



Contribution of both catalytic constant and Michaelis constant to CALB enantioselectivity: Use of FEP calculations for prediction studies

L. Chaput^{a,b}, Y.-H. Sanejouand^b, A. Balloumi^a, V. Tran^b, M. Graber^{a,*}

^a Université La Rochelle, UMR 6250 LIENSs CNRS – ULR, Pôle Sciences et Technologies, Bâtiment Marie Curie, Avenue Michel Crépeau, 17042 La Rochelle, France

^b Unité Biotechnologie, Biocatalyse et Biorégulation (U3B), UMR CNRS 6204 – Faculté des Sciences et des Techniques de Nantes, 2, rue de la Houssinière, BP 92208, 44322 Nantes Cedex 3, France

ARTICLE INFO

Article history:

Received 21 July 2011

Received in revised form

29 November 2011

Accepted 30 November 2011

Available online 16 December 2011

Keywords:

Enantioselectivity

Enzyme catalysis

Free energy perturbation

Candida antarctica lipase B

Molecular modelling

ABSTRACT

Candida antarctica lipase B (CALB) is characterized by its stability and ease of production and is widely used in the pharmaceutical industry. Here we report on the enantioselectivity of the enzyme using both experimental and computational methods. The apparent kinetic parameters were first experimentally determined for enantiopure butan-2-ol and pentan-2-ol substrates. We demonstrate that enantiopreference for the *R* form of butan-2-ol arises mainly from a lower apparent K_M . This corresponds to a major contribution of $\Delta \Delta G_{ES}$, the free energy difference between the *ES* complex formed with the *R* and *S* enantiomers, to $\Delta \Delta G^\ddagger$, the free energy difference between both transition states, in comparison with $\Delta \Delta G_{cat}$, the activation free energy difference. In the case of pentan-2-ol, we show that the enantiopreference for the *R* form comes from both a lower K_M and a higher k_{cat} . In addition, we used, for the first time, the free energy perturbation method to evaluate the free energy difference between tetrahedral intermediates formed with *R* and *S* alcohol enantiomers for a series of secondary alcohols. This is a valid model for $\Delta \Delta G^\ddagger$. Computational results were found to be in qualitative agreement with experimental data, and enable the determination of substrate orientation in the active site with fair confidence.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Lipase B from *Candida antarctica*, CALB, is α/β hydrolase (EC 3.1.1.3) which catalyses *in vivo* the hydrolysis of triglycerides. It also has a high activity and specificity for a wide range of synthetic esters of primary and secondary alcohols [1]. Due to its stability in organic media and its large-scale availability, it has found widespread applications in the enantioselective synthesis of molecules of pharmaceutical interest and in the resolution of racemic mixtures [2]. CALB displays an enantiopreference for the *R* form, according to Kazlauskas rules [3], and its enantioselectivity towards secondary alcohols is a property that has often been explored [4]. Enantioselectivity is considered to be related to the energy barrier difference between enantiomers, associated with the transition states formed during the second step of the bi-bi ping pong enzymatic mechanism [5]. The energetic determinants of CALB enantioselectivity is presently incompletely understood, despite many attempts to rationalize selectivity.

An initial explanation of CALB enantioselectivity was provided in 1998 by the team of Hult [4,6–9], who suggested that the orientation of *R* and *S* enantiomers are significantly different in the

alcohol-binding part of the active site, which contains the stereospecificity pocket [10], defined by Thr42, Ser47 and Trp104. More precisely, the model assumes that the slow-reacting *S* enantiomer has to orient its large substituent into the stereospecificity pocket in order to conserve the hydrogen bond between the oxygen of the alcohol moiety and His224-N ϵ , which is essential for catalysis. Although the large substituent does not fit easily into the stereospecificity pocket, this is the only binding mode which allows all the essential hydrogen bonds to form in the *S* enantiomer transition state and which leads to catalysis. The *R* enantiomer, on the other hand, positions its large substituent towards the surface of the protein and this does not lead to steric limitation. As a consequence, CALB has an enantiopreference for the *R* form.

In the past, there have been many other attempts to evaluate enantiomeric ratio by molecular modelling. For instance, enantioselectivity has been correlated, with more or less success, to the difference in the potential energy part ΔU of the free energy difference $\Delta \Delta G^\ddagger$ between the two enantiomer transition states [11–13]. One of the obvious limitations of this kind of calculation is that, whilst the potential energy of a protein in explicit solvent is typically of the order of several thousand kilocalories per mole, the energy difference between *R* and *S* tetrahedral intermediates, which are good models for the transition states [14], is expected to be $<5 \text{ kcal mol}^{-1}$, as the difference in free energy $\Delta \Delta G^\ddagger$ is related to the enantioselectivity, expressed by the enantiomeric ratio E ,

* Corresponding author. Tel.: +33 5 46 45 86 30; fax: +33 5 46 45 82 65.

E-mail address: mgrab@univ-lr.fr (M. Graber).

as follows: $\Delta\Delta G^\ddagger = -RT \ln E$ [15–17]. In addition, when the calculations are based on energy-minimized structures, they face the multi-minima problem, although this may be partly overcome by simulated annealing protocols [6]. Finally potential energy calculations assume that the entropy contribution to enantioselectivity is of minor importance, whereas a pronounced contribution of entropy to CALB enantioselectivity has been demonstrated in several cases [18,19]. Indeed it can represent as much as 50 % of the differential activation free energy in absolute value [20].

Other types of computational approaches have also been used for quantitative prediction of lipase enantioselectivity. For instance, Braiuca et al. [21] adapted the 3D-QSAR to quantitatively predict CALB enantioselectivity towards a wide set of substrates with relatively good accuracy. In 2000, Schultz et al. indirectly assessed the instability of the tetrahedral intermediate formed by the lipase from *Pseudomonas cepacia* with the *S* enantiomer of different secondary alcohols, compared to the *R* form, by measuring the hydrogen bond distance between the oxygen atom of the free alcohol in the active site and His224-N ϵ [22]. It appeared that the *S* enantiomer has greater difficulty to form this essential-to-catalysis hydrogen bond, suggesting that the energy for the reaction to proceed towards the formation of the tetrahedral intermediate is higher for the *S* substrate. Moreover, a correlation was established between the length of the *S* enantiomer hydrogen bond and the enantiomeric ratio [22]. More recently, Juhl et al. [23] used a customised docking method, where enzyme induced-fit was included by minimization of the enzyme–substrate complex to predict CALB and W104A mutant enantioselectivities for 1-phenylethyl butyrate. García-Urdiales et al. [24] used a different strategy and measured the steric constraints due to the nucleophile observed during standard molecular dynamics simulations. They showed that, in the case of the *S* form, higher CALB enantioselectivity correlates with a higher number of van der Waals unfavourable contacts [24].

There are existing methods to calculate free energy differences. One possibility is to use the thermodynamic integration method to quantitatively predict CALB enantioselectivity. Zhou [25] calculated free energy differences for butan-2-ol and four other *sec*-alcohols with either a bulky group or a bromide as substituent. Calculations were performed by modifying the dihedral angle value defined around the chiral carbon to transform, in a stepwise manner, from the *R* enantiomer to the *S* enantiomer. The free energy perturbation (FEP) method is an alternative approach. It has been applied for several decades [26–28] and remains a powerful tool today. This method is based on measurements of the progressive transformation of the system from an initial state to a final state, usually by following a non-physical (often coined alchemical) path. Throughout this process, free energy variation is measured. As a consequence, FEP calculations provide both internal energy and entropy contributions to the free energy. FEP has been extensively employed to compute free energy differences in numerous applications [29], including enzymatic stereospecificity [30] and allows for the calculation of differences as small as 1 kcal mol⁻¹ [31]. In practice, however, FEP calculations remain a difficult challenge, with well-known limitations [32]. Firstly, the molecular model and the force field used to describe the system thermodynamics must yield realistic probabilities for its most representative conformations. Secondly, low frequency motions and long time relaxations of the system must be handled properly to obtain a fair sampling of the relevant conformational space [33]. This is often done by restricting the number of degrees of freedom chosen to model the system but such a choice then affects the sets of sampled conformations. Another difficulty is finding a good convergence for the calculations [31].

Nevertheless, the FEP method has been successfully used to study enzyme enantioselectivity. Colombo et al. [34] obtained free energy differences in good agreement with experimental results.

They described the enantioselectivity of subtilisin in the resolution of a racemic mixture of *sec*-phenethyl alcohol by a transesterification reaction with vinyl acetate acyl donor [34].

In addition to these computational studies, there has been a lot of experimental work showing enantiopreference of wild type CALB for the *R* form of secondary alcohols. The classical method to quantify enantioselectivity is to measure enantiomeric ratio *E*, by using a formula with enantiomeric excess of substrates and products and conversion ratio [35], instead of using the original definition of $E = (k_{cat}/K_M)_R / (k_{cat}/K_M)_S$, which involves expensive and fastidious kinetic studies with enantiopure substrates. *E* values obtained by these rapid methods are useful for optimizing reaction yield and product purity, but fail to give information about the reaction step at which the enantiopreference occurs.

In the present work, we have highlighted the contribution of both catalytic constant and Michaelis constant to CALB enantioselectivity. Kinetic parameters for CALB-catalysed acylation of enantiopure butan-2-ol and pentan-2-ol were first experimentally determined. This allows the determination of the difference in reaction free energy profile for the two enantiomers, which leads to a better understanding of the origin of enantioselectivity. FEP calculations were then performed for tetrahedral intermediates formed with CALB and five different secondary alcohols with relatively high structural similarity (butan-2-ol, pentan-2-ol, hexan-3-ol, 3-methylbutan-2-ol, and 4-methylpentan-3-ol). A system in which the side chains of both *R* and *S* alcohols simultaneously exist was therefore built and the alchemical transformation consisted of going step by step from an interaction of the enzyme with the *R* form to an interaction with the *S* enantiomer. It is worthwhile stating that such an approach also provides concrete predictions of the substrate orientation within the CALB active site.

2. Experimental

2.1. Enantioselectivity measurements

Enantiomeric ratio *E* was calculated from apparent kinetic parameters obtained with enantiopure butan-2-ol and pentan-2-ol, according to the equation $E = (k_{cat}/K_M)_R / (k_{cat}/K_M)_S$. Enantiomeric ratio values for butan-2-ol, pentan-2-ol and hexan-3-ol, reported in Section 4, were previously obtained in our laboratory, from enantiomeric excess of substrates and products [35], in a continuous solid–gas reactor with methylpropanoate as acyl donor, and immobilized CALB, as previously described [20,36]. For branched substrates, values were picked from the literature [4], which were obtained at 39 °C in hexane, with *S*-ethyl thiooctanoate as acyl donor.

2.2. Enzyme and chemicals for kinetics

Novozym[®] 435 (immobilized *C. antarctica* lipase B) was kindly provided by Novozymes A/S, Bagsvaerd, Denmark. *R* and *S* pure enantiomers (99 %) of butan-2-ol and pentan-2-ol were purchased from Sigma–Aldrich (St. Louis, USA), whilst methylpropanoate was from Fluka (St. Quentin-Fallavier, Switzerland).

2.3. Enzymatic reactions with enantiopure butan-2-ol and pentan-2-ol

Initial rate measurements were performed at 45 °C in 2-methylbutan-2-ol. 4 ml of the reaction mixtures containing various amounts of butan-2-ol (100–4000 mM for the *S* form and 100–1000 mM for the *R* form) or pentan-2-ol (100–900 for the *S* form and 100–1500 mM for the *R* form) were incubated for 10 min prior to addition of 10 mg of Novozym[®] 435 for the acylation of butan-2-ol or *R* pentan-2-ol with 430 mM methylpropanoate, or

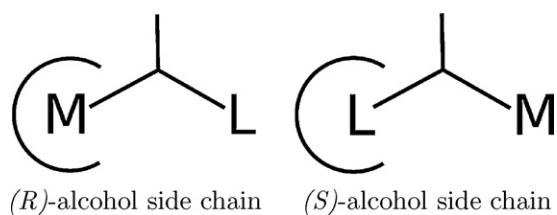


Fig. 1. Orientation of the *R* and *S* side chain of a pair of enantiomers within the stereospecificity pocket of CALB, sketched as a truncated circle. L and M are the large and medium chains of the substrate, respectively.

20 mg of Novozym[®] 435 for reaction with *S* pentan-2-ol. 200 μ l samples were taken at intervals and centrifuged at 14,000 \times g. The supernatant was analysed by gas chromatography (GC), after two times dilution with 2-methylbutan-2-ol.

2.4. GC analysis

Quantitative analysis of reaction products were conducted using a 7890 GC system from Agilent for the analysis of 1-methylpropyl propanoate (55 $^{\circ}$ C, 20 min) and of 1-methylbutyl propanoate (55 $^{\circ}$ C 15 min, 3 $^{\circ}$ C min⁻¹, 85 $^{\circ}$ C 5 min), at a flow rate of 1.5 ml min⁻¹ with a Chirasil-Dex CB (25 m, 0.25 mm i.d., 0.25 μ m β -cyclodextrin, Chrompack, France) column. Products were detected by FID and quantified using HP Chemstation software. External calibration was performed with chemically synthesized esters from the corresponding alcohol and propanoic anhydride in pyridine at room temperature.

3. Computational methods

3.1. Enzyme and tetrahedral intermediate structures

The starting CALB enzyme was the $R=1.55$ \AA crystallographic structure solved by Uppenberg et al. [37] (PDB entry 1TCA). A transition state analogue crystal structure, obtained with phosphonate irreversible inhibitor (PDB entry 1LBS) was used to build the tetrahedral part of the reaction intermediate. The acyl part is a propanoyl group, to allow for the correct location of the central part of the tetrahedral intermediate. The negatively charged oxygen was oriented towards the oxyanion hole to establish hydrogen bonds with Thr40 and Gln106. The position of the substituents of *R* and *S* alcohol enantiomers in the active site were orientated according to the model postulated by Haefner et al. [4], so that the large chain of the *S* alcohol was oriented into the stereospecificity pocket, and the medium chain towards the active site entry, as shown in Fig. 1. Inversely, the medium chain of the *R* alcohol was located in the stereospecificity pocket, and the large chain was oriented towards the active site entry. Note that this was done in order to, firstly, define a starting point for the simulations and, secondly, during the course of the calculations, the substrates were thus free to reorient themselves in a different way.

However, the essential hydrogen bond involved in the transition state between His224-N ϵ and the alcohol oxygen of the tetrahedral intermediate, was prevented from disruption during the molecular dynamics simulations by a harmonic constraint of 100 kcal mol⁻¹ applied on the distance between heavy atoms exceeding 2.8 \AA . Without this constraint, the hydrogen bond would probably have been disrupted during the dynamics. When hydrogen bond disruption is observed in the case of the *S* enantiomer [38], it is usually considered as a measure of the difficulty this enantiomer experiences before reaching an orientation which allows the reaction to proceed. In the present work, our choice has been to maintain the system as close to the transition state as possible, in order to obtain estimates for free energy differences at the top of the

rate-limiting energy barrier. From a technical point of view, the energy restraint was included in the potential energy of the system, and was effective all along the simulation. Free energy was calculated as the summation of energy differences, therefore the energetic restraint was balanced. The only way, then, for the restraint to affect the free energy is by modifying the conformation sampling, and that is exactly the point which was improved by using H-bond restraint, which enhances the sampling around the transition state conformation.

3.2. Molecular dynamics simulations

CHARMM c35 program [39] and the CHARMM22 all-atom force field were used. Force field parameters for the tetrahedral intermediate were taken from the literature [40]. These parameters were obtained from *ab initio* calculations and were specifically developed for CHARMM22 force field. Other parameters required for modelling the alkyl side chains of alcohols were defined by homology with available CHARMM22 parameters.

In order to mimic our experimental conditions, within a solid-gas reactor, a 7 \AA water layer was added to the CALB structure, all crystallographic water molecules were retained for correct solvation of the active site. Then all water molecules were energy-minimized, with 10,000 steps of conjugated gradient followed by 2000 steps with the ABNR (Adopted Basis Newton-Raphson) algorithm.

Next, the whole system was energy-minimized again, except for atoms kept fixed throughout this study, namely, all atoms more than 24 \AA away from the oxygen of the alcohol moiety, and all α -carbon atoms more than 22 \AA away. As a result, 75% of the protein was left totally free to move. Our criterium is expected to maintain the enzyme closer to the crystallographic structure, as well as to improve the convergence of FEP calculations by preventing the unwanted contribution of remote events, such as drifts of loops or of residues far from the active site. It is, therefore, a compromise between having a good convergence whilst leaving significant possibility for the active site to freely accommodate the tetrahedral intermediate and the side chains of the substrates. In addition, the protocol was designed to evaluate small conformational rearrangements such as those occurring when going from the *R* to the *S* enantiomer. Sampling larger conformation movements, like α -helix would require much longer sampling duration at each λ , which was not the aim of the article. Other authors introduced similar types of restraints on atoms far from the region where the study focuses on FEP methods [41,42]. In another example, Allouche et al. [43] defined free atoms in a radius of 9 \AA centred on the alchemical transformation of one cation to another (Ca²⁺ towards Mg²⁺). In our study, however, alchemical transformations concern more atoms, hence the choice of a much larger radius for the free part of the system. On the other hand, Trodler and Pleiss [44] showed by molecular modelling that the structure of *C. antarctica* lipase B during molecular dynamics simulations, in water or in five different organic solvents, exhibits a low deviation from the crystal structure. Finally, McCabe et al. [45] made circular dichroism measurements, using synchrotron radiation, to study conformational changes in water and various organic solvents and concluded that the secondary structure of CALB in aqueous buffer is close to its crystal structure. These studies suggest that CALB does not exhibit significant domain movements and that fixing atoms far away from the active site may not perturb CALB significantly.

An initial heating stage of 50 ps, from 100 to 300 K, was performed before every FEP calculation and then the temperature was increased by 1 K every 100 steps. For accurate results, the system must be well equilibrated with long initial equilibration periods, corresponding to 800 ps (trajectory 1) and 1000 ps (trajectories 2 and 3). During all molecular dynamics simulations, the non-bonded

pairs list was updated every 20 steps and the temperature was checked every 5000 steps.

3.3. Free energy perturbation protocol

A system was built with distinct, and simultaneously existing, side chains of both *R* and *S* secondary alcohols, including the chiral carbon, also called the dual topology approach. In accordance with the alchemical transformation principle, the oxygen atom is shared by the two enantiomers. A hybrid potential energy $U(r, \lambda)$ is associated with such a system, which is a function of r , the atom coordinates, and λ , a coefficient scale to quantify the interaction energy of the *R* enantiomer with the rest of the enzyme. The corresponding coefficient for the *S* enantiomer is $1 - \lambda$. The free energy difference, ΔF along the path starting at $\lambda_i = 0$ and ending at $\lambda_i = 1$ was calculated using the exponential formula [46]:

$$\Delta F_{R \rightarrow S} = F_S - F_R = \sum_{i=0}^{n-1} -\beta^{-1} \ln \langle e^{-\beta \Delta U_{\lambda_i \rightarrow \lambda_{i+1}}} \rangle_{\lambda_i}$$

where F_S and F_R represent the free energy of the enzyme with, respectively, the *R* and *S* alcohol moiety linked to the tetrahedral intermediate. $\langle \rangle_{\lambda_i}$ represents the canonical ensemble average obtained with λ_i , where β is the inverse of the thermal energy, namely, $k_B T$, k_B being the Boltzmann constant and T , the temperature. The sum of the n successive terms calculated along the *R* to *S* transformation leads to the total free energy difference between the complex with *R* and *S* enantiomers. The value obtained with the sum computed for the reverse transformation of *S* to *R* should rigorously give the opposite value to the *R* to *S* transformation. Such a comparison is a standard indicator of the quality of a FEP calculation.

In practice, λ_i was linearly incremented by steps of 0.05, along twenty ($n = 20$) successive equilibration and productive dynamics, of 25 and 50 ps each. The decision to divide the transformation into twenty steps yields a small free energy perturbation at each step and, as a consequence, better accuracy. Too high a number of sub-trajectories is not useful because if λ_i is too close to 0 (or to 1), weak interactions between alcohol side chains and the enzyme yield unrealistic behaviour characterized by wide movements of the side chains and, therefore, by a poor sampling of the conformational space. This is why the λ_i value was initially set at 0.975 and 0.025 for the interaction between, respectively, the *R* and *S* alcohol side chain and the enzyme. The transformation of the substrate from *R* to *S* was followed by a 100 ps equilibration step and, then, by the reverse transformation (*S* to *R*), in order to assess the convergence of each FEP calculation. Twenty productive sub-trajectories were thus obtained for each transformation of an *R* enantiomer to an *S* enantiomer, and the same number obtained for the reverse path. In addition, each FEP calculation was made twice, starting from two sets of initial conditions.

The free energy perturbation (FEP) protocol is presumed to be reliable when the energy computed for the *R* to *S* transformation and for the reverse path give similar values with opposite signs. This presumption cannot, in fact, be taken for granted. Indeed, reaching convergence in free energy perturbation calculations is a challenge that depends on several parameters involved in the process. The FEP method assumes that the part of the phase space where the energies of both compared systems are significantly different can be extensively explored. This can only be achieved by using enough computer time to sample the conformations with significant energy differences between two neighbouring states along the path followed during each transformation. In practice, FEP calculations were performed using the BLOCK command available in CHARMM [39], which allows one to set the value of λ_i . The

Table 1

Kinetic constants for *R* and *S* enantiopure butan-2-ol and pentan-2-ol (V_{\max}^{app} expressed in $\text{mmol min}^{-1} \text{mg}^{-1}$ of immobilized enzyme, and K_M expressed in mM).

	Butan-2-ol	Pentan-2-ol
K_M		
<i>R</i>	258	70
<i>S</i>	1284	>472
V_{\max}^{app}		
<i>R</i>	0.062	0.048
<i>S</i>	0.051	≈0.001
<i>E</i>	6	>324

perturbation at each step was then calculated using the FREE module of CHARMM and the wide sampling method. Moreover, the trajectory at λ_i was used to compute incremental ΔF values from $\lambda_i - 0.025$ to λ_i and from λ_i to $\lambda_i + 0.025$. The molecular dynamics transformation of *R* to *S*, followed by the reverse path, took about 18 h when run on 24 parallelized Bi-Xeon processors.

4. Results and discussion

4.1. Experimental determination of apparent kinetic constants for enantiopure butan-2-ol and pentan-2-ol

Experimental apparent kinetic parameters obtained for enantiopure butan-2-ol and pentan-2-ol are presented in Table 1. The K_M^{app} for the *R* form of butan-2-ol is 258 mM and for the *S* form is 1284 mM. In the case of pentan-2-ol, we obtained a K_M^{app} of 70 mM for the *R* form. However, CALB was not saturated when using *S* pentan-2-ol concentrations up to 1500 mM, consequently the K_M^{app} value was above 472 mM, the alcohol concentration giving half of the reaction rate obtained at 1500 mM. Concentrations above 1500 mM of *S* pentan-2-ol were not used owing to the exorbitant price of this enantiopure alcohol. The V_{\max}^{app} for butan-2-ol is 0.062 for the *R* form and 0.051 for the *S* form, in $\text{mmol min}^{-1} \text{mg}^{-1}$ of immobilized enzyme. Pentan-2-ol gave values of 0.048 and around 0.001 mmol min^{-1} , for the *R* and *S* forms respectively, results obtained by non linear regression of the Michaelis–Menten equation. The data demonstrate that for butan-2-ol, the enantio-preference for the *R* substrate can be attributed mainly to the K_M difference, which is lower, by a factor of 5 for the *R* alcohol, suggesting that the enzyme affinity for the *R* alcohol is greater than for the *S* form. In contrast, the apparent V_{\max} is only slightly higher for the *R* enantiomer. However, in the case of pentan-2-ol, the enantio-preference arises from both a lower K_M and a much higher V_{\max} for the *R* form. Assuming that 1 g of Novozym® 435 contains 1 μmol of active CALB [47], k_{cat} can be calculated from V_{\max} values and are equal to 1033 s^{-1} and 850 s^{-1} for the *R* and *S* form of butan-2-ol, and to 800 s^{-1} and 17 s^{-1} for the *R* and *S* form of pentan-2-ol.

Hypothetical energy profile diagrams for the CALB-catalysed acylation reactions acting on *R* and *S* enantiomers for both butan-2-ol and pentan-2-ol are shown in Fig. 2.

Apparent K_M can be used as an apparent dissociation constant, K_S , under conditions of steady state assumption. The lower K_M for the *R* enantiomer corresponds to a ground state stabilization of the *ES*-complex, which has a lower energy compared to *S* enantiomer. k_{catR} and k_{catS} correspond to the difference between *ES* and *ES*[‡] for *R* and *S* enantiomers. In the case of both pentan-2-ol and butan-2-ol, k_{catR} is higher than k_{catS} . The second order rate constant for the enzymatic reaction, starting with free enzyme and substrate is given by k_{cat}/K_M and corresponds to the difference between *E* + *S* and *ES*[‡]. The enantio-preference for the *R* form is linked to the energy barrier difference between enantiomers, associated with the transition states (*ES*_{*S*}[‡]–*ES*_{*R*}[‡]). We demonstrate here that this difference is primarily due to the K_M difference between the *R* and *S* form,

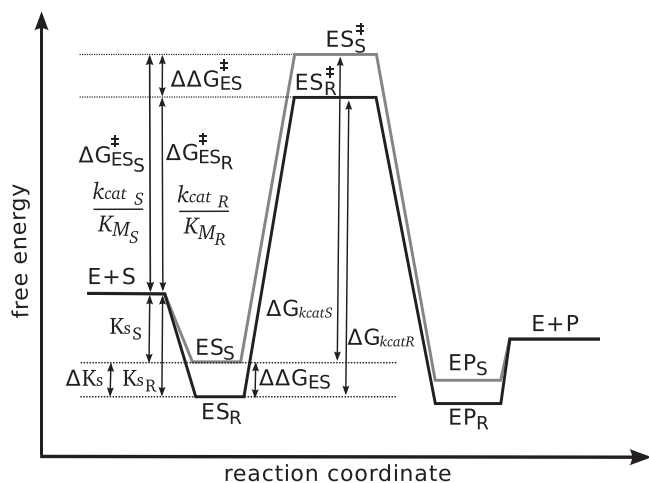


Fig. 2. Energy profile diagrams for the CALB-catalysed acylation reaction acting on R (black line) and S (grey line), with *E* the enzyme, *S_S* and *S_R* the S and R enantiopure substrates, *ES_S* and *ES_R*, the enzyme–enantiopure substrate complexes, *ES_S[‡]* and *ES_R[‡]*, the enzyme–enantiopure substrate transition state complexes, *P_S* and *P_R*, the S and R enantiopure products. $\Delta\Delta G_{ES}^{\ddagger}$ represents the energy difference between bound enantiopure substrates, ΔG_{kcatS} and ΔG_{kcatR} represent the energy activation for each enantiomer, $\Delta\Delta G_{ES}^{\ddagger}$ represents the free energy difference between R and S enzyme–enantiopure substrate transition state complexes.

rather than to the k_{cat} in the case of butan-2-ol. In contrast, the energy barrier difference for pentan-2-ol is associated with both K_M and k_{cat} .

These results are significantly different from previous experimental results obtained by Magnusson et al. [48] (2005) for CALB catalysed acylation of pure enantiomers of 1-phenylethanol, a very bulky secondary alcohol. The large enantiomeric ratio ($E = 1, 300, 000$) was almost entirely due to the difference in apparent k_{cat} between the enantiomers, 570 s^{-1} and 0.00053 s^{-1} for the R and S forms, respectively. We find that the respective contributions of K_M and k_{cat} to enantioselectivity are highly variable and probably depend on substrate structure. In the particular case of a secondary alcohol with a very bulky large substituent, like 1-phenylethanol, Magnusson et al. obtained an enantioselectivity almost entirely due to the difference in k_{cat} . The inverse trend was obtained in the present work, with another particular case: butan-2-ol which has only an ethyl as “large” substituent and for which enantioselectivity is almost entirely due to the difference in K_M . In case of pentan-2-ol an intermediate result was obtained: enantioselectivity arises from both a lower K_M and a much higher k_{cat} for the R form.

E values (Table 1) calculated using the formula $E = (k_{cat}/K_M)_R / (k_{cat}/K_M)_S$ give a value of 6 for butan-2-ol and a value above 324 for pentan-2-ol. These values are similar to those obtained previously in our laboratory at the same temperature, in a continuous solid–gas reactor with the same acyl donor at 20 % relative humidity, 6 for butan-2-ol and 330 for pentan-2-ol [20], using enantiomeric excess to calculate E . Due to the similarity of the results obtained using different approaches, we chose to add hexan-3-ol as an additional substrate for our computer simulations, a substrate for which we had previously determined an E value of 80 under the same conditions (Table 2). Two supplementary substrates containing branched substituents (3-methylbutan-2-ol and 4-methylpentan-3-ol) were also added as extra challenges for our computational modelling. For these last substrates, E values were taken from the literature [4]. They were obtained at the same temperature but with a different acyl donor: S-ethyl thiooctanoate (Table 2).






4.2. Comparison of FEP results from this study with experimental data

Table 2 represents the values obtained by FEP methods from calculations of $\Delta F_{R \rightarrow S}$ and $\Delta F_{S \rightarrow R}$, the free energy difference going from the R form of the enzyme to the S form of the enzyme and the reverse path. We obtained positive values for ΔF when going from the R to the S enantiomer, except for one trajectory, and obtained consistently negative values along the reverse path. This corresponds to a lower free energy with the R form than with the S form, as expected from experimental data. Thus from a qualitative point of view, FEP results correctly predict the enantioselectivity of CALB for the R form. From a quantitative point of view and for linear secondary alcohols, average ΔF values are equal to $1.38 \text{ kcal mol}^{-1}$ for butan-2-ol, $3.62 \text{ kcal mol}^{-1}$ for pentan-2-ol and $2.58 \text{ kcal mol}^{-1}$ for the first trajectory of hexan-3-ol (the exclusion of trajectories 2 and 3 is explained below). These values are in good agreement with experimental $\Delta\Delta G^{\ddagger}$. For ramified alcohols, the average ΔF value is equal to 2.72 for 3-methylbutan-2-ol and the unique ΔF value obtained with correct signs for 4-methylpentan-3-ol is equal to $1.97 \text{ kcal mol}^{-1}$. FEP calculations correctly rank these two substrates, but fail to rank linear and ramified alcohols. The ΔF values of branched alcohols are lower than experimental $\Delta\Delta G^{\ddagger}$, and the standard deviation obtained for 3-methylbutan-2-ol is higher than for linear alcohols. Calculations with 4-methylpentan-3-ol gave the weakest convergence, perhaps due to the large bulky substituent of this alcohol, which may require a pronounced rearrangement of the enzyme structure in order to accommodate both enantiomers.

It is worth noticing that substrates with very different enantiomeric ratios are expected to have transition states with small free energy differences between both enantiomers, an important point when comparing experimental and calculated free energy differences, $\Delta\Delta G_{exp}^{\ddagger}$. Indeed, for experimental E values ranging between 6.0 and 705 (Table 2), calculated free energy differences are expected to range between 1 and 4 kcal mol^{-1} , which is far less than the energy of a typical hydrogen bond, which is about $4.8 \text{ kcal mol}^{-1}$.

Hexan-3-ol is a special case in point, in the case of trajectory 1, FEP results are close to the experimental data ($2.61 \text{ kcal mol}^{-1}$), but trajectory 2 had significantly higher ΔF values, therefore a third trajectory was performed. Trajectories 2 and 3 both yield an average ΔF value equal to $4.21 \text{ kcal mol}^{-1}$, instead of $2.58 \text{ kcal mol}^{-1}$ found from trajectory 1. Comprehensive analysis of the global structure of CALB during the three independent trajectories provides a good indication of the putative implication of secondary structure elements surrounding the active site. During the first trajectory, which displays better ΔF , a small shift of the α -helix 10 occurred, compared to the crystallographic structure. This shift was not observed during the two other trajectories. This α -helix shift may be involved in a better fit between the enzyme and the substrate, explaining the smaller free energy change during the first trajectory. Interestingly, several groups previously observed the movement of this α -helix. Recently, Trodler and Pleiss [44] observed the flexibility of this helix during molecular dynamics simulations. Skjøt et al. [49] also demonstrated that during a 10 ns dynamic trajectory in a water box with periodic conditions, α -helix 5 and 10 of CALB displayed significant mobilities. α -helix 5 is far from the active site, therefore, in the case of the substrates described here, it can be assumed that this helix has a negligible effect on substrate conformation, contrary to α -helix 10, which is directly in contact with the side chain of the alcohol, particularly through residues Ile189, Ala278 and Ala282. Unfortunately, the FEP/MD protocol was not designed to handle events like α helix movement, as observed by Zhao et al. [50]. These authors assessed the free energies for antidepressant binding to protein receptor by FEP/MD and mentioned that the receptor was sampled in only one conformation, due to the fact that large

Table 2
 Calculated $\Delta F_{R \rightarrow S}$ and $\Delta F_{S \rightarrow R}$ compared to the experimentally determined $\Delta \Delta G_{R \rightarrow S}^{\ddagger \text{exp}}$, as derived from the experimentally determined enantioselectivity E . All free energy differences are expressed in kcal mol⁻¹.

	Substrate	E	$\Delta \Delta G_{R \rightarrow S}^{\ddagger \text{exp}}$	Traj.	$\Delta F_{R \rightarrow S}$	$\Delta F_{S \rightarrow R}$	$ \Delta F^{\text{avg}} \pm \text{sd}^a$
	Butan-2-ol	6.0	1.07	1	+1.10	-1.33	1.38 ± 0.66
				2	+0.78	-2.30	
	Pentan-2-ol	330	3.46	1	+4.04	-3.94	3.62 ± 0.5
				2	+3.57	-2.94	
	Hexan-3-ol	80	2.61	1	+2.06	-3.09	2.58
				2	+4.90	-4.84	
				3	+4.29	-4.79	
	3-Methylbutan-2-ol	705	3.91	1	+3.50	-2.03	2.72 ± 0.8
				2	+3.32	-2.02	
	4-Methylpentan-3-ol	109	2.80	1	+2.86	-1.07	1.97
				2	-0.07	+1.60	

^a Standard deviation.

^b No convergence.

movements occurred during a larger time scale, exceeding the time frame used in the protocol.

Our results are in concordance with the alternative protocol of thermodynamic integration used by Zhou [25] for his free energy calculations for two bulky chiral alcohols as well as for butan-2-ol. He failed, however, to predict enantioselectivity for chiral alcohols containing bromide and the reproducibility of his results was not addressed, as only a single trajectory per substrate was performed [25].

4.3. Substrate orientations

In our study, the starting orientation of the alcohol substrates was defined according to the model suggested by Haeffner et al. [6]. For all substrates, except pentan-2-ol, this orientation was well conserved all along the dynamic trajectory for both *S* and *R* alcohol side chains. So we think that initial orientations we used for the substrates are acceptable, because we did not observe any rearrangement for them, even when substrates get the opportunity to reorient themselves, when they are weakly interacting with the enzyme (λ near 0 and 1 for *S* and *R* enantiomers, respectively). As seen in Fig. 3A, in the case of butan-2-ol, the large chain (ethyl group) of the *S* alcohol is located in the stereospecificity pocket, close to amino acids Gly39, Thr40 and Thr42. The large chain of the *R* alcohol is oriented towards the entry of the active site. In the case of pentan-2-ol, a re-orientation of the large substituent of the *S* enantiomer alcohol occurs for both trajectories 1 and 2. It can be observed in Fig. 3B that the large chain bends towards the active site entry. This is probably due to steric constraints within the stereospecificity pocket. To our knowledge, this orientation was not observed previously. Kwon et al. [11] did demonstrate, however, that for some alcohols, side chains may adopt other orientations than those described by the Haeffner model.

4.4. Relation between free energy profile of enzymatic catalysis and experimental and computational results

The free energy differences calculated here, using the FEP method, are expected to match those for the transition states ES^\ddagger of the reaction, as catalytic rates depend on the energy barrier associated with the formation of the transition state. The

experimental data display slightly higher apparent k_{cat} for the *R* enantiomer in the case of butan-2-ol and high apparent k_{cat} in the case of pentan-2-ol, and show that enantioselectivity is also linked to the difference between apparent K_M . The data can be understood by detailing the energy profile diagram of the reaction (Fig. 2) by using Eyring's transition state theory, which links the reaction rate to the activation energy. The Eyring law defines the rate k of a reaction as a function of temperature and of the activation energy: $k = k_B T / h e^{-\Delta G^\ddagger / RT}$.

This energy corresponds to the free energy difference between a ground state and the top of the energy barrier. In the present case, the top of this barrier has been modelled by the tetrahedral intermediate ES^\ddagger . The ground state can be defined as the complex enzyme–substrate ES . Activation energy is equal to ΔG_{kcat} , corresponding to k_{cat} (see Fig. 2), and for given concentrations of *E* and *S*, the rate constant for the enzymatic reaction, starting with free enzyme and substrate, is given by k_{cat}/K_M and corresponds to the energy ΔG^\ddagger . In the Michaelis–Menten model, the reaction proceeds through a classical ligand–enzyme interaction model, and the K_M can be assimilated with the inverse of the affinity constant, under conditions of steady state. A small K_M corresponds to a good affinity of the substrate for the enzyme, characterized by a low free energy ΔG_{ES} stabilizing the complex ES . Therefore, as shown in Fig. 2, the K_M is linked to the energy difference ΔG_{ES} between free substrate and bound substrate.

Here, it was experimentally observed that the K_M is much lower for the *R* enantiomer. This means that the energy of the ES complex is lower for the *R* enantiomer, and the difference in affinity for the enzyme between the two enantiomers is characterized by the free energy difference $\Delta \Delta G_{ES}$. In addition, experimental data demonstrate that the apparent V_{max} is higher for the *R* enantiomer, indicating that the free energy, ΔG_{kcat} , is smaller for the *R* enantiomer. Thus, the difference between transition state energy $\Delta \Delta G^\ddagger$ depends on both $\Delta \Delta G_{ES}$ and $\Delta \Delta G_{kcat}$, with $\Delta \Delta G_{kcat} = \Delta G_{kcatS} - \Delta G_{kcatR}$. Consequently, enantioselectivity is characterized by the free energy difference $\Delta \Delta G^\ddagger$ that can be expressed as $\Delta \Delta G^\ddagger = \Delta \Delta G_{ES} + \Delta \Delta G_{kcat}$. The resolution of butan-2-ol by CALB-catalysed acylation demonstrates that the contribution of the first energetic term ΔG_{ES} is predominant compared to $\Delta \Delta G_{kcat}$, whereas for pentan-2-ol, the two factors are important.

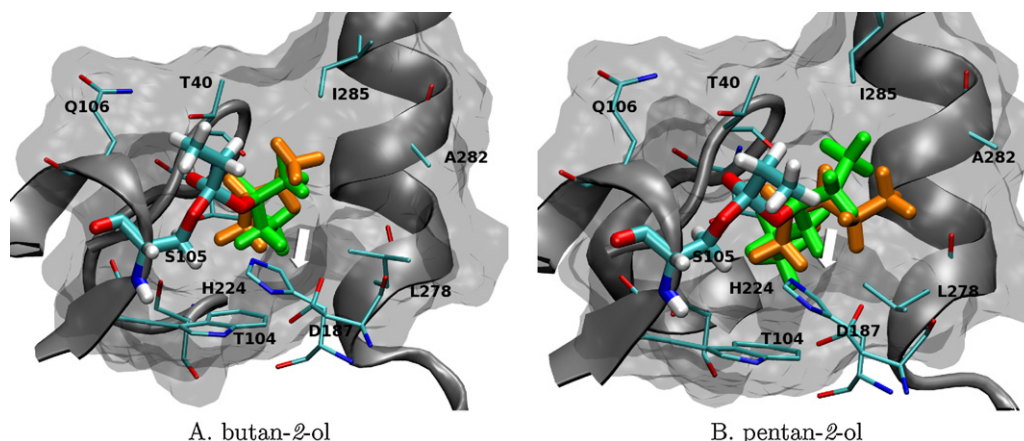


Fig. 3. Tetrahedral intermediate (TI) and side chains orientation of butan-2-ol and pentan-2-ol enantiomers. *S* enantiomer in green and *R* enantiomer in orange. The white arrow points out the stereospecificity pocket.

4.5. Conclusions

Experimental data obtained with enantiopure butan-2-ol shows that enantioselectivity of CALB for the *R* form arises mainly from a lower apparent K_M and, to a much lesser extent, from a higher k_{cat} for this enantiomer. With pentan-2-ol, enantioselectivity arises from both a lower K_M and a much higher k_{cat} for the *R* enantiomer, suggesting that no general rule can be defined for all substrates, as far as the contribution of the various kinetic parameters to enantioselectivity is concerned. FEP calculations presented in this study successfully provided qualitative prediction of the enantioselectivity of CALB for *R* enantiomers in the case of four of the five substrates tested. However, the quantitative prediction of the enantiomeric ratio itself proved challenging. In the best cases, namely, butan-2-ol and pentan-2-ol, the corresponding free energy difference was overestimated by 0.2–0.3 kcal mol⁻¹, on average. For hexan-3-ol, only one trajectory amongst the three which were performed, gave almost the same result as the experimental result. This corresponds also to the only case where a small shift of the α -helix 10 occurs, and this is certainly a case where a better accommodation of the enzyme occurred. On the whole, FEP calculations provide a much more efficient evaluation of energy difference between enantiomers than potential energy evaluation. Differences in absolute values between calculated and experimental $\Delta\Delta G^\ddagger$ may be attributed to the limitations of our approach for modelling global enzyme accommodation, including, for example, the possible movement of the α -helix 10 which may allow CALB to adapt the shape of its active site to large substrates like hexan-3-ol. We are fully aware of the fact that the protocol can be improved, in order to give a better reproducibility and enabling at the same time, a global enzyme accommodation during simulation. Another possible source of error is the approximate description of the transition state of the reaction as a tetrahedral intermediate. In this respect, QM/MM calculations may prove useful, to better define both the geometry of the transition state and the distribution of charges around the chiral centre and also perhaps including a few key neighbouring amino acid residues [51–53]. Notwithstanding, FEP calculations can provide results for novel substrates, without the need for a significant number of experimental data to adjust the model, as is the case, for instance, with 3D-QSAR methods. Moreover, FEP calculations can provide clues about tetrahedral intermediate geometries. Indeed, the analysis of the trajectories performed during this study strongly supports Haefner's model of the orientation of the *S* substrates within the active site of CALB: the *S* orientation was preserved in all cases, except that of pentan-2-ol. The data presented here help in understanding the origin of

enantioselectivity through differences in energetic pathways between enantiomers. This is a first attempt to quantify free energy difference using the FEP method to study CALB enantioselectivity. In the future this could become a very interesting tool for the pharmaceutical industry.

Acknowledgements

This study was supported by the French ANR (National Research Agency) through the EXPENANTIO project (program CP2D). CINES and IDRIS from GENCI (Grand Equipement National de Calcul Intensif) are acknowledged for giving access to super-computing facilities. The manuscript was corrected by a native English scientific translator (<http://traduction.lefevere-laotide.net>).

References

- [1] R.D. Schmid, R. Verger, *Angew. Chem. Int. Ed.* 37 (1998) 1608–1633.
- [2] M.T. Reetz, *Curr. Opin. Chem. Biol.* 6 (2002) 145–150.
- [3] R.J. Kazlauskas, A.N.E. Weissfloch, A.T. Rappaport, L.A. Cuccia, *J. Org. Chem.* 56 (1991) 2656–2665.
- [4] D. Rotticci, F. Haefner, C. Orrenius, T. Norin, K. Hult, *J. Mol. Catal. B: Enzym.* 5 (1998) 267–272.
- [5] M.-P. Bousquet-Dubouch, M. Graber, N. Sousa, S. Lamare, M.-D. Legoy, *Biochim. Biophys. Acta* 26 (2001) 90–99.
- [6] F. Haefner, T. Norin, K. Hult, *Biophys. J.* 74 (1998) 1251–1262.
- [7] C. Orrenius, F. Haefner, D. Rotticci, N. Öhrner, T. Norin, K. Hult, *Biocatal. Biotransform.* 16 (1998) 1–15.
- [8] F. Haefner, T. Norin, *Chem. Pharm. Bull.* 47 (1999) 591–600.
- [9] D. Rotticci, J.C. Rotticci-Mulder, S. Denman, T. Norin, K. Hult, *ChemBioChem* 2 (2001) 766–770.
- [10] J. Uppenberg, N. Öhrner, M. Norin, K. Hult, G.J. Kleywegt, S. Patkar, V. Waagen, T. Anthonsen, J.T. Alwyn, *Biochemistry* 34 (1995) 16581–16838.
- [11] C.H. Kwon, D.Y. Shin, J.H. Lee, K.S. Wook, K.J. Won, *J. Microbiol. Biotechnol.* 17 (2007) 1098–1105.
- [12] C. Orrenius, C. van Heusden, J. van Ruiten, P. Overbeeke, H. Kierkels, J.A. Duine, A.J. Jongejan, *Protein Eng.* 11 (1998) 1147–1153.
- [13] S. Raza, L. Fransson, K. Hult, *Protein Sci.* 10 (2001) 329–338.
- [14] A. Warshel, G. Naray-Szabo, F. Sussman, J.K. Hwang, *Biochemistry* 28 (1989) 3629–3637.
- [15] H. Eyring, M. Polanyi, *Z. Phys. Chem.* 12 (1931) 279.
- [16] R.S. Philipps, *Enzyme Microb. Technol.* 14 (1992) 417–419.
- [17] R.S. Philipps, *Trends Biotechnol.* 14 (1996) 13–16.
- [18] A.P.L. Overbeeke, C. Orrenius, J.A. Jongejan, J.A. Duine, *Chem. Phys. Lipids* 93 (1998) 81–93.
- [19] J. Ottosson, L. Fransson, K. Hult, *Protein Sci.* 11 (2002) 1462–1471.
- [20] V. Leonard-Nevers, Z. Marton, S. Lamare, K. Hult, M. Graber, *J. Mol. Catal. B: Enzym.* 59 (2009) 90–95.
- [21] P. Braiuca, K. Lorena, V. Ferrario, C. Ebert, L. Gardossi, *Adv. Synth. Catal.* 351 (2009) 1293–1302.
- [22] T. Schulz, J. Pleiss, R.D. Schmid, *Protein Sci.* 9 (2000) 1053–1062.
- [23] P.B. Juhl, P. Trodler, S. Tyagi, J. Pleiss, *BMC Struct. Biol.* 9 (2009) 39–55.
- [24] E. García-Urdiales, N. Ríos-Lombardía, J. Mangas-Sánchez, V. Gotor-Fernández, V. Gotor, *ChemBioChem* 10 (2009) 1830–1838.

- [25] Y. Zhou, Ph.D. thesis, Computational Study of Enzyme Enantioselectivity, University of Delft, NL, 2006.
- [26] W.L. Jorgensen, *Acc. Chem. Res.* 22 (1989) 184–189.
- [27] J. Aqvist, *J. Phys. Chem.* 94 (1990) 8021–8024.
- [28] T. Simonson, G. Archontis, M. Karplus, *Acc. Chem. Res.* 35 (2002) 430–437.
- [29] D.Q. McDonald, W.C. Still, *J. Comput. Chem.* 118 (1996) 2073–2077.
- [30] C. de Graaf, C. Oostenbrink, P.H.J. Keizers, B.M.A. van Vugt-Lussenburg, J.N.M. Commandeur, N.P.E. Vermeulen, *Eur. Biophys. J.* 36 (2007) 589–599.
- [31] A. Aleksandrov, D. Thompson, T. Simonson, *J. Mol. Recognit.* 23 (2010) 117–127.
- [32] C.D. Christ, A.E. Mark, W.F. Van Gunsteren, *J. Comput. Chem.* 31 (2009) 1569–1582.
- [33] A. Di Nola, A.T. Brünger, *J. Comput. Chem.* 19 (1998) 1229–1240.
- [34] G. Colombo, S. Toba, J.K.M. Merz, *J. Am. Chem. Soc.* 121 (1999) 3486–3493.
- [35] C.-S. Chen, Y. Fujimoto, G. Girdaukas, C.J. Sih, *J. Am. Chem. Soc.* 104 (1982) 7294–7299.
- [36] V. Leonard, S. Lamare, M.-D. Legoy, M. Graber, *J. Mol. Catal. B: Enzym.* 32 (2004) 53–59.
- [37] J. Uppenberg, M.T. Hansen, S. Patkar, T.A. Jones, *Structure* 2 (1994) 293–308.
- [38] J. Nyhln, B. Martn-Matute, A.G. Sandstrm, M. Bocola, J.-E. Bckvall, *ChemBioChem* 9 (2008) 1968–1974.
- [39] B.R. Brooks, C.L. Brooks III, A.D. Mackerell, L. Nilsson Jr., R.J. Petrella, B. Roux, Y. Won, G. Archontis, B. C. S. Boresch, A. Caffisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoseck, W. Im, K. Kuczera, T. Lazaridis, J. MA, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaeffer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, *J. Comput. Chem.* 30 (2009) 1545–1614.
- [40] N. Otte, M. Bocola, W. Thiel, *J. Comput. Chem.* 30 (2009) 154–162.
- [41] R. Mutyala, R.N. Reddy, M. Sumakanth, P. Reddanna, M.R. Reddy, *J. Comput. Chem.* 28 (2007) 932–937.
- [42] J. Wang, Y. Deng, B. Roux, *Biophys. J.* 91 (2006) 2798–2814.
- [43] D. Allouche, J. Parello, Y. Sanejouand, *J. Mol. Biol.* 285 (1999) 857–873.
- [44] P. Trodler, J. Pleiss, *BMC Struct. Biol.* 8 (2008) 1–10.
- [45] R. McCabe, A. Rodger, A. Taylor, *Enzyme Microb. Technol.* 36 (2005) 70–74.
- [46] R.W. Zwanzig, *J. Chem. Phys.* 22 (1954) 1420–1426.
- [47] C. Hedfors, K. Hult, M. Martinelle, *J. Mol. Catal. B: Enzym.* 66 (2010) 120–123.
- [48] A.O. Magnusson, M. Takwa, A. Hamberg, K. Hult, *Angew. Chem. Int. Ed.* 44 (2005) 4582–4585.
- [49] M. Skjøt, L. De Maria, R. Chatterjee, A. Svendsen, S.A. Patkar, P.R. Østergaard, J. Brask, *ChemBioChem* 10 (2009) 520–527.
- [50] C. Zhao, D.A. Caplan, S.Y. Noskov, *J. Chem. Theory Comput.* 6 (2010) 1900–1914.
- [51] R. Murphy, D. Philipp, R. Friesner, *J. Comput. Chem.* 21 (16) (2000) 1442–1457.
- [52] Y. Zhang, J. Kua, J. McCammon, *J. Am. Chem. Soc.* 124 (35) (2002) 10572–10577.
- [53] C. Hensen, J. Hermann, K. Nam, S. Ma, J. Gao, H. Holtje, *J. Med. Chem.* 47 (27) (2004) 6673–6680.