



## LIPASE-CATALYSED ACYL TRANSFER IN DEOXY SUGARS: A COMPUTATIONAL STUDY

<sup>1</sup>TOOMAS KAEVAND, <sup>2</sup>LY VILLO, <sup>3</sup>ANDRUS METSALA, <sup>4</sup>OMAR PARVE,  
<sup>5</sup>ÜLO LILLE

<sup>1</sup> HPC service manager, Information Technology Services, Tallinn University of Technology, Ehitajate tee 5,

19086 Tallinn, Estonia

<sup>2</sup>Associated Prof., Department of Chemistry, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

<sup>3</sup>Senior researcher, Department of Chemistry, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

<sup>4</sup>Leading researcher, Department of Chemistry, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

<sup>5</sup>Prof. Emeritus, Department of Chemistry, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

E-mail: <sup>1</sup>[toomaskaevand@gmail.com](mailto:toomaskaevand@gmail.com), <sup>2</sup>[lee@chemnet.ee](mailto:lee@chemnet.ee), <sup>3</sup>[metsala@chemnet.ee](mailto:metsala@chemnet.ee), <sup>4</sup>[omar@chemnet.ee](mailto:omar@chemnet.ee), <sup>5</sup>[lille@chemnet.ee](mailto:lille@chemnet.ee)

### ABSTRACT

*CALB-catalyzed acyl transfer from vinyl acetate to bromo- and ester-derivatives of tri- and tetra-deoxysugars has been studied by Amber biomolecular simulation packet. In comparison to the experimental data the best results have been obtained via an analysis of catalytically essential hydrogen bonds.*

**Keywords:** CALB, acyl transfer, deoxy sugars, hydrogen bonds, Amber package

### 1. INTRODUCTION

Lipases are the most widely used class of enzymes for biochemical applications and their structure, mechanism of action and properties are reviewed in [1]. Basic features of its structure have been clarified many years ago [2, 3]. The active site, substrate binding features, and flexibility are the objects of the ongoing research [4],[5],[6, 7].

Lipase-catalyzed acyl transfer has been used by authors research circle since 1997 [8]. Most results on the chemoenzymatic synthesis of  $\alpha$ -bromo- $\omega$ -hydroxyaldehydes in cyclic and open-chain forms (tri- and tetra-deoxy sugars termed as DOS) have been generalized in [9] (and citation therein). Some examples of the ongoing research are given in [10], [11]. In these studies *Candida antarctica* lipase B (CALB) has been used, certain enantiomerically pure compounds were obtained, covalent trapping of the open-chain with the active site nucleophile has been proposed.

CALB catalyses enantioselective acyl transfer from the acyl donor (first substrate, an activated ester) to the acceptor (a racemic second substrate, e.g. a secondary alcohol to be resolved) by prototypical

ping-pong reaction mechanism of the serine protease trypsin [3],[12]. This mechanism is based on nucleophilic catalysis. It includes the catalytic triad Asp-His-Ser, acyl donor and acceptor. The acyl transfer takes place via transition state (TS) with two consecutive tetrahedral intermediates (TI) formed in nucleophilic SerO<sub>y</sub> and the acyl acceptor attacks respectively. The TS is stabilized by 6 catalytically essential hydrogen bonds (HB-s). In case of sugars as substrates the situation is more complicated due to the dynamic equilibrium between the cyclic and acyclic forms of substrates: other nucleophiles present at the active center of the enzyme can act as trapping agents by the third nucleophilic attack to electrophilic carbonyl group in the acyclic substrate.

The formation of TS is the rate limiting step of the acyl transfer and different computational procedures have been used for its characterization and prediction in this way the transfer selectivity. The estimated parameters have been the respective differences of the activation free energy values [13], hydrogen-bonding networks [14], conformational states [15].

The aim of the present report is to clarify the main features of the achieved enzymatic resolution results by 3 different the molecular modeling approaches shown above to gain the understanding of the chiral selection process on the molecular level. We use the cyclic furanose, pyranose and respective open-chain forms of DOS in the role of acyl acceptors which enter in the reaction of its second step.

## 2. METHODOLOGY

Under study is a number of cyclic (2R3R), (2R3S), (2S3R), (2S3S) and open-chain (2R), (2S)

stereoisomeric hemiacetals (termed further as isomers) consisting two and one stereocenter respectively in shown configurations, see Fig. 1. In accordance with IUPAC names the carbonylic open-chain carbon and anomeric carbon in the cycle are numbered 1 and 2 respectively. According to QC calculations the equilibrium shown on Fig.1 is fully shifted to the cyclic isomers [16]. There exist 6 isomers in each of three groups of studied compounds resulting in 18 isomers. Among them are anomeric pairs 1,3; 2,4; 7,9; 8,10.

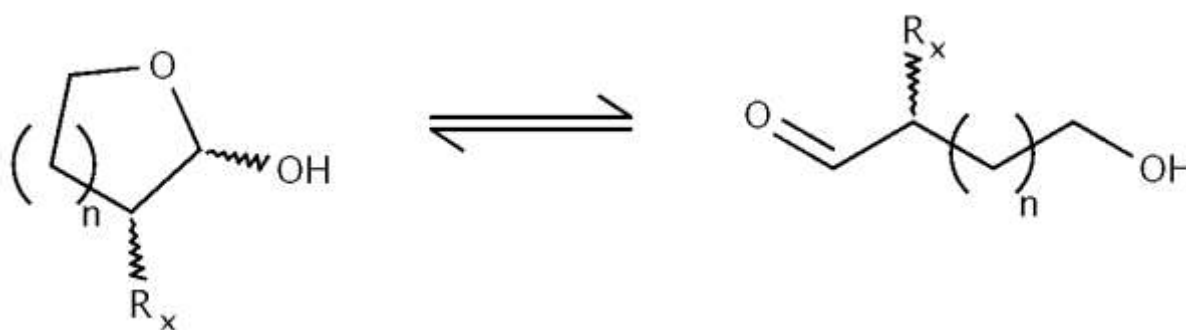


Fig.1 Equilibrium between cyclic and open-chain stereoisomers of the studied hemiacetals. Designation of stereoisomers: Bromo-derivatives,  $R_x = Br$ ,  $n = 1$ , cyclic 1-4(2R3R), (2R3S), (2S3R), 2S3S), acyclic 5-6 (2R), (2S);  $n = 2$ , analogously cyclic 7-10, acyclic 11-12; Esters,  $R_x = OC(O)R^*$ ,  $R^*$  is  $-C(NHBoc)CH_2-p$ -hydroxyphenyl;  $n = 1$ , analogously cyclic 13-16, acyclic 17-18.

From the enzyme (PDB-entry 1TCA) inhibitor is removed. From the active site of the enzyme the water molecules within 6 Å of Ser105  $O_\gamma$  are removed as well (in the whole system exists ca 400 water molecules). Note, that 286 water molecules were included in the organic solvent simulations [17]. Then from the molecule of the vinylacetate (acyl donor) the tetrahedral center is formed via an imaginary attack of Ser105- $O_\gamma$ . Acyl acceptor, i.e. a nucleophile (i.e. certain isomer of a cyclic hemiacetal or the respective open-chain  $\alpha$ -bromo- $\omega$ -hydroxy-aldehyde in putative dynamic equilibrium) is attached via the respective oxygen atom ( $O_n$ ) to the tetrahedral center, see Fig. 2.

The formed TI is docked by AutoDock Vina software [18], which generates proper coordinates for atoms of TI. A proper TI conformation in the active centre of the enzyme is selected manually from ca 10 conformers proposed by program using the lowest-scoring function. The free energy of binding ( $E_b$ ) is calculated for all selected TI-s using the same software.

PDB files of the enzyme and TI are concatenated into a single PDB file so, that the coordinates of the formed TI atoms will be at the end of the new file. Then the atom numeration in the file will be corrected (by hand or with a special program). A file with substrate atoms and bond parameters are generated using Antechamber software (part of Amber package nr 12 [19]).

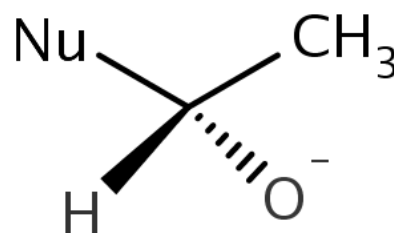


Fig.2. Tetrahedral intermediate of the acyl transfer from vinyl acetate to the hemiacetals as acyl acceptors (Nu). Ser 105  $O_\gamma$  is replaced with an hydrogen atom.



The missing force field (ff) parameters are generated by FRCMOD and GAFF (General Amber ff). The single PDB file is then loaded along with the ff information from previous steps to Leap (tLeap or xLeap) program. Then His 224 is protonated, covalent bond is formed between Ser 105 O<sub>γ</sub> and the tetrahedral center of the acyl donor. Analogically the carbonyl carbon of acyclic stereoisomers is attached to Thr40-O<sub>γ</sub> and the new hydroxyl group is formed at the end of the open chain. So the TI-s of the acyl transfer studied by molecular dynamics has been obtained. Then the system charge is revised and counterions (Na<sup>+</sup>) are added up to the zero charge. The system is solvated by the creation of the boundary box filled with chloroform ( $\epsilon_r$  4.8) and vinylacetate molecules (3:1).

Molecular dynamics simulations (NPT) are performed at 300 K and 1 atm. PRMTOP and INPCRD files (run Leap) from previous steps are used. The minimization, heating to 300 °K and equilibration of the systems are performed before the data production during 2 ns. Ensemble potential energy of systems has been averaged over the full length of MD trajectories. Sample structure was extracted after every 50 thousand steps per 2fs (in total 10 ps). In this way 200 frames over the whole trajectory are generated. These data used for further analysis by subprogram pmemd.. The presence of HB-s is clarified in all 200 frames via the calculated respective averaged geometric parameters. Valid HB is defined as this with distance from donor hydrogen to the acceptor (termed as  $d_{DH...A}$ ) <3.5 Å and angle >120 degrees as used in substrate selectivity study of CALB [14].

Energy ( $\Delta H_f$ ) values of HB-s in gas phase are calculated using Gaussian-2,ROHF, 6-31G(d). The difference in the free energy of activation between the two isomers  $\Delta\Delta G^\ddagger$  is calculated as the difference of the average free energy of isomers to be resolved [13].

The active site structure of CALB 1TCA was visualized using UCSF Chimera (see Fig 3 as an example).

### 3. RESULTS AND DISCUSSION

#### 3.1. Analysis of HB-s.

In the catalytic process seven donor/acceptor pairs are involved resulting in the formation in the cyclic and acyclic TIs 14 and 16 HB-s respectively. Here we pay the main attention to the catalytically most essential HB-s which stabilize the TS in serine proteases [12] (Table, to avoid the landscape

presentation the standard deviations are not shown, in our context their role is not critical).



Table

Averaged geometric parameters of catalytically essential HB-s<sup>a</sup>

Nr.(type of isomer) <sup>b</sup>	Nr. of HB and d <sub>DH...A</sub> , Å /angels, degrees					
	4	5	9	11	12	14
1(2R3R)	1.7/64	<b>1.8/161</b>	3.5/114	4.8/164	4.6/124	4.6/168
2(2R3S)	<b>1.6/161</b>	<b>1.7/146</b>	3.1/n.d. <sup>c</sup>	5.0/116	5.4/102	<b>1.8/143</b>
3(2S3R)	<b>1.7/163</b>	<b>2.3/136</b>	3.2/n.d. <sup>c</sup>	5.0/135	5.0/103	3.9/148
4(2S3S)	<b>1.7/160</b>	<b>2.1/143</b>	<b>3.2/147</b>	4.7/118	5.6/86	4.1/150
5(2R)*	6.4/46	<b>2.0/136</b>	1.8/n.d. <sup>c</sup>	3.7/169	<b>1.4/168</b>	<b>1.7/151</b>
6(2S)*	<b>1.7/166</b>	<b>1.8/136</b>	1.8/n.d. <sup>c</sup>	3.7/170	<b>1.8/168</b>	<b>1.7/150</b>
7(2R3R)*	<b>1.7/166</b>	<b>1.9/160</b>	<b>1.8/133</b>	<b>1.8/163</b>	<b>1.8/165</b>	<b>1.9/165</b>
7(2R3R) <sup>m</sup>	<b>1.7/166</b>	<b>1.7/160</b>	<b>3.4/133</b>	-	-	<b>1.9/165</b>
8(2R3S)	<b>1.7/163</b>	<b>1.7/150</b>	<b>3.4/127</b>	<b>1.8/125</b>	4.5/139	<b>1.8/135</b>
9(2S3R)	<b>1.7/159</b>	<b>1.7/145</b>	<b>3.4/129</b>	4.8/128	4.1/154	<b>1.9/144</b>
10(2S2S)	<b>1.7/165</b>	<b>1.8/142</b>	3.1/117	5.0/119	4.3/134	4.3/147
11(2R)*	<b>1.7/164</b>	<b>1.8/149</b>	<b>2.0/135</b>	<b>1.8/167</b>	<b>1.8/138</b>	<b>1.7/165</b>
12(2S)*	<b>1.7/156</b>	<b>1.8/138</b>	6.8/149	6.8/106	6.6/115	<b>2.1/141</b>
13(2R3R)*	<b>1.7/163</b>	<b>2.0/166</b>	<b>3.1/142</b>	<b>1.9/145</b>	<b>2.8/164</b>	<b>1.8/168</b>
14(2R3S)	<b>1.6/164</b>	3.8/165	3.8/160	4.4/147	4.2/138	4.0/125
15(2S3R)	<b>1.7/169</b>	<b>1.8/154</b>	<b>2.0/153</b>	<b>1.8/159</b>	3.8/152	<b>1.7/164</b>
16(2S3S)	<b>1.7/160</b>	<b>1.8/165</b>	<b>3.1/145</b>	5.4/109	<b>3.2/133</b>	<b>2.1/129</b>
17(2R)	<b>1.6/156</b>	<b>1.7/150</b>	<b>3.1/135</b>	<b>1.9/162</b>	<b>1.7/166</b>	<b>1.8/161</b>
18(2S)	<b>1.6/163</b>	<b>2.1/126</b>	<b>2.3/127</b>	4.3/120	<b>2.2/146</b>	n.d/143 <sup>c</sup>

a) designation of HB-s: 4) His224-N<sub>δ</sub>H/Asp187O<sub>2</sub>; 5) His224-N<sub>ε</sub>H/Ser105-O<sub>γ</sub>;

9) His224-N<sub>ε</sub>H/TI-O<sub>n</sub>; 11) Thr40-NH/TIO<sup>(-)</sup>; 12) Thr40-O<sub>γ</sub>H/TIO<sup>(-)</sup>;

14) Gln106-NH/TI-O<sup>(-)</sup> (values of valid H-bonds with distance <3.5 Å and angle >120 degrees are shown in bold)

b) Isomers preferentially recognized by CALB are marked \*; m- mutant

c) considered as <120 degrees

The geometric characteristics shown in Tab in bold correspond mostly to the HB-s of medium or moderate strength [20]. In accordance with this statement the calculated ΔH<sub>f</sub> of shown HB-s are in the range of 6.4-14.4 kcal/mol. In comparison to the water these relatively high values originate apparently from the system media, i.e. the aprotic solvent [21]. Notable that the hydroxyl group formed in the trapping process of the open chain via the attack of Thr40-O<sub>γ</sub> on the carbonyl group does not form a valid HB with the oxyanion O<sup>(-)</sup>. The reason for this is the absence of the linearity in the putative HB. Instead of HB formation (numbered as HB 13) a strong repulsive interaction is present. The origin of these interactions need further study. In general, the highest amount of catalytically essential HB-s is observed in TI-s 7, 11 and 13 and they are also preferred by CALB. This is an essential result.

In mutated enzyme Thr40Val the oxyanion hole is less stabilized since HB-s 11 and 12 are absent (the respective donors do not exist). The effect of this mutation on other HB-s is negligible and is not shown in Tab. There is one exception and it is shown in Tab: HB 9 in the isomer nr.7. In mutated enzyme it is weak HB. This HB plays critical role in the stereoselection process of secondary alcohols [22]. The stereocenter of hemiacetals has certain similarity to this of secondary alcohols and probably this mutation can decrease the stereoselectivity. Notable that any Thr-O<sub>γ</sub>H atoms from the other side-chain residues which could form the HB with aldehyde oxygen and then allow for the covalent bond formation in the trapping process of the open chain have not been found. Hence, the ring opening process probably will not take place. However, experimental evidence is needed to check this recommendation.

### 3. 2. Calculations of the potential energy (ΔG<sup>#</sup>) values of TI-s

The average potential values of the whole systems are rather high, ca -19000± ca100 kcal/mol. The shown large fluctuations are probably caused by the general protein motions. However, the use of averaged values the effect of them is minimized. Remarkable, that systems consisting of acyclic forms have higher negative values relative to these



of cyclic forms, for furanose- and pyranose- type TI-s ca -540 and -100 kcal/mol respectively. It seems that this effect reflects of the enzymatic environment on the potential energy of the docked TI. Among cyclic pyranoses highest negative value has TI nr. 7.  $\Delta\Delta G^\ddagger$  value of enantiomeric pair 7/10 is estimated 31 kcal/mol and these of diastereomeric pairs 7/8 and 7/9 29 and 25 kcal/mol respectively. These high values lead to very high enantioselectivity values as well.

These results point to 1) the more favorable ring-opening accompanying the acyl transfer process in furanoses and 2) preferential formation of pure pyranose 2R3R among other pyranose cyclic stereoisomers as shown in the physical experiment [16]. Exceptionally the higher negative value for acyclic form nr 11 contradicts to the slight experimental preference of acyclic form nr 12.

### 3. 3. Calculation of the binding energy

The absolute  $E_b$  values of TI-s are ca 5 kcal/mol (the standard error of the used software is 2.85 kcal/mol [18]). The lower border of  $E_b$  values given in this paper for protein-ligands complexes with ligands containing 5 rotatable bonds is ca 5 kcal/mol). Note that AutoDock Vina approximates the  $E_b$  values to this of the distance-dependent scoring function consisting of steric, hydrophobic, HB and number of active rotatable bond terms. Receptors are treated here as rigid. Covalent binding is not included and therefore the obtained values are quite logical. However, these values did not give any information about the resolution process.

According to [23] the  $E_b$  values of the TS reach 38 kcal/mol. The reason of this is the covalent binding, for example the acylenzyme formation [24]. In our study the  $E_b$  values have been calculated in the docking step. The acylenzyme is formed later, in the preparation step of MD simulation.

### 4. CONCLUDING REMARKS

The potential energy calculation and HB network gives useful information on the stereoselection of studied isomers. However, the prediction of this process is so far rather questionable. According to [25] and [26] molecular recognition process seems to be a complex and emergent property resulting from an interplay between all inter- and intramolecular interactions and cannot be predicted even assuming that each H-bonds strength is exactly known.

From the other side, to obtain quantitative results based on relatively incomplete experimental facts (no kinetic data, dynamic equilibrium between the

starting compounds, possible transformations on silica as shown in [27] etc.) is not easy.

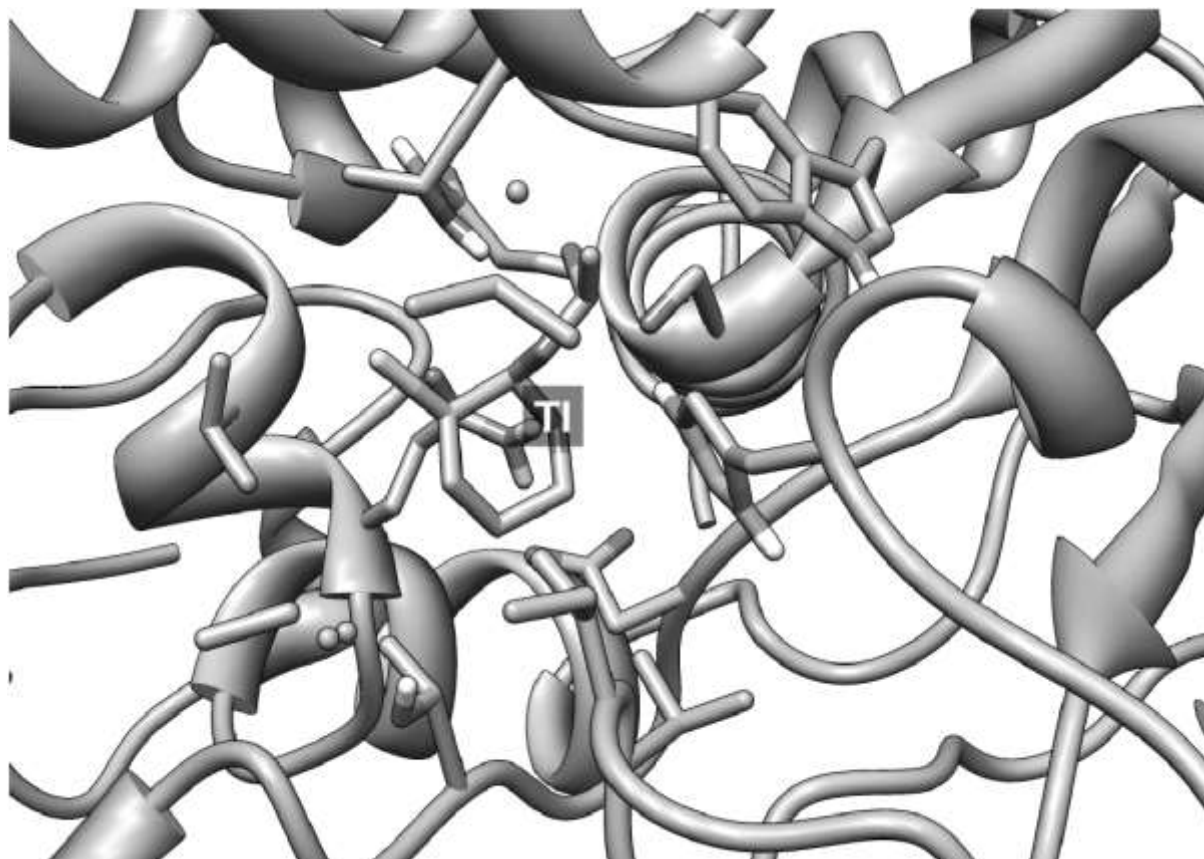


Fig. 3 Tetrahedral intermediate (TI) of the isomer 7(2R3R) in the active center of CALB.

## REFERENCES

- [1] M. Kapoor, M.N. Gupta, Lipase promiscuity and its biochemical applications, *Process Biochemistry*, 47 (2012) 555-569.
- [2] J. Uppenberg, e. al, The sequence, crystal structure determination and refinement of two crystal forms of lipase B from *Candida antarctica*, *Structure*, 2 (1994) 293-308.
- [3] M. Martinelle, K. Hult, Kinetics of acyl transfer reactions in organic media catalysed by *Candida antarctica* lipase B., *Biochim. et Biophys. Acta* 1251 (1995) 191-197.
- [4] M.R. Ganjalikhany, e. al, Functional Motions of *Candida antarctica* Lipase B: A Survey through Open-Close Conformations, *PLoS One*, 7 (2012) e 40327.
- [5] M. Housaindokhta, M. Bozorgmehr, H. Monhemi, Structural behavior of *Candida antarctica* lipase B in water and supercritical carbon dioxide: A molecular dynamic simulation study, *Journal of Supercritical Fluids*, 63 (2012) 180-186.
- [6] H.B. Idris, A., Immobilized *Candida antarctica* lipase B: Hydration, stripping off and application in ring opening polyester synthesis, *Biotechnology Advances*, 30 (2012) 551-563.
- [7] S.-Z. Wang, J.-P. Wu, G. Xu, L.-B. Yang, A kinetic and thermodynamic study of the lipase-catalyzed remote resolution of a chiral tertiary alcohol, *Biochemical Engineering Journal*, 65 (2012) 57-62.
- [8] O. Parve, I. Vallikivi, L. Lahe, P. Sikk, T. Käämbre, Ü. Lille, A chemoenzymatic approach to the preparation of optically active  $\alpha$ -bromo- $\omega$ -hydroxy aldehyde hemiacetals, *Proc.Estonian Acad.Sci.Chem.*, 46 (1997) 186-190.
- [9] L. Villo, Stereoselective Chemoenzymatic Synthesis of Deoxy Sugar Esters Involving *Candida antarctica* Lipase B in: Faculty of Science, Department of Chemistry, Tallinn



- University of Technology, Tallinn, 2008, pp. 116, Thesis.
- [10] L. Villo, M. Kreen, M. Kudryashova, A. Metsala, S. Tamp, Ü. Lille, T. Pehk, O. Parve, A chemoenzymatic synthesis of deoxy sugar esters involving stereoselective acylation of hemiacetals catalyzed by CALB, *J. Molecular Catalysis B: Enzymatic*, 68 (2011) 44-51.
- [11] L. Villo, M. Kreen, M. Kudryashova, A. Metsala, S. Tamp, Ü. Lille, T. Pehk, O. Parve, A chemoenzymatic synthesis of a deoxy sugar ester of N-Boc-protected L-tyrosine, in: J. Whittall, Sutton, P.W. (Ed.) *Practical methods for Biocatalysis and Biotransformations 2*, Wiley, 2012, pp. 335-339.
- [12] T. Ishida, S. Kato, Role of Asp102 in the Catalytic Relay System of Serine Proteases: A Theoretical Study, *JACS*, 126 (2004) 7111-7118.
- [13] S. Raza, L. Fransson, K. Hult, Enantioselectivity in *Candida antarctica* lipase B: A molecular dynamics study, *Protein science*, 10 (2001) 329-338.
- [14] W. Li, e. al, Molecular modeling of substrate selectivity of *Candida antarctica* lipase B and *Candida rugosa* lipase towards c9, t11- and t10, c12-conjugated linoleic acid, *J. Mol. Catal. B: Enzymatic*, 57 (2009) 299-303.
- [15] G. Fuentes, A.V. Ballesteros, C., Specificity in lipases: A computational study of transesterification of sucrose, *Protein science*, 13 (2004) 3092-3103.
- [16] L. Villo, A. Metsala, O. Parve, T. Pehk, Chemical versus enzymatic acetylation of alpha-bromo- -hydroxyaldehydes: decyclization of hemiacetals by lipase, *Tetrahedron Lett.*, 43 (2002) 3203-3207.
- [17] P. Trodler, J. Pleiss, Modeling structure and flexibility of *Candida antarctica* lipase B in organic solvents, *BMC Structural Biology*, 8:9 (2008) 1-10.
- [18] O. Trott, A. Olson, Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading, *Journal of Computational Chemistry*, 31 (2009) 455-461.
- [19] R. Salomon-Ferrer, D.A. Case, R.C. Walker, An overview of the Amber biomolecular simulation package, *WIREs Comput Mol Sci* 00 (2012) 1-13.
- [20] E. Arunan, e. al, Defining the hydrogen bond: An account (IUPAC Technical Report)\*, *Pure Appl. Chem.*, Vol. 83, No. 8, pp. 1619–1636, 2011., 83 (2011) 1619-1636.
- [21] P. Sigala, e. al, Determination of Hydrogen Bond Structure in Water versus Aprotic Environments To Test the Relationship Between Length and Stability, *JACS*, 137 (2015) 5730-5740.
- [22] J. Uppenberg, e. al, Crystallographic and Molecular-Modeling Studies of Lipase B from *Candida antarctica* Reveal a Stereospecificity Pocket for Secondary Alcohols, *Biochemistry*, 34 (1995) 16838-16851.
- [23] A. Smith, X. Zhang, A. Leach, K. Houk, Beyond Picomolar Affinities: Quantitative Aspects of Noncovalent and Covalent Binding of Drugs to Proteins, *J. Med. Chem.*, 52 (2009) 225-233.
- [24] V. Lafaquire, e. al, Control of Lipase Enantioselectivity by Engineering the Substrate Binding Site and Access Channel, *Chem. Bio. Chem.*, 10 (2009) 2760-2771.
- [25] J. Grunenberg, Complexity in molecular recognition, *Phys. Chem. Chem. Phys.*, 13 (2011) 10136-10146.
- [26] R. Baron, J. McCammon, Molecular Recognition and Ligand Association, *Ann. Rev. Phys. Chem.*, 64 (2013) 151-175.
- [27] S. Brand, e. al, The First Examples of Dynamic Kinetic Resolution by Enantioselective Acetylation of Hemithioacetals: An Efficient Synthesis of Homochiral o-Acetoxy sulfide, *Tetrahedron Lett.*, 36 (1995) 8493-8497.