CHEMBIOCHEM

Supporting Information

DOI: 10.1002/cbic.200900128

Supporting Information

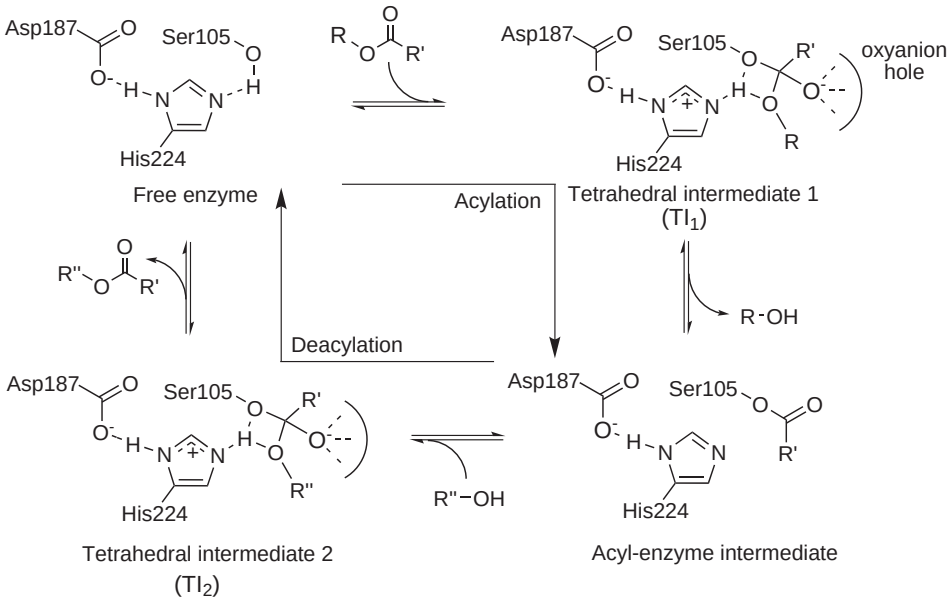
for

Lactone Size Dependent Reactivity in CALB: A Molecular Dynamics and Docking Study

Martijn. A. J. Veld, Linda Fransson, Anja R. A. Palmans*, E. W. Meijer, and Karl Hult*

Table of Contents

S1 CALB MECHANISM



Scheme 1 *Common reaction mechanism[1] of serine hydrolases (residue numbering according to CALB structure) showing the presence of an acyl-enzyme intermediate (bottom right) and two tetrahedral intermediates: TI_1 (top right) and TI_2 (bottom left).*

After attack of the serine 105 nucleophile onto the substrate in the Michaelis-Menten (enzyme-substrate) complex a first tetrahedral intermediate (TI_1) is formed (Scheme 1, top right). Subsequently, the alcohol part of the substrate is liberated by histidine 224 assisted proton transfer to the alkyl oxygen of the substrate giving an acyl-enzyme intermediate (Scheme 1, bottom right). This acyl-enzyme intermediate subsequently reacts with a nucleophile through a second tetrahedral intermediate (TI_2, Scheme 1, bottom left). After dissociation of the enzyme-product complex the free enzyme is regenerated (Scheme 1, top left). The formation of the acyl-enzyme intermediate is generally accepted to be the determining step for the ring opening of unsubstituted lactones.[2]

S2 RING OPENING KINETICS MEASUREMENTS

S2.1 Experimental details

Two separate one-pot experiments were performed. The ring opening kinetics of δ-valerolactone (VL), ε-caprolactone (CL), and dodecanolactone (DDL) were measured simultaneously (table 1, experiment A) and heptanolactone (HL), DDL, and pentadecanolactone (PDL) were measured in one pot (table 1, experiment B).

S2.1.1 Materials and equipment

All used glassware was dried at 120 °C for at least 2 h before use and directly put under an Argon atmosphere afterwards.

VL, CL, cycloheptanone, cyclododecanone, *m*-CPBA, PDL were purchased from Sigma-Aldrich and used as received unless otherwise stated. All solvents were obtained from Biosolve (The Netherlands). Novozym 435, immobilized *Candida antarctica* lipase B on a polyacrylic resin, was obtained from Novozymes A/S.

Toluene was distilled over Na before use and stored over 4 Å molecular sieves under an Argon atmosphere. Novozym 435 was dried overnight at 50 °C under vacuum over P_2O_5 and the vacuum was released by backfilling the oven with nitrogen.

HL and DDL were synthesized according to a literature procedure.[3] VL, CL, and HL were purified by distillation over CaH_2 and subsequently stored under an argon atmosphere in a closed flask. Freshly distilled VL was stored in the refrigerator at 4 °C to avoid spontaneous polymerization. The synthesized HL was contaminated with residual cycloheptanone, which could not be removed successfully. A correction for the presence of cycloheptanone (11 mol% according to ^{1}H-NMR analysis based on the integral values at δ = 4.4 ppm and δ = 2.6 ppm) was made during stock solution preparation. The influence of the presence of cycloheptanone on the enzyme kinetics was negligible.

GC-FID analysis was performed on a Shimadzu GC-17A equipped with a FID, an AOC-20i autosampler and a Zebron ZB-5MS column (30 m, id 0.25 mm, 0.25 μm film thickness) using helium as carrier gas. The following settings were used: T_{inj} = 300 °C, T_{det} = 300 °C, column velocity = 30 cm/s, V_{inj} = 1 μL). Temperature program: 100 °C | 2 min → 260 °C 20 °C/min | 4 min. An overview of the retention times is given in Table 1. Data analysis of the GC-traces was performed using DAX 8.0 data acquisition and data analysis software. Lactone conversion (X) was calculated from the initial relative peak areas and the relative peak areas with respect to the internal standard. The apparent reaction rates for the individual components were obtained by linear least-squares curve fitting of ln(1-X) versus time plots through the origin assuming first order reaction kinetics.

Table 1: *Overview of retention times for the GC-FID analysis of the used lactones and the corresponding ring-opening products with 1-propanol.*

Component[a]	t_R lactone [min]	t_R product [min]
VL (A)	3.89	5.71
CL (A)	4.75	6.60
HL (B)	4.15	7.41
DDL (A/B)	8.15	10.78
PDL (B)	10.11	13.11
1,3,5-tri-t-butyl-benzene (IS)	7.05	-
cycloheptanone	3.56	-

a (A): lactones used for one-pot experiment A;
(B): lactones used for one-pot experiment B;
(IS) internal standard.

S2.1.2 Ring opening kinetics measurement

Stock solutions were prepared by weighing the desired amounts of the three lactones (2.5 mmol), 1,3,5-tri-*t*-butylbenzene (0.75 mmol, 0.10 eq., internal standard) and 1-propanol (15.0 mmol) into a volumetric flask after which the volume was adjusted to 25.00 mL by addition of dry toluene.

Stock solution (8.0 mL) was transferred by syringe into a dried Schlenk tube and placed under an argon atmosphere. After stirring for 15 minutes at 60 °C, dried Novozym 435 (25 mg) was added. Small samples of the reaction mixture (~0.1 mL) were taken at regular intervals for a period of 3 hours (experiment A) or 90 minutes (experiment B) and directly filtered over a piece of cotton in a pipette to remove the enzyme from the sample. The pipette was carefully rinsed with DCM (1.5 mL) and the samples were subsequently analyzed by GC-FID using the procedure described above.

S2.2 Results for ring opening competition experiments

Samples in experiment A (VL, CL and DDL) were analyzed in triplo by GC-FID, samples in experiment B (HL, DDL and PDL) were measured in duplo.

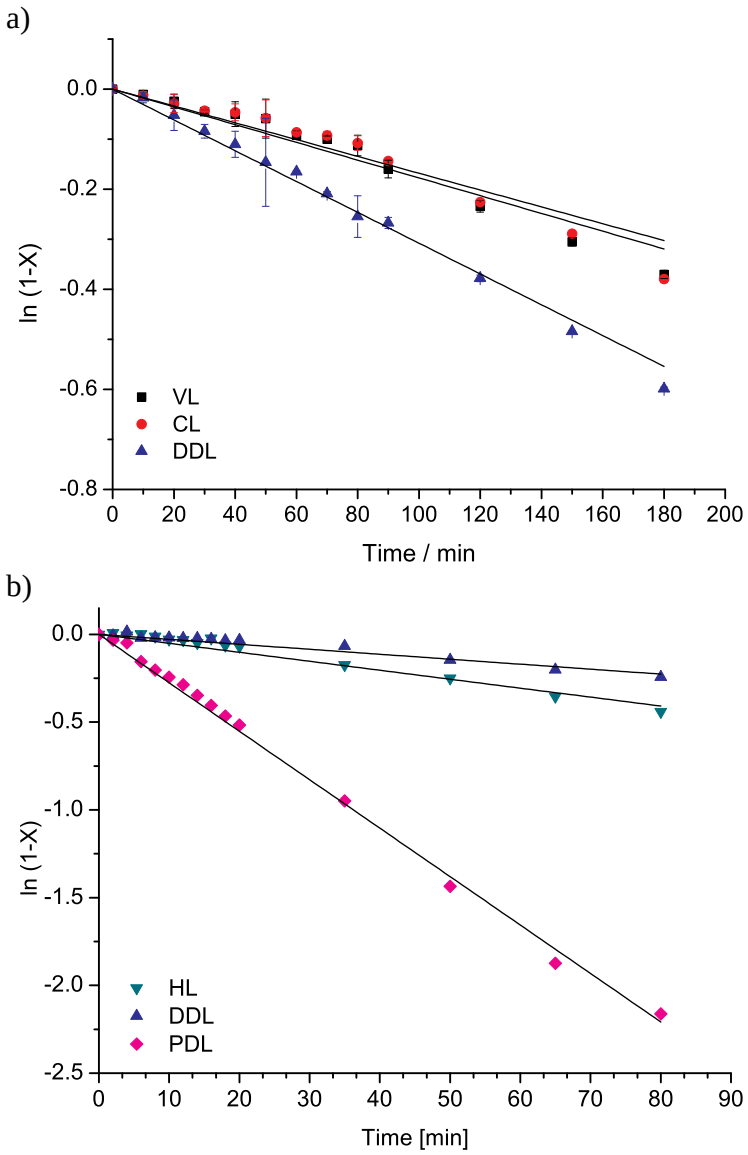


Figure 1: *Kinetic plots showing curve fits of time vs ln(1-X) for competition experiments using Novozym 435 ($C_{lactone}$=0.10 M, $C_{1\text{-}propanol}$=0.6 M, T=60 °C, V=8.0 mL). a) VL, CL, and DDL (24.0 mg dry N435) b) HL, DDL, and PDL (24.5 mg dry N435).*

S3 DOCKING STUDIES

S3.1 Experimental details

S3.1.1 General information

Preparation of the CALB receptor structure was performed using Sybyl 7.2[4] on a SGI-IRIX workstation. All energy calculations were performed using a distance dependent dielectric constant and a non-bonding (NB) cutoff of 8 Å using Kollman All atom type definitions. All MD calculations were performed using random initial velocity (Boltzmann distributed), a temperature of 300 K, a step size of 1 fs, non-bonded update every 25 fs, and a coupling factor of 100 fs. All energy minimizations were done using an initial Simplex optimization when required, followed by a Powell minimization with a termination gradient of 0.05 kcal/mol.

The docking calculations were performed on a UNIX cluster with a total of 32 nodes (Dual 2.8GHz Intel XEON processors, 2 GB internal memory, 1 main node) using a pre-compiled version of Autodock 4.0 and Autogrid 4.0. Docking runs were run in parallel and submitted using the Portable Batch Scheduler (PBS). Preparation of the 3D input files and data analysis were performed using a desktop PC running under windows XP SP2.

S3.1.2 Receptor preparation

Candida antarctica lipase B X-ray structure[5] with PDB-code 1TCA was used as a starting structure and was obtained from the RCSB Protein Databank.[6] The NAG401/NAG402 sugar groups were removed from the structure after which the chain ends were fixed to charged residues. Randomly oriented hydrogen atoms were added to the enzyme and water molecules and residue Asp134 was protonated. Subsequently, Kollmann All atom type definitions were assigned to the complete structure.

The position of the protons was optimized using a stepwise procedure consisting of short (1.2 ps) MD simulations followed by energy minimization. First the position of the protons on the water molecules were optimized, followed by the protons on the enzyme structure. This series of steps was repeated until the energy was stable. Four water residues (HOH130, HOH149, HOH238, and HOH285) were removed from the active site to make space for the substrate intermediates. Finally, the whole system including heavy atoms was energy minimized. RMSD value of backbone heavy atoms: 0.439 Å compared with the starting structure in PDB-file 1TCA.

Blocking groups at the chain ends from the completely minimized empty enzyme structure were removed. The enzyme structure was transferred to AutoDocktools (ADT) 1.3.4.[7] and the present Kollman All atom charges were kept. Non-polar hydrogen atoms in the enzyme structure were merged into the connected heavy atoms. The energy grid (60×60×60 grid points, 0.375 Å spacing) was centered over the active site of the enzyme and covered the complete active site, neighboring residues and active site entrance.

Calculations were performed for the A C H HD N NA OA SA receptor atom types and the C OA H ligand atom types using a smoothing factor of 0.5 and a default value of -0.1456 for the distance dependent dielectric constant. The used grid parameter file for Autogrid4 is given in section S3.1.7.

S3.1.3 Preparation of ligand files with the lactone conformations

The low energy lactone conformations of VL, CL and HL were built using Cambridgesoft Chem3D 11.0[8] according to the structures described in literature.[9] Protons were added automatically after which the structures were minimized in a stepwise manner using the MM2 force field. Default energy minimization settings (gradient limit: 0.100) were used during all steps. First all protons were minimized in energy, followed by minimization of all protons and the carbonyl group and a final minimization of the complete structure. Structures were centered on the origin and saved in the .pdb-file format. After minimization all CL and HL conformations were in good agreement with the described structures (maximum deviations in bond length ~0.018 Å, bond angle 0.5°, and dihedral angle 3.0°).[9] Small differences in energy with the published data were observed (Table 3), which is due to the use of slightly different parameters in different versions of the MM2 force field.

Ligand files for docking were prepared using the prepare_ligandv4.py script from the ADT 1.3.4-package using the previously made MM2 energy-minimized ligand conformations. Non-polar hydrogen atoms were kept and Gasteiger charges were assigned to the ligand files.

Energetic and structural information on the various lactone conformations of VL, CL and HL is given in section S3.1.7.

S3.1.4 Docking process

The preparation of the receptor structure and calculation of the corresponding affinity grids is described in section S3.1.1. Used ligand files were all prepared as described in section S3.1.3. The built in Lamarckian Genetic Algorithm present in Autodock 4.0 was used for the actual docking process.

Optimization of the parameters for the genetic algorithm lead to the following parameters that have been used for all 256 docking runs per conformation: initial population size 150, maximum number of energy evaluations: $5 \cdot 10^6$, maximum number of generations: $27 \cdot 10^3$. Parameters for the Solis and Wets local search algorithm were set to their default values. A default docking parameter (.dpf) file is given in section S3.1.8.

S3.1.5 Results analysis

Log files with docking results were processed using self-written Unix shell scripts. Desired information such as ligand coordinates, energy and orientation was extracted to generate multi-model .pdb-files for further analysis with VMD 1.8.6.[10] Information on distances, angles and dihedral angles (Section S3.1.6) between the enzyme and the docked substrate were exported using a self-written TCL-script in VMD. These exported data were used to calculate the fitness of a docked position.

S3.1.6 Scoring of docking results

Average scores and docking energies were calculated using an UNIX shell script for the docked states in cluster 1 only, using a 0.5 Å cutoff value for the RMSD of the docked state with respect to the docked state with the lowest energy. For the calculation of the fitness a total of five conditions were taken into account (Table 2).

All distances mentioned shown in Figure 2a were assigned a score with a linear dependence between 2.5 Å and 3.5 Å (Table 2; conditions 1.1, 1.2, 1.3, 2, and 3). A full score was given for distances shorter than 2.5 Å. Possible hydrogen bonding in the oxyanion hole (Table 2; condition 1) was only taken into account if at least 2 distances of the possible hydrogen bonds (Table 2; conditions 1.1, 1.2 and 1.3) were shorter than 3.5 Å.

Bond angles (Table 2 condition 4) and dihedral angles (Table 2 condition 5) were assigned a score that quadratically depended on the deviation from the optimum value. The overall score (S_{total}) was obtained by multiplication of the partial scores S_1-S_5 followed by normalization. An overall score of 1 means a perfect match with the expected NAC and indicates a possibility for rapid transformation of the docked state into the acyl-enzyme intermediate due to very similar positioning of the atoms in both cases. A zero score means that at least one of the prerequisites is not fulfilled correctly.

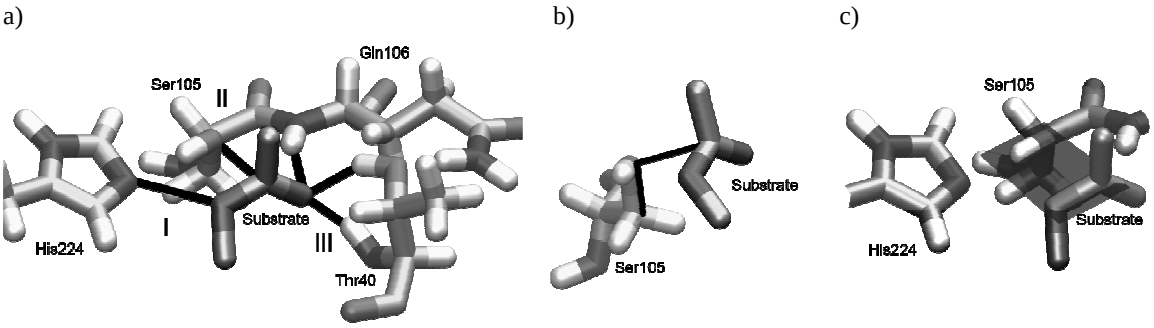


Figure 2: *Definition of scoring function: a) distance requirements between substrate(only partly shown for clarity) and active site residues (in all cases < 3.5 Å) I: His224N-substrate alkyl oxygen II: Ser105 oxygen-substrate carbonyl C; III: Distance Carbonyl O-oxyanion hole (2 out of three possible H-bonds required) b) Angular dependence for angle of attack of Ser105 O on substrate carbonyl C (109 ° ± 20 °). c)Dihedral angle between His224N-Ser105H-Ser105O-substrate alkyl (0° ±45°).*

Table 2: *Overview of score calculation for docked ligand positions.*

Condition[a]		Requirement	Score
1[b]	H-bonding in oxyanion hole	at least two of three possible H-bonds present	$S_1 = \frac{1}{3}\sum_{i=1}^{3} S_{hb(i)}$
1.1[c]	Distance (d) L.$O_{carbonyl}$ - THR40.H	0 < d < 3.0 A	$\begin{cases} d < 2.5\ Å \rightarrow S_{hb1} = 1 \\ 2.5\ Å < d < 3.5\ Å \rightarrow S_{hb1} = (3.5 - d) \end{cases}$
1.2[c]	Distance (d) L.$O_{carbonyl}$ - THR40.HG1	0 < d < 3.0 A	$\begin{cases} d < 2.5\ Å \rightarrow S_{hb2} = 1 \\ 2.5\ Å < d < 3.5\ Å \rightarrow S_{hb2} = (3.5 - d) \end{cases}$
1.3[c]	Distance (d) L.$O_{carbonyl}$ - GLN106.H	0 < d < 3.0 A	$\begin{cases} d < 2.5\ Å \rightarrow S_{hb3} = 1 \\ 2.5\ Å < d < 3.5\ Å \rightarrow S_{hb3} = (3.5 - d) \end{cases}$
2	Distance (d) L.$C_{carbonyl}$ - SER105.OG	0 < d < 3.0 A	$\begin{cases} d < 2.5\ Å \rightarrow S_2 = 1 \\ 2.5\ Å < d < 3.5\ Å \rightarrow S_2 = (3.5 - d) \end{cases}$
3	Distance (d) L.O_{alkyl} - HIS224.NE2	0 < d < 3.0 A	$\begin{cases} d < 2.5\ Å \rightarrow S_3 = 1 \\ 2.5\ Å < d < 3.5\ Å \rightarrow S_3 = (3.5 - d) \end{cases}$
4	Angle (a) SER105.CB - SER105.OG - L.$O_{carbonyl}$	89 < a < 129	$S_4 = 1 - \left(\frac{a - 109}{20}\right)^2$
5	Dihedral (d) L.O_{alkyl} - L.$C_{carbonyl}$ - SER105.OG - SER105.HG	-45 < d < 45	$S_5 = 1 - \left(\frac{d}{45}\right)^2$
	Overall score		$S_{total} = \sum_{i=1}^{5} S_i$

[a] Atoms are indicated as follows: RESIDUE.ATOM with the atom numbering following the Yasara defnitions. Amide protons in the backbone are indicated by H and atoms in the residue side chain have a two-letter abbreviation (A=Alpha, B=beta, C=gamma etc). Atoms in the ligand are indicated by L.*

[b] At least two possible H-bond distances need to be less than 3 Å

[c] Score per H-bond, normalized over all three possible hydrogen bonds

S3.1.7 Grid parameter file (.gpf)

```
npts 60 60 60                       # num.grid points in xyz
gridfld GRIDFILE_D.maps.fld         # grid_data_file
spacing 0.375                       # spacing(A)
receptor_types A C H HD N NA OA SA  # receptor atom types
ligand_types C OA H                 # ligand atom types
receptor RECEPTOR_FILE.pdbqt        # macromolecule
gridcenter 8.203 2.306 7.383        # xyz-coordinates or auto
smooth 0.5                          # store minimum energy w/in rad(A)
map GRIDFILE.C.map                  # atom-specific affinity map
map GRIDFILE.OA.map                 # atom-specific affinity map
map GRIDFILE.H.map                  # atom-specific affinity map
elecmap GRIDFILE.e.map              # electrostatic potential map
dsolvmap GRIDFILE.d.map             # desolvation potential map
dielectric -0.1465                  # <0, AD4 distance-dep.diel;>0, constant
```

S3.1.8 Default docking parameter file (.dpf)

```
outlev 1                            # diagnostic output level
intelec                             # calculate internal electrostatics
seed pid time                       # seeds for random generator
ligand_types C H OA                 # atoms types in ligand
fld GRIDFILE.maps.fld               # grid_data_file
map GRIDFILE.C.map                  # atom-specific affinity map
map GRIDFILE.H.map                  # atom-specific affinity map
map GRIDFILE.OA.map                 # atom-specific affinity map
elecmap GRIDFILE.e.map              # electrostatics map
desolvmap GRIDFILE.d.map            # desolvation map
move LIGANDFILENAME                 # small molecule
tran0 random                        # initial coordinates/A or random
about 0.0 0.0 0.0                   # small molecule center
quat0 random                        # initial quaternion
ndihe 0                             # number of active torsions
dihe0 random                        # initial dihedrals (relative) or random
tstep 2.0                           # translation step/A
qstep 50.0                          # quaternion step/deg
dstep 50.0                          # torsion step/deg
torsdof 0 0.274000                  # torsional degrees of freedom and coefficient
rmstol 2.0                          # cluster_tolerance/A
extnrg 1000.0                       # external grid energy
e0max 0.0 10000                     # max initial energy; max number of retries
ga_pop_size 150                     # number of individuals in population
ga_num_evals 5000000                # maximum number of energy evaluations
ga_num_generations 27000            # maximum number of generations
ga_elitism 1                        # number of top individuals to survive to next
generation
ga_mutation_rate 0.02               # rate of gene mutation
ga_crossover_rate 0.8               # rate of crossover
ga_window_size 10                   #
ga_cauchy_alpha 0.0                 # Alpha parameter of Cauchy distribution
ga_cauchy_beta 1.0                  # Beta parameter Cauchy distribution
set_ga                              # set the above parameters for GA or LGA
sw_max_its 300                      # iterations of Solis & Wets local search
sw_max_succ 4                       # consecutive successes before changing rho
sw_max_fail 4                       # consecutive failures before changing rho
sw_rho 1.0                          # size of local search space to sample
sw_lb_rho 0.01                      # lower bound on rho
ls_search_freq 0.06                 # probability of performing local search on individual
set_sw1                             # set the above Solis & Wets parameters
compute_unbound_extended            # compute extended ligand energy
ga_run 256                          # do this many hybrid GA-LS runs
analysis                            # perform a ranked cluster analysis
```

S4 SCHEMATIC REPRESENTATION OF DOCKING RESULTS

S4.1 Overview of total energies and docking scores

A schematic overview of all total energies and corresponding docking scores for the investigated lactone conformations is given in Figure 3. All VL and CL conformations show little resemblance with the ideal near attack complex, whereas the low energy docked states HL1 and HL2 show better correspondence.

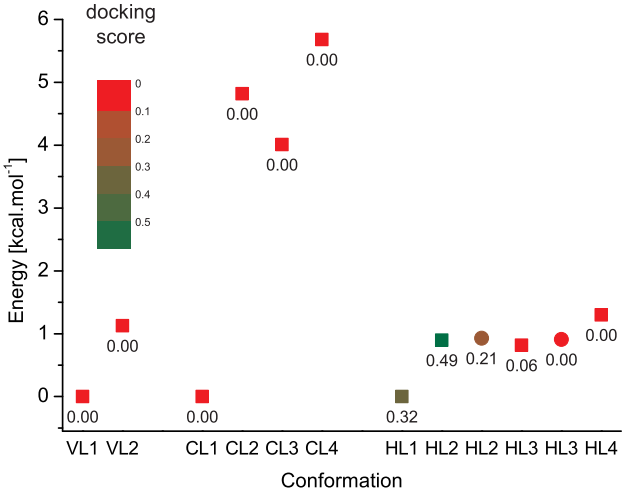


Figure 3: *Overview of docking results for all investigated VL, CL and HL conformations. Results for the first cluster of solutions are shown in squares, second clusters larger than 100 results are shown in circles. The numbers below the data points are the corresponding docking scores.*

S4.2 Dependence of docking score on dihedral angle

A large negative dihedral angle around the lactone ester bond is required for a good fit in the CALB active site indicating a preference for transoid lactones (Figure 4). All cisoid ester conformations with dihedral angles around 0 ° show a docking score of 0.00 corresponding to a clear mismatch between the substrate and the enzyme active site.

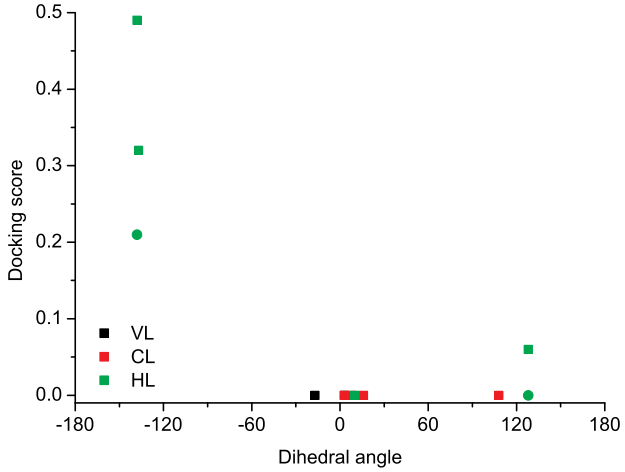


Figure 4:*Dependence of docking score on the dihedral angle around the ester bond for all investigated lactone conformations. The docking score for clusters with more than 100 docked positions are given in circles.*

S5 MD SIMULATIONS

S5.1 Experimental details

S5.1.1 General information

All molecular dynamics (MD) simulations were performed using Yasara[11] on a Linux-based computer cluster using 4 CPUs and 1.0 Gb internal memory per run.

All energy calculations were performed using the AMBER 96 force field[12] using a 7.86 Å cutoff and the Particle Mesh Ewald algorithm[13] was used to treat longrange electrostatic interactions. A timestep of 1 fs for intramolecular and 2 fs for intramolecular interactions was used and the temperature was set to 300 K during the simulations unless otherwise noted. No explicit solvent was used during the simulations. An automated assignment of atom types and charges was used for the complete structure.

S5.1.2 Preparation of free CALB structure

Candida antarctica lipase B X-ray structure[5] with PDB-code 1TCA was used as a starting structure and was obtained from the RCSB Protein Databank.[6] An N-terminal acetyl and a C-terminal N-methyl group were added to the chain end and hydrogen atoms were added to the complete structure using the add hydrogens command in Yasara.

The complete enzyme structure was optimized in a stepwise manner using a series of energy minimization experiments. To remove possible bumps in the covalent geometry each minimization step was started with a steepest descent minimization, followed by a simulated annealing (time step 2 fs, atom velocities scaled down by 0.9 every 10th step) until convergence was reached, i.e. the energy improved by less than 0.1% during 200 steps.

First, the position of the protons on the water molecules were optimized, followed by the protons on the enzyme structure. Then the position of the water oxygen atoms was optimized followed by minimization of the complete water molecules. Subsequently the position of all hydrogen atoms in the structure was optimized. Finally, the position of the side chains was optimized followed by a minimization of the complete structure including the enzyme backbone. RMSD value of backbone atoms: 0.444 Å compared with the starting structure in PDB-file 1TCA. This energy-minimized structure was used as a starting point for the preparation of the tetrahedral intermediate structures.

S5.1.3 Preparation of Enzyme with Tetrahedral Intermediates Structures

Four water residues (HOH130, HOH149, HOH238, and HOH285) were removed from the active site to make space for the lactone tetrahedral intermediates. Protons were added to residues Asp134 (0 net charge on residue) and His224 (+1 net charge on residue). The tetrahedral intermediates were made by manual addition of heavy atoms to the Ser105 O_γ atom. The alkyl backbone of the lactone substrate was positioned such that anti- or gauche-conformations were obtained as much as possible. The tetrahedral center had an

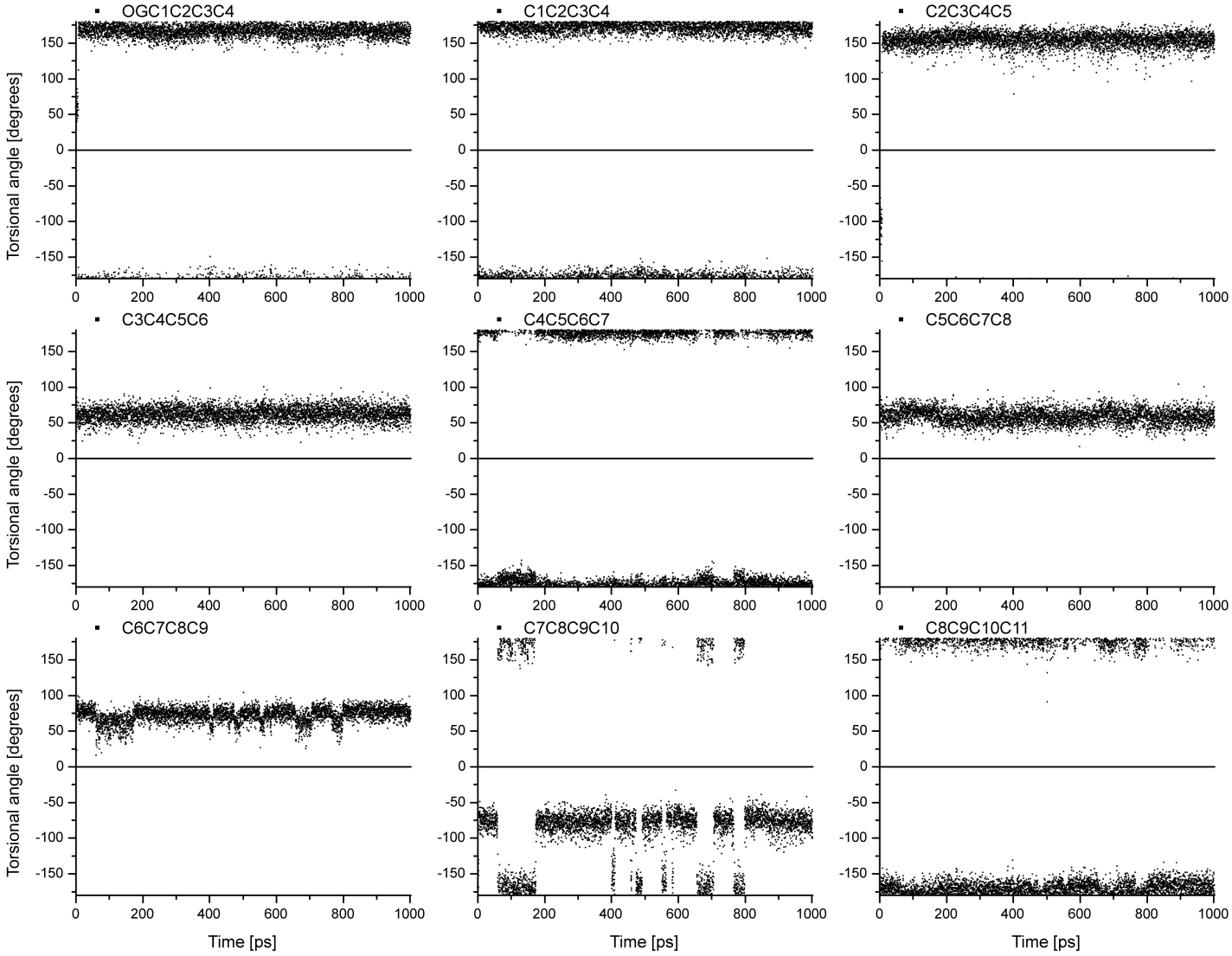
R-configuration. Positioning was such that distances between the oxyanion of the substrate and the oxyanion hole (consisting of atoms Gln106.H, Thr40.H, and Thr40.HG1) were < 3 Å. Distances between His224.HE2 and the alkyl oxygen atom of the tetrahedral intermediate and the Ser104.OG atom were also < 3 Å in order to have the possibility of hydrogen bonding. After addition of the tetrahedral intermediate heavy atoms, hydrogen atoms were added. The hydrogen atom at the oxyanion was removed, which resulted in a -1 net charge on the complete residue (-0.79 charge on the oxyanion atom).

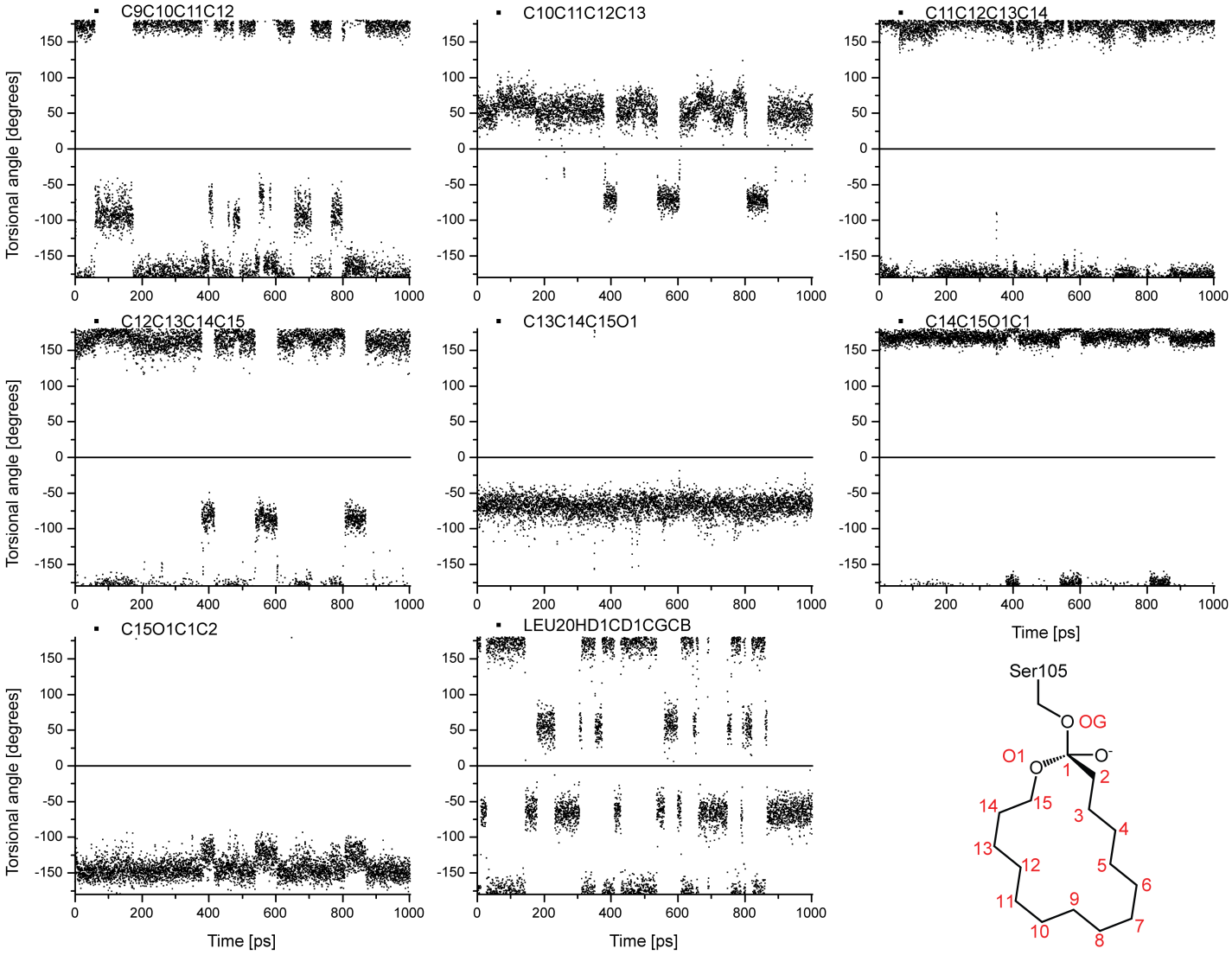
The geometry of the structure was optimized in a stepwise manner as described above in section S5.1.2. First the newly generated tetrahedral carbon atom was energy minimized, followed by energy minimization of the complete tetrahedral centre (1 C and 3 O atoms). Then, the position of the hydrogen atoms on the tetrahedral intermediate was optimized. Subsequently, the complete side chain of residue 105 was energy minimized followed by a minimization including the sidechains of residues Thr40 and His224. In a final step, the complete enzyme structure was energy minimized. RMSD values for the backbone atoms between the empty enzyme structure and the tetrahedral intermediate structures were less than 0.11 Å for all structures.

S5.1.4 Molecular Dynamics Simulations

The completely energy minimized enzyme-tetrahedral intermediate structures as prepared in section S5.1.3 were used in the MD simulations. Parameters were set to their default values as described in section S5.1.1. The simulation was started using a pre-defined Yasara macro. The enzyme structure was slowly heated from 0 K to 300 K in 10K increments by rescaling of the temperature over a total time of 15 fs. Then the actual 1 ns MD simulation was run and structural information was saved every 200 fs (total of 5000 structures per MD run). Energies, RMSD values and hydrogen bond distances between the tetrahedral intermediate and the enzyme were exported using a self-written script for further data analysis.

Figure 5 (pages SI 13 and SI 14): *overview of dihedral angles of PDL backbone during the 1 ns MD simulation. Numbering of the atoms of the PDL backbone is given at the bottom right. Succeeding graphs show dihedral angles around the next dihedral angle in a clockwise direction. The dihedral angle given by $O_GC_1C_2C_3C4$ is shown at the top left of the figure. The change in dihedral angle by free rotation of a methylgroup of residue Leu23 is shown as a reference.*


OGC1C2C3C4
C1C2C3C4
C2C3C4C5
C3C4C5C6
C4C5C6C7
C5C6C7C8
C6C7C8C9
C7C8C9C10
C8C9C10C11
Torsional angle [degrees]
Time [ps]
150
100
50
0
-50
-100
-150
0
200
400
600
800
1000


C9C10C11C12
C10C11C12C13
C11C12C13C14
C12C13C14C15
C13C14C15O1
C14C15O1C1
C15O1C1C2
LEU20HD1CD1CGCB
Torsional angle [degrees]
Time [ps]
Ser105
OG
O1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

S6 DOCKING RESULTS

S6.1 Low Energy Lactone Conformations

Relative ground state energies for the most important lactone conformations of VL, CL, and HL are given in Table 3. Most of the small lactone conformations are cisoid in nature as indicated by the small dihedral angles around atoms C_N-O_{alkyl}-C_1-C_2 (Table 3). For HL three more transoid lactone conformations were observed together with a single cisoid conformation (HL4) which is highest in energy. However, in polar environments this more cisoid conformation can become more stable due to stronger solvation.[9]

Table 3: *Overview of energies for the MM2 minimized lactone conformations of VL, CL, HL.*

Conformation[a]	Relative Energy MM2 / literature [kcal/mol][b]	Dihedral angle [°][c]	cisoid / transoid
VL 1	+0.00 / +0.00	+3.5	cisoid
VL 2	+1.23 / +0.50	-16.6	cisoid
CL 1	+0.00 / +0.00	+8.0	cisoid
CL 2	+2.57 / +2.72	+3.2	cisoid
CL 3	+4.28 / +4.24	+16.0	cisoid
CL 4	+4.98 / +5.32	+107.8	transoid
HL 1	+0.00 / +0.00	-136.8	transoid
HL 2	+0.81 / +1.06	-137.5	transoid
HL 3	+0.91 / +1.02	+127.5	transoid
HL 4	+1.27 / +1.31	+9.7	cisoid

[a] Conformations are numbered in increasing order of relative ground state energy as determined after MM2 energy minimization in the gas phase.

[b] Energies were obtained using the Cambridgesoft Chem3D MM2 force field using a minimum RMS gradient of 0.01. Calculated data correspond to literature data.[9]

[c] Dihedral angle C_N-O_{alkyl}-C_1-C_2. Carbon atoms are numbered starting from the carbonyl carbon atom (C_1). The terminal carbon atom is indicated with C_N.

No influence of the presence of water molecules in the enzyme structure on interaction energies or docking positions was found for any of the lactone conformations. Furthermore, only minor differences in interaction energy and docking score were observed for docking of the various lactone conformations into non-energy minimized enzyme structures obtained from a MD run.

S6.2 Graphical representation of docked lactone conformations

The docked lactone conformations of VL, CL and HL in the active site of CALB are given in Figure 6 - Figure 8 from two different viewpoints for each conformation. Clusters with RMSD $\leq$ 0.50 Å from the best observed docking position are shown in cluster 1 for each of the substrates. Docking solutions with RMSD $\geq$ 0.50 Å have been put in cluster 2 which is also shown if applicable. For each of the clusters, the number of dockings in the cluster, the average RMSD value with respect to the best solution, docking and relative total energies are given in the caption. Numbering of conformations is in order of increasing ground state energy (E_{intra}).

In the figures at the right side, the side chain of residue Gln106 is hidden to give a clearer

view on the docking position of the substrates. To make comparison between various conformations with respect to the enzyme active site easier a yellow sphere was added to the carbonyl oxygen atom of all substrates. A purple sphere was added to the alkyl oxygen atom. To guide the eye, dotted lines have been added between the three atoms forming the oxyanion hole and the carbonyl oxygen atom and between the alkyl oxygen atom and the N_ε atom of His224.

Figure 6 shows the docking of both conformations of VL. The relative energies are relatively close together and are largely determined by the ground state energy of the lactone. All docked positions for both VL conformations show incorrect positioning of the alkyl oxygen in the oxyanion hole region.

Docking of CL conformation 1 (Figure 7a) shows rather good positioning of the ligand, although the distances between the carbonyl carbon and Ser105OH and between the alkyl oxygen atom and His224 are rather large. CL conformation 2 (Figure 7b) shows completely wrong binding of the substrate as the alkyl oxygen shows interaction with the oxyanion hole residues instead of the carbonyl oxygen atom. CL conformations 3 and 4 (Figure 7c and Figure 7d) are high in energy and reveal an incorrect angle of attack of the Ser105OH onto the carbonyl carbon atom. Moreover, for CL conformation 3, the position of the alkyl oxygen atom is too far away from His224.

All relative energies for the docking of the 4 HL conformations are close together (less than 1.5 kcal/mol difference). Both HL conformations 1 (Figure 8a) and 2 (Figure 8b) show very favorable positioning of the substrate with respect to the enzyme. Conformation HL 4 (Figure 8d) shows unproductive binding due to incorrect placement of the carbonyl oxygen atom with respect to the oxyanion hole.

6.2.1.1 δ-Valerolactone (VL) conformation 1 E_{intra}= 0.00 kcal.mol^{-1}

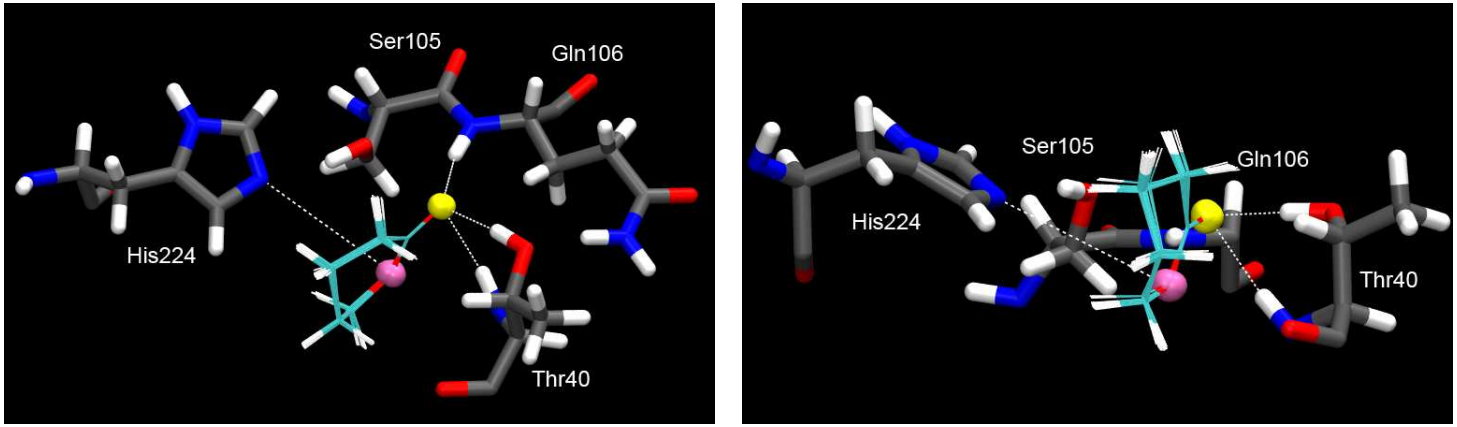


Figure 6a1: *Cluster 1: 220 structures, <RMSD> 0.04 Å, E_{inter} -3.94 kcal/mol, $E_{total,rel}$ 0.00 kcal/mol, <docking score> 0.00*

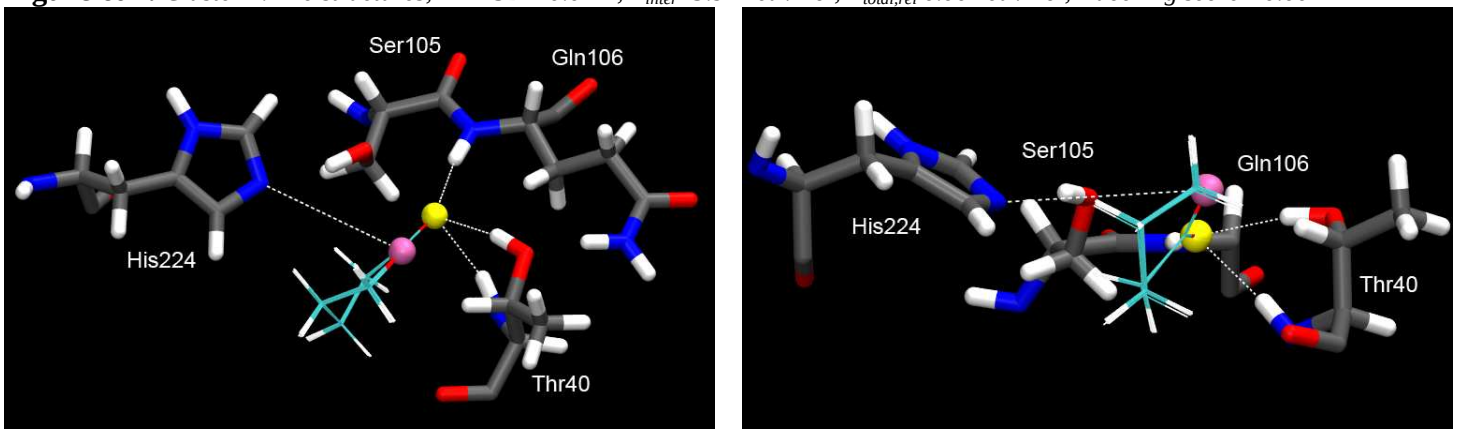


Figure 6a2: *Cluster 2: 36 structures, <RMSD> 1.17 Å, E_{inter} -3.90 kcal/mol, $E_{total,rel}$ 0.04 kcal/mol, <docking score> 0.00*

6.2.1.2 δ-Valerolactone (VL) conformation 2 E_{intra}= 0.13 kcal.mol^{-1}

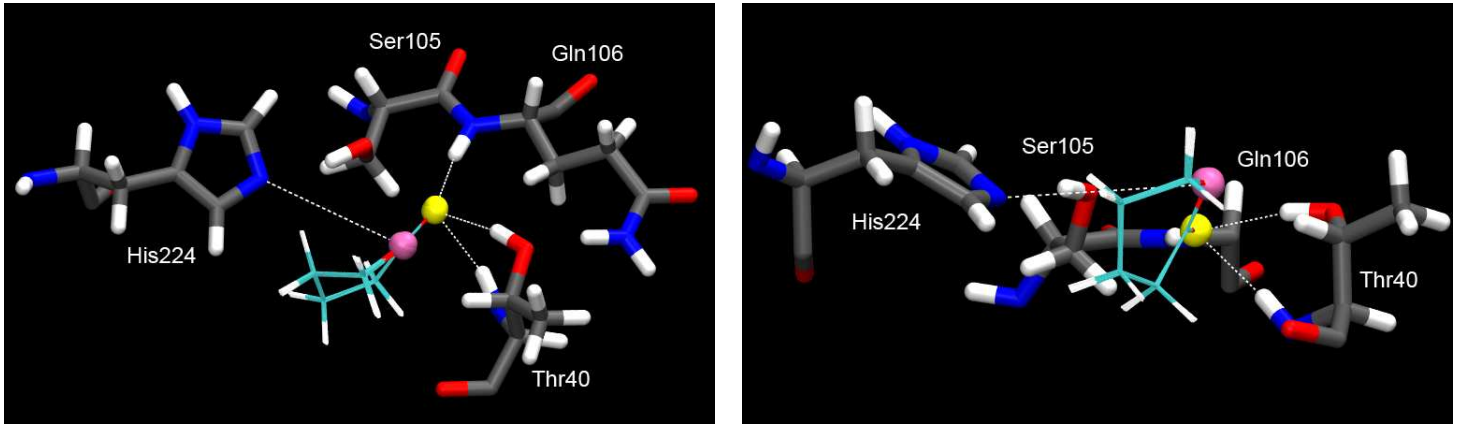


Figure 6b1: *Cluster 1: 182 structures, <RMSD> 0.03 Å, E_{inter} -3.86 kcal/mol, $E_{total,rel}$ 1.31 kcal/mol, <docking score> 0.00*

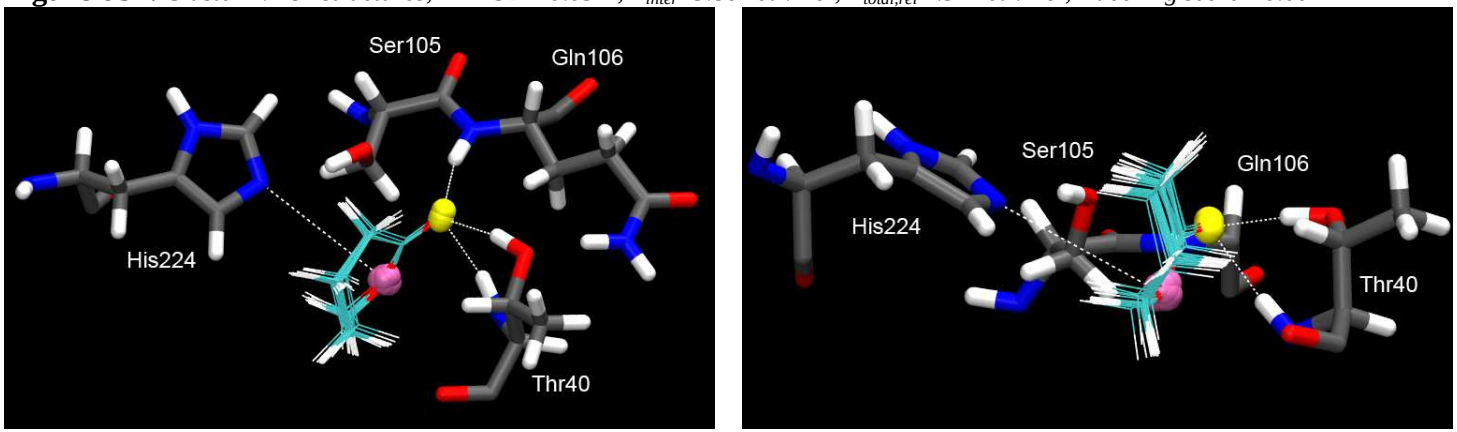


Figure 6b2: *Cluster 2: 74 structures, <RMSD> 1.18 Å, E_{inter} -3.82 kcal/mol, $E_{total,rel}$ 1.35 kcal/mol, <docking score> 0.00*

6.2.1.3 ε-Caprolactone (CL) Conformation 1 E_{intra}= 0.00 kcal.mol^{-1}

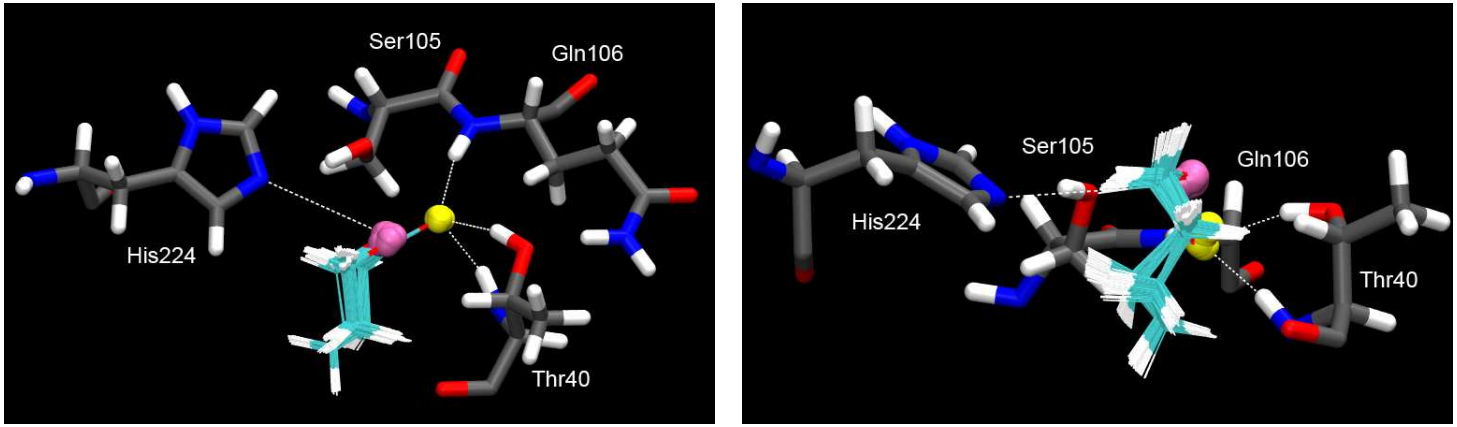


Figure 7a1: *Cluster 1: 226 structures, <RMSD> 0.24 Å, E_{inter} -4.00 kcal/mol, $E_{total,rel}$ 0.00 kcal/mol, <docking score> 0.00*

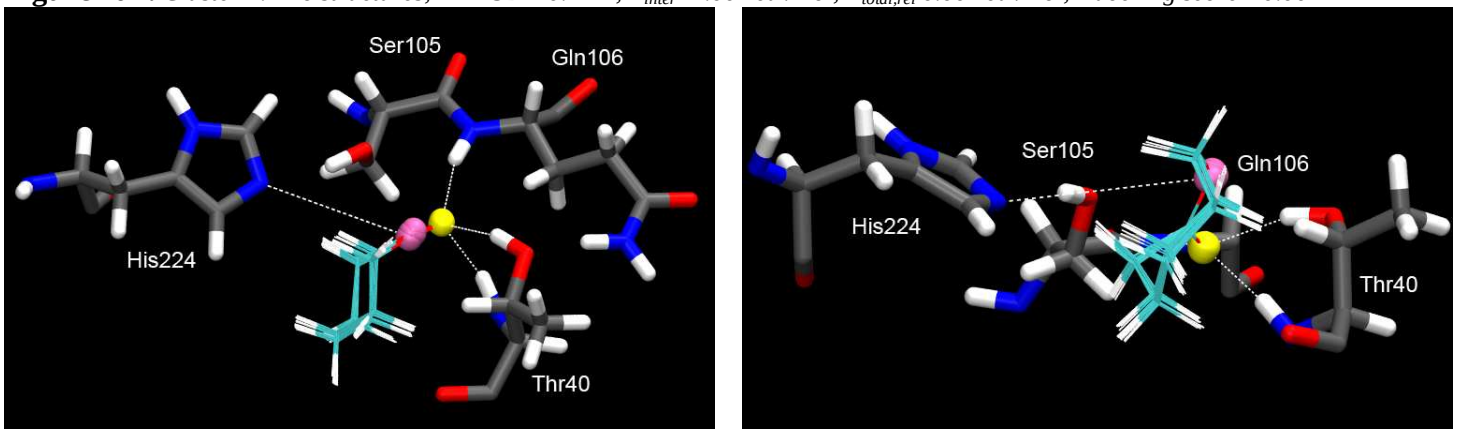


Figure 7a2: *Cluster 2:30 structures, <RMSD> 0.82 Å, E_{inter} -3.96 kcal/mol, $E_{total,rel}$ 0.04 kcal/mol, <docking score> 0.00*

6.2.1.4 ε-Caprolactone (CL) Conformation 2 E_{intra}= 2.57 kcal.mol^{-1}

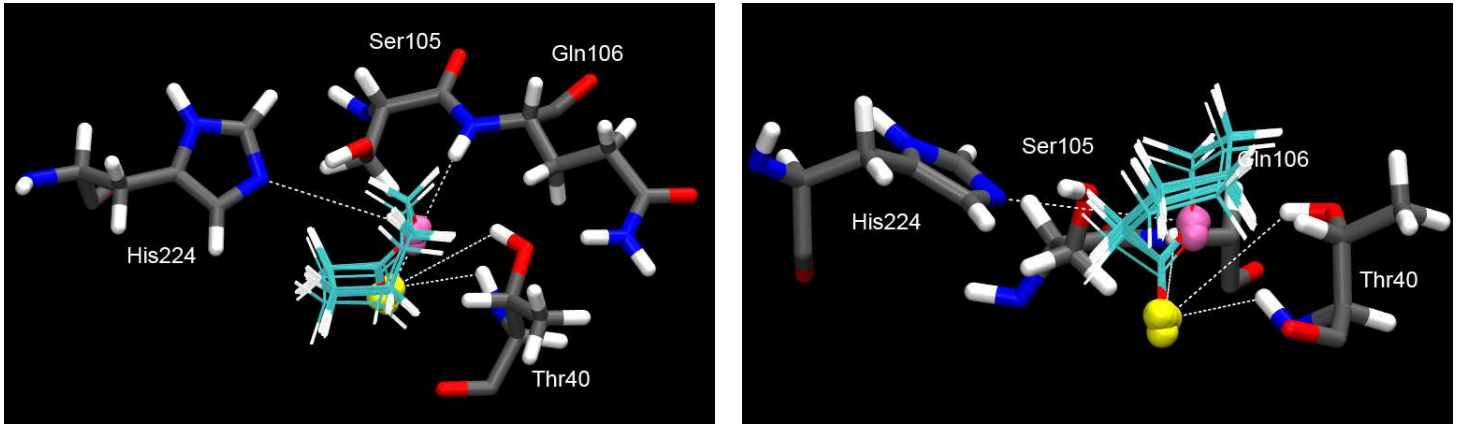


Figure 7b: *Cluster 1: 256 structures, <RMSD> 0.22 Å, E_{inter} -3.39 kcal/mol, $E_{total,rel}$ 3.18 kcal/mol, <docking score> 0.00*

6.2.1.5 ε-Caprolactone (CL) Conformation 3 E_{intra}= 4.28 kcal.mol^{-1}

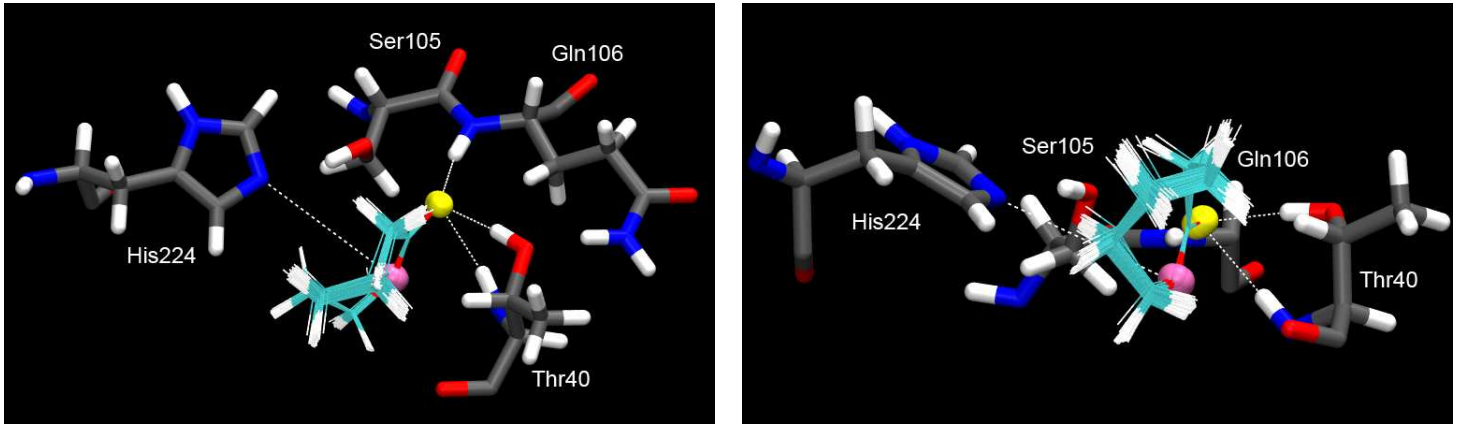


Figure 7c: *Cluster 1: 256 structures, <RMSD> 0.12 Å, E_{inter} -4.27 kcal/mol, $E_{total,rel}$ 4.01 kcal/mol, <docking score> 0.00*

6.2.1.6 ε-Caprolactone (CL) Conformation 4 E_{intra}= 4.97 kcal.mol^{-1}

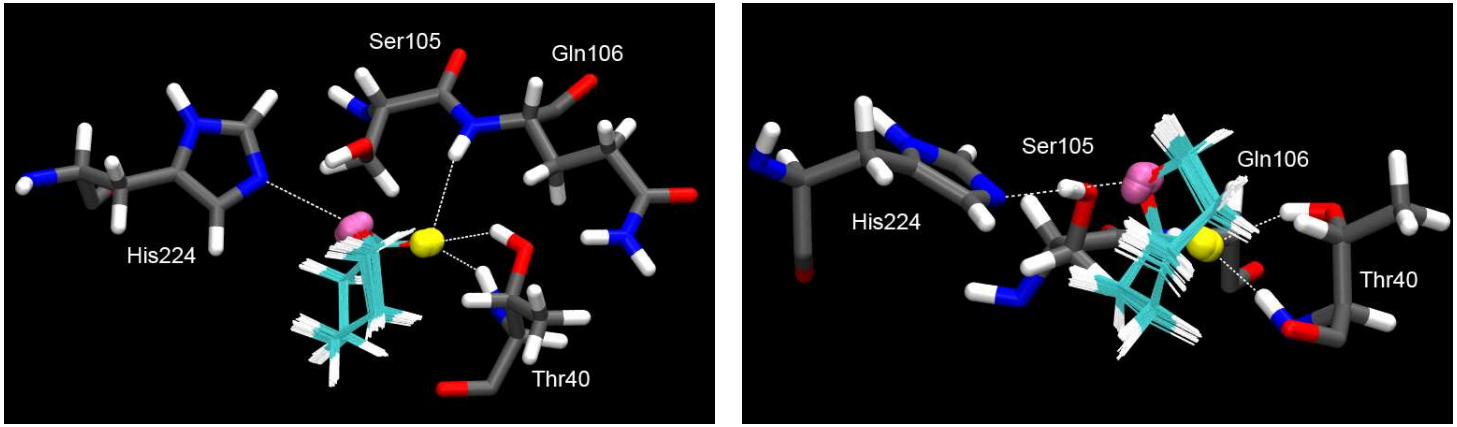


Figure 7d1: *Cluster 1: 249 structures, <RMSD> 0.16 Å, E_{inter} -3.29 kcal/mol, $E_{total,rel}$ 5.68 kcal/mol, <docking score> 0.00*

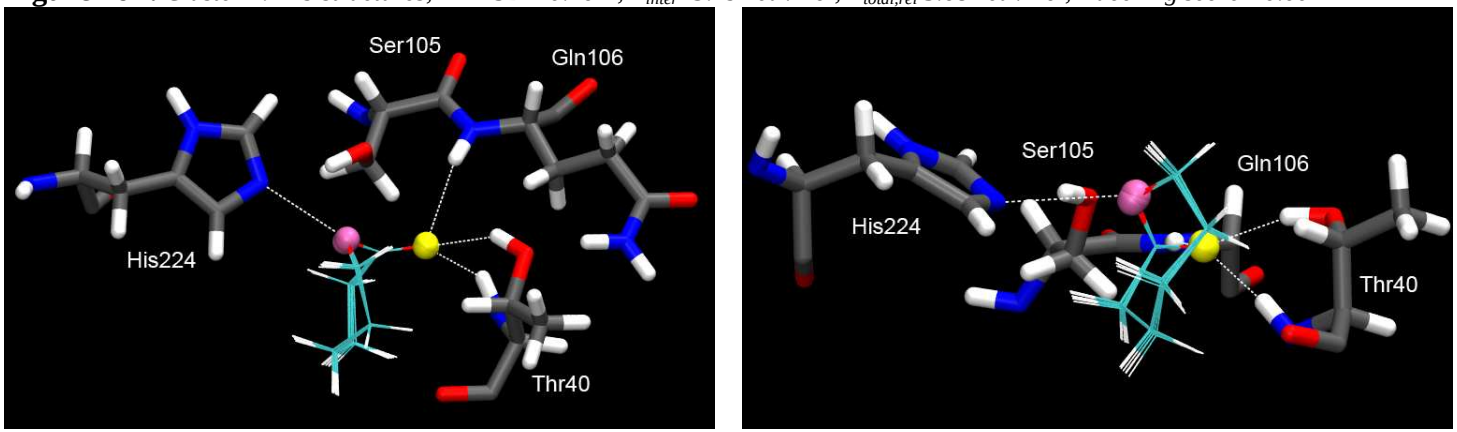


Figure 7d2: *Cluster 2: 7 structures, <RMSD> 0.58 Å, E_{inter} -3.26 kcal/mol, $E_{total,rel}$ 5.71 kcal/mol, <docking score> 0.00*

6.2.1.7 Heptanolactone (HL) Conformation 1 E_{intra}= 0.00 kcal.mol^{-1}

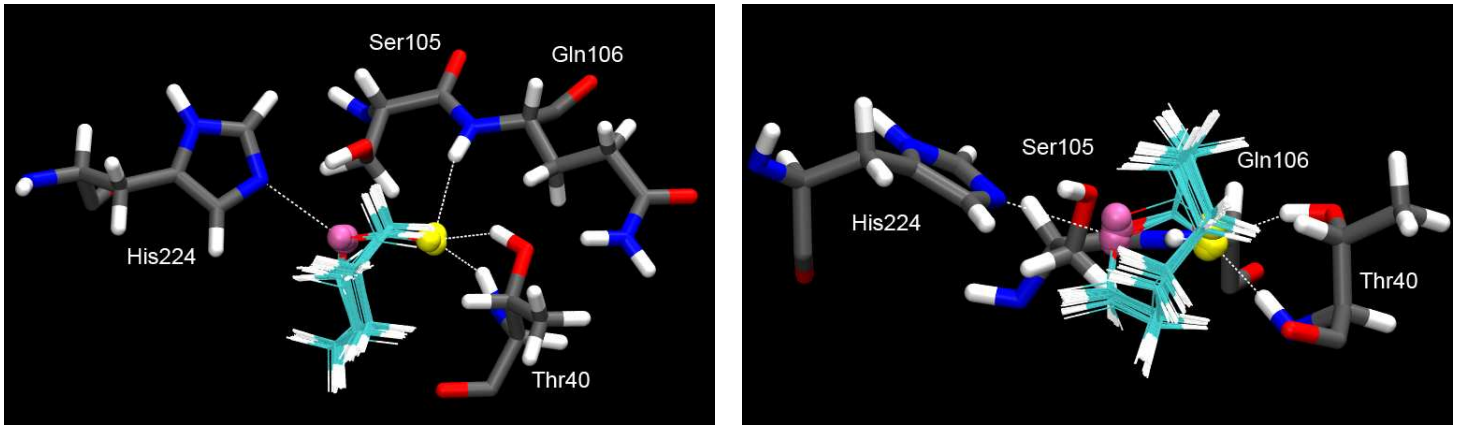


Figure 8a1: *Cluster 1: 255 structures, <RMSD> 0.26 Å, E_{inter} -3.48 kcal/mol, $E_{total,rel}$ 0.00 kcal/mol, <docking score> 0.32*

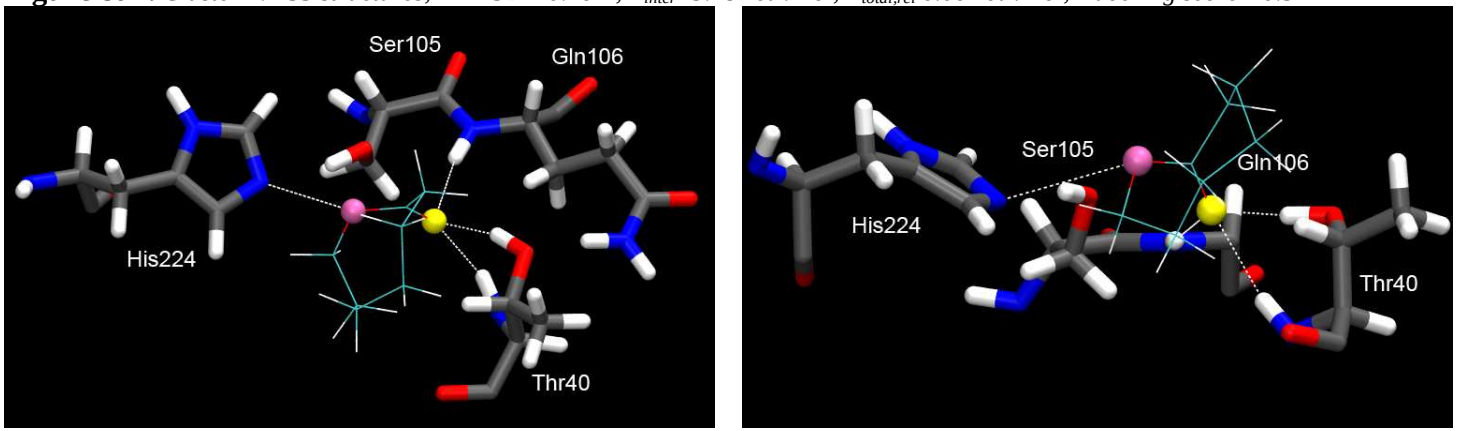


Figure 8a2: *Cluster 2: 1 structure, <RMSD> 1.54 Å, E_{inter} -3.22 kcal/mol, $E_{total,rel}$ 0.26 kcal/mol, <docking score> 0.00*

6.2.1.8 Heptanolactone (HL) Conformation 2 E_{intra}= 0.81 kcal.mol^{-1}

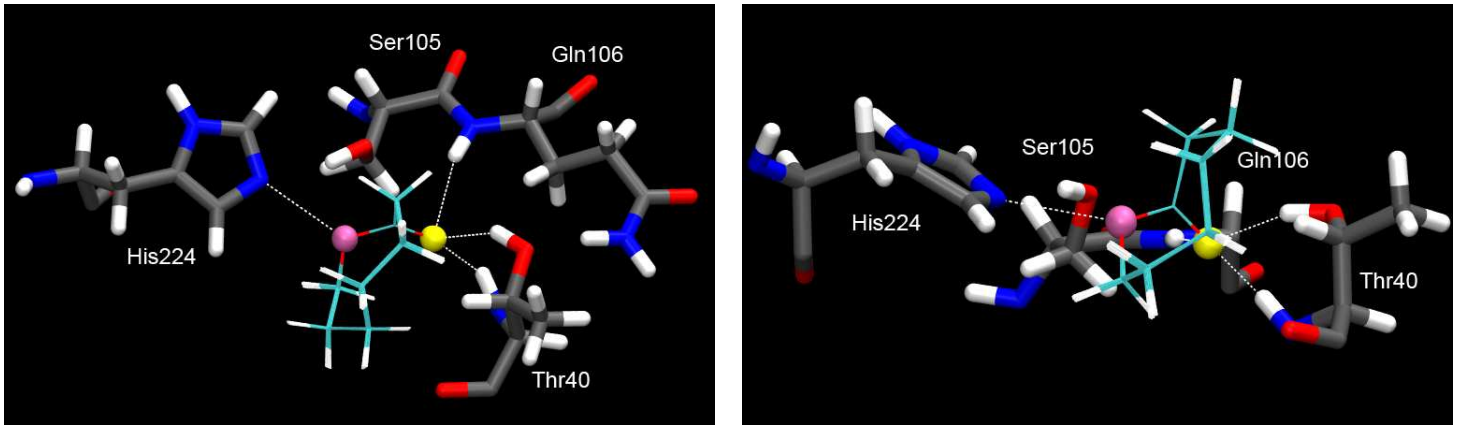


Figure 8c1: *Cluster 1: 129 structures, <RMSD> 0.03 Å, E_{inter} -3.39 kcal/mol, $E_{total,rel}$ 0.90 kcal/mol, <docking score> 0.49*

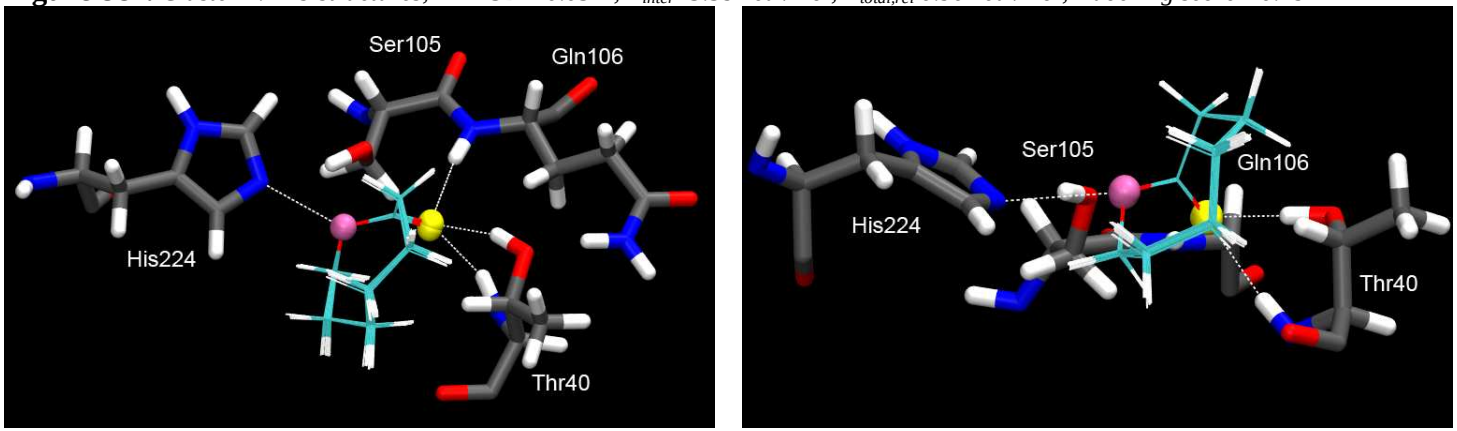


Figure 8c2: *Cluster 2: 127 structures, <RMSD> 0.56 Å, E_{inter} -3.36 kcal/mol, $E_{total,rel}$ 0.93 kcal/mol, <docking score> 0.21*

6.2.1.9 Heptanolactone (HL) Conformation 3 E_{intra}= 0.91 kcal.mol^{-1}

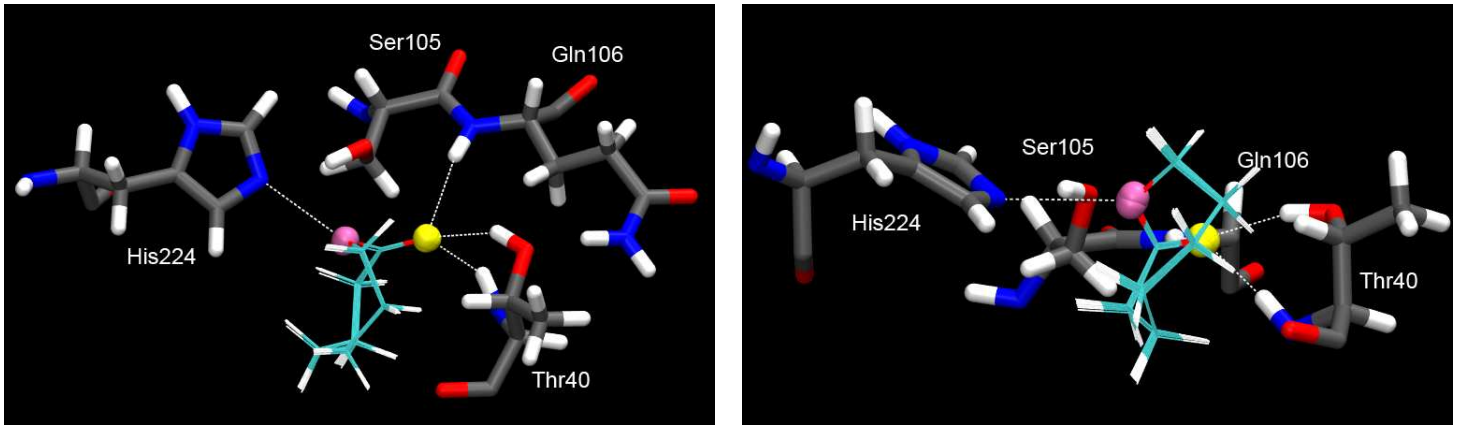


Figure 8b1: *Cluster 1: 150 structures, <RMSD> 0.03 Å, E_{inter} -3.57 kcal/mol, $E_{total,rel}$ 0.82 kcal/mol, <docking score> 0.06*

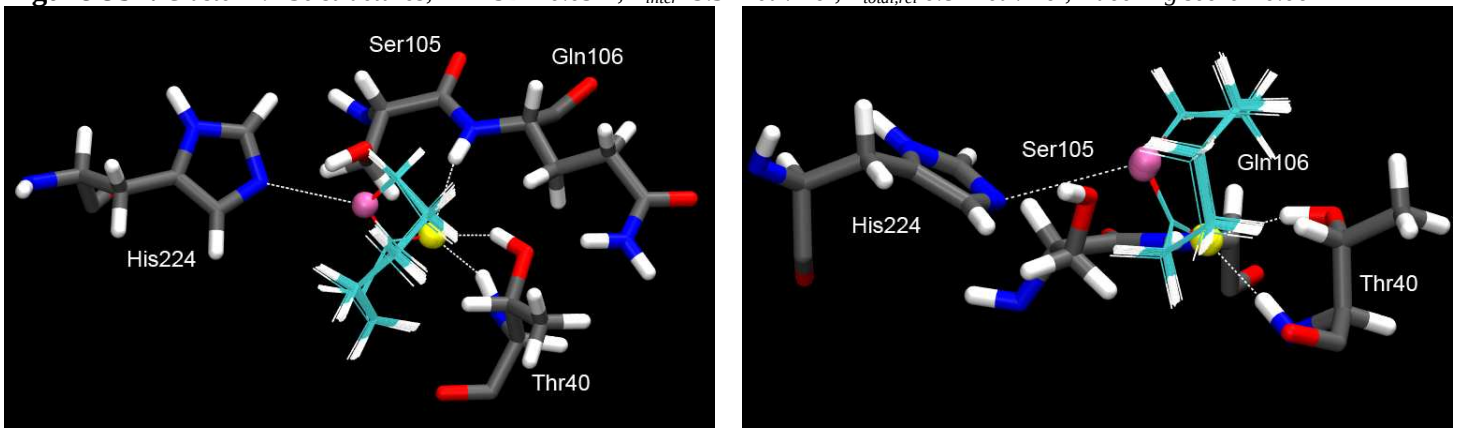


Figure 8b2: *Cluster 2: 106 structures, <RMSD> 1.52 Å, E_{inter} -3.48 kcal/mol, $E_{total,rel}$ 0.91 kcal/mol, <docking score> 0.00*

6.2.1.10 Heptanolactone (HL) Conformation 4 E_{intra}= 1.27 kcal.mol^{-1}

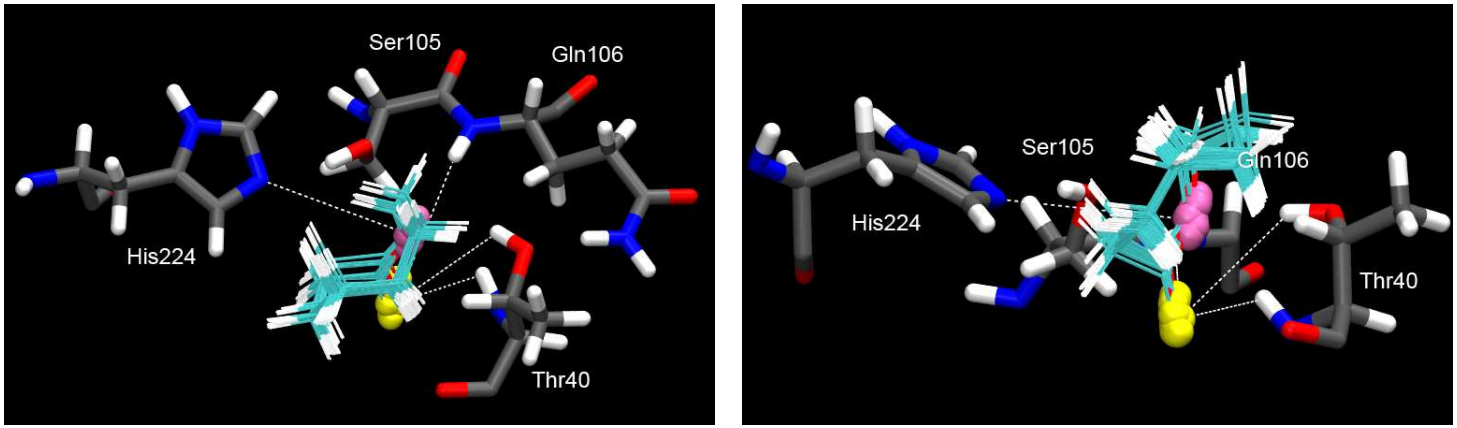


Figure 8d: *Cluster 1: 256 structures, <RMSD> 0.19 Å, E_{inter} -3.45 kcal/mol, $E_{total,rel}$ 1.30 kcal/mol, <docking score> 0.00*

S7 REFERENCES

[1] E. M. Anderson, M. Karin, O. Kirk, *Biocatal. Biotransform.* **1998**, *16*, 181-204.

[2] H. Uyama; S. Namekawa; S. Kobayashi, *Polymer Journal* **1997**, *29*, 299-301.

[3] L. van der Mee, F. Helmich, R. de Bruijn, J. Vekemans, A. R. A. Palmans, E. W. Meijer, *Macromolecules* **2006**, *39*, 5021-5027.

[4] Sybyl 7.2, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.

[5] J. Uppenberg, N. Oehrner, M. Norin, K. Hult, G. J. Kleywegt, S. Patkar, V. Waagen, T. Anthonsen, T. A. Jones, *Biochemistry* **1995**, *34*, 16838-16851.

[6] http://www.rcsb.org DOI: 10.2210/pdb1TCA/pdb

[7] M. F. Sanner, *Journal of Molecular Graphics & Modelling* **1999**, *17*, 57-61.

[8] Cambrigesoft Chem3D Pro 11.0, **2007**.

[9] N. L. Allinger, *Pure Appl. Chem.* **1982**, *54*, 2515-2522.

[10] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33-38, 27-38.

[11] E. Krieger, T. Darden, S. B. Nabuurs, A. Finkelstein, G. Vriend, *Proteins-Structure Function and Bioinformatics* **2004**, *57*, 678-683.

[12] P. Kollman, R. Dixon, W. Cornell , T. Fox, C. Chipot, A. Pohorille, in *Computer Simulation of Biomolecular Systems, Vol. 3* (Eds.: W. van Gunsteren, P. Weiner, A. Wilkinson), **1997**, pp. 83-96.

[13] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577-8593.

[14] http://www.rcsb.org DOI: 10.2210/pdb1TCA/pdb