Journal of Molecular Catalysis B: Enzymatic 59 (2009) 96–105

Contents lists available at ScienceDirect

Journal of Molecular Catalysis B: Enzymatic

journal homepage: www.elsevier.com/locate/molcatb

A molecular modelling study to rationalize the regioselectivity in acylation of flavonoid glycosides catalyzed by *Candida antarctica* lipase B

Eduardo B. De Oliveira[a], Catherine Humeau[a,*], Latifa Chebil[a], Elaine R. Maia[b], François Dehez[c], Bernard Maigret[d], Mohamed Ghoul[a], Jean-Marc Engasser[a]

[a] *Laboratoire Biocatalyse Bioprocédés (LBB), ENSAIA-INPL, Nancy Université, 2 av. de la Forêt d'Haye, 54500, Vandoeuvre-lès-Nancy, France*
[b] *Laboratorio de Estudos Estruturais Moleculares (LEEM), Instituto de Quimica, Universidade de Brasilia, CP 4478, 70904-970, Brasília-DF, Brazil*
[c] *Equipe de Dynamique des Assemblages Membranaires (EDAM), CNRS, Université Henri Poincaré, 7565, 54500, Vandoeuvre-lès-Nancy, France*
[d] *Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA), CNRS, Université Henri Poincaré, BP 239, 54506, Vandoeuvre-lès-Nancy, France*

ARTICLE INFO

Article history:
Received 2 December 2008
Received in revised form 23 January 2009
Accepted 27 January 2009
Available online 12 February 2009

Keywords:
Candida antarctica lipase B
Docking
Flavonoid acylation
Molecular modelling
Regioselectivity

ABSTRACT

The regioselective behaviour of the *Candida antarctica* lipase B (CALB) towards two flavonoid glycosides, rutin and isoquercitrin, in the acetylation reaction was investigated through molecular modelling. A protocol constituted by a Monte Carlo-based docking procedure and classical force fields calculations was applied to find probable binding modes of the substrates inside the catalytic cavity and optimize the corresponding complexes. The analysis of these complexes allowed identifying productive ones (that means, those able to lead to the formation of the ester product) according to three parameters: (1) protein distortion; (2) stability of hydrogen bond interactions with the oxyanion hole residues; (3) localization of hydroxyl groups with regard to the region comprised between the catalytic histidine and serine residues. Results showed that the aglycon part of both rutin and isoquercitrin was localized at the entrance of the binding pocket, stabilized by hydrogen bond and hydrophobic interactions. The sugar part of the flavonoids was placed close to the pocket bottom. In particular, only the primary 6″-OH of the isoquercitrin glucose and the secondary 4‴-OH of the rutin rhamnose were expected to be acetylated, as they were the only ones to stabilize simultaneously near to the catalytic histidine and the acetate bound to the catalytic serine. These findings are in accordance with experimental data and give a suitable explanation, at an atomic level, of the regioselectivity of CALB in the flavonoid glycosides acetylation.

1. Introduction

Flavonoids are a broad class of low-molecular weight polyphenols, produced as secondary metabolites by the greater part of vegetal species. Chemically, they are benzo-γ-pyrone derivatives, with a phenyl-type substituent on the C2 or C3 position. Frequently, but not obligatory, they are found in nature covalently attached to a glycoside part. These compounds exhibit a wide range of biological, pharmacological and medicinal properties, such as anti-oxidant, anti-allergenic, anti-viral and cancer preventive effects, making them very attractive for use as functional ingredient in cosmetic, pharmaceutical and food products [1,2]. However, most of them present a poor solubility in both polar and non-polar media [3], which strongly restricts their incorporation in many formulations. A solution to improve the hydrophobic nature (lipophilization) of flavonoids consists in their acylation [4], which can be accomplished by chemical or enzymatic processes. With the chemical acylation, many of the hydroxyl groups present on the flavan skeleton and on the sugar moieties can be esterified [5], which can lead to the acylation of some phenol groups that are directly implicated in the beneficial properties of these molecules [6].

On the contrary, using enzymes, such as lipases, only a single or a few hydroxyl groups were reported to be esterified [7]. In this case the number and position of the reactive hydroxyls depend on the structures of both the flavonoid and the lipase. With *Candida antarctica* lipase B (CALB, EC 3.1.1.3), the most frequently used enzyme to perform esterification reactions in non-aqueous media [4], the diglycosylated flavonoid rutin (quercetin-3-O-glucose-rhamnose, Fig. 1), which carries 10 hydroxyl groups, was found to be acylated only on the secondary 4‴-OH of rhamnose. This strict regioselectivity was reported for a wide range of acyl donors, such as fatty acids with different chain length and their vinyl or ethyl derivatives [8–11]. When CALB was used for the acylation of the monoglycosylated flavonoid isoquercitrin (quercetin-3-O-glucose, Fig. 1), which exhibits eight hydroxyl groups, it is the primary 6″-OH of the glucose moiety that was esterified. With vinyl acetate as the acyl donor,

* Corresponding author. Tel.: +33 3 83 59 57 84; fax: +33 3 83 59 57 78.
E-mail addresses: eduardobasilio@yahoo.com.br (E.B. De Oliveira), catherine.humeau@ensaia.inpl-nancy.fr (C. Humeau).

1381-1177/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.molcatb.2009.01.011

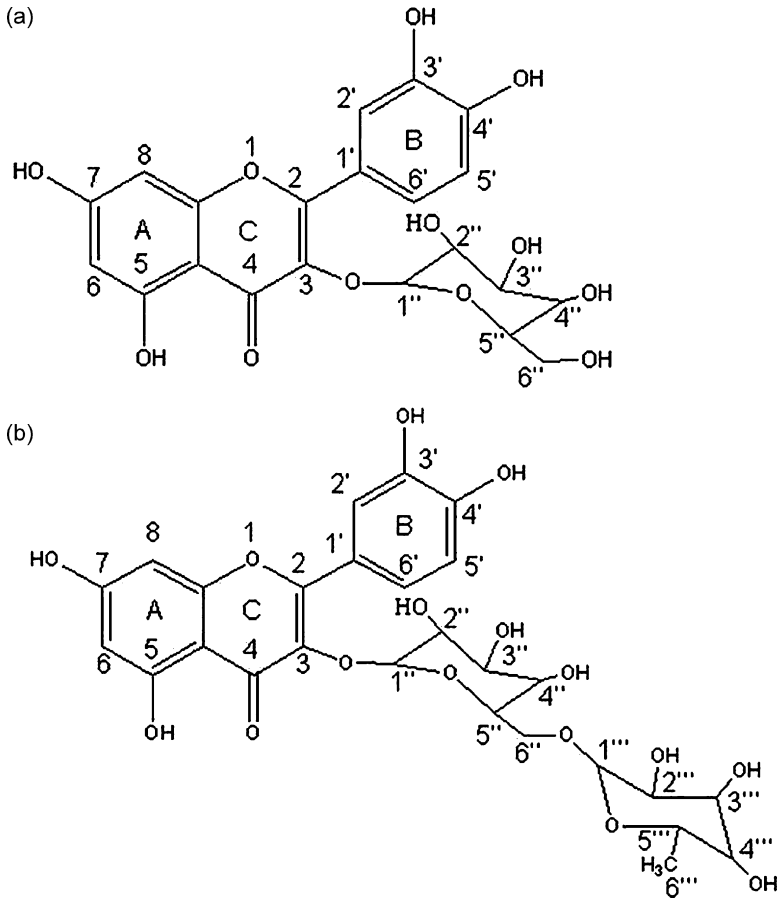

Fig. 1. Chemical structures of (a) isoquercitrin and (b) rutin, with the systematic numbering of carbon atoms and nomenclature of aglycon rings.

the obtained isoquercitrin-6″-acetate could be further esterified to produce isoquercitrin-3″,6″-diacetate [12]. With other acyl donors, the formation of isoquercitrin-6″-monoester as the sole product has been reported [13–15]. In these studies no explanation on a molecular level was provided to the observed regioselectivity of the CALB-catalyzed acylation of either rutin or isoquercitrin.

Recently, computer-based modelling methodology has been increasingly used to understand the mechanism and selectivity of enzyme-catalyzed reactions and to suggest how to change the selectivity of these reactions by modification of the enzyme or the substrate structure [16,17]. With lipases, many studies address the use of molecular modelling tools to furnish a rational explanation for their enantioselectivity towards a large set of chiral substrates, in both hydrolysis and esterification reactions [18–23]. Also, some works deal with the regioselectivity of lipase-catalyzed acylation of poly-hydroxylated compounds, such as resorcinarenes [24], glycosides [25–27] and prostaglandins [28]. To our knowledge, there is no published work on the application of computer-based modelling to describe the molecular interactions involved in flavonoids recognition by lipases. Thus, the present study aims to provide a molecular-level explanation for the observed regioselectivity of CALB towards the 6″-OH and 4‴-OH groups of isoquercitrin and rutin, respectively. A combined docking, molecular mechanics and molecular dynamics approach was used to study the positions, the orientations, the interactions of these two compounds in the active site of CALB and the accessibility of their hydroxyl groups to the catalytic residues. All these factors are important to determine if a given substrate binding mode is productive or not, thus allowing to elucidate the regioselectivity in the reaction.

2. Methodology

All simulations were carried out on a bi-processor AMD Dual Core 280 2.4 GHz. Molecular mechanics and molecular dynamics calculations and docking simulations were performed with CHARMm force field [29], adopting a 12 Å non-bound spherical cut-off, using the program-package Discovery Studio version 1.7 (Accelrys Inc.). In molecular dynamics calculations, the isothermal-isochoric ensemble (NVT) was used, under periodic boundary conditions (PBC). A time step of 1 fs and the Verlet-Leapfrog algorithm were systematically applied for the integration of the equations of motion. Connolly surfaces [30] were calculated using Insight II LS (Accelrys Inc.). The study comprises of the following steps: preparation of starting structures; modelling of the acetyl–enzyme (docking target); docking of the flavonoid substrate and subsequent scoring of generated poses; optimization and structural analysis of the best-scored poses; verification of the reliability of the final models.

2.1. Target preparation

The CALB crystal structure was taken from the Protein Data Bank (PDB entry: 1LBS; resolution: 2.60 Å) [31]. The complex has six chains per unit cell of dimensions $a = 229.5$ Å, $b = 95.6$ Å, $c = 86.8$ Å; $\alpha = \beta = \gamma = 90.0^{\circ}$. Each chain system is composed of 317 amino acids, one ethyl-hexylphosphonate inhibitor (HEE) covalently attached to the Ser105 residue, a dimer of *N*-acetyl-glucosamine covalently bound to the Asn74 residue and 92 adsorbed water molecules. Only one protein chain and its water molecules were kept and used as initial atomic coordinates set for the simulations; the oth-

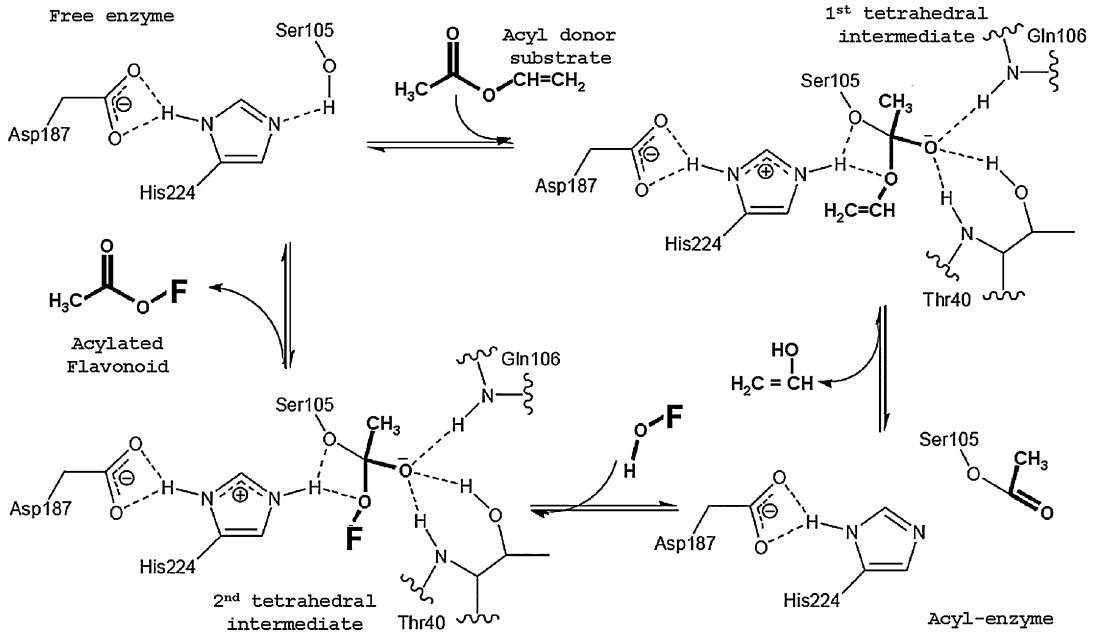

Fig. 2. The reaction mechanism of the CALB-catalyzed acylation of a flavonoid, evidencing the two tetrahedral intermediates and the acyl–enzyme complex. In this study, the acyl donor substrate is vinyl acetate, F–OH is the flavonoid to be acetylated (docked ligand) and the docking target is the acetyl–CALB.

ers were erased. The inhibitor and the two *N*-acetyl-glucosamines were also removed. All hydrogen atoms were added in their theoretical positions. The catalytic residues Asp187 and His224 were assigned as deprotonated. Next, this new system was resolvated with 6145 water molecules in a box with previous optimized dimensions of $a = 64$ Å, $b = 59$ Å, $c = 63$ Å; $\alpha = \beta = \gamma = 90.0^{\circ}$. A structure relaxation procedure was executed. In this procedure, constraints and restraints were added then progressively removed [32], in order to preserve the crystalline organization of protein atoms: (a) only protein hydrogen atoms and water molecules were allowed to move, keeping all other heavy atoms fixed, in order to eliminate initial strains; (b) the protein backbone was kept fixed and harmonic restraints were imposed on the protein side chains; (c) harmonic restraints were applied only on the protein backbone, allowing the side chains free to move; (d) finally, the whole system was minimized without any constraints. The first step was done by performing 1000 Steepest Descent iterations. The steps (b)–(d) were performed using the Conjugate Gradient method until the root mean square deviation (*rmsd*) gradient fell below 0.01, 0.001 and 0.0001 kcal mol^{-1} Å^{-1}, respectively. In order to equilibrate the optimized system, a standard dynamic trajectory was carried out at 300 K, for 1 ns. The protein structure in the last frame was taken for subsequent modelling steps.

The binding pocket of CALB consists of an elliptical and steep funnel whose walls predominantly constitute by hydrophobic aliphatic amino acid residues [42]. The catalytic machinery is located at the bottom of the pocket, a more hydrophilic region, and constitutes by the triad Ser105. . .His224. . .Asp187 [31]. The well-known reaction mechanism is based on the acylation and deacylation of the Ser105 residue and involves two tetrahedral intermediates (Fig. 2). The first one results from the nucleophilic attack of the catalytic serine on the acylating substrate and originates the acyl–enzyme complex. The second tetrahedral intermediate derives from the nucleophilic attack of the acyl acceptor substrate on the acyl–enzyme and leads to the release of the ester product [31,42]. Therefore, the Ser105 side chain hydroxyl was replaced by acetate in order to mimic the acyl–enzyme complex [24,27]. To allow the adjustment of the acetate, further 1000 steps of Conjugate Gradient energy minimization were performed with a fix constraint applied to the protein backbone. This optimized acetyl–CALB complex was taken as target in the docking procedure.

2.2. Docking and scoring

Isoquercitrin and rutin were non-covalently docked against the acetyl–CALB. This approach allows exploring all the different orientations of one ligand within the catalytic pocket in one single simulation, contrarily to another approach which consists in docking the ligands covalently attached to the acyl–enzyme, mimicking the second tetrahedral intermediate. In this case, all the potentially reacting functions must be separately modelled. Docking simulations were performed with LigandFit [33], which combines a shape comparison filter with a Monte Carlo conformational search algorithm. Moreover, LigandFit owns an internal scoring function (DOCKSCORE) that estimates iteratively the interaction energy (E) between each docked pose and the receptor, taken as the sum of van der Waals and electrostatic energies, represented in the equation below:

$$E = \sum_{i,j} \varepsilon_{ij} \left[2\left(\frac{r_{ij}^{*}}{r_{ij}}\right)^{9} - 3\left(\frac{r_{ij}^{*}}{r_{ij}}\right)^{6} \right] + \frac{332.0716}{\varepsilon} \sum_{i,j} \frac{q_i q_j}{r_{ij}}$$

where ε is the dielectric constant, and (r_i^* is the van der Waals radius and ε_i is the energy parameter of the ith ligand atom, while r_j^* and ε_j are similarly parameters for the jth protein atom). The distance between the *i-th* ligand atom and the *j*th protein atom is represented by r_{ij} and their respective charges (in atomic units) are q_i and q_j. Thus, in addition to the geometric fit, the docking procedure assures to keep only the ligand poses having favourable energy of interaction with the protein. For each substrate 15,000 Monte Carlo trials were performed and only the 50 most favourable poses were retained. Docked poses were scored by six different scoring functions: LigScore 1 and 2 [34], PLP 1 and 2 [35], PMF [36] and Jain [37], followed by a consensus scoring. In fact, the consensus combines data obtained from different scores to compensate for inaccuracy from individual scoring functions and has been demonstrated to improve the probability of finding correct solutions [38]. The three

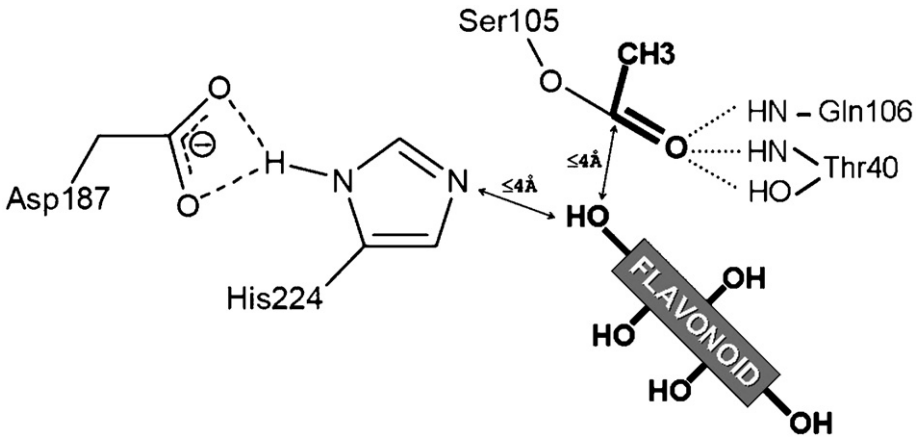

Fig. 3. Scheme of structural requirements that flavonoid hydroxyl groups must satisfy in the models in order to be considered as reactive.

best-scored poses according to the consensus scoring were selected and the corresponding complexes were optimized and further analyzed.

2.3. Post-docking optimization

The docking/scoring procedure of LigandFit is based on both shape fitting and energetic stability of the protein–ligand system, guaranteeing that only viable complexes are saved [33]. As this docking algorithm treats the protein as a rigid body, enzyme–substrate complexes were submitted to a careful post-docking optimization, in order to take into account a potential induced fit effect; that means, little displacements in the protein structure due to the presence of the ligand [39]. In the selected complexes, the acetate oxygen was orientated to the oxyanion hole, so that it can form hydrogen bonds interactions with the oxyanion hole residues (Thr40 and Gln106), mimicking the oxyanion orientation in the second tetrahedral intermediate [19]. Then three consecutive energy minimizations were performed: (a) a fix constraint was applied on the protein backbone and all the remaining parts of the systems were tethered with a harmonic restraint; (b) the protein backbone was kept fix, the side chains were kept tethered and the docked ligand was free to move; (c) only the backbone was kept fix, with side chains and the ligands free to move. The Conjugate Gradient algorithm was used in the three steps, with root mean square deviation (*rmsd*) gradients of 0.01, 0.001 and 0.0001 kcal mol^{-1} $Å^{-1}$, respectively.

2.4. Analysis of the optimized complexes

In this step, in order to determine if optimized complexes corresponded to productive binding modes, they were analyzed by combining three criteria (Fig. 3) used in previous studies dealing with molecular modelling of lipases:

The first criterion is the *rmsd* between the protein structure in the complex and its initial crystal structure. Even though some enzymes may undergo relatively large conformational changes upon the binding of substrates (induced fit effect), characterized by *rmsd* values up to 6 Å [39], we chose a *rmsd* value around 3 Å for CALB as the limit between reactive and unreactive complexes, as found by Vallikivi et al. [28] for prostaglandins acetylation.

The second one is that complexes must conserve at least two of the three hydrogen bonds between the acetate oxygen (namely Ace:O) and the oxyanion hole residuesThr40 and Gln106, postulated to stabilize the oxyanion in the second tetrahedral intermediate [31].

The third one is the distance between the flavonoid hydroxyl group and the acetate sp^2 carbon (namely Ace:C) and the Nε atom of His224 (namely His224:Nε). We assumed that, in order to be considered as reactive, a hydroxyl group must be simultaneously placed at 4 Å maximum from the Ace:C and His224:Nε atoms, as stated by Palloci et al. [27] for acetylation of trytilated glycosides catalyzed by *Candida rugosa* lipase.

The breach of any of these conditions would be indicative that the complex is not able to lead to the formation of the second tetrahedral intermediate; that means, the binding mode is non-productive.

2.5. Molecular dynamics simulations on the productive complexes

After the identification of the productive binding modes, further 100 ps MD trajectory was executed on the corresponding complexes, maintaining the backbone fix, in order to evaluate the dynamics of the docked substrates. In particular, the stability of the essential hydrogen bond interactions and the critical distances for reactivity were verified. The trajectories were divided in 5 ps of heating from 50 to 300 K, 5 ps of equilibration at 300 K and 90 ps of production at 300 K. Frames were registered every 500 fs. Only production phases were considered for analysis.

3. Results

3.1. Target preparation and analysis

Prior to docking simulations, the CALB structure was submitted to a careful structure relaxation protocol [32] to optimize the system without harming the disposition of the atoms from the crystal structure. In order to monitor eventual distortions in the lipase structure in the procedure, the root mean square deviation (*rmsd*) was computed all along the equilibration step (MD trajectory of 1 ns at 300 K), considering the superimposition of all the heavy atoms. The relaxation procedure was performed in presence of explicit water molecules, rather than to choose a particular solvent. This decision was based on two arguments: first, even though an enzymatic reaction is carried out in organic media, the presence of a certain amount of water molecules in the reaction medium contributes to the structural integrity of the enzyme and thus to its optimal activity [40]. Secondly, Trodler and Pleiss [41] showed, using molecular dynamics simulations with explicit solvent molecules, that CALB is a little more flexible in water than in organic solvents. In that way, if CALB structure is well conserved after the simulations in water, it is reasonable to assume that it

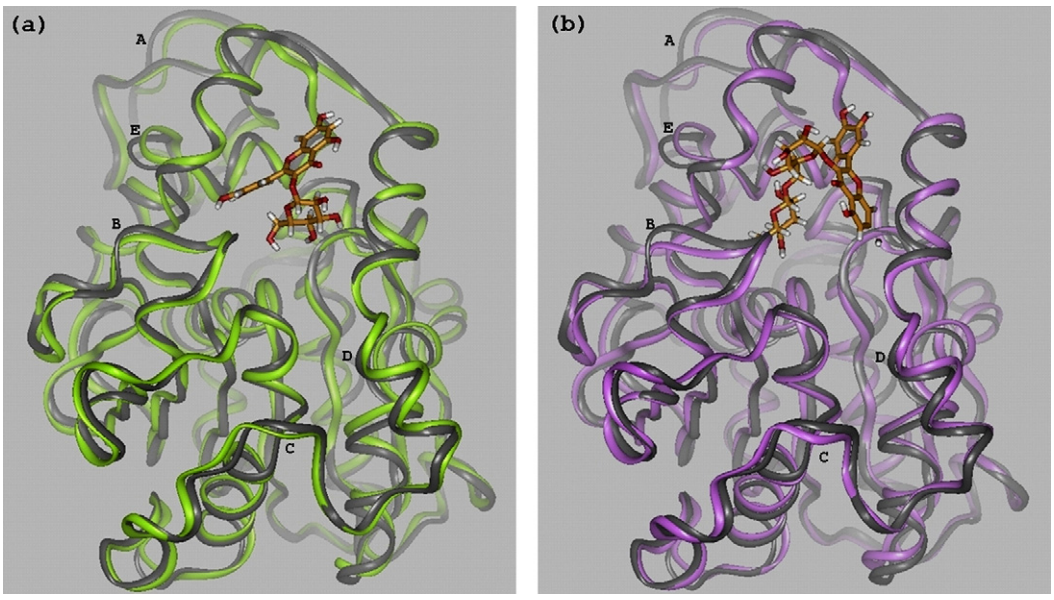

Fig. 4. Ribbons showing CALB crystal structure superimposed to its structure after the modelling procedure complexed with (a) isoquercitrin and (b) rutin. Docked flavonoids are represented in sticks. The crystal structure is colored in gray in both cases; that after the modelling procedure is colored in green in (a) and in violet in (b). Main chain differences are found in the sequences: (A) Cys311-Gly313, corresponding to the loop between the sheets β8 and β9; (B) Val190-Pro192, contained in the loop that connects the sheets β6 and β7; (C) Asn259-Pro262, contained in the loop that connects the helices α9 and α10; (D) Asn85-Ile87, contained in the helix α3; (E) Thr158-Ala162, contained in the loop between the helices α6 and α7. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

should be also preserved in organic solvents. The *rmsd* values (all heavy atoms) were stabilized at about 1.1 Å, after approximately 600 ps of equilibration. This result indicates a high-structural conservation, evidencing a successful relaxation procedure. Then, starting from the relaxed CALB structure, the acetyl–enzyme intermediate was manually built and optimized. When superimposing it on the relaxed CALB structure, no structural changes were observed, as reflected by the negligible *rmsd* value (all protein heavy atoms). Moreover, the hydrogen bond interaction between the side chains of Asp187 and His224 was maintained and the distance between the Ser105:Oγ and His224:Nε atoms remained at 3.2 Å. This indicates that the initial disposition of the catalytic triad residues was not damaged. Therefore, this acetyl–CALB structure was taken as target for the docking of the two flavonoid substrates.

3.2. Docking results

In the case of isoquercitrin the 50 complexes presented the substrate in two orientations: the sugar part or the aglycon part (ring A) orientated towards the enzyme core. In the case of rutin, all the retained complexes presented the substrate with its sugar part orientated towards the interior of the cavity, with only small conformational differences among them. For both flavonoids, the three best-scored complexes, according to the consensus scoring, were representative of these trends and thus were optimized and analyzed in order to determine if they corresponded to productive binding modes or not, according to three criteria: protein distortion; content of essential interactions between the acetate and the oxyanion hole residues; proximity of the hydroxyl groups to the His224:Nε and Ace:C atoms.

Although some protein conformational rearrangements may be needed for a better ligand affinity, if the protein distortion is too high the orientation of some essential residues may be disturbed, leading to a loss of activity. Vallikivi et al. [28] showed that, in complexes between CALB and prostaglandins (also polyhydroxylated compounds), *rmsd* values higher than 3.0 Å were experimentally well correlated with unreactive enzyme–substrate complexes. Therefore, we also adopted the maximal value of 3.0 Å as a first criterion for a productive complex. In our case, for all the retained CALB–flavonoid complexes, the superimposition of the CALB backbone to the protein crystal structure showed only minor displacements (*rmsd* values about 0.9 Å). Minimal chain distortions were found mainly in loop regions, which are usually the more mobile regions in proteins [43]. Sequences presenting some distortion due to the simulations are specified in Fig. 4. When all heavy atoms were superimposed, *rmsd* values went up to about 3.0 Å, as side chains were free to move in the last step of the optimization procedure. In spite of this, the cavity shape and the general visual appearance of the protein structure were not substantially altered.

To achieve the reaction, the Ace:O atom must form hydrogen bonds interactions in the oxyanion hole, with the backbone –NH of the residues Thr40 and Gln106 and the side chain –OH of Thr40. In fact, these hydrogen bonds enhance the partial positive charge on the Ace:C, enabling the nucleophilic attack of the acyl acceptor substrate. Therefore, as previously mentioned, another criterion to consider a complex as productive was that at least two of these three hydrogen bonds must be conserved after the post-docking optimization. Hydrogen bond distances and angles values are given in Table 1.

Another imperative criterion that a complex must obey in order to be considered as productive is the proximity of at least one flavonoid hydroxyl group from catalytic residues. More specifically, it must be positioned simultaneously close to His224:Nε to transfer its proton and enhance its nucleophilicity and close to Ace:C to execute the nucleophilic attack itself. We decided to adopt for these distances a maximal value of 4 Å, as did by Palocci et al. [27]. The distance values obtained for the most buried substrate hydroxyl groups in each complex are noted in Table 2.

3.2.1. Docking of isoquercitrin

Among the three retained complexes for isoquercitrin (namely I1, I2 and I3; Fig. 5), two possible orientations were found: the sugar part or the aglycon part (ring A) orientated towards the enzyme core. In I2, the isoquercitrin sugar moiety is orientated towards the exterior of the cavity and the aglycon moiety is orientated inwards. The hydroxyl groups closer to the active site are the 5-OH and 7-OH phenolic groups, with respective distances of 5.6 Å and 5.2 Å from

Table 1
Hydrogen bond distances and angles involving the substrates in the three best-scored complexes after optimization.

Ligand	Complex	Donor	Acceptor	Distance (Å)	Angle (°)
Isoquercitrin	I1	Isoq:HO7	Val149:O	2.1	137.2
		Isoq:HO5	Isoq:O4	1.8	149.7
		Isoq:HO4′	Ile189:O	1.8	164.2
		Thr40:HN	Ace:O	1.9	125.5
		Thr40:HG1	Ace:O	1.4	123.4
		Gln106:HN	Ace:O	2.2	8.5
	I2	Thr40:HN	Ace:O	2.1	169.7
	I3	Thr40:HG1	Ace:O	1.5	136.5
		Gln106:HN	Ace:O	2.1	97.0
Rutin	R1	Rut:HO7	Leu278:O	2.1	146.3
		Rut:HO4′	Val149:O	1.9	149.6
		Rut:HO5	Rut:O4	2.0	159.3
		Rut:HO3‴	Gln157:OE1	1.9	126.2
		Thr40:HN	Ace:O	1.9	121.6
		Thr40:HG1	Ace:O	1.7	152.6
		Gln106:HN	Ace:O	2.2	15.7
	R2	Rut:HO4′	Val149:O	2.0	148.7
		Rut:HO7	Leu278:O	2.1	137.6
		Thr40:HN	Ace:O	2.4	110.2
	R3	Rut:HO4″	Rut:O6″	2.1	143.7
		Rut:HO4‴	Gln157:OE1	2.0	150.4
		Thr40:HN	Ace:O	2.4	119.9
		Thr40:HG1	Ace:O	1.8	132.3
		Gln106:HN	Ace:O	1.9	14.7

Ace:C and 8.3 Å and 7.1 Å from the His224:Nε. In addition, in this complex the acetate forms only one hydrogen bond with the oxyanion hole residues (Thr40:NH. . .Ace:O). This may be attributed to a hydrophobic contact between the acetate methyl chain and the C6 of the ring A of the substrate (distance of 3.4 Å), turning the acetate oxygen away from the Thr40 and Gln106 residues. For these reasons, I2 cannot represent a productive binding mode and was rejected. In I1 and I3, the aglycon moiety lodges at the entrance of the binding pocket, whereas the sugar is oriented inwards. In I3 the acetate interacts with the oxyanion hole residues through two hydrogen bonds (Thr40:OH. . .Ace:O and Gln106:NH. . .AceO). However, the closest substrate hydroxyl group to the catalytic residues (6″-OH) is placed at 5.8 Å from His224:Nε and 4.2 Å from Ace:C, which exceeds the maximal values admitted as to allow the proton transfer and the nucleophilic attack. Therefore, I3 was also discarded. In the complex I1, the acetate conserved the three possible hydrogen bonds interactions with the Thr40 and Gln106 residues. In addition, the nearest hydroxyl group to the catalytic residues (6″-OH) is positioned simultaneously at 4.0 Å from His224:Nε and 3.5 Å from Ace:C, fulfilling the distance criteria to react. Thus, I1 may be considered as a productive binding mode.

More specifically, in I1 isoquercitrin docked within the target with the aglycon moiety anchored at the entrance of the binding pocket, where it is stabilized by two types of interactions: hydrogen bond interactions involving the phenolic groups (7-OH. . .Val149:O and 4′-OH. . .Ile189:O) and nonpolar interactions between the aromatic cycles and hydrophobic aliphatic residues, which abound in this region of the enzyme (Leu140, Ala141, Leu144, Val 149, Val154, Ile189 and Val 286). The glucose moiety is placed within the cavity. The three most buried hydroxyl groups are the 3″-OH, 4″-OH and 6″-OH, but only this last is at a distance inferior to 4.0 Å from the Ace:C and His224:Nε atoms (Table 1). In fact, in this primary alcohol function the –CH_2– alkyl group proceeds as an "extension piece" that allows the –OH to draw near the catalytic residues. Therefore, the molecular model hints that only the 6″-OH group of isoquercitrin is expected to react. These findings agree with exper-

Table 2
Distances[a] of buried hydroxyl groups of the docked flavonoids from the His224:Nε and the acetate carbonyl carbon (Ace:C) atoms in the three best-scored complexes after optimization.

Ligand	Complex	Buried –OH groups	Distance from Ace:C (Å)	Distance from His224:Nε (Å)
Isoquercitrin	I1	3″-OH	7.4	8.7
		4″-OH	4.6	6.2
		6″-OH	3.5	4.0
	I2	5-OH	5.6	8.3
		7-OH	5.2	7.1
	I3	3″-OH	8.6	8.7
		4″-OH	5.9	6.5
		6″-OH	4.2	5.9
Rutin	R1	2‴-OH	6.3	8.2
		3‴-OH	3.9	7.0
		4‴-OH	2.9	3.9
	R2	2‴-OH	6.7	7.8
		3‴-OH	5.9	8.5
		4‴-OH	4.1	7.8
	R3	2‴-OH	7.1	12.5
		3‴-OH	4.7	10.1
		4‴-OH	3.7	9.4

[a] Distance measured considering the hydroxyl oxygen atoms.

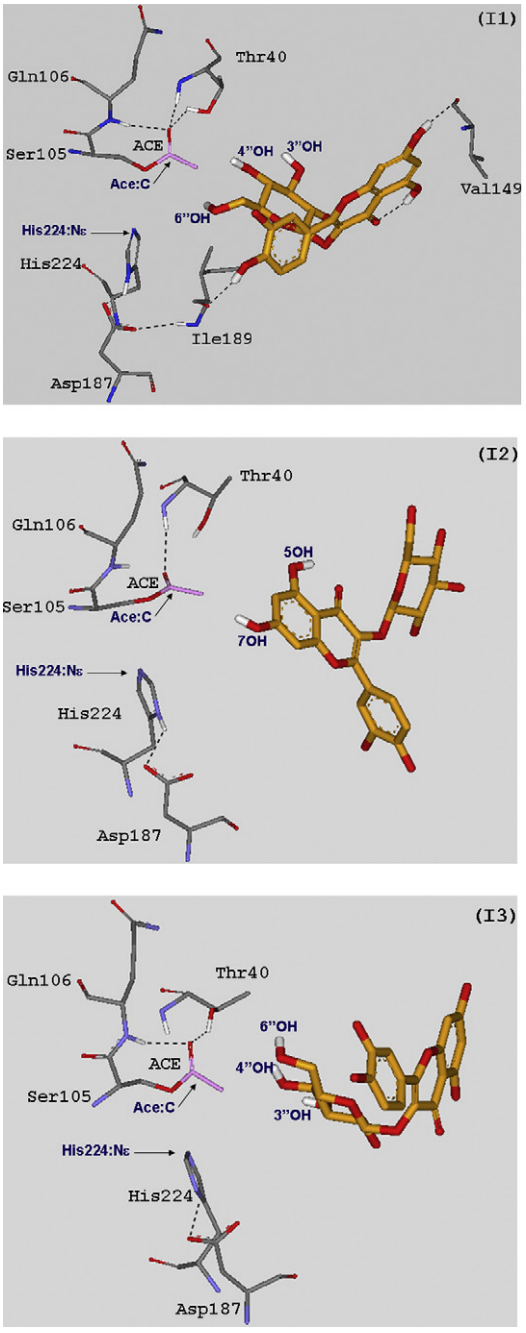

Fig. 5. Orientations and hydrogen bonds interactions (dashed lines) of the substrates in the three best-scored acetyl–CALB isoquercitrin complexes (I1, I2 and I3). For clarity, only the residues useful for discussion are shown and non-interacting hydrogen atoms are hidden.

imental studies showing that isoquercitrin-6″-monoester was the sole product directly obtained in the CALB-catalyzed acetylation of isoquercitrin [12]. The additional formation of isoquercitrin-3″,6″-diacetate was already reported [44], but further investigations indicated that it is likely to result from the acetylation of the initially synthesized isoquercitrin-6″-acetate, in a sequential way [12]. Several other authors also showed that CALB catalyzes the monoacylation of isoquercitrin with phenolic acids as cinnamic acid [13], *p*-coumaric acid [14] and phenyl propionic acid [15] as acyl donors. In all cases the preference of this enzyme for the 6″-OH group stood unaffected, which further supports our model for the binding of isoquercitrin within the CALB active site.

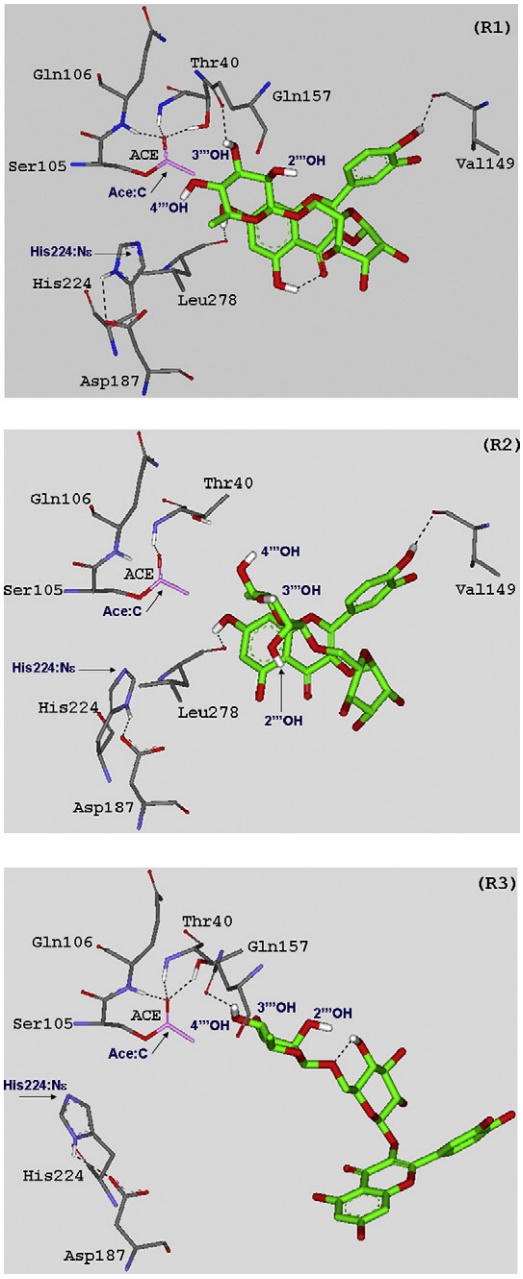

Fig. 6. Orientations and hydrogen bonds interactions (dashed lines) of the substrates in the three best-scored acetyl–CALB rutin complexes (R1, R2 and R3). For clarity, only the residues useful for discussion are shown and non-interacting hydrogen atoms are hidden.

3.2.2. *Docking of rutin*

In the case of rutin, the three retained complexes (namely R1, R2 and R3; Fig. 6) presented the flavonoid docked in similar orientations. In all of them, the aglycon moiety stabilizes at the entrance of the cavity, while the sugar part of the substrate points towards the bottom. Notably, the 4‴-OH group of the rhamnose unit is the closest one to the catalytic residues. However, the three complexes present rutin in different conformations, in which the dihedral angle between the two sugars is more or less bent (162° for R1, 149° for R2 and 102° for R3). In R2, only one hydrogen bond between the residues of the oxyanion hole and the acetate is present (Thr40:NH. . .Ace:O). Indeed, the orientation of the acetate may be influenced by the 6‴-CH_3 of the rutin rhamnose through a hydrophobic attraction (distance of 3.6 Å to the 6‴-C atom). Furthermore, in this complex the distance from the 4‴-OH to the

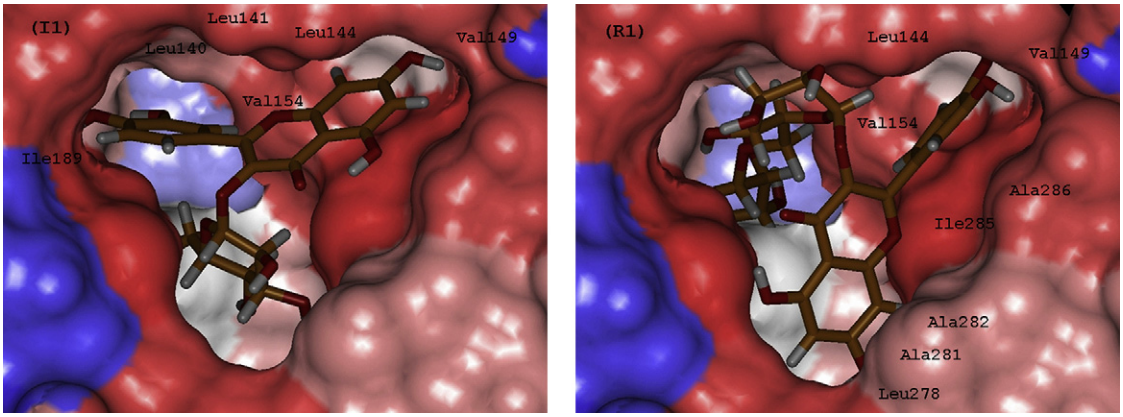

Fig. 7. Connolly accessible surface of CALB binding pocket evidencing the hydrophobic interactions of isoquercitrin in I1 and rutin in R1. Hydrophobic regions of the enzyme pocket are colored in red and hydrophilic ones in white/blue.

His224:Nε is 7.7 Å, exceeding the value admitted as to permit the proton transfer. Consequently, R2 was rejected. In R3, the acetate presents all the three possible hydrogen bond interactions with the oxyanion hole residues (Thr40:OH. . .Ace:O, Thr40:NH. . .Ace:O and Gln106:NH. . .Ace:O). The 4‴-OH is positioned at 3.6 Å from Ace:C, but at 9.4 Å from the His224:Nε. Therefore, rutin in R3 is also unable to react and the complex was also discarded. In R1 all the three possible hydrogen bond interactions with the oxyanion hole are also maintained. The rutin conformation is quite different from that in R3: the sugar moiety is somewhat benter, so that the 4‴-OH is placed at 3.9 Å from His224:Nε and 2.9 Å from Ace:C. So, this complex may be considered as a productive binding mode.

In this complex, the rutin aglycon part stabilized at the entrance of the cavity thanks to hydrogen bonds interactions involving phenolic groups (4′-OH. . .Val149:O and 7-OH. . .Leu278:O) and nonpolar interactions between its aromatic cycles and the side chains of residues Leu144, Val149, Val154, Ala281, Ala282, Ile285 and Val286. The rhamnose part, that has three hydroxyl groups (2‴-OH, 3‴-OH and 4‴-OH), is located near the little hydrophilic region at the bottom of the cavity, neighbouring the catalytic residues. Among these three hydroxyl groups, only the 4‴-OH satisfies the criteria of distance from His224:Nε and Ace:C atoms to be considered as reactive. The 3‴-OH group forms a hydrogen bond interaction with the Gln157 residue (3‴-OH. . .Gln157:OE1), which confers additional stability to the rhamnose cycle in this orientation. All these observations hint that the acylation reaction can only occur on the 4‴-OH. This model is well corroborated by experimental studies showing that the CALB-catalyzed acylation of rutin with different acyl donors and under diverse reaction conditions led to the sole formation of the 4‴-monoester [8–11].

4. Discussion

In order to understand the orientations of rutin and isoquercitrin within the CALB binding pocket, the observed interactions between these two substrates and the enzyme residues were compared with already published data about protein–flavonoid interactions. A theoretical affinity order between flavonoids and 20 amino acids was proposed by Cordoniu-Hernandez et al. [45,46], based on quantum mechanics semi-empirical calculations and the analysis of crystal structures of protein–flavonoid complexes from PDB. These authors stated that flavonoids display a more pronounced affinity towards hydrophilic and aromatic residues rather than hydrophobic ones. Indeed, the abundance of hydroxyl groups in flavonoid structure enables interactions of hydrogen bond type with polar residues, whereas their aromatic rings can easily form π–π interactions with aromatic residues side chains. In the case of CALB, major part of the binding pocket area is nonpolar, enfolded mainly by hydrophobic aliphatic residues of the helices α5 and α10 and a loop region that projects the Ile189 into the channel [31]. This can justify the fact that the aglycon moiety of both studied substrates is stabilized at the cavity entrance, despite the absence of aromatic amino acid residues (Fig. 7). Favourable interactions between phenolic substrates and the nonpolar residues Leu144, Ile189 and Leu278 were already evidenced by Otto et al. [25], in a computational study of the CALB-catalyzed synthesis of arylaliphatic glycolipids. Only a little region at the bottom of the cavity, around the catalytic Ser105, presents a character somewhat polar, due to the side chains of the polar residues Thr40, Asp134 and Gln157. This justifies the orientation of the glycoside moiety of the two substrates towards the active site.

In the productive complexes, the aglycon parts of the two substrates are located at the cavity's doorway, but surprisingly, with distinct orientations, one perpendicular to the other (Fig. 7). In order to explain such behaviour, we docked isoquercitrin and rutin employing another approach that is frequently used when docking structurally similar ligands against one given target protein: one ligand is initially docked and the others are superimposed over the analogue parts of the first one. Then, the resulting systems are sequentially optimized, analyzed and compared [47]. In our case, the rutin aglycon part was superimposed over its counterpart in the previously docked isoquercitrin and vice versa. In the first case, the rutin disaccharide had not enough space to lodge within the cavity, engendering an important sterical clash with the acetate and the Ser105. Conversely, in the second case, the isoquercitrin glucose is not long enough to reach the catalytic residues, resulting in a nonproductive binding mode (see Supplementary Data). Therefore, it can be concluded that, although these two substrates have similar structures, they effectively present distinct binding modes within the CALB active site, as determined by LigandFit.

We complemented our study by additional molecular dynamics simulations on the productive complexes (100 ps at 300 K), to verify the stability of the hydrogen bond interactions and to evaluate the dynamic behaviour of the docked substrates. In fact, applying molecular dynamics on complexes obtained by docking was shown to be efficient to refine the models, accounting for the flexibility of both receptor and ligand, and to check the complex temporal stability [48]. Results showed that, in any case, the proteins did not undergo significant distortions, as revealed by the maximal *rmsd* values (about 0.2–0.5 Å, considering all heavy atoms). The presence of the essential hydrogen bond interactions between the acetate carbonyl oxygen (Ace:O) and Thr40 and Gln106 was monitored. The visual analysis of trajectories shows that at least two of them are

Table 3
Minimal and maximal transacetylating distances[a] during the MD production phase carried out on the productive complexes.

Ligand	Buried –OH groups	Minimal distance (Å)	Maximal distance (Å)	Maximal variation (Å)
Isoquercitrin	3″-OH	8.0	8.4	0.4
	4″-OH	4.6	5.8	1.2
	6″-OH	4.0	4.6	0.6
Rutin	2‴-OH	7.6	8.2	0.6
	3‴-OH	5.6	6.5	0.9
	4‴-OH	3.9	4.3	0.4

[a] Distance between the oxygen atoms of the flavonoid hydroxyl groups and the carbonyl carbon atom of the acetate (Ace:C) bound to Ser105.

present in most of trajectory frames during the production phase. In order to evaluate the stability of the hydroxyl groups positions, their transacetylating distances (distance between the hydroxyl oxygen atoms and the carbonyl carbon of acetate, Ace:C) were monitored all along the trajectory. Among them, only the 6″-OH of isoquercitrin and the 4‴-OH of rutin reached distances inferior to 4 Å from Ace:C. Furthermore, maximal fluctuations in these distances were weak (about 0.4–0.6 Å), reflecting the good stability of the docked flavonoids (Table 3). These results from molecular dynamics simulations on the productive complexes reinforce our previous conclusions about regioselectivity in CALB-catalyzed acylation of isoquercitrin and rutin.

5. Summary and conclusions

Docking simulations combined with molecular mechanics and molecular dynamics calculations were used to study CALB-catalyzed acetylation of the two glycosylated flavonoids isoquercitrin and rutin. Both flavonoids were found to dock with their aglycon part at the cavity entrance, stabilized by hydrogen bond interactions between their phenolic groups and the backbone carbonyl of hydrophobic residues and hydrophobic interactions with the side chains of several Ala, Ile, Leu and Val residues enfolding the pocket. In spite of the aromatic character of the substrates, no π–π interactions were possible due to the lack of aromatic residues in the binding pocket walls. Their sugar part was stabilized in the neighbourhood of the catalytic residues, a little polar region at the bottom of the cavity. Further analyses of the complexes showed that the 4‴-OH of rutin and the 6″-OH of isoquercitrin, both localized at the extremity of the substrates glycosidic part, were the only hydroxyl groups satisfying the proximity criteria to the catalytic residues. This strict regioselectivity suggested by the models corroborated previous published experimental results on the CALB-catalyzed acetylation of isoquercitrin and rutin.

Combined theoretical and experimental studies are now pursued to investigate the capacity of our molecular modelling procedure to predict the regioselectivity of CALB for the acetylation of other glycosylated and aglycon flavonoids. In fact, as flavonoids show a very large structural diversity (number and positions of glycoside units and hydroxyl groups, degree of oxidation of the C-ring. . .), they have different sizes, flexibilities, charge distributions, hydrophilic–hydrophobic balance properties, and are likely to yield different interaction modes with the lipase. Thus the entire modelling protocol outlined here must be applied to a broader range of flavonoid substrates, in order to verify if the geometrical cut-off between reactive and unreactive complexes (protein *rmsd* and proximity of flavonoid hydroxyl groups to the lipase catalytic residues) can be generalized to predict the regioselectivity of CALB-catalyzed acetylation of other members of this class of compounds.

Acknowledgment

We are thankful to Evelyne Ronat for her precious and essential technical support all along this work.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molcatb.2009.01.011.

References

[1] B.H. Havsteen, Pharmacol. Ther. 96 (2002) 67–202.
[2] A.M. Boudet, Phytochemistry 68 (2007) 2722–2735.
[3] L. Chebil, C. Humeau, J. Anthoni, F. Dehez, J.M. Engasser, M. Ghoul, J. Chem. Eng. Data 52 (2007) 1552–1556.
[4] P. Villeneuve, Biotechnol. Adv. 25 (2007) 515–536.
[5] A.M. Mariotte, E. Perrier, A. Boumendjel, D. Bresson-Rival, Nouveaux esters de flavonoïdes et leur utilisation en cosmétique, en dermopharmacie, en pharmacie et en agro-alimentaire, in FR 2778663, France, 1999.
[6] Z.Y. Chen, P.T. Chan, K.Y. Ho, K.P. Fung, J. Wang, Chem. Phys. Lip. 79 (1996) 157–163.
[7] L. Chebil, C. Humeau, A. Falcimaigne, J.M. Engasser, M. Ghoul, Process Biochem. 41 (2006) 2237–2251.
[8] F. Mellou, H. Loutrari, H. Stamatis, C. Roussos, F.N. Kolisis, Process Biochem. 41 (2006) 2029–2034.
[9] M.H. Katsoura, A.C. Polydera, L. Tsironis, A.D. Tselepis, H. Stamatis, J. Biotechnol. 123 (2006) 491–503.
[10] M. Ardhaoui, A. Falcimaigne, J.M. Engasser, P. Moussou, G. Pauly, M. Ghoul, J. Mol. Catal. B: Enzym. 29 (2004) 63–67.
[11] M. Ardhaoui, A. Falcimaigne, S. Ognier, J.M. Engasser, P. Moussou, G. Pauly, M. Ghoul, J. Biotechnol. 110 (2004) 265–271.
[12] L. Chebil, J. Anthoni, C. Humeau, C. Gerardin, J.M. Engasser, M. Ghoul, J. Agric. Food Chem. 53 (2007) 9496–9502.
[13] N. Nakajima, K. Ishihara, T. Itoh, T. Furuya, H. Hamada, J. Biosci. Bioeng. 87 (1999) 105–107.
[14] K. Ishihara, N. Nakajima, J. Mol. Catal. B: Enzym. 23 (2003) 411–417.
[15] D.E. Stevenson, R. Wibisono, D.J. Jensen, R.A. Stanley, J.M. Cooney, Enzyme Microbial Technol. 39 (2006) 1236–1241.
[16] R.J. Kazlauskas, Curr. Opin. Chem. Biol. 4 (2000) 81–88.
[17] P. Braiuca, C. Ebert, A. Basso, P. Linda, L. Gardossi, Trends Biotechnol. 24 (2006) 419–425.
[18] C. Orrenius, F. Haeffner, D. Rotticci, N. Ohrner, T. Norin, K. Hult, Biocatal. Biotransf. 16 (1998) 1–15.
[19] T. Schultz, J. Pleiss, R.D. Schmid, Prot. Sci. 9 (2000) 1053–1062.
[20] A. Tafi, A. Van Almsick, F. Corelli, M. Crusco, K.E. Laumen, M.P. Schneider, M. Botta, J. Org. Chem. 65 (2000) 3659–3665.
[21] S. Raza, L. Fransson, K. Hult, Prot. Sci. 10 (2001) 329–338.
[22] J. Ottosson, L. Fransson, K. Hult, Prot. Sci. 11 (2002) 1462–1471.
[23] V. Léonard, L. Fransson, S. Lamare, K. Hult, M. Graber, ChemBioChem 8 (2007) 662–667.
[24] B. Botta, G. Zappia, A. Tafi, M. Botta, F. Manetti, E. Cernia, G. Milana, C. Palocci, S. Soro, G. Delle Monache, J. Mol. Catal. B: Enzym. 16 (2002) 241–247.
[25] R.T. Otto, H. Scheib, U.T. Bornscheuer, J. Pleiss, C. Syldatk, R.D. Schmid, J. Mol. Catal. B: Enzym. 8 (2000) 201–211.
[26] L. Cipolla, M. Lotti, L. De Gioia, F. Nicotra, J. Carbohydr. Chem. 22 (2003) 631–644.
[27] C. Palocci, M. Falconi, S. Alcaro, A. Tafi, R. Puglisi, F. Ortuso, M. Botta, L. Alberghina, E. Cernia, J. Biotechnol. 128 (2007) 908–918.
[28] I. Vallikivi, L. Fransson, K. Hult, I. Järving, T. Pehk, N. Samel, V. Tõugu, L. Villo, O. Parve, J. Mol. Catal. B: Enzym. 35 (2005) 62–69.
[29] A.D. MacKarrel Jr., D. Bashford, M. Bellot, R.L. Dunbrack Jr., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. 102 (1998) 3586–3616.
[30] M.L. Connolly, Science 221 (1983) 709–713.
[31] J. Uppenberg, M.T. Hansen, S. Patkar, T.A. Jones, Structure 2 (1994) 293–308.
[32] D.H.J. Mackay, A.J. Cross, A.T. Hagler, in: G.D. Fasman (Ed.), Prediction of Protein Structure and the Principles of Protein Conformation, Plenum Press, New York & London, 1990, pp. 317–358.
[33] C.M. Venkatachalam, X. Jiang, T. Oldfield, M. Waldman, J. Mol. Graph. Modell. 21 (2003) 289–307.
[34] A. Krammer, P.D. Kirchhoff, X. Jiang, C.M. Venkatachalam, M. Waldman, J. Mol. Graph. Modell. 23 (2005) 395–407.

[35] X.D. Li, T.J. Hou, X.J. Xu, Acta Phys. Chim. Sin. 21 (2005) 504–507.
[36] I. Muegge, J. Med. Chem. 49 (2006) 5895–5902.
[37] A. Jain, Curr. Prot. Pept. Sci. 7 (2006) 407–420.
[38] M. Feher, Drug Discov. Today 11 (2006) 421–428.
[39] A.M. Davis, S.J. Teague, Angew. Chem. Int. Ed. 38 (1999) 736–749.
[40] A.M. Kilbanov, Tibtech 15 (1997) 97–100.
[41] P. Trodler, J. Pleiss, BMC Struct. Biol. 8 (2008), art. no. 9.
[42] J. Pleiss, M. Fischer, R.D. Schmid, Chem. Phys. Lip. 93 (1998) 67–80.
[43] A. May, F. Sieker, M. Zacharias, Curr. Comput. Aided Drug Des. 4 (2008) 143–153.
[44] B. Danieli, M. Luisetti, G. Sampognaro, G. Carrea, S. Riva, J. Mol. Catal. B: Enzym. 3 (1997) 193–201.
[45] E. Codorniu-Hernandez, A. Mesa-Ibirico, R. Hernandez-Santiesteban, L.A. Montero-Cabrera, F. Martinez-Luzardo, J.L. Santana-Romero, T. Borrmann, W.D. Stohrer, Int. J. Quant. Chem. 103 (2005) 82–104.
[46] E. Codorniu-Hernandez, A. Rolo-Naranjo, L.A. Montero-Cabrera, THEOCHEM 819 (2007) 121–129.
[47] C.K.Z. Andrade, W.A. Silva, E.R. Maia, J. Biomol. Struct. Dyn. 25 (2007) 35–48.
[48] H. Alonso, A.A. Bliznyuk, J.E. Gready, Med. Res. Rev. 26 (2006) 531–568.