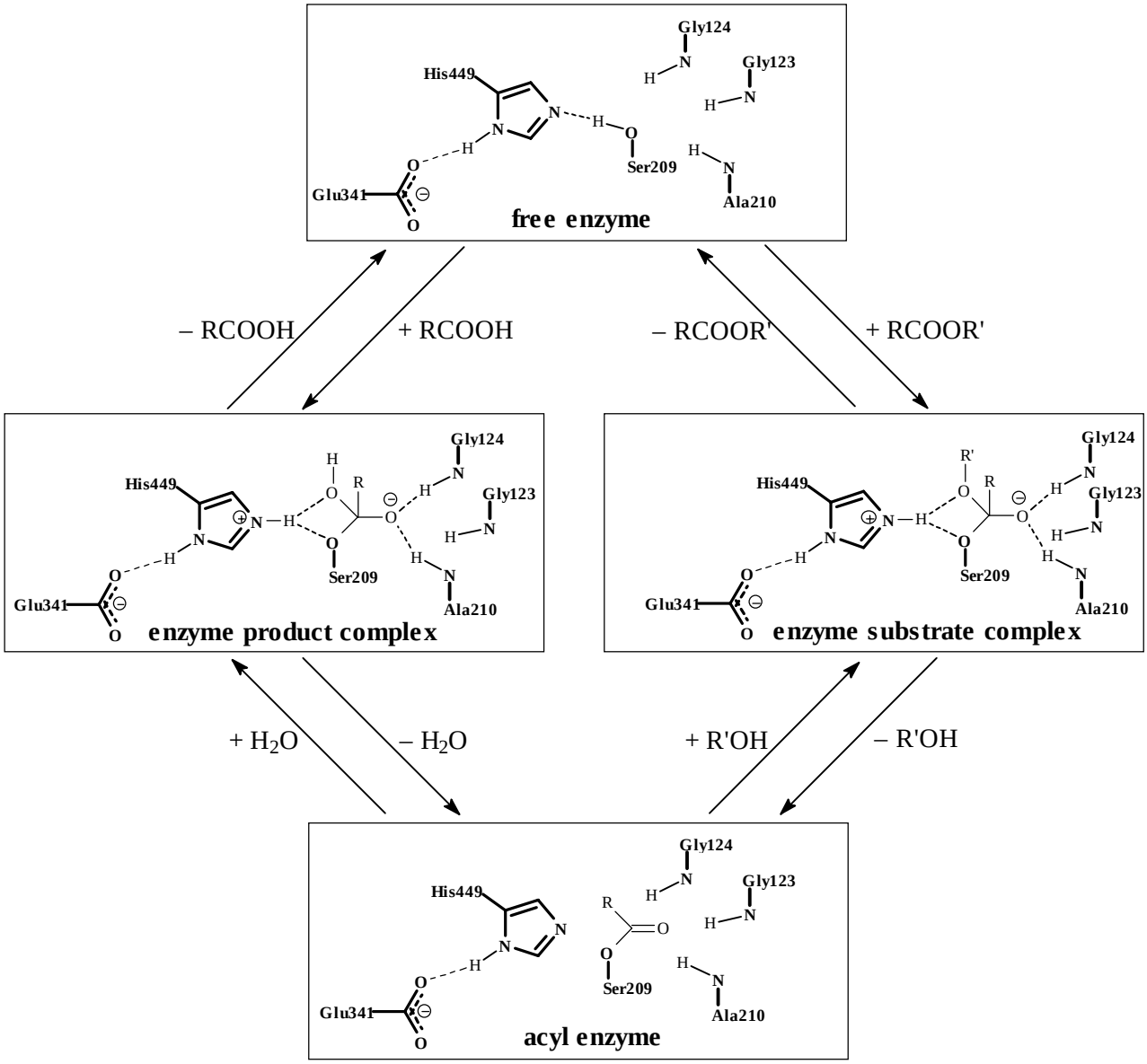


Figure 2 *Suggested catalytic cycle of ester hydrolysis by CRL*

The lipase from *Candida rugosa* (CRL) has been found to be particularly well suited [4, 5]. CRL shows a high enantioselectivity not only to the native substrate but also to other carboxylic acids and esters. It was found recently that the rate of hydrolysis of (±)-cis-4-acetamido-cyclopent-2-ene-1-carboxylic esters (1a-f) in Fig. 1 by CRL mainly depends on their absolute configuration and on the chain length of the alcohol component [6]. These esters serve as precursors for the synthesis of aristeromycin, the carbocyclic analogue of adenosine. Aristeromycin is of interest due to its cytostatic and antiviral activities.

To understand why CRL shows a high enantioselectivity over a broad range of substrates, some insights into the mechanism and the structural basis for the chiral preferences are necessary. Therefore, we have performed PM3 and *ab initio* HF-SCF calculations on the reaction pathway using simplified model systems for the CRL substrate species. The proposed scheme [4] of the catalytic cycle is shown in Fig. 2 for the hydrolysis of esters by CRL considering only the active site. The starting structures for the calculation of the simplified CRL substrate adducts were adapted from X-ray data of covalent complexes of CRL with the inhibitor O-menthyl hexylphosphonochloridate [4,7,8,9]. The simplified model includes the essential groups of the active site of CRL, i.e. the side chains of the catalytic triad (Ser 209, Glu 341, His 449) and the oxyanion hole, which consists of the backbone of the amino acid residues Gly 123, Glu 124 and Ala 210, as well as the model substrate acetic acid methyl ester. The

Table 1 *Simplifications of the active site in CRL substrate complex for the model system A*

Enzyme part	Model system A
Gly 122-Gly 123-Gly 124	2-acetylamino-*N*-methyl-acetamide
Ser 209-Ala 210	3-hydroxy-*N*-methyl-propionamide
Glu 341	acetate
His 449	5-methyl-1*H*-imidazole
menthyl hexylphosphonate	acetic acid methyl ester

simplifications of the enzyme and substrate parts in the model system A are illustrated in Table 1 and Fig. 3. A complete PM3 pathway for the hydrolysis was calculated for the model system A. The qualitative correctness of the pathway was verified by *ab initio* HF-SCF calculations with a 6-31G basis of significant points on the energy profile. The contribution of the relevant amino acid residues of the active site of CRL was investigated by calculations of alternative model systems.

MD simulations using the AMBER 4.1 program were performed on the diastereomeric CRL substrate adducts with (±)-cis-4-acetamido-cyclopent-2-ene-1-carboxylic acid hexyl ester ((±)-1e). All amino acid residues and water molecules located in a cut-off radius of 1500 pm from the substrate were considered in the model CRL-(±)1e complexes. Especially, the trajectories and histograms of significant hydrogen bonds in the active site of the diastereomeric CRL-(±)-1e model complexes were analysed in order to obtain hints about the enantiodifferentiation in the enzymatic hydrolysis. Finally, MD simulations on CRL model complexes with the esters (±)-1a-e (Fig. 1) as substrates were performed. The aim of these comparative studies was to explain the chain dependence of the substrates on the rate of hydrolysis found by kinetic measurements [6].

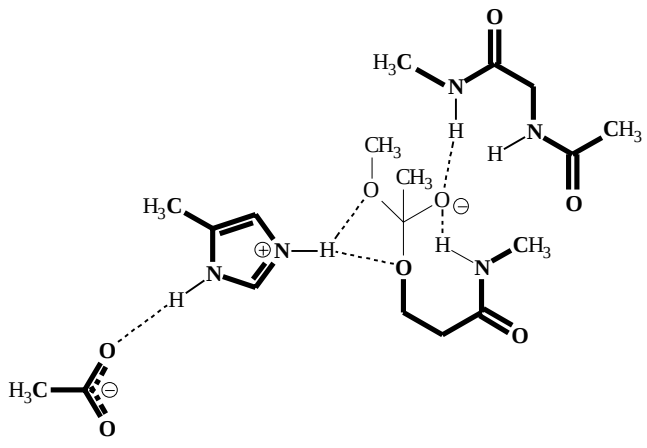


Figure 3 *Simplified active site of the model system A (bold printed systems were not optimized)*

Methods

The semiempirical PM3 method [10] within the MOPAC 6.0 software package [11] was used for the calculations of the complete hydrolysis pathway on the simplified CRL substrate systems. The procedure used in our calculations for a careful searching of the transition state structure will be described in the next section of this paper. The comparative *ab initio* HF-SCF 6-31G calculations of significant structures on the reaction profile were performed using the program TURBOMOLE [12] implemented in the MSI software package InsightII 4.0.0 [13].

For the MD simulations, we used the AMBER 4.1 software package [14]. Atomic net charges for the substrate molecules were adapted from the fit to reproduce the electrostatic potential within the PM3 method. This procedure is frequently used to evaluate atomic charges in force field methods [5,15,16] for comprehending the electrostatic interactions and is consistent with the defined atomic charges for the used amino acid fragments in the AMBER program [17,18]. Classical MD studies with the nVp ensemble were performed at a temperature of 300 K and with a simulation time of 300 ps using our experiences on simulations of biochemical and mesogenic systems [19,20]. In the first period (60 ps) the temperature was raised stepwise from 0 to 300 K. A time step of 0.5 fs was used in all simulations. Interactions between atoms at distances longer than 1000 pm were not taken into account to reduce the simulation time. The starting structures of the model enzyme complexes for the MD simulations are based on X-ray data of the CRL inhibitor complex with O-(1R, 2S, 5R)-menthyl hexylphosphonate group (PDB-file 1LPM) [4,9].

The calculations are carried out on IBM RISC 6000 (PM3), Indigo^2 (TURBOMOLE), and HP SPP 2000 (AMBER) workstations.

Results and discussion

Calculations on the reaction mechanism of hydrolysis

For the PM3 calculation on the catalytic mechanism of the hydrolysis of esters by CRL, the model system A defined in Fig. 3 and Tab. 1 was considered. Some constraints were introduced in the optimization process to keep the amino acid

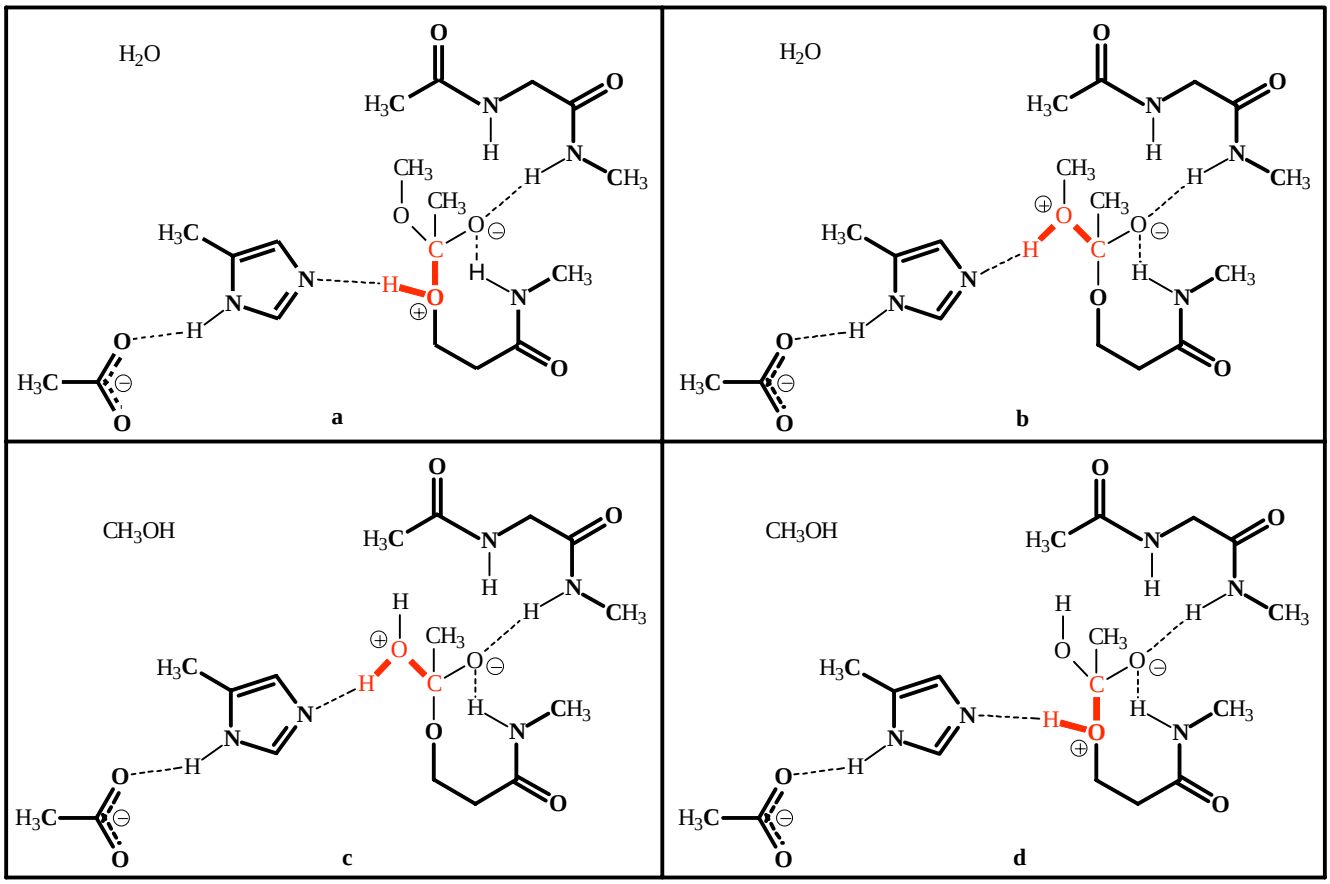


Figure 4 *Partial reactions of the enzymatic hydrolysis for the model system A (**a:** Attack of the ester, **b**: Cleavage of the alcohol, **c**: Attack of the water molecule, **d**: Cleavage of the acid)*

residues of the model system in their enzyme relevant positions. The heavy atoms of the model enzyme maintained their original positions. All hydrogen atoms and the atoms of the model substrate acetic acid methyl ester were fully optimized. The hydrolysis process was divided in four partial reactions illustrated in Fig. 4. In each of the partial reactions (a-d), two bond lengths (red marked in Fig. 4) were varied and the corresponding energy hypersurface was calculated. The direct localization of the transition states was not possible within the MOPAC 6.0 transition state optimizer because of the fixed constraints. The positions of the transition states were carefully determined as highest points on the two-dimensional reaction paths using the programs Gridview and Re_view [21] by a successive decrease of the step size up to 1 pm.

The results of the partial reactions are summarised and illustrated as a complete reaction profile for the enzymatic hydrolysis in Fig. 5. The reaction profile contains nine significant points, five minima and four transition states. The obtained energies of the transition states are very high for the catalytic reaction. Obviously, this is due to the fact that a constraint optimization was used in the calculation on the model system. The introduced constraints on the heavy atoms of the model enzyme were necessary to keep them in their enzyme relevant positions. Moreover, the enzyme environment was only considered in a limited way. Therefore, the results should be seen as a qualitative estimation of the discussed reaction mechanism. The appropriate structures are shown in Fig. 6 together with the specification of relevant

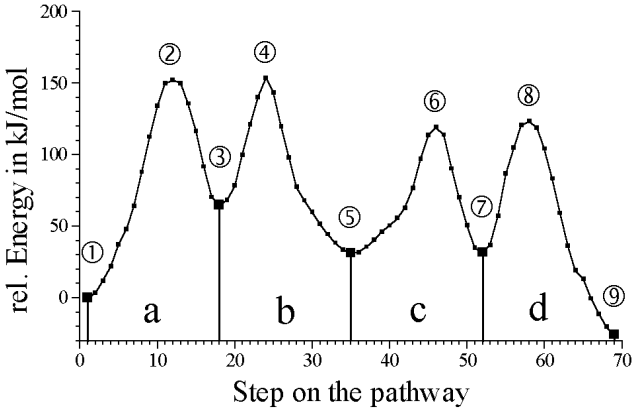


Figure 5 *Complete reaction profile of the enzymatic hydrolysis of acetic acid methyl ester (model system A, PM3)*

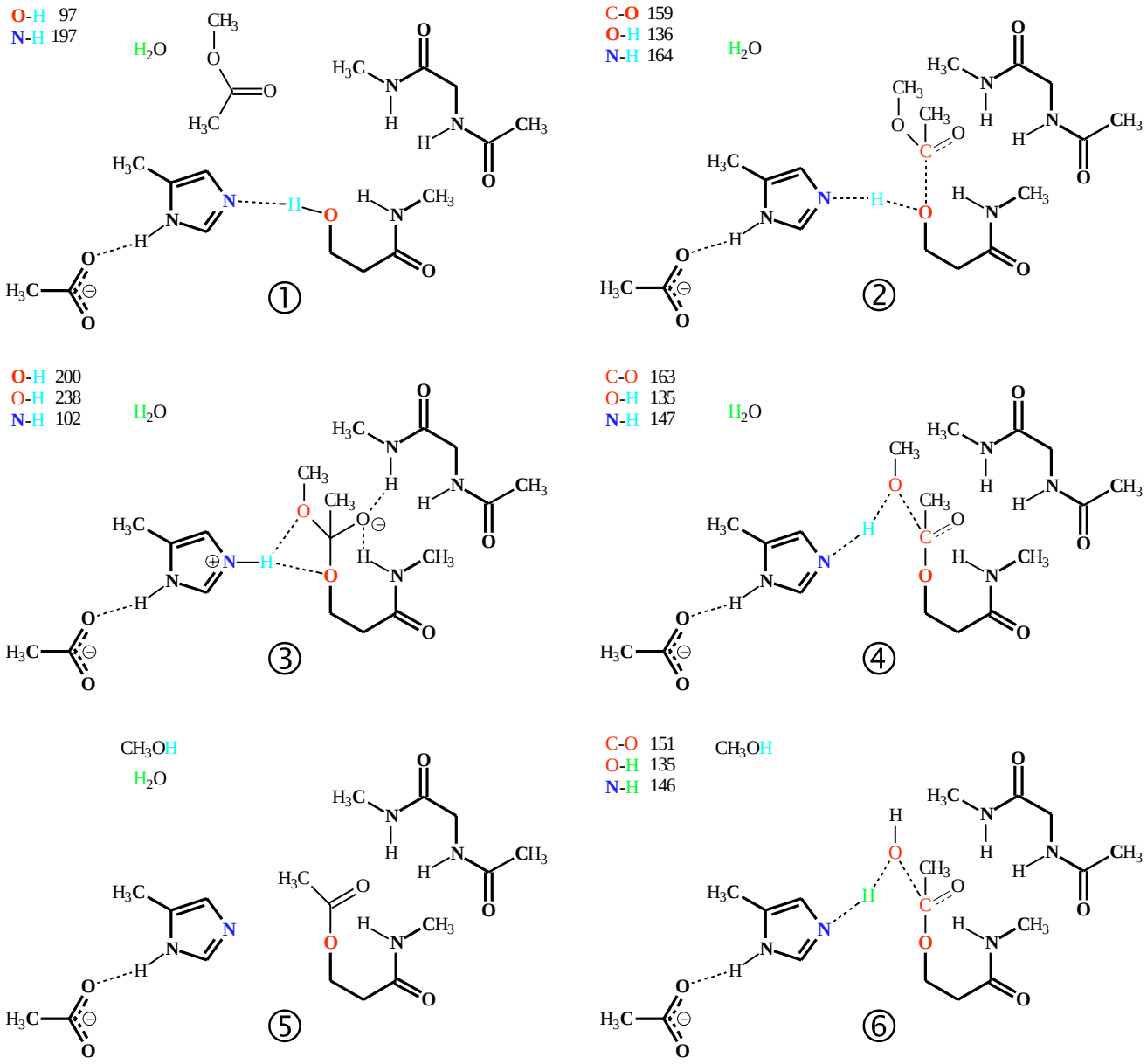


Figure 6 *PM3 structures of minima and transition states (TS) on the pathway of the enzymatic hydrolysis (significant bond lengths in pm)*

bond lengths. They illustrate the function of the amino acid residues of the catalytic triad and the oxyanion hole in CRL. The amino acid residues of the oxyanion hole stabilise the negative charge of the oxyanion and the acetate (Glu 341 in the enzyme) causes a delocalisation of the formal positive charge formed at the N3 atom of the 5-methyl-imidazole (His 449 in the enzyme) in the tetrahedral intermediates of the enzyme substrate and enzyme product complexes. The nearly symmetric shape of the reaction profile supports the findings that CRL catalyses both saponification and esterification. The acylation (a, b) and the deacylation (c, d) in Fig. 5 show remarkable structural and energetic similarities.

Further model calculations with modified active sites of CRL were performed to study the function of the oxyanion hole and Glu 341 on the reaction profile of the enzymatic hydrolysis in more detail. In the model system B, the amino acid residues of the oxyanion hole were not included, whereas in the model system C the acetate was not considered. The nine significant points of the reaction profile for the two model systems are shown together with the complete pathway of the system A in Fig. 7. A comparison of the relative energies of the significant structures indicates that larger deviations occur especially in the enzyme substrate and enzyme product complexes (steps 18 and 52). Obviously, these charged

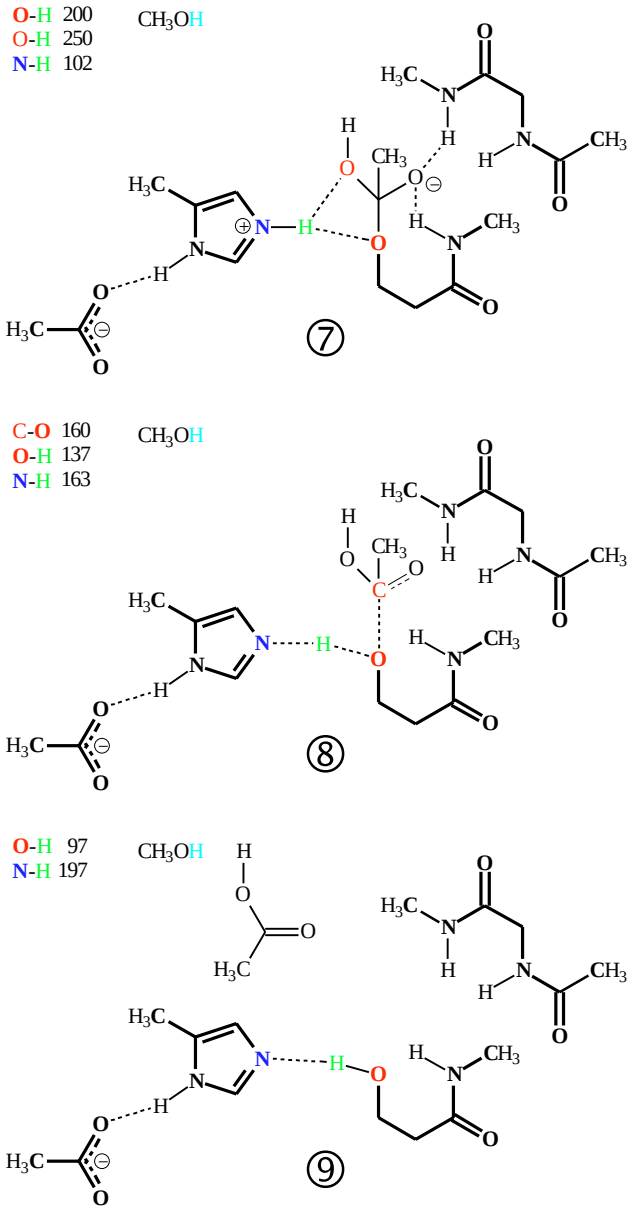


Figure 6 (continued) *PM3 structures of minima and transition states (TS) on the pathway of the enzymatic hydrolysis (significant bond lengths in pm)*

tetrahedral intermediates are stabilised in a special way by the enzyme environment. In the transition states (steps 12, 24, 46, and 58) the effect is lower. The smallest influence was found in the free enzyme structures (steps 1 and 69) as well as the acyl enzyme (step 35) in Fig. 7. It is remarkable that in all structures along the pathway a hydrogen atom remains on the N1 atom of the 5-methyl-imidazole. This is in agreement with neutron diffraction studies on a trypsin monoisopropylphosphonyl adduct [22]. Thus, the calculations support the charge transfer and proton relay function of His 449 in the catalytic triad of CRL.

First *ab initio* HF-SCF 6-31G calculations were performed for the nine significant points using the PM3 structures to verify the PM3 results on the reaction profile for the model system A in a qualitative way. The results are illustrated in Fig. 8 . They indicate that the two methods show a qualitatively comparable trend in the relative energies of the significant points of the reaction profile. Especially the nearly symmetric shape of the reaction profile was confirmed by the *ab initio* calculations. For a more sophisticated comparison further *ab initio* studies and density functional calculations including geometry optimization will be peformed.

Molecular dynamics studies

Molecular dynamics simulations using the AMBER 4.1 program were performed on diasteromeric CRL substrate adducts to obtain information about on the enantiodifferentiation of the enzyme in the catalytic hydrolysis. Since structural

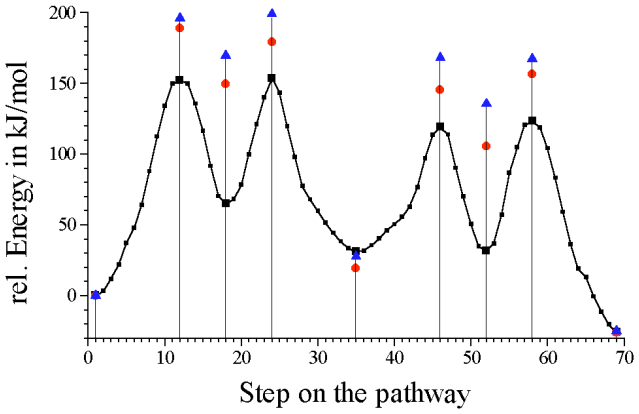


Figure 7 *Complete reaction profile and significant points of the enzymatic hydrolysis for the systems A (■), B (•) and C (▲)*

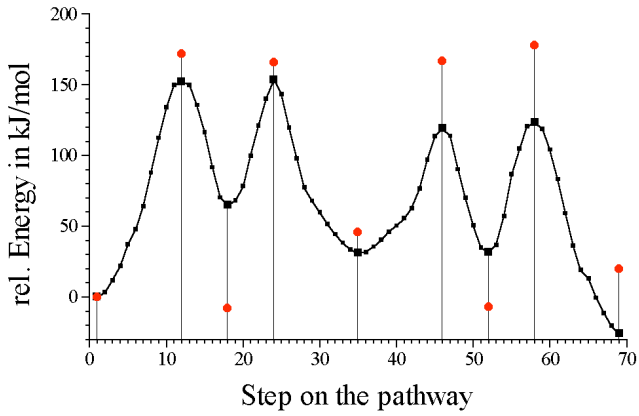


Figure 8 *Comparison for the PM3 (■) and ab initio (•) results for the model system A*

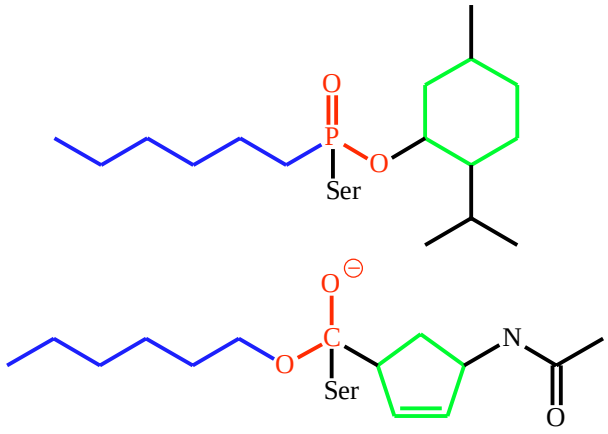


Figure 9 *Analogy of the serine adducts of menthyl hexylphosphonate (above) and 4-acetamidocyclopent-2-ene carboxylic hexylester (below)*

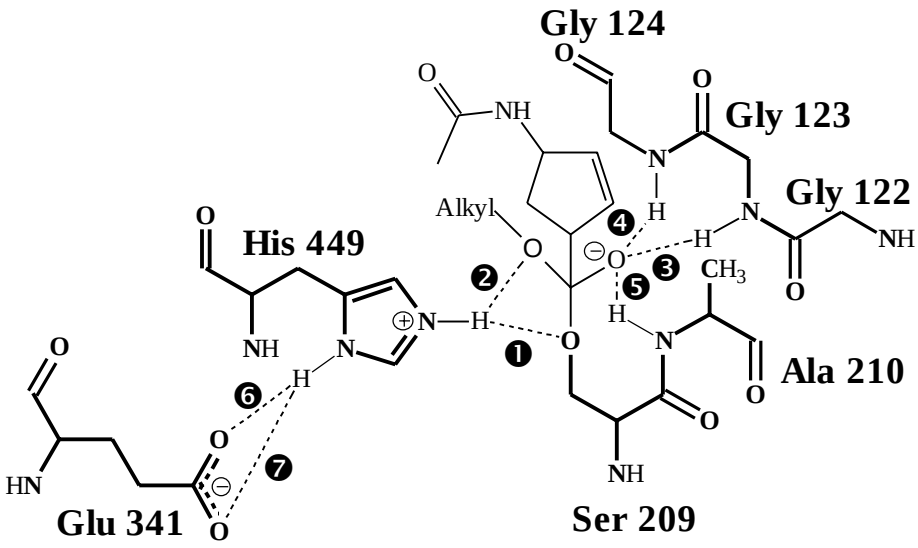
data were not available for the CRL substrate complexes with the 4-acetamido-cyclopent-2-ene-1-carboxylic esters, suitable starting structures had to be generated for the MD studies. Based on the similarity of the serine adducts with O-menthyl hexylphosphonochloridate and cis-4-acetamido-cyclopent-2-ene-1-carboxylic hexyl esters (Fig. 9) and the available X-ray data for the enzyme inhibitor complex [4,9], the starting structures were adapted for the model complexes of CRL with the esters 1a-e (Fig. 1). In the CRL model complexes an enzyme part with a cut-off radius of 1500 pm related to the atoms of the substrate was used to handle the simulations. The regarded enzyme part corresponds to about 36 % of the atoms of CRL.

The atomic charges for the substrate part that are not defined in the fragment library of the AMBER program were adapted by charges reproduced from the electrostatic potential (ESP) from PM3 calculations on the corresponding systems. This method is commonly used to define the atomic charges in force field methods [15,16]. Moreover, the applicability of the procedure was verified by calculations of the atomic charges of the tetrapeptide H-Gly-Phe-Ser-Gly-OH. For the tetrapeptide the atomic charges were calculated within the Mulliken population analysis as well as from ESP derived point charges using the PM3 method (MOPAC 6.0) and the *ab initio* method with STO-31G, 6-31G, and 6-31G* basis sets (TURBOMOLE). The results indicated that the atomic charges (ESP) obtained by the PM3 method correlate quite well with the AMBER reference charges of the tetrapeptide. Therefore, the ESP derived PM3 atomic charges were used in the AMBER MD studies on the model enzyme systems.

At first, classical MD simulations are performed on the diastereomeric model complexes of CRL with the hexyl esters (±)-1e (Fig.1). The MD parameters are given in the previous section. In order to avoid instabilities during the MD run, the amino acid residues of the enzyme part were kept in their enzyme-relevant positions. This was realised by fixing the coordinates of the N, C, and Cα atoms of the protein backbone and the oxygen atoms of the internal water molecules to their locations in the enzyme. From the analysis of the MD simulations, data were obtained on the conformational behaviour of the diastereomeric CRL substrate complexes. Especially the seven relevant hydrogen bonds in the active site of the CRL-(±)-1e model complexes (Fig. 10) were investigated in more detail. Their relevance and stability was illustrated by the corresponding trajectories and histograms.

The trajectories and histograms for the hydrogen bonds Nε(His 449)-H—Oγ(Ser 209) ❶ and Nε(His 449)-H—O(ester) ❷ are summarized in Fig. 11 for the complexes CRL-(+)-1e and CRL-(-)-1e. The horizontal and vertical lines at 280 pm in the diagrams are considered as a distance limit for the formation of a hydrogen bond XH—Y [23]. The results for the hydrogen bonds ❶ and ❷ of the complexes CRL-(+)-1e (blue) and CRL-(-)-1e (red) in Fig. 11 indicate a different

Figure 10 *Numbering of significant hydrogen bonds formed in the active site of the CRL substrate complexes*



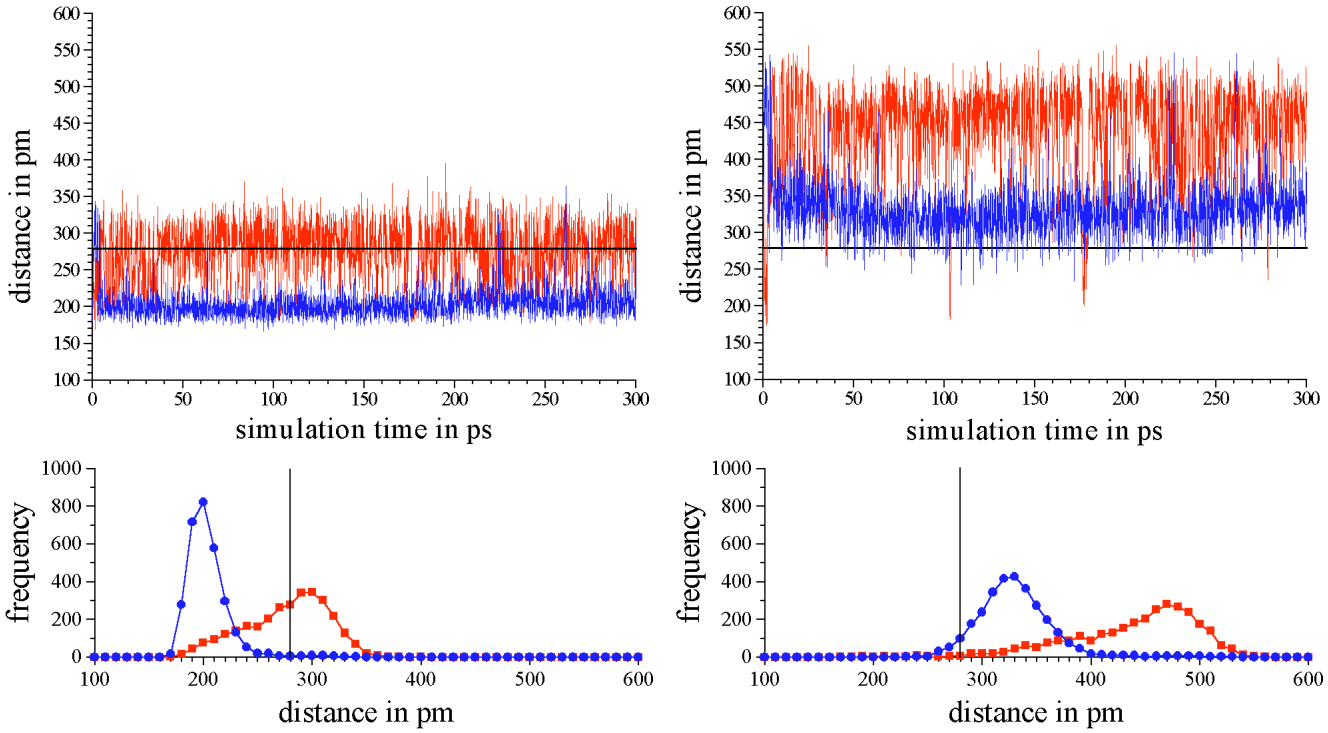


Figure 11 *Trajectories and histograms of the hydrogen bonds ❶ (left) and ❷ (right) for the model enzyme adducts with the esters (+)-1e and (-)-1e, respectively*

behaviour of the H bonds in the diasteromeric complexes. For the CRL-(+)-1e complex in comparison to the CRL-(-)-1e complex the mean values of the hydrogen bond lengths are generally smaller. The results predict that the formation of hydrogen bonds in the enzyme substrate complexes should be favoured for the (+)-1e ester. These findings are also reflected by the histograms. The hydrogen bonds ❸, ❹, ❺ in Fig.10, which represent the interactions of the oxyanion hole, show no significant differences in the two diastereomeric complexes. In both complexes the three hydrogen bonds are formed during the complete MD run. Therefore, the trajectories and histograms of these hydrogen bonds are not illustrated.

The trajectories and histograms of the two possible hydrogen bonds ❻ and ❼ between His 449 and Glu 341 are shown in Fig.12 for the CRL-(±)-1e complexes. The results show that during the total MD run only one of the possible hydrogen bonds is realised at a time in both complexes. There are no hints for the formation of symmetrical bifurcated hydrogen bonds between His 449 and the carboxylate group of Glu341. It is remarkable that for the CRL-(-)-1e complex an exchange is noticed in the formation of the hydrogen bond after a simulation time of about 10 ps (Fig. 12). A possible explanation could be the higher flexibility of the imidazole ring of His 449 in the CRL-(-)-1e complex caused by the lower stability of its hydrogen bonds to Ser 209 and the substrate in comparison to CRL-(+)-1e.

Finally, the analysis of the MD results of the diastereomeric CRL-(±)-1e systems indicate that the significant hydrogen bonds in the active site are generally more stable in the CRL-(+)-1e model complex. Especially the reduced possibility of the formation of the hydrogen bond ❷ (Fig. 10) which is important for the cleavage of the alcohol component of the substrate could be seen as a reason for the slower saponification of the (-)-1e in comparison to the (+)-1e ester by CRL [6].

Kinetic measurements of the enzymatic hydrolysis on 4-acetamido-cyclopent-2-ene-1-carboxylic alkyl esters (n = 1,2,3,4,6,8) by CRL have shown that the rate of hydrolysis reaches a maximum for the hexyl ester [6]. Therefore, we have performed MD simulations on the corresponding CRL model complexes with the substrates (+)-1a-e. From the analysis of the trajectories and histograms no remarkable deviations were found in the stability of the seven hydrogen bonds of the active site for the CRL-(+)-1a-e complexes. Obviously, more sophisticated studies are necessary on CRL-(+)-1a-e complexes including all significant steps of the mechanism of the enzymatic hydrolysis to gain more insights on the relation between the rate of hydrolysis and the length of the alcohol component of these systems. Nevertheless, the molecular modelling studies on the catalytic mechanism of CRL represent a contribution for the energetic and structural characterisation of significant points on the reaction path way of the ester hydrolysis. Moreover, the calculations should be seen as a first step for an explanation of the enantiodifferentiation of the substrates 1a-e by CRL.

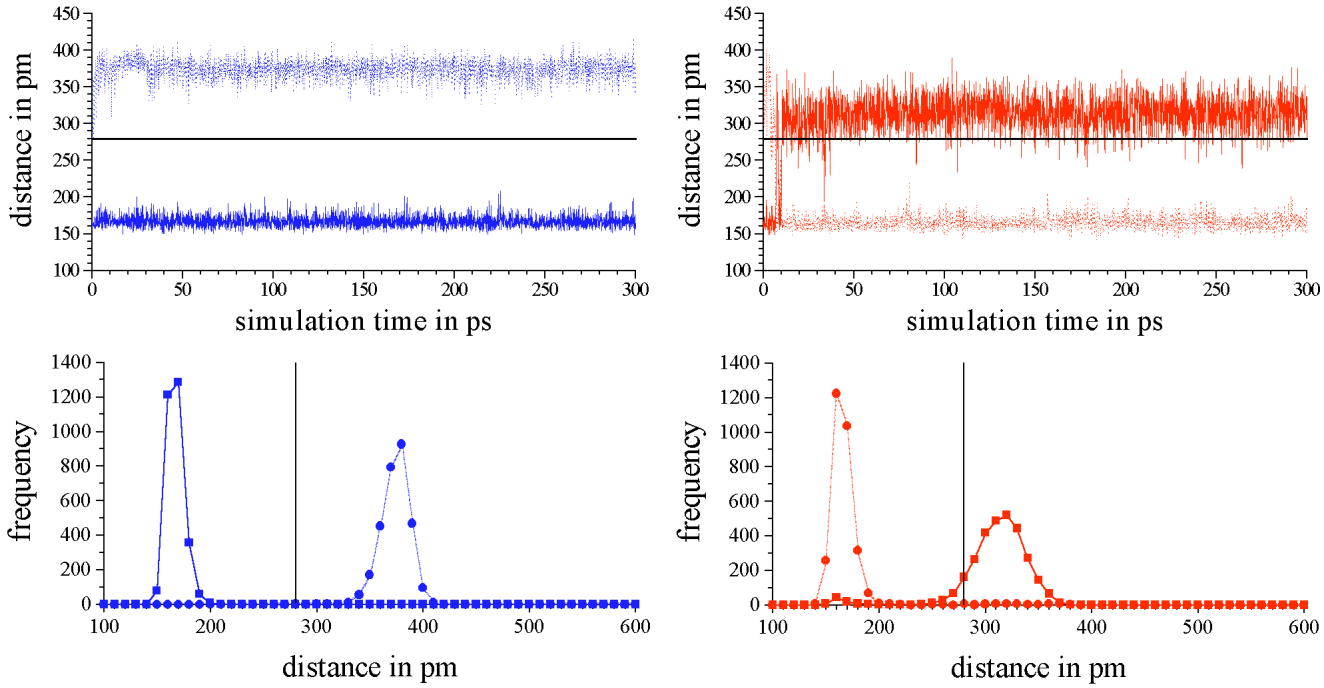


Figure 12 *Trajectories and histograms of the hydrogen bonds ❻ (bold) and ❼ (thin) of the model enzyme adducts with the esters (+)-1e and (–)-1e, respectively*

Acknowledgement We thank the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie for financial support and access to new computer facilities.

Supplementary material available statement PDB files for some species are attached to this publication. Moreover, the animation files for the graphical illustration of the calculated PM3 pathway including diagrams of the changes of significant bond lengths and atomic charges along reaction pathway are given.

0395s01.pdb: Optimized PM3 structure of the model system A (enzyme + ester)

0395s02.pdb: Starting structure (AMBER) of the CRL-(+)-1e model complex

0395s03.pdb: AMBER structure of the CRL-(+)-1e model complex after a simulation time of 300 ps.

0395s04.pdb: Starting structure (AMBER) of the CRL-(-)-1e model complex

0395s05.pdb: AMBER structure of the CRL-(-)-1e model complex after a simulation time of 300 ps.

0395v01.flc: Created file for the animation of the mechanism of the enzymatic hydrolysis for the model system A. It can be viewed with FLC video players, e.g. AAWIN and QuickTime Version3.0 or higher.

References

1. Wong, C.; Whitesides, G.M. *Enzymes in Synthetic Organic Chemistry*, Elsevier, Oxford 1994.
2. Drauz, K. [Ed.] *Enzyme Catalysis in Organic Synthesis: a Comprehensive Handbook*, VCH, Weinheim 1995.
3. Brockmann, H. L.; Law, J. H.; Kezdy, F. J. *J. Biol. Chem.* **1973**, *248*, 4965.
4. Cygler, M.; Grochulski, P.; Kazlauskas, R. J.; Schrag, J. D.; Bouthillier, F.; Rubin, B.; Serreqi, A. N.; Gupta, A. K. *J. Am. Chem. Soc.* **1994**, *116*, 3180.
5. Holmquist, M.; Haefner, F.; Norin, T.; Hult, K. *Protein Sci.* **1996**, *5*, 83.
6. Csuk, R.; Dörr, P. *Tetrahedron* **1995**, *51*, 5789.
7. Grochulski, P.; Li, Y.; Schrag, J. D.; Bouthillier, F.; Smith, P.; Harrison, D.; Rubin, B.; Cygler, M. *J. Biol. Chem.* **1993**, *268*, 12843.
8. Grochulski, P.; Li, Y.; Schrag, J. D.; Cygler, M. *Protein Sci.* **1994**, *3*, 82.
9. Grochulski, P.; Bouthillier, F.; Kazlauskas, R. J.; Serreqi, A. N.; Schrag, J. D.; Ziomek, E.; Cygler, M. *Biochemistry* **1994**, *33*, 3494.
10. Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.
11. Stewart, J. J. P., MOPAC 6.0, QCPE 455, Department of Chemistry, Indiana University, Bloomington, IN, 990.
12. Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, *C. Chem. Phys. Lett.* **1989**, *162*, 165.

13. Insight II 4.0.0 Turbomole, Molecular Simulations, Inc.
14. Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham; T. E. III; Ferguson,D. M.; Seibel, G. L.; Sing, U. C.; Weiner, P. K.; Kollman, P. A. AMBER 4.1, University of California, San Francisco, 1995.
15. Norin, M.; Hult, K.; Mattson, K.; Norin T. *Biocatalysis* **1993**, *7*, 131.
16. Norin, M.; Haefner, F.; Achour, A.; Norin, T.; Hult, K. *Protein Sci.* **1994**, *3*, 1493.
17. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
18. Cornell, W. D.; Cieplak, P. A.; Bayley, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.
19. Von Fircks, A.; Naumann, S.; Friedemann, R.; Koenig, S. *J. Mol. Model.* **1996**, *2*, 312.
20. Friedemann, R.; Naumann, S. *J. Mol. Struct. (Theochem)* **1997**, 398-399, 405.
21. http1.brunel.ac.uk:8080/depts/chem/ch241s/re_view/re_view.htm.
22. Kossiakoff, A. A.; Spencer, S. A. *Biochemistry* **1981**, *20*, 6462.
23. Smith, A. E.; Lindner, H.-J. *J. Comp. Aided Mol. Design* **1991**, *5*, 235.