Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 29, Issue Number 5, (2012)

Potential Human Cholesterol Esterase Inhibitor Design: Benefits from the Molecular Dynamics Simulations and Pharmacophore Modeling Studies

http://www.jbsdonline.com

Shalini John Sundarapandian Thangapandian Keun Woo Lee*

Division of Applied Life Science (BK21 Program), Systems and Synthetic Agrobiotech Center (SSAC), Research Institute of Natural Science (RINS), Plant Molecular Biology and Biotechnology Research Center (PMBBRC), Gyeongsang National University (GNU), 501 Jinju-daero, Gazha-dong, Jinju 660-701, Republic of Korea

Abstract

Human pancreatic cholesterol esterase (hCEase) is one of the lipases found to involve in the digestion of large and broad spectrum of substrates including triglycerides, phospholipids, cholesteryl esters, *etc*. The presence of bile salts is found to be very important for the activation of hCEase. Molecular dynamic simulations were performed for the apoform and bile salt complexed form of hCEase using the co-ordinates of two bile salts from bovine CEase. The stability of the systems throughout the simulation time was checked and two representative structures from the highly populated regions were selected using cluster analysis. These two representative structures were used in pharmacophore model generation. The generated pharmacophore models were validated and used in database screening. The screened hits were refined for their drug-like properties based on Lipinski's rule of five and ADMET properties. The drug-like compounds were further refined by molecular docking simulation using GOLD program based on the GOLD fitness score, mode of binding, and molecular interactions with the active site amino acids. Finally, three hits of novel scaffolds were selected as potential leads to be used in novel and potent hCEase inhibitor design. The stability of binding modes and molecular interactions of these final hits were re-assured by molecular dynamics simulations.

Key words: Pancreatic cholesterol esterase; Molecular dynamic simulation; Structure-based pharmacophore; Database screening; Lipinski's rule; Molecular docking.

Introduction

There is a proven connection between primary hypercholesterolemia, a risk factor of atherosclerosis, and coronary heart diseases (CHD). Various methods have been used to control the plasma cholesterol because cholesterol is the product of biosynthesis as well as dietary intake. Controlling of plasma cholesterol level especially large low-density lipoprotein (LDL)-cholesterol becomes a major health objective due to its participation in cancer, diabetes, and obesity (1). Statins are proved to be a great success in controlling the plasma cholesterol levels (2, 3). Even though statins are more effective they showed many side effects including muscle weakness, increases in serum levels of hepatic transaminases, headache, and sleep disorders in controlled

Abbreviations: CHD: Coronary heart diseases; LDL: Low density lipoprotein; HMG-CoA: 3-hydroxy-3-methyl-glutaryl-CoA; hCEase: Human pancreatic cholesterol esterase; MD: Molecular dynamics; bCEase: Bovine cholesterol esterase; TCH: Taurocholate; PDB: Protein data bank; PME: Particle mesh ewald; RMSD: Root mean square deviation; RMSF: Root mean square fluctuation; DS: Discovery studio; CLR-I: Cluster I; CLR-II: Cluster II; HY: Hydrophobic; HBA: Hydrogen bond acceptor; HBD: Hydrogen bond donor; ADMET: Absorption, distribution, metabolism, elimination and toxicity; GOLD: Genetic optimization for ligand docking.

*Corresponding author:
Keun Woo Lee
Phone: +82-55-772-1360
Fax: +82-55-772-1359
E-mail: kwlee@gnu.ac.kr

tests. The side effects like sensorimotor neuropathy, depression, and eczema have been observed on prolonged use of statins (4-6). In addition, the sensitivity of plasma cholesterol levels to dietary changes in cholesterol is highly inconstant and hence statins cannot be tolerated by all hypercholesterolemic individuals. This increases the interest towards other targets that involved in the control of plasma cholesterol level.

Human pancreatic cholesterol esterase (hCEase) also known as bile salt-activated lipase (BAL) is one of the two lipases found in the secretion of the vertebrate pancreas (7, 8). It is involved in the digestion of broad spectrum of substrates including triacylglycerides, phospholipids, cholesteryl esters, esters of lipid-soluble vitamins and fatty acids (9). It is present in few mammals including humans and it plays a crucial role in the hydrolysis of fatty acids present in the breast milk and in the development of pancreas in infants. Gene knock-out studies in transgenic mice have showed the reduced uptake of cholesterol ester revealing that CEase is responsible for the intestinal absorption of cholesteryl esters (10). The activity of CEase goes beyond just hydrolyzing the dietary lipids.

CEase initially secreted from pancreatic acinar cells and lactating glands of higher mammals including humans and released into small intestine. It remains inactive during its travel to small intestine and it was also observed in endothelial cells, in macrophages, eosinophils, and in the liver (11). It has been revealed that the CEase and chaperone Grp94 interaction plays a very important role in the processes of CEase folding and secretion (12). This complex persists during its travel in order to protect itself from the proteolytic enzymes in the duodenal environment. When it reaches the intestinal lumen, the complex will dissociate and CEase becomes active in presence of the bile salts to digest dietary lipids. The bile salts have several effects on CEase as a cofactor (13). Because of its broad substrate specificity, the activation by primary bile salts is important for the hydrolysis of long chain fatty acid moieties such as esters of glycerol and cholesterol, as well as esters of the fat soluble vitamins (A, D and E). The hydrolysis of water-soluble substrates like esters of fatty acids containing carbon chain shorter than eight atoms can be performed in the absence or without the activation by bile salts. The hydrolytic activity of CEase towards emulsified long-chain triglycerol or cholesteryl ester is accomplished exclusively with primary bile salts containing a 7α-hydroxyl group indicating specific interactions are necessary for the activation of the enzyme (14, 15). There are two bile salt binding regions in CEase, one is near the active site named proximal site and another is at the back side of the catalytic domain named remote site (Figure 1). The bile salt binds to CEase through the electrostatic interaction of arginine and other residues. The activation of CEase depends on the presence of bile salts. This dependency of CEase on bile salts permits the enzyme to accommodate both hydrophobic and hydrophilic substrates, thus showing its function between that of triacylglycerol lipases and esterases (13, 16). The hydrolysis of cholesterol ester by CEase releases free cholesterol, which combines with cholesterol comprised in bile secretions to form the collection of absorbable cholesterol. The bile salts are essential not only for the activation of CEase but also for the solubilization of the free cholesterols which are less soluble in the absence of bile salts. Along with this, both the primary and secondary bile salts may also protect CEase from intestinal proteolysis (17, 18). The transport proteins are required to carry free cholesterol from micelles to the enterocytes for absorption. It has been suggested that CEase involved in the transport of free cholesterol to the enterocytes and act as cholesterol transfer protein but there are some contradictory reports regarding this role and has not been clearly established (19). CEase is also reported to be involved in lipoprotein metabolism, in which it converts large low-density lipoprotein (LDL) to smaller and denser more cholesteryl ester-rich lipoproteins, thereby it may regulate serum cholesterol levels. Many evidences have indicated that CEase may have deleterious effects in atherosclerosis because of the conversion of larger and less atherogenic low density lipoprotein to the smaller and more atherogenic low density lipoprotein subspecies (20-23). The function of plasmatic CEase in atherogenesis and its relationship with various pathological conditions are

not clearly established (24). Recently, more interest is shown towards the inhibition of CEase as potential target especially for the development of hypocholesterolemic agents because it involved in the control of plasma cholesterol level. The inhibition of CEase may render a way to limit the bioavailability of dietary cholesterol derived from cholesterol esters and also limit the absorption of free cholesterol (25).

The CEase is a serine hydrolase and belongs to the α/β-hydrolyase fold family. The members of this family share secondary and tertiary structural characteristics (26-28) and utilize Ser194, Asp320, and His435 catalytic triad mechanism. The CEase, serine lipases, and serine proteases possess a conserved Ser-Asp-His catalytic triad and share the catalytic mechanism. This catalytic triad serves as general acid-base and nucelophilic catalytic entity along with an oxyanion hole made of Gly107, Ala108, and Ala195 residues (Figure 1) (7). CEase catalyzes the ester hydrolysis by the operation of catalytic triad which is stereochemically confluent with the catalytic triads of serine proteases and uses the acylation-deacylation mechanism (29, 30). Therefore, CEase and other proteases such as serine proteases and serine lipases are expected to be inhibited by the same class of mechanism-based inhibitors by involving common catalytic mechanism for substrate hydrolysis (31, 32). The mechanism-based inhibitors are borinic acids, beta-lactams, hexadecylsulfonyl fluoride, beta-lactones, aryl phosphate, aryl haloketones, choloroisocoumarins, and carbamates (33-38). It has also been used in the biocatalytic resolution of spirobiindanols, binapthols, and indole derivatives (39, 40). Though many inhibitors are available for the inhibition of CEase still there is a need to develop new inhibitors because of their low potency and adverse effects.

The main aim of this study is to generate a bile salt bound complex of hCEase using the coordinates of bile salts from bovine CEase and observe the changes at the catalytic site to accommodate long chain fatty acids. Molecular dynamics (MD) simulations were performed to obtain the conformational changes upon the binding of bile salts. Two structure-based pharmacophore models were generated using the representative structures snapped from the highly populated regions of simulated trajectories. These models, upon validation, were used in database screening to identify the chemical compounds that are structurally compatible with the active site of the enzyme. Molecular docking results were also used as final filter to assess the binding behavior of the identified hits whereas MD simulations were used to re-evaluate the stability of the binding of the identified hits.

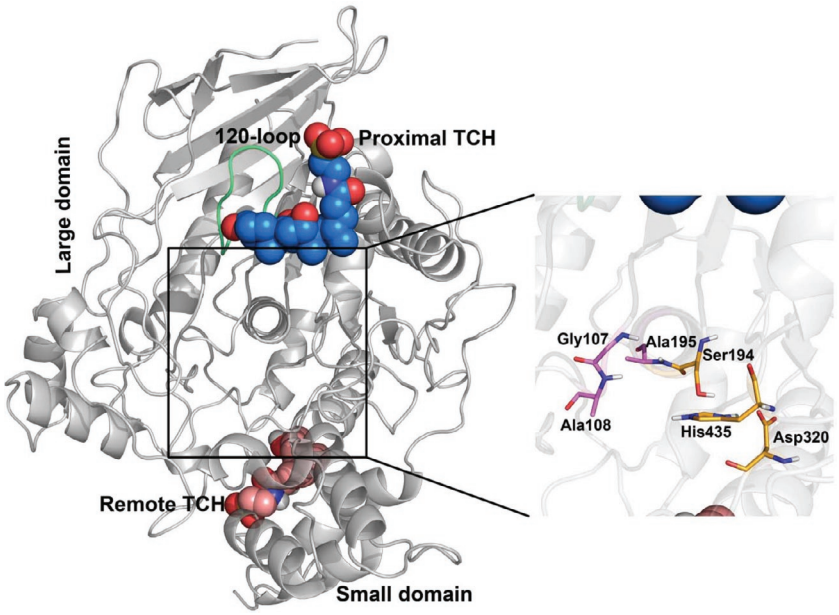

Figure 1: Structural details of the hCEase with TCH in their proximal and remote regions. The right side of the Figure shows the zoomed view of the catalytic site amino acids.

Materials and Methods

The hCEase-TCH Complex Preparation

Among the crystal structures of various lipases, only one crystal structure of catalytic domain of hCEase (PDB ID: 1F6W) was solved. The crystal structure of bovine CEase (bCEase) bound with two bile salts is also available in (PDB ID: 1AQL). As mentioned earlier, CEase has two bile salt binding regions, namely, proximal and remote sites. The bile salts available in bCEase are taurocholates (TCH) (Figure S1). The presence of bile salt is found to be very important for the activation of CEase, which enables the enzyme capable of digesting long chain fatty acids. To build the 3D structure of hCEase complexed with bile salts, we have used the 3D coordinates of bile salts from the crystal structure of bCEase by superimposing both the structures. Surprisingly, the bCEase has shown more than 90% similarity with hCEase and thereby it was appropriate to use the 3D coordinates of bile salts from bCEase.

Molecular Dynamics Simulations

Two systems including the apoform of hCEase with no bound bile salts and a hCEase-TCH complex with two bile salts were considered in MD simulation study because molecular dynamics studies have proven to provide at the atomic level information on the structural and dynamics of biological macromolecules (41-61). The MD simulations of both the systems were performed using the GROMACS Version 4.0.7 program running on a high performance linux cluster computer with GROMOS87 forcefield (62, 63). Initially hydrogen atoms were added to all the ionizable groups of the protein appropriate to a neutral pH 7.0. Topology files and other forcefield parameters including the charges of the ligand atoms were generated for TCH using the PRODRG web-server (64). The two systems were immersed in a cubic box of water molecules with the size of 1 nm from the surface of the protein to carry out the simulations in an aqueous environment. Both the systems were neutralized by adding Cl^- counter-ions replacing the water molecules. After neutralization, the energy of the systems was minimized using steepest descent approach for 10000 steps available in the GROMACS package. Later, a 100 ps position restraining MD run was performed to restrain the backbone of the protein by a 1000 kJ/mol Å^2 harmonic constraint, in order to relieve close contacts before the actual simulation. After equilibration, 5 ns MD simulation production runs were performed for two systems with a constant number of particles (*N*), pressure (*P*), and temperature (*T*), *i.e.,* NPT canonical ensemble and the periodic boundary conditions were used in all three dimensions to avoid the edge effects. The water molecules, receptor, and ligands were coupled in a temperature bath at 300 K using Berendsen's coupling algorithm with a coupling constant $T_t = 0.1$ ps. The pressure coupling was set with a constant pressure of 1 bar and a coupling constant Tp of 1 ps. For long-range electrostatics the particle mesh Ewald (PME) method (65) with a van der Waals interactions cutoff of 14 Å and a Coulomb interaction cutoff of 12 Å was updated for every 10 steps. All bonds involving at least one hydrogen atom were constrained to their equilibrium position using LINCS algorithm. The coordinates of the simulated systems were collected every 1 ps during the production MD run in the trajectories. The basic analyses like atomic root mean square deviations (RMSD) of the enzyme and the bile salts, convergence of energies, hydrogen bond interaction, and root mean square fluctuation (RMSF) were used to check the stability of the systems.

Overall Stability and Cluster Analyses

The trajectories obtained from 5 ns MD simulations were used to investigate the flexibility of the systems. The major conformations of the hCEase-TCH complex protein are studied by performing conformational cluster analysis (66, 67). Prior to the cluster analysis, the MD simulation trajectories were analyzed using many auxiliary programs rendered with the GROMACS package (68). The programs include

g_rms to calculate the RMSD by comparing any of the two structures, g_rmsf for the root mean square fluctuation (RMSF), g_hbond to calculate the inter and intra molecular H-bond interactions, g_energy to calculate potential energy, and g_gyrate for the radius of gyration (Rg) of the structure. Another important analysis to measure the distance between active site amino acids was computed using g_mindist program. Totally, 5000 conformations were saved in the MD simulation trajectory and hence the selection of few representative structures from the trajectory is considered vital. The g_cluster function available in gromacs was used to find out the representative structures from the trajectory. Conformations saved from MD simulation trajectories were initially clustered based on the RMSD and the obtained clusters were further clustered using gromos method. Ten clusters were generated using gromos method, which adds a structure to particular cluster when its distance to any element of the cluster is less than the cutoff value of 0.13 nm.

Pharmacophore Model Generation

The Discovery Studio (DS) 2.5 program (Accelrys Inc., San Diego, USA) was used in the construction and visualization of pharmacophore models and further studies. The two representative structures from cluster I (CLR-I) and cluster II (CLR-II) were used in the generation of structure-based pharmacophore models. In the active site of CLR-I and CLR-II, spheres that include all the essential amino acids were generated using *Binding Site Analysis* tool available in DS. The pharmacophoric features that are complementary to the active site amino acids enclosed within the sphere were generated using the *Interaction Generation* protocol. The lipophilic and hydrophilic features like hydrophobic (HY), hydrogen bond acceptor (HBA), and hydrogen bond donor (HBD) features were generated during the pharmacophore generation. The pharmacophoric features in the form of cluster were refined using *Edit Cluster Pharmacophore* tool available in DS. During pharmacophore clustering the most representative pharmacophore features were selected to be present in the pharmacophore model. Hence, the pharmacophore model was developed to complement the essential amino acids in the active site of the protein by taking protein's flexibility into account. The two generated pharmacophore models Hypo I and Hypo II from the top two MD clusters, CLR-I and CLR-II, were carried out for pharmacophore validation process before utilizing them in database screening.

Validation

A set of 90 compounds which are reported for their CEase inhibitory activities, expressed in K_i values, was collected from the literature (26, 69-75). The 2D structure of all of these compounds were built using ChemSketch version 12 program. The built 2D structures were converted to 3D structures and consequently a database has been developed using DS. This set containing 90 compounds was used to validate the generated pharmacophore model. All the compounds were minimized to the closest local minimum using CHARMM forcefield. Diverse conformations were generated using *Diverse Conformation Generation* protocol available in DS with 'Best conformation analysis' option and a maximum number of 255 conformations with a constraint of energy threshold of 20 kcal/mol above the global energy minimum. A number of energetically reasonable conformational models were generated to represent the flexibility of each compound. The compounds with the generated conformations were mapped on both the generated pharmacophore models Hypo I and II separately using *Ligand Pharmacophore Mapping* protocol *with Best/Flexible Search* option as available in DS.

In Silico *Screening of Databases*

The validated pharmacophore models Hypo I and II were used as 3D structural queries in database screening against chemical databases, namely, Maybridge and Asinex to retrieve hit compounds mapping the identified pharmacophoric features. The database screening was performed using the *Best Flexible* search option in

Ligand Pharmacophore Mapping protocol available in DS. The retrieved hits were refined using various drug-like filters like Lipinski's rule of five (76, 77) and ADMET (78, 79) properties in order to select only drug-like compounds.

Molecular Docking

The molecular docking study was performed using GOLD (Genetic Optimization for Ligand Docking) program version 4.1 (80, 81). The GOLD uses genetic algorithm for docking flexible ligands into rigid or partially flexible protein active site. The protein structure CLR-I that was taken from cluster I was used in the molecular docking study. The water molecules were removed and hydrogen atoms were added to the protein structure. In CEase the ester hydrolysis is catalyzed by the action of catalytic triad and oxyanion hole residues. Thus interaction with hydroxyl group of Ser194 that acts as nucleophile is essential for the ester hydrolysis reaction. Protein residues within the radius of 10 Å around the hydroxyl group of Ser194 were defined to form the active site of the enzyme during molecular docking. Except the early termination option, which was changed from its default value 3 to 5, all other parameters were kept at their default values. Top 10 docked poses were allowed to be saved using the early termination option of the program. This early termination option quits the genetic optimization calculation of a ligand if the RMSD between any of its 5 conformations is less than 1.5 Å. The compounds with best GOLD fitness score, mode of binding, and interaction with essential amino acids were considered during the selection of final hit compounds. Molecular interactions were also observed using *Molegro virtual docker* (82) and *Ligplot* programs (83). The Novelty of the final hit compounds was confirmed using *SciFinder Scholar* (84) and *PubChem* (85) search.

Result and Discussion

Molecular Dynamics Simulations

The MD simulations have been extensively used to get the possible bioactive conformation when there is no crystal structure available for protein-ligand complex. The activated CEase is found to be very important for the digestion of cholesterol esters, fatty acids and triglyceryl lipases. The crystal structure of human activated CEase with bound bile salts is not available in PDB. Thus the coordinates of bile salts were obtained from the structurally similar bCEase complexed with bile salts. In order to get the possible bioactive conformation of hCEase complexed with two molecules of TCH, a 5 ns MD simulation was performed. In the other hand the apoform of hCEase (hCEase-Apo) was also subjected to MD simulation for 5 ns under the same conditions to compare the structural changes brought upon TCH binding. These MD simulations were performed to explore the behavior of the active site region of hCEase in presence of two TCH molecules in its allosteric site and also to validate the stability of both the systems.

The RMSD of two systems with respect to their starting structures were calculated to examine the conformational deviations of both the systems within the aqueous environment and also to observe the deviation of the hCEase-TCH complex structure with respect to that of apoform. The calculations of the backbone RMSD of hCEase-TCH and hCEase-Apo forms as a function of simulation time revealed that the complex has reduced flexibility to that of apoform (Figure 2A). The RMSD of the apoform was slightly high when compared to that of complex. This observation from the RMSD comparison showed that the presence of TCH stabilizes the system. The RMSF plot was used to identify the highly flexible or fluctuating parts along with the stable parts of the structures by examining the quantity of fluctuation of different parts of the polypeptide chain of the complex structure with respect to the apoform. From the plot it was observed that the highly flexible regions belong

to the loop structure and the stable ones belong to the secondary structure elements. The complex structure highly fluctuated when compared to the empty structure (Figure 2B). This fluctuation was mainly due to the binding of proximal and remote TCH molecules. The residues Gln93, Asn120, Leu277, Ala325, Ala372, and Pro428 responsible for the fluctuation were present in the highly flexible loop regions. Loop residues Gln93, Asn120, and Pro428 are present in the large domain and Leu277, Ala325, and Ala372 are present in the small domain regions. Though there are high fluctuations observed in the loop region in presence of TCH, the active site residues were well stabilized when compared to the apoform. The conformation of the loop residues 115-125, known as 120-loop, was considered very important in the activation step of CEase upon binding of bile salts. The 120-loop is in its closed conformation in the inactive form *i.e.* in the absence of bile salts whereas it remains open in the presence of bile salts especially the proximal bile salt (7, 86). This is because the proximal bile salt binding region is present close to the 120-loop. In case of proximal bile salt binding, the loop opens in order to accommodate the bile salt thereby provides the space for the long substrates whereas in the absence of proximal bile salt the loop will be in the closed form in that way there is no enough space for the long substrates to bind. In this study, we have observed the conformational changes of CEase upon TCH binding through MD simulations. The 120-loop was found neither closed nor opened in the crystal structure of hCEase but the loop moved to form a closed conformation thereby it closed the bile salt binding cavity as well as the part of the active site after 5 ns MD simulation of apoform. In hCEase-TCH complex, the 120-loop conformation was changed by moving 9.5 Å away from the closed conformation of apoform to the open conformation (Figure S2). This depicts the structural evidence of the importance of the proximal bile salt in the activation of CEase. Thereby, the TCH binding position probably stabilizes the open conformation of the CEase. Hence, the binding of bile salts not only brought the conformational changes in the particular region but it also influenced the overall structure elements. The proximal TCH that binds near the 120-loop is responsible for the fluctuation of Gln93, Asn120, and Pro428 as well as other loop residues present in the large domain and it also has very less influence on the small domain. The remote TCH was bound in a cleft between the large and the small domains. The remote TCH binding influenced the conformational changes of the protein but not like proximal TCH, because the remote bile salt doesn't participate

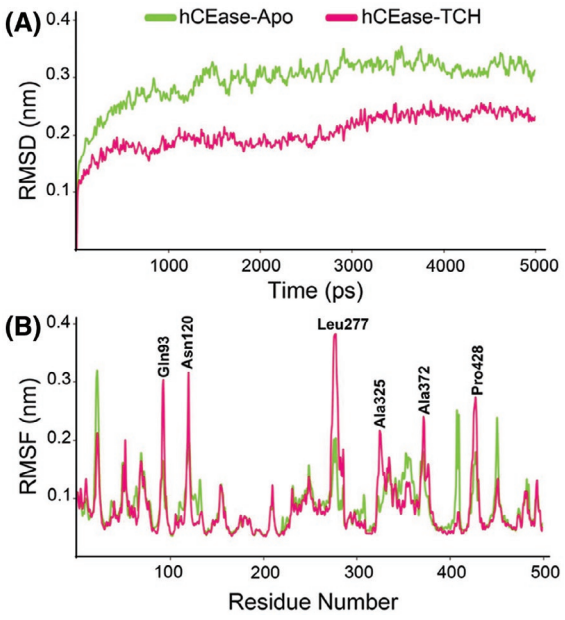

Figure 2: (**A**) The backbone root mean square deviation (RMSD) and (**B**) Cα root mean square fluctuation of the hCEase-Apo and hCEase-TCH systems obtained during 5 ns MD simulations.

in any of the catalytic function, but still its presence affects the structural conformation. As a result of binding of the remote TCH the fluctuation in the loop residues such as Leu277, Ala325, and Ala372 was observed.

The number of hydrogen bond (H-bond) interactions between the TCH and protein was also measured throughout the time of simulations. It was observed that the proximal TCH which is very important for the catalytic activity found to have less H-bond interactions with the protein whereas the remote TCH which is less important for the catalytic activity is found to have strong H-bond interactions with the protein (Figure S3). The proximal TCH has formed important H-bond interactions necessary for the activation of the protein even though it has less number of H-bonds whereas the remote TCH did not form any H-bond interactions required for the activation of the enzyme. The potential energy plot indicated that the energy of both the systems has reduced initially to some extent to be stable in the simulation environment and became unchanged throughout the simulation (Figure S4A). The potential energy difference between hCEase-TCH complex and hCEase-Apo form was very less. The radius of gyration (Rg) was measured to observe the compactability of the systems throughout the simulation as another measure of the system stability. The observation from the Rg graph (Figure S4B) proved that the compactness of the complex structure slightly decreased when compared to the apoform.

Cluster Analysis

The structure of hCEase-TCH after 5 ns MD simulation was selected for further studies using cluster analysis instead of average structure of the protein. The trajectories of hCEase-TCH system were used in cluster analysis. From the cluster analysis results, it was observed that the most populated cluster I and II were containing 1226 and 941 structures, respectively, among the 10 generated clusters using the MD trajectory (Figure S5). The gromos method was used in the generation of clusters with the RMSD cut off value of 0.13 nm. The middle structures from cluster I and II were used in structure-based pharmacophore model generation.

Structure-Based Pharmacophore Models

The main aim of this work is to develop structure-based pharmacophore models from the activated hCEase. The structures selected from the dynamic trajectories of 5 ns MD simulations were used in the generation of structure-based pharmacophore models. Clusters of features complementing all the essential amino acids in the active site were generated. The *Edit/Cluster option* available in DS was used to edit the clustered pharmacophoric features. Totally six features from CLR-I (Hypo I) and CLR-II (Hypo II) were finalized based on the important amino acids in the active site. The pharmacophoric features of both Hypo I and II were made of 3 HBA, 1 HBD, and 2 HY features complementing the same set of amino acid residues. The HBD feature was generated complementary to the hydroxyl group of Ser194 whereas the HBA feature was generated as a complementary feature to His435 because these two residues are important to initiate the catalytic mechanism. Two more HBAs were generated complementing Ala195 and Ala108 residues. In case of the HBA created for Ala108, the feature cage also includes the functional groups of Gly107 residue. Two HY features complementing Ala108 and the hydrophobic region of Ser194 were generated (Figure 3A and 3B). Though Hypo I and II pharmacophore models were made of same features, the inter-feature distance constraints of each pharmacophore models were different (Figure 3C and 3D). This difference in the distance constraints explains the conformational flexibility of the vital interaction points available at the active site. Therefore the compounds that are retrieved through both the pharmacophore models in the database screening are supposed to have the conformational flexibility to interact the key residues in the active site.

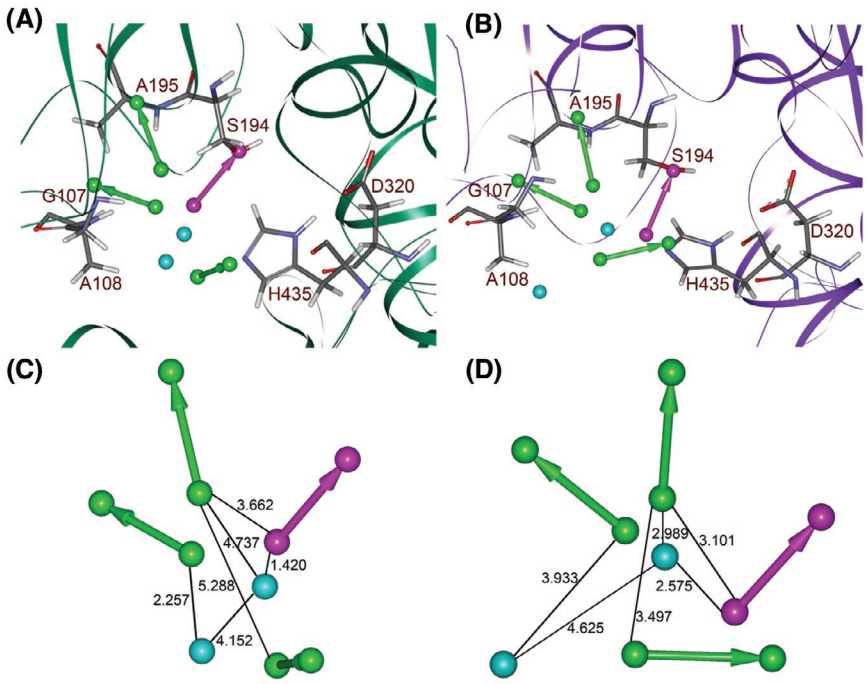

Figure 3: Structure-based pharmacophore models in the active sites of (**A**) Hypo I and (**B**) Hypo II. The distance constraints between the chemical features of the pharmacophore models (**C**) Hypo I (**D**) Hypo II. Green color represents HBA, magenta color represents HBD and cyan color represents HY features.

Pharmacohore Validation

The quality of the generated pharmacophore models needs to be ascertained using a set of compounds containing experimentally known CEase inhibitors that were collected from various literature resources. This validation was performed in order to check the ability of the generated pharmacophore models to select and classify correctly the actives from the inactive molecules during the screening process. A set of 90 compounds including 33 active inhibitor compounds was considered in this validation process (Table SI). The active and inactive compounds were categorized based on the experimental activity values. The compounds with $<1\,\mu M$ was considered as active and remaining as inactives. Hypo I and II were used to screen the set of known CEase inhibitors using *Ligand Pharmacophore Mapping* protocol running with the *BEST/flexible* search option as available in DS. The calculated percentage yield of Hypo I and II were 70.27% and 68.57%, respectively. The same numbers of false positives, 11, were identified in the screening using Hypo I and Hypo II whereas the false negatives were 7 and 9, respectively (Table I). This indicates the efficiency of the pharmacophore models in the database screening and in turn the quality of the models.

Table I

Validation results for two pharmacophore models.

Parameter	Hypo I	Hypo II
Total molecules	90	90
Total numbers of actives (A)	33	33
Total hits (Ht)	37	35
Active hits (Ha)	26	24
% Yield of actives [(Ha/Ht) × 100]	70.27	68.57
False Negatives [A-Ha]	7	9
False Positives [Ht-Ha]	11	11

In Silico *Screening of Databases*

The validated pharmacophore models, Hypo I and II, were used as 3D queries in the screening of chemical databases. The *Best Flexible* search option was used and the *Maximum Omit Feature* option was changed from the default value of 0 to 1 to retrieve the potential compounds equivalent to the known most active compounds. Pharmacophore models Hypo I and II have retrieved 18748 and 23742 compounds, respectively, from Maybridge database whereas 82751 and 101933 compounds, respectively, from Asinex database. The retrieved compounds were refined using various filters. As a first filter the compounds containing fit value >3 were selected as the known most active CEase inhibitor has scored a fit value of 2.416. As second and third filters, Lipinski's rule of five and ADMET properties were used to refine the retrieved hits by removing the compounds with non-druglike properties. A total of 327 and 1260 compounds from Hypo I and II, respectively, satisfying the drug-like filters were carried out for molecular docking studies.

Molecular Docking Studies

In order to further refine and also to reduce the false positives, the retrieved hits that passed all the drug-like filters were carried out for molecular docking. The retrieved 327 and 1260 compounds from Hypo I and II along with 6 most active compounds were docked into the active site of CLR-I & II. The docking was carried out using GOLD 4.1 program. Ten distinct poses for each ligand in the active site of CLR-I & II were generated. The docked compounds were refined based on GOLD fitness score, binding mode, and molecular interactions with the active site residues. The most active compounds have shown high GOLD fitness score compared to the least active compounds indicating a high correlation with biological activity. It was also found that the hydrogen bond interactions play a significant role in determining the fitness score of the compounds. The GOLD fitness score for the most active compound 1 is 54.74. Hence the hit compounds screened from chemical databases having GOLD fitness score greater than 50 were selected. The selected compounds were further refined based on the binding modes, molecular interactions with the active site amino acid residues, and diverse chemical scaffolds. Therefore, based on all these criteria, we carefully analyzed the docking results and 3 compounds of new scaffolds with high GOLD fitness score, favorable binding mode, and strong interactions with essential amino acids were selected.

Binding Mode of Compound 1

Compound 1 (Figure S6) scored a GOLD fitness value of 54.74 and had formed H-bond interactions with Gly107, Ala108, and His435 (Figure 4A). The oxygen atom fused inside the isochromen-1-one ring was found to have H-bond interactions with Ala108 and His435. The carbonyl group and the chlorine atom of the same moiety were having H-bond interactions with His435. The overlay of compound 1 over pharmacophoric features was similar to its binding mode (Figure 5A). The fused oxygen atom of isochromen-1-one moiety overlaid on the HBA feature that compliments to the Ala108 residue. In a similar way the side chain oxygen atom overlaid on the HBA feature showed H-bond interaction with the His435. The aliphatic side chain of compound 1 interacted hydrophobically with the hydrophobic residues at the active site of the protein in the same way it was overlaid on the hydrophobic features. The compound 1 has missed the HBD feature of the developed pharmacophore model. Hence any hit compound with a HBD group and able to overlay on the HBD feature at the same time with strong molecular interactions with the active site amino acids can act as an effective inhibitor of CEase. The binding mode of compound 1 was used as a reference to compare the binding modes of the final hit compounds.

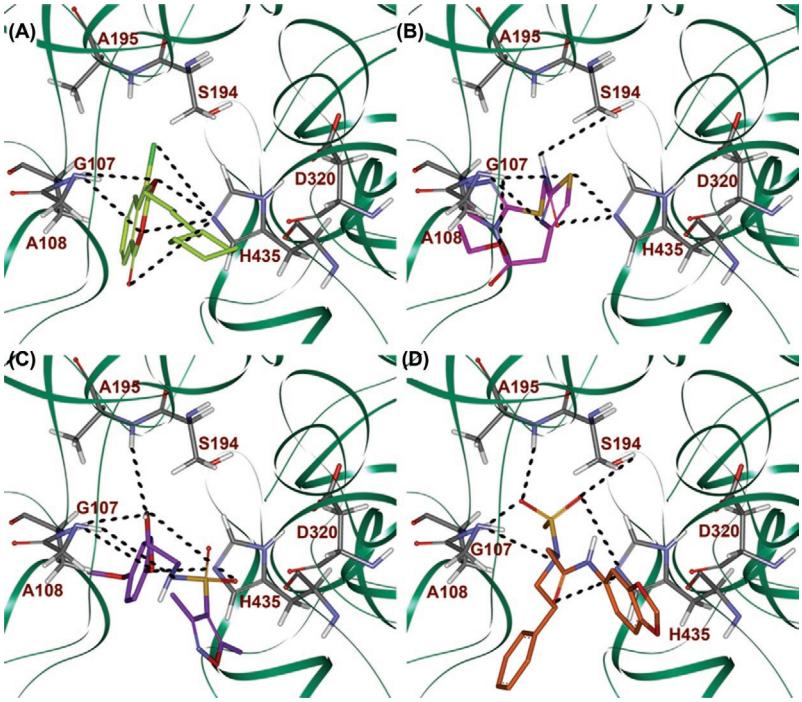

Figure 4: Binding orientations of (**A**) most active compound 1 (**B**) Hit 1 (**C**) Hit 2 and (**D**) Hit 3 are shown in green, magenta, violet and orange, respectively, at the active site of CLR-I. Hydrogen bonds are shown in black dashed lines.

Binding Mode of Hit 1

The Hit 1 shows a GOLD fitness score of 60.90 and its binding mode within the active site of the protein was given in Figure 4B. The sulphur atom present in thiazole ring forms H-bonds with His435 and Gly107. The NH and carbonyl group of the central amide group interact with Ser194 and His435, respectively. This NH group present in the central amide group donates a H-bond to the hydroxyl group of Ser194. The oxygen atom of ester group and nitrogen atom of thiazole ring showed H-bond interactions with Gly107 and Ala108, respectively. This hit compound also shows hydrophobic interactions with Ala108 and Asp320. The mode of binding is similar with the pharmacophore overlay of the compound (Figure 5B and 6). The sulphur

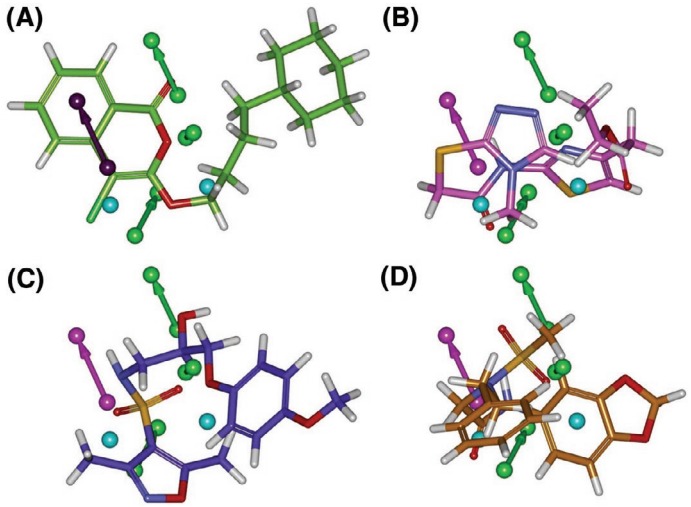

Figure 5: Ligand pharmacophore mapping. The cages were removed for clear view. (**A**) Compound 1 (green) (**B**) Hit 1 (magenta) (**C**) Hit 2 (violet) and (**D**) Hit 3 (orange) mapped over Hypo I. In the pharmacophore hypothesis green represents HBA, magenta represents HBD and cyan represents HY-AR features.

atom of the thiazole ring and the oxygen atom of the ester group that are overlaid on the acceptor features accept H-bonds from the amino group of Gly107 and His435, respectively. The NH group overlaid on the HBD feature donates a H-bond to the hydroxyl group of the catalytically important Ser194 residue. This compound, retrieved from Asinex database, is a derivative of 4H-1,2,4-triazole-3-thiol chemical scaffold.

Binding Mode of Hit 2

The Hit 2 with GOLD fitness score of 59.56 has shown interactions with Gly107, Ala108, Ala195, and His435. One of the oxygen atoms of SO_2 group, the NH group present adjacent to it and the oxygen atom of the phenol moiety have formed H-bond interactions with His435 and Ala108, respectively. The OH group present in the aliphatic chain has shown H-bond interactions with Gly107, Ala108, Ala195, and His435 (Figure 4C). The carbons present in the isoxazole ring and in the aliphatic chain showing hydrophobic interactions with Ala108, Ser194, and His435. The pharmacophore overlay of Hit 2 upon the pharmacophore model is similar to its binding mode at the active site (Figure 5C and 6). The SO_2 and NH group present adjacent to the isoxazole moiety overlaid on the acceptor features and accepts H-bonds from Ala108 and His435 residues. The oxygen atoms present in the hydroxy and phenoxy groups overlaid on the acceptor features accepts H-bond from NH group of Gly107, Ala108, and Ala195 residues. This compound, retrieved from Maybridge database, is a derivative of p-methoxyphenol with long substitutions at the phenoxy group.

Binding Mode of Hit 3

The GOLD fitness score of Hit 3 is 55.623 and its binding mode within the active site of CLR-I showed strong molecular interactions with the important active site residues. The H-bond network formed by Hit 3, which is a derivative of substituted benzodioxole, included interactions with Gly107, Ala108, Ser194, Ala195, and His435 residues (Figure 4D). The carbonyl oxygen, the nitrogen atoms of the only amide moiety and one of the oxygen atoms of sulfonamido (SO_2N) at the center of the compound have shown H-bond interactions with His435. The one of the oxygen atoms of SO_2 interacts with Ser194 and another interacts with Ala195 and Ala108. The nitrogen atom of the sulfonamide group has shown H-bond interaction with Ala108. One of the oxygen atoms of SO_2 overlaid on donor feature of the pharmacophore formed H-bond interactions with Ser194 and His435. Another oxygen atom overlaid on the acceptor feature interacts with Gly107, Ala108, and Ala195 (Figure 5D and 6). The carbonyl oxygen and nitrogen atoms of the amide moiety overlaid on the acceptor features have shown H-bond interactions with His435. The six-membered part of 1,3-benzodioxole ring overlaid on hydrophobic feature of the pharmacophore model has shown hydrophobic interactions with His435 and phenyl ring and the attached aliphatic side chain has shown hydrophobic interactions with Ala108. This hit compound, retrieved from Asinex database, is a derivative of N-(1,3-benzodioxol-5yl) acetamide. The pharmacophore overlay of the hit compounds along with most active compound 1 over the pharmacophoric feature with cage is given in Figure S7. The *SciFinder Scholar* and *PubChem* compound search confirmed that these hit

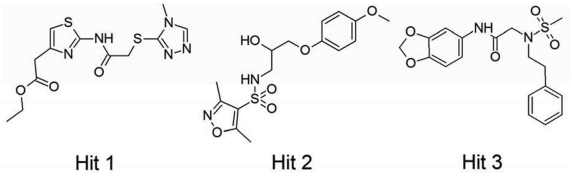

Figure 6: The 2D representation of final hit compounds.

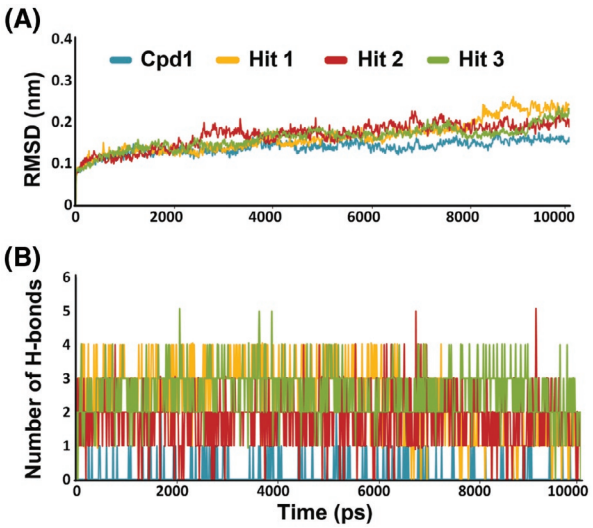

Figure 7: (**A**) The backbone root mean square deviation (RMSD) and (**B**) The number of protein-ligand hydrogen bond interactions formed in four systems during 10 ns MD simulations. Blue, yellow, maroon, and green colors represent compound 1 (Cpd 1), Hit 1, Hit 2, and Hit 3, respectively.

compounds are novel and were not reported elsewhere earlier for CEase inhibitory activities.

Binding Mode Validation

The binding modes of the 3 final hit compounds and the most active compound 1 were carried out for 10 ns MD simulations. Not all the docking programs available to date are effective to all protein targets. Hence we used MD simulation as an endorsement to validate the binding orientations of the docked final hit compounds. The four enzyme-TCH-inhibitor complexes, Cpd 1, Hit 1, Hit 2, and Hit 3, were prepared using the same procedure the hCEase_TCH system was prepared earlier in this study. The calculation of backbone RMSD of four systems revealed that all the systems have converged well during 10 ns MD simulations (Figure 7A). The H-bond interaction network between protein and ligand molecules was analyzed throughout the simulations. Interestingly, all the final hit compounds have shown more number of H-bonds (maximum of four or five) than the most active compound1. The number of H-bond interactions formed by the most active compound 1 was very low (none or one) throughout the simulation when compare to the hit compounds (Figure 7B). The final structure of every simulation was taken as representative structure and utilized to visualize the H-bond interactions formed with the key residues in the active site of the enzyme using DS (Figure S8). The number of H-bond interactions formed with the important catalytic residues was more with the final 3 hit compounds compared to the most active compound 1. These MD results have validated the final hit compounds to be used as potential leads in the designing of novel hCEase inhibitors.

Conclusion

Inhibition of hCEase has become very important strategy in treating hypercholesterolemia due to its role in atherocleorsis process. The presence of bile salts is very important for the activation of this enzyme. Till date there is no structure for hCEase complexed with bile salt was determined crystallographically. But the crystal structure of bCEase complexed with bile salts was determined and available in PDB. This bCEase shows more than 90% similarity with hCEase and thus the coordinates of the bile salts of bCEase were utilized to generate a hCEase model complexed with bile salts. The apoform and the complex protein were subjected to 5 ns MD simulations in order to observe the structural changes during the simulation. The

representative structures of MD simulation were selected using cluster analysis and used in structure-based pharmacophore generation study. The generated pharmacopore models were validated and used as a 3D query in database screening. The retrieved compounds were refined by various drug-likeness filters like Lipinki's rule of five and ADMET properties. The compounds that satisfied all the drug-like filters were subjected to molecular docking in order to refine the hits further based on the binding modes and molecular interactions with the active site residues. The selection of final compounds from the molecular docking study was restricted to remove the false positives that could arise from the database screening. The GOLD fitness score, mode of binding, and interaction with active site amino acids were used as the basic criteria to select the final compounds. The molecular docking results were carefully analyzed and 3 compounds, namely, Hit 1, Hit 2, and Hit 3, with GOLD fitness score values of 60.901, 59.558 and 55.623, respectively, were selected as final hits. The binding modes of the final hit compounds were further validated by performing 10 ns MD simulations. The MD simulation results have shown that the identified hit compounds binds the active site of the enzyme stronger than the most active compound 1 through more and stable number of hydrogen bonds. The final hits were subjected to novelty study using *SciFinder Scholar and PubChem Structure* search tools, which confirmed that these final hit compounds were not reported earlier for their CEase inhibition. The final hits as such or upon further optimization can be used as potential leads in novel CEase inhibitor designing.

Supplementary Material

Supplementary material dealing with this manuscript is available at no charge from the authors directly; the supplementary data can also be purchased from Adenine Press for US $50.00. It can also be downloaded free of charge from the author's server at the URL http://bio.gnu.ac.kr/JBSD_JS.

Acknowledgement

This research was supported by Basic Science Research Program (2009-0073267), Pioneer Research Center Program (2009-0081539), and Management of Climate Change Program (2010-0029084) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST) of Republic of Korea. And this work was also supported by the Next-Generation BioGreen 21 Program (PJ008038) from Rural Development Administration (RDA) of Republic of Korea.

References

1. T. Gordon, W. B. Kannel, W. P. Castelli, and T. R. Dawber. *Arch Intern Med 141*, 1128-1131 (1981).
2. S. M. Grundy, J. I. Cleeman, C. N. Merz, H. B. Brewer, Jr., L. T. Clark, D. B. Hunninghake, R. C. Pasternak, S. C. Smith, Jr., and N. J. Stone. *Arterioscler Throm Vasc Biol 24*, e149-e161 (2004).
3. I. Hsu, S. A. Spinler, and N. E. Johnson. *Ann Pharmacotherapy 29*, 743-759 (1995).
4. E. Boumendil and P. Tubert-Bitter. *Epidemiology 6*, 322-325 (1995).
5. T. Phan, J. G. McLeod, J. D. Pollard, O. Peiris, A. Rohan, and J. P. Halpern. *J Neurol Neurosurg Psychiatry 58*, 625-628 (1995).
6. E. Proksch. *Hautartz 46*, 76-80 (1995).
7. S. Terzyan, C. S. Wang, D. Downs, B. Hunter, and X. C. Zhang. *Protein Sci 9*, 1783-1790 (2000).
8. C. S. Wang and J. A. Hartsuck. *Biochim Biophys Acta 166*, 1-19 (1993).
9. G. Lin, W. C. Liao, and S. Y. Chiou. *Bioorg Med Chem 8*, 2601-2607 (2000).
10. X. Wang, C. S. Wang, J. Tang, F. Dyda, and X. C. Zhang. *Structure 5*, 1209-1218 (1997).
11. D. Y. Hui and P. N. Howles. *J Lipid Res 43*, 2017-2030 (2002).
12. N. Bruneau, D. Lombardo, and M. Bendayan. *J Cell Sci 111*, 2665-2679 (1998).
13. O Hernell. *Eur J Clin Invest 5*, 267-272 (1975).
14. L. Blackberg and O. Hernell. *FEBS Lett 323*, 207-210 (1993).
15. C. S. Wang and J. A. Hartsuck. *Biochim Biophys Acta 1166*, 1-19 (1993).

16. G. W. Vahouny, S. Weersing, and C. R. Treadwell. *Biochim Biophys Acta 98*, 607-616 (1965).
17. S. C. Myers-Payne, D. Y. Hui, H. L. Brockman, and F. Schroeder. *Biochemistry 34*, 3942-3947 (1995).
18. P. N. Howles, C. P. Carter, and D. Y. Hui. *J Biol Chem 271*, 7196-7202 (1996).
19. R. Shamir, W. J. Johnson, R. Zolfaghari, H. S. Lee, and E. A. Fisher. *Biochemistry 34*, 6351-6358 (1995).
20. D. Y. Hui and P. N. Howles. *J Lipid Res 43*, 2017-2030 (2002).
21. J. E. Heidrich, L. M. Contos, L. A. Hunsaker, L. M. Deck, and D. L. Vander Jagt. *BMC Pharmacol 4*, 5 (2004).
22. M. Pietsch and M. Gutschow. *J Biol Chem 277*, 24006-24013 (2002).
23. J. Brodt-Eppley, P. White, S. Jenkins, and D. Y. Hui. *Biochim Biophys Acta 1272*, 69-72 (1995).
24. D. Lombardo. *Biochim Biophys Acta 1533*, 1-28 (2001).
25. S. John, S. Thangapandian, S. Sakkiah, and K. W. Lee. *J Enzyme Inhib Med Chem 26*, 535-545 (2010).
26. J. J. Heynekamp, L. A. Hunsaker, T. A. Vander Jagt, R. E. Royer, L. M. Deck, and D. L. Vander Jagt. *Bioorg Med Chem 16*, 5285-5294 (2008).
27. O. Rebai, J. Le Petit-Thevenin, N. Bruneau, D. Lombardo, and A. Verine. *Arterioscler Thromb Vasc Biol 25*, 359-364 (2005).
28. M. Cygler, J. D. Schrag, J. L. Sussman, M. Harel, I. Silman, M. K. Gentry, and B. P. Doctor. *Protein Sci 2*, 366-382 (1993).
29. D. L. Ollis, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Reington, I. Silman, J. Schrag, J. L. Sussman, K. H. G. Verschueren, and A. Goldman. *Protein Eng 5*, 197-211 (1992).
30. D. M. Quinn and S. R. Feaster. In comprehensive biological catalysis. A mechanistic reference, M. Sinnot, Ed., Academic Press, San Diego, CA, pp. 455-482 (1998).
31. W. Boland, C. Frößl, and M. Lorenz. *Synthesis 12*, 1049-1072 (1991).
32. A. Svendsen. Sequence comparisons within the lipase family. In: lipases their structure, Biochemistry and Application, Woolley, P. and S. B. Petersen, Eds., Cambridge University Press, Cambridge, pp. 1-21 (1994).
33. S. R. Feaster and D. M. Quinn. *Methods Enzymol 286*, 231-252 (1997).
34. S. R. Feaster, K. Lee, N. Baker, D. Y. Hui, and D. M. Quinn. *Biochemistry 35*, 16723-16734 (1996).
35. G. Lin, C. T. Shieh, Y. C. Tsai, C. I. Hwang, C. P. Lu, and G. H. Cheng. *Biochim Biophys Acta 1431*, 500-511 (1999).
36. F. Batoli, H. K. Lin, F. Ghomashchi, M. H. Gelb, M. K. Jain, and R. Apitz-Castro. *J Biol Chem 269*, 15625-15630 (1994).
37. L. D. Sutton, J. S. Stout, L. Hosie, P. S. Spencer, and D. M. Quinn. *Biochem Biophys Res Commun 134*, 386-392 (1986).
38. D. G. Tew, H. F. Boyd, S. Ashman, C. Theobald, and C. A. Leach. *Biochemistry 37*, 10087-10093 (1994).
39. R. J. Kazlauskas. *J Am Chem Soc 111*, 4953-4959 (1989).
40. G. Lin, S. H. Liu, S. J. Chen, F. C. Wu, and H. L. Sun. *Tetrahedron Lett 34*, 6057-6058 (1993).
41. Z. Gong, Y. Zhao, and Y. Xiao. *J Biomol Struct Dyn 28*, 431-441 (2010).
42. J. Wiesner, Z. Kriz, K. Kuca, D. Jun, and J. Koca. *J Biomol Struct Dyn 28*, 393-403 (2010).
43. F. Mehrnejad and M. Zarei. *J Biomol Struct Dyn 27*, 551-559 (2010).
44. B. Jin, H. M. Lee, and S. K. Kim. *J Biomol Struct Dyn 27*, 457-464 (2010).
45. S. Roy and A. R. Thakur. *J Biomol Struct Dyn 27*, 443-455 (2010).
46. C. Carra and F. A. Cucinotta. *J Biomol Struct Dyn 27*, 407-427 (2010).
47. Y. Yu, Y. Wang, J. He, Y. Liu, H. Li, H. Zhang, and Y. Song. *J Biomol Struct Dyn 27*, 641-649 (2010).
48. Z. Cao and J. Wang. *J Biomol Struct Dyn 27*, 651-661 (2010).
49. S. Sharma, U. B. Sonavane, and R. R. Joshi. *J Biomol Struct Dyn 27*, 663-676 (2010).
50. M. J. Aman, H. Karauzum, M. G. Bowden, and T. L. Nguyen. *J Biomol Struct Dyn 28*, 1-12 (2010).
51. L. K. Chang, J. H. Zhao, H. L. Liu, J. W. Wu, C. K. Chuang, K. T. Liu, J. T. Chen, W. B. Tsai, and Y. Ho. *J Biomol Struct Dyn 28*, 39-50 (2010).
52. Y. Yuan, M. H. Knaggs, L. B. Poole, J. S. Fetrow, and F. R. Salsbury, Jr., *J Biomol Struct Dyn 28*, 51-70 (2010).
53. C. Koshy, M. Parthiban, and R. Sowdhamini. *J Biomol Struct Dyn 28*, 71-83 (2010).
54. A. Sharadadevi and R. Nagaraj. *J Biomol Struct Dyn 27*, 541-550 (2010).
55. Tuszynska and J. M. Bujnicki. *J Biomol Struct Dyn 27*, 511-520 (2010).
56. P. Sklenovsky and M. Otyepka. *J Biomol Struct Dyn 27*, 521-539 (2010).
57. J. F. Varughese, J. M. Chalovich, and Y. Li. *J Biomol Struct Dyn 28*, 159-173 (2010).
58. R. Nasiri, H. Bahrami, M. Zahedi, A. A. Moosavi-Movahedi, and N. Sattarahmady. *J Biomol Struct Dyn 28*, 211-226 (2010).
59. Z. Cao, L. Liu, and J. Wang. *J Biomol Struct Dyn 28*, 343-353 (2010).

60. L. Zhong. *J Biomol Struct Dyn 28*, 355-361 (2010).
61. H. M. Lee, B. Jin, S. W. Han, and S. K. Kim. *J Biomol Struct Dyn 28*, 421-430 (2010).
62. H. J. C. Berendsen, V. D. Spoel, and R. V. Drunen. *Comput Phys Commun 95*, 43-56 (1995).
63. E. Lindahl, B. Hess, and V. D. Spoel. *J Mol Model 7*, 306-317 (2001).
64. A. W. Schuttelkopf and D. M. F. van Aalten. *Acta Cryst D60*, 1355-1363 (2004).
65. E. Fadrná, K. Hladecková, and J. Koca. *J Biomol Struct Dyn 23*, 151-162 (2005).
66. J. Varady, X. Wu, X. Fang, J. Min, Z. Hu, B. Levant, and S. J. Wang. *J Med Chem 46*, 4377-4392 (2003).
67. S. Karkola, H. D. Höltje, and K. A. Wähälä. *J Steroid Biochem Mol Biol 105*, 63-70 (2007).
68. Y. Tao, Z. H. Rao, and S. Q. Liu. *J Biomol Struct Dyn 28*, 143-158 (2010).
69. G. Lin, W. C. Liao, and S. Y. Chiou. *Bioorg Med Chem 8*, 2601-2607 (2000).
70. G. Lin, C. T. Shieh, Y. C. Tsai, C. I. Hwang, C. P. Lu, and G. H. Chen. *Biochim Biophys Acta 1431*, 500-511 (1999).
71. L. M. Deck, M. L. Baca, S. L. Salas, A. Lucy, L. A. Hunsaker, and D. L. Vander Jagt. *J Med Chem 42*, 4250-4256 (1999).
72. G. Lin, C. T. Shieh, H. C. Ho, J. Y. Chouhwang, W. Y. Lin, and C. P. Lu. *Biochemistry 38*, 9971-9981 (1999).
73. M. Pietsch and M. Gutschow. *J Med Chem 48*, 8270-8288 (2005).
74. K. Sakai, K. Watanabe, K. Masuda, M. Tsuji, K. Hasumi, and A. *J Antibiot 48*, 447-456 (1995).
75. A. Eilfeld, C. M. Gonzalez Tanarro, M. Frizler, J. Sieler, B. Schulze, and M. Gutschow. *Bioorg Med Chem 16*, 8127-8135 (2008).
76. S. Thangapandian, S. John, S. Sakkiah, and K. W. Lee. *Eur J Med Chem 45*, 4409-4417 (2010).
77. C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. *Adv Drug Deliv Rev 46*, 3-26 (2001).
78. W. P. Walters and M. A. Murcko. *Adv Drug Deliv Rev 54*, 255-271 (2002).
79. S. John, S. Thangapandian, S. Sakkiah, and K. W. Lee. *Eur J Med Chem 45*, 4004-4012 (2010).
80. G. Jones, P. Willett, R. C. Glen, A. R. Leach, and R. J. Taylor. *Mol Biol 267*, 727-748 (1997).
81. M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor. *Proteins 52*, 609-623 (2003).
82. R. Thomsen and M. H. Christensen. *J Med Chem 49*, 3315-3321 (2006).
83. A. C. Wallace, R. A. Laskowski, and J. M. Thornton. *Protein Eng 8*, 127-134 (1995).
84. A. B. Wagner. *J Chem Inf Model 46*, 767-774 (2006).
85. Y. Wang, E. Bolton, S. Dracheva, K. Karapetyan, B. A. Shoemaker, T. O. Suzek, J. Wang, J. Xiao, J. Zhang, and S. H. Bryant. *Nucleic Acids Res 38*, D255-D266 (2010).
86. S. A. Moore, R. L. Kingston, K. M. Loomes, O. Hernell, L. Bläckberg, H. M. Baker, and E. N. Baker. *J Mol Biol 312*, 511-523 (2001).

Date Received: July 18, 2011

Communicated by the Editor Ramaswamy H. Sarma