Journal of Bioinformatics and Computational Biology
Vol. 9, Suppl. 1 (2011) 15–35

DOI: 10.1142/S0219720011005732

Imperial College Press
www.icpress.co.uk

DISCOVERY AND EVALUATION OF POTENTIAL SONIC HEDGEHOG SIGNALING PATHWAY INHIBITORS USING PHARMACOPHORE MODELING AND MOLECULAR DYNAMICS SIMULATIONS

SWAN HWANG*, SUNDARAPANDIAN THANGAPANDIAN†, YUNO LEE‡, SUGUNADEVI SAKKIAH§, SHALINI JOHN¶ and KEUN WOO LEE‖

Division of Applied Life Science (BK21 program)
Systems and Synthetic Agrobiotech Center (SSAC)
Plant Molecular Biology and Biotechnology Research Center (PMBBRC)
Research Institute of Natural Science (RINS)
Gyeongsang National University (GNU)
501 Jinju-daero, Gazha-dong, Jinju, 660-701 Republic of Korea
**swan@bio.gnu.ac.kr*
†*sunder@bio.gnu.ac.kr*
‡*youknow@bio.gnu.ac.kr*
§*suguna@bio.gnu.ac.kr*
¶*shalini@bio.gnu.ac.kr*
‖*kwlee@gnu.ac.kr*

Received 25 July 2011
Revised 12 August 2011
Accepted 12 August 2011

Sonic hedgehog (Shh) plays an important role in the activation of Shh signaling pathway that regulates preservation and rebirth of adult tissues. An abnormal activation of this pathway has been identified in hyperplasia and various tumorigenesis. Hence the inhibition of this pathway using a Shh inhibitor might be an efficient way to treat a wide range of malignancies. This study was done in order to develop a lead chemical candidate that has an inhibitory function in the Shh signaling pathway. We have generated common feature pharmacophore models using three-dimensional (3D) structural information of robotnikinin, an inhibitor of the Shh signaling pathway, and its analogs. These models have been validated with fit values of robotnikinin and its analogs, and the best model was used as a 3D structural query to screen chemical databases. The hit compounds resulted from the screening docked into a proposed binding site of the Shh named pseudo-active site. Molecular dynamics (MD) simulations were performed to investigate detailed binding modes and molecular interactions between the hit compounds and functional residues of the pseudo-active site. The results of the MD simulation analyses revealed that the hit compounds can bind the pseudo-active site with high

‖Corresponding author.

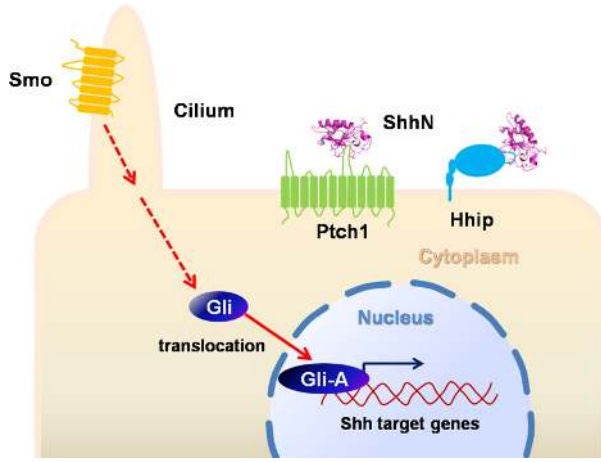
affinity than robotnikinin. As a result of this study, a candidate inhibitor (GK 03795) was selected as a potential lead to be employed in future Shh inhibitor design.

Keywords: Sonic hedgehog; robotnikinin; pharmacophore; molecular dynamics simulation; drug design.

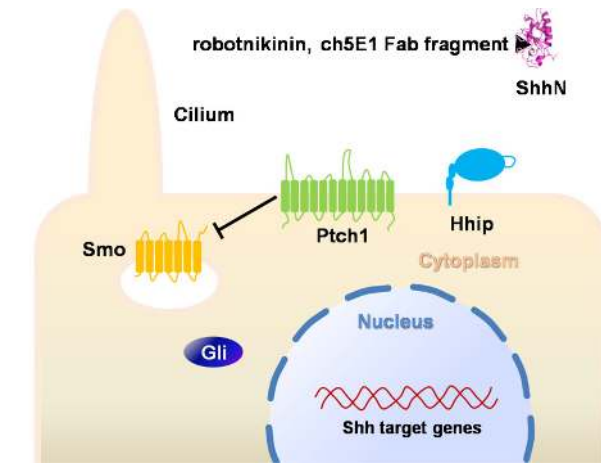
1. Introduction

The hedgehog signaling pathway has a myriad of important functions related to body patterning, organ development, central nervous system patterning during embryogenesis, and epithelial cell proliferation.[1–3] Additional functions of this pathway are reported as maintaining adult coronary vasculature,[4] regenerating neural tissues,[5] and preserving and repairing other tissues.[2] However, an abnormal activation of this pathway can lead to medulloblastoma,[6] prostate cancer proliferation,[7] basal cell carcinomas,[8] myloid leukemia,[9] basal cell carcinomas, gliomas, sarcomas, tumors of the digestive tract, small cell lung cancers, and pancreatic carcinomas.[7] The hedgehog signaling pathway is extensively being researched because of these crucial roles. The hedgehog protein diffuses from a source to target area up to 30-cell-diameter and it acts as morphogen.[10] The hedgehog signaling pathway is initiated by binding of hedgehog onto 12-transmembrane cell surface receptor Patched (Ptch). This interaction alleviates Ptch's catalytic inhibition of G-protein–coupled receptor (GPCR)-like protein Smoothened (Smo). An activation of Smo recruits downstream components into primary cilium for delivering signals. These signals stimulate glioma-associated oncogene homolog zinc finger protein (Gli) transcription factors that provide assistance during the transcription of hedgehog target genes (Fig. 1(a)). In the absence of hedgehog protein, the hedgehog signaling pathway is maintained in inactive state (Fig. 1(b)).[2,11,12]

The hedgehog protein family group consists of three members in mammals: Sonic (Shh), Indian (Ihh), and Desert (Dhh) hedgehogs. Shh stimulates cell proliferation of neural tissue, hair, tooth, whisker, and gut as well as specifies cell fate of neural tissues, limb, and somites in morphogenesis of lung branching, teeth, hair, and prostate. The function of Ihh is to breed cells in cartilage and gut. It is also concerned with bone development and endothelial cell induction. The Dhh is related to gonad and peripheral nerve organization.[9] The Shh is the best-studied protein among the members of the hedgehog protein families. A Shh precursor undergoes signal sequence cleavage, which divides it into 20 kDa Shh N-terminal peptide (ShhN) and 25 kDa Shh C-terminal peptide (ShhC).[13,14] Through this reaction, cholesterol is covalently attached to the ShhN, and this process is essential for the long-range diffusion of the ShhN from the source cell.[9] The Shh signaling pathway can be suppressed by macromolecules such as Cell adhesion molecule–related/downregulated by oncogenes (Cdo),[15] Hedgehog-interacting protein (Hhip),[16] and murine:human chimeric 5E1 (ch5E1).[17] The Cdo and Hhip reside in membrane of cell, while the ch5E1 is moving as an inhibitory antibody. All of these macromolecules share similar binding area of the Shh, which is the periphery



(a)



(b)

Fig. 1. A schematic representation of the Shh signaling pathway. (a) The activation of the Shh signaling pathway, where Gli-A represents the activated form of Gli. (b) The inhibition of the Shh signaling pathway.

or the core of the pseudo-active site that coordinates zinc and two calcium ions. Though the binding area of these macromolecules looks similar, their binding with the metal ions present in the binding pocket is different.

To date, a number of Shh signaling pathway inhibitors have been reported,. These inhibitors include Smo inhibitors such as cyclopamine and its analogs or synthetic antagonists, Gli-transcription or ShhN inhibitors such as GANT61, GANT58, or robotnikinin.[18] Among these several inhibitors, robotnikinin is the only small-molecule inhibitor that has a function of acting on the Shh itself. The ShhN may

be a proper target to develop the inhibitory lead candidates, which can halt the Shh signaling pathway because formation of ShhN–Ptch complex is the initial and at the same time the essential step in this pathway. ShhN has an apparent catalytic site which has structural homology to bacterial carboxypeptidase. This apparent catalytic site was presumably responsible for the hydrolytic activity. Contrary to this expectation, ShhN variants generated by mutating important residues of carboxypeptidase domain did not result in loss of the Shh signaling.[19] Alternatively, we hypothesized that a physical interruption between the Shh and Ptch can be a valuable measure to block the Shh signaling pathway and can be used in the treatment of various cancers, not by the inhibition of the hydrolytic activity of the Shh.

Accordingly, we have generated a pharmacophore model consisting of common chemical features of robotnikinin and its analogs. This pharmacophore model was used as a three-dimensional (3D) structural query in virtual screening of chemical databases. Prediction of structural binding modes of robotnikinin and hit compounds resulted from the virtual screening at the pseudo-active site of the Shh was facilitated by molecular docking simulation. The interactions of these ligand–protein complexes were evaluated by MD simulations of 4 ns.

2. Methods and Materials

2.1. *Collection of known inhibitors*

Robotnikinin and its analogs with inhibitory profiles toward the Shh signaling pathway were collected from the literature to construct a training set.[20] Each training set compound has 12-, 13- or 14-membered macrocyclic scaffold with different substitutions on C2 and C6 positions (Fig. 2).

These training set compounds were classified into highly active, moderately active, and inactive compounds, based on their EC_{50} values against Gli. The compounds showing an EC_{50} value of less than 10 μM were classified as highly active compounds, whereas the compounds with EC_{50} value between 10 μM and 20 μM were classified as moderately active compounds. The compounds with EC_{50} value greater than 20 μM were classified as inactive compounds.

2.2. *Molecular modeling of conformational models*

The chemical structures of the training set compounds were sketched using *ChemSketch* version 12 program.[21] The two-dimensional (2D) structures of these compounds were exported and converted into 3D structures using Accelrys Discovery Studio 2.5 (DS) program, followed by energy minimization. The *Smart Minimizer* option available in *Energy Minimization* protocol was used in this step. In order to ensure a wide coverage of spatial accessibility of the training set compounds, diverse conformational models for each compound were generated by poling algorithm[22] employing *Generate Conformations* module of DS. In this process,

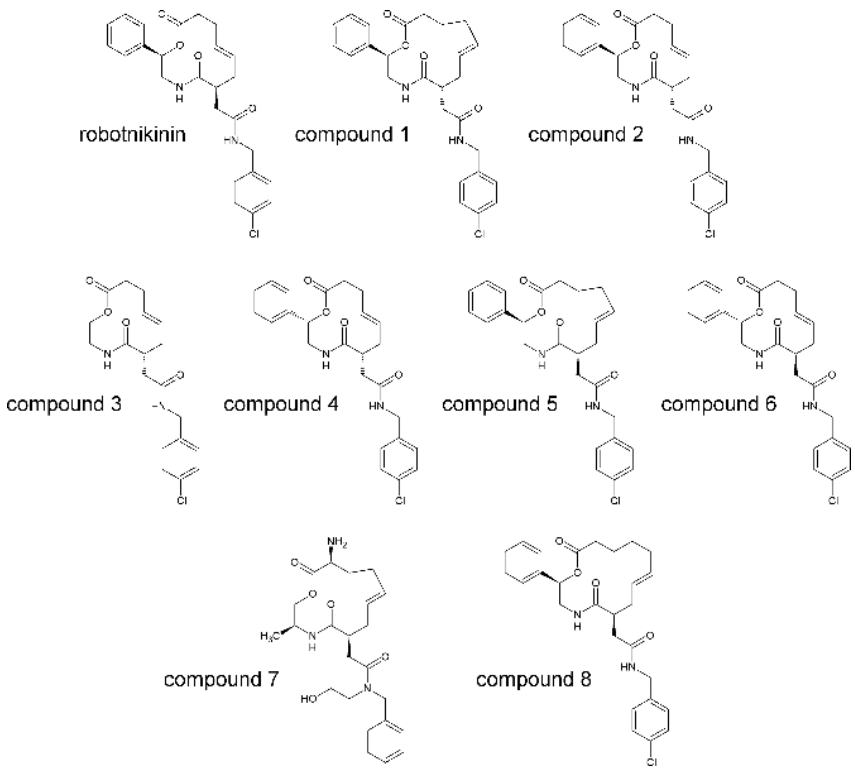


Fig. 2. Two-dimensional chemical structures of the training set compounds.

Principal and *MaxOmitFeat* values of 2 and 0 were set to highly active compounds, whereas 1 and 1, respectively, were set for moderately active compounds. The inactive compounds were given 0 and 2 for *Principal* and *MaxOmitFeat* values, respectively. A maximum limitation of the number of conformations produced was set to 255.

2.3. *Pharmacophore model generation*

The definition of a pharmacophore was recently modified by IUPAC, which stated that it is "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response".[23] All the conformational models of the training set compounds were provided as sources to generate the pharmacophore models. Additionally, the necessary chemical features for constructing the pharmacophore model such as hydrogen bond acceptor (HBA), hydrogen bond donor (HBD), and hydrophobic (HY) and ring aromatic (RA) features were predicted based on the training set compounds by applying *Feature Mapping* module from DS. Multiple pharmacophore models were created based on the conformational models and the

chemical features of the training set compounds using *Common Feature Pharmacophore Generation* module of DS. The *Maximum Excluded Volumes* value of 3 was included to apply conformational restrictions reflecting structural information from the inactive compounds and *Maximum Features* value was assigned to 6 in this process.

2.4. *Database searching and drug-likeness prediction*

The virtual screening process was performed using *Ligand Pharmacophore Mapping* module of DS to retrieve the potential candidate inhibitors from databases. The best pharmacophore model that resulted from the *Common Feature Pharmacophore Generation* procedure was used as a query for retrieving structurally diverse compounds from chemical databases mapping the features of the generated pharmacophore model. The *Maybridge* (59,652 compounds) and *Chembridge* (50,000 compounds) databases were used in the virtual screening. The *MaxOmitFeat* option was assigned to 1 and *Fitting Method* option was set to *Best/Flexible* in this process. In order to determine whether the compounds resulted from the *Ligand Pharmacophore Mapping* possess drug-like properties, Lipinski's rule of five and ADMET (absorption, distribution, metabolism, elimination and toxicology) filtrations were introduced.[24] The *Lipinski Filter* option of DS was used to evaluate drug-likeness of the retrieved compounds in this process. The ADMET properties were subsequently calculated using *ADMET Descriptors* module of DS.[25–27]

2.5. *Molecular docking*

In order to dock robotnikinin and the drug-like hit compounds into the pseudo-active site of the Shh, the 3D structure of the ShhN (residues 38–191) was acquired from the X-ray crystal structure of the human Shh-Hhip complex structure (PDB ID : 3HO5). This crystal structure was determined at a resolution of 3.01 Å containing one zinc and two calcium ions at the pseudo-active site. In this study, *LigandFit* module as implemented in DS was used to perform molecular docking simulations. Before the docking process, the pseudo-active site of the Shh structure was specified as a binding site and the energy grid option was set to PLP1. The Shh structure was kept rigid in subsequent process and diverse conformations of the small molecules generated in the state of bond lengths and bond angles were untouched and torsion angles were randomized by employing Monte Carlo algorithm. The conformations of the drug-like compounds were energy minimized by the *Smart Minimizer* to retrieve candidate conformations of the small molecules that are accommodative in the grid of the binding site. *Smart Minimizer* performed energy minimization with 1,000 steps of *Steepest Descent* with a minimization gradient tolerance of 0.001 using *CHARMm* force field, followed by the *Conjugate Gradient* minimization until the energy of the conformation of a drug-like compound converged to a local minimum. The minimized conformations of the drug-like compounds were matched to the shape of the binding site and were docked into the binding site. Fitness of each

conformation was calculated using various scoring functions (LigScore1, LigScore2, -PLP1, -PLP2, Jain, -PMF, Ludi_1, Ludi_2, Ludi_3, -PMF04, DOCK_SCORE). The *Consensus Score* module of DS was applied to compute the consensus scores for the multiple scores resulted from *LigandFit* docking simulation. The hit compounds that scored top using the consensus scoring function were selected as final lead compounds.

2.6. *Molecular dynamics simulation*

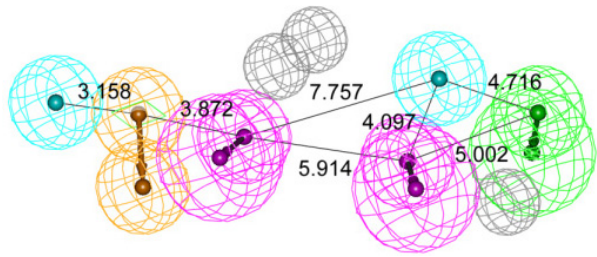
A series of MD simulations were performed for complexes of Shh-robotnikinin and Shh-hits with *GROMOS96 43a1* force field using *GROMACS 4.5.3* program with high parallel computing.[28,29] The properties of robotnikinin and the hit compounds such as topologies, net charges, force field parameters, and correct hybridization state of each atom were prepared using *PRODRG 2.5* server.[30] To solvate the protein–ligand complexes, a rectangular water box that is 1.5 nm longer from the surface of the protein in each dimension was created and explicit Simple Point Charge (SPC) water models were added around each protein–ligand complexes within the water box. In order to neutralize the net charge of each system, four chloride anions were added replacing four SPC water models. Subsequently, to reach a local minimum conformation, each system was energetically minimized by *Steepest Descent* method until the energy reaches the tolerance value of 2,000 kJ/mol·nm^{-1} in a maximum of 10,000 steps. In system equilibration process, the peptide heavy atoms of the Shh were fixed by Linear Constraint Solver (LINCS) constraint algorithm[31] and the SPC water models with chloride anions were allowed to move freely for 100 ps under constant pressure of 1 bar and temperature of 300 K. The production MD simulations were performed for 4 ns without any restraints of movement under condition of 1 bar and 300 K. The coordinates of all the atoms in each system were saved at every 1 ps during the MD simulations. Periodic boundary conditions were given in all directions by constant pressure-temperature (NPT) ensemble and direct-space interactions were calculated by particle mesh Ewald (PME) method.[32] The Lennard-Jones and Coulombic interactions were calculated within cut-off distances of 1.4 and 0.9 nm.

3. Results and Discussion

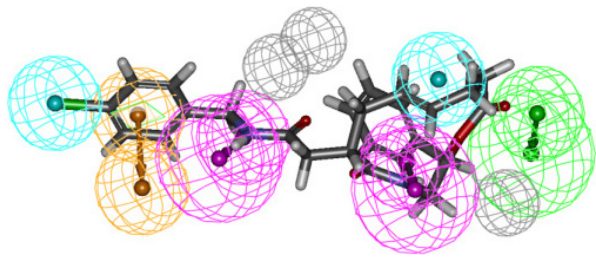
3.1. *Generation of pharmacophore models*

The pharmacophore models were generated by the combination of the common chemical features of the training set compounds that are structurally diverse with good range of biological activities and excluded volumes that provide structural information from the inactive compounds (Fig. 3(a)).

The chemical features of the training set compounds suggested by the *Feature Mapping* protocol were considered for constructing pharmacophore hypotheses. Ten



(a)



(b)

Fig. 3. The generated pharmacophore model. Features are color coded as follows: HBD, magenta; HBA, green; HY, cyan; RA, orange; excluded volume, gray. (a) The pharmacophore model, HYPO4, with its interfeature distance constraints. (b) Mapping of robotnikinin upon HYPO4.

Table 1. Results of common feature pharmacophore model generation.

Hypothesis	Features[a]	Rank	Direct Hit[b]	Partial Hit[c]
HYPO1	RHHDAA	124.995	111111	0000000
HYPO2	RHHDDA	124.397	111111	0000000
HYPO3	RHHDDA	124.397	111111	0000000
HYPO4	RHHDDA	124.103	111111	0000000
HYPO5	RHHDDA	123.474	111111	0000000
HYPO6	RHHAAA	123.243	111111	0000000
HYPO7	RHHDAA	122.747	111111	0000000
HYPO8	RHHDAA	122.747	111111	0000000
HYPO9	RHHDDA	122.744	111111	0000000
HYPO10	RHHDAA	122.522	111111	0000000

[a]R, ring aromatic; H, hydrophobic; D, hydrogen bond donor; A, hydrogen bond acceptor.
[b,c]Direct hit or Partial hit means direct or partial match of compound to the features. 1, match and 0, no match.

hypotheses (HYPO1 to HYPO10) were obtained from the *Common Feature Pharmacophore Generation* method (Table 1).

The highest rank score implies the best hypothesis. Among these 10 hypotheses, HYPO4 was considered as the best and selected for further study because it could optimally discriminate the compounds of training set into active, moderately active,

and inactive compounds correlating with their experimental activities. The HYPO4 scored 124.103 and comprised one RA, two HY, two HBD, and one HBA features.

The fit value that is measured based on how well a chemical compound maps over the pharmacophore model was calculated for all the training set and hit compounds using ligand pharmacophore mapping method (Table 2).

The fit values of the training set compounds had shown a positive correlation with corresponding Gli EC_{50}. In this ligand pharmacophore mapping, the only chloride atom of robotnikinin mapped on to the HY feature of HYPO4, whereas the phenyl ring attached to it mapped on to the RA feature of HYPO4. The NH of amide group mapped on to one of the HBD features of HYPO4. The hydrocarbon portion, NH and one of the cyclic carbonyl groups of the macrocyclic ring mapped over the HY, second HBD and HBA features, respectively (Fig 3(b)). On the other hand, the inactive compounds were not mapped on to the chemical features of HYPO4 well and indicated the reason for their low active nature. Hence, we concluded that the HYPO4 is valid enough to retrieve potential compounds from the chemical databases that can be utilized in potent Shh inhibitor design.

3.2. *Database searching and drug-likeness prediction*

The HYPO4 was used as a 3D query during *Ligand Pharmacophore Mapping* procedure for finding Shh inhibitor-like chemical compounds from the databases. The query returned 2,915 and 47 hit compounds from *Maybridge* and *Chembridge* databases, respectively (Fig. 4).

In order to choose the compounds with drug-like properties from the results of the virtual screening, the Lipinski's rule of five and ADMET filters were applied. According to Lipinski's rule of five, a drug-like compound should contain less than 5 hydrogen bond donors, less than 10 hydrogen bond acceptors, a molecular weight of less than 500 kDa, and a log P value of less than 5. Subsequently, the ADMET

Table 2. The training set compounds along with their EC50 values against Gli and fit values based on the selected pharmacophore model, HYPO4.

Compound	EC_{50} (μM)	Fit value	Group
Robotnikinin	4	3.98	Highly active
1	4	5.99	
2	7	4.33	
3	15	3.18	Moderately active
4	15	2.66	
5	15	1.99	
6	10	0.22	
7	n.a[a]	0.10	Inactive
8	n.a	$1.76e^{-06}$	

[a]No detectable activity.

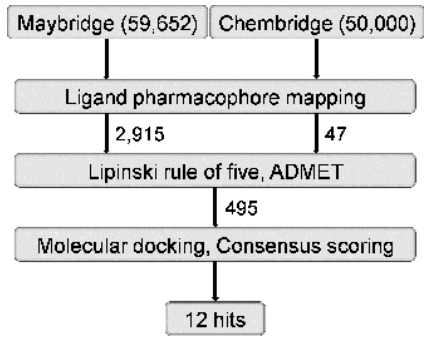


Fig. 4. Flowchart representing the procedure followed in virtual screening coupled with drug-likeness filtering, molecular docking, and consensus scoring.

filters were also utilized for selecting the compounds with good or moderate human intestinal absorption, low blood–brain barrier (BBB) penetration, no inhibition of *CYP2D6*, and no hepatotoxicity from the hit compounds filtered by the Lipinski's rule of five. Finally, 495 drug-like compounds resulted from the drug-like filters were subjected to molecular docking simulations to observe protein–ligand interactions.

3.3. *Molecular docking*

We investigated the necessary interactions between the Shh and Ptch prior to the molecular docking study, but the 3D structure of the Ptch was not solved yet on the account of difficulty in solving the 3D structure of a transmembrane protein. Thus instead of Ptch, we have studied about the 3D structural information of the human Shh-Hhip complex (Figs. 5(a) and 5(b)).

In order to investigate the sequence similarity between the Ptch and Hhip from the amino acids sequences, a multiple sequence alignment reported by Bosanac *et al.*,[33] was reproduced with the several species of Ptchs and human Hhip using Clustal X version 2 program (Fig. 5(c)).[34] The conserved sequences resulted from the multiple sequence alignment were detected on the Ptch loop 2 (L2) peptides that are the second large extracellular loop region located between transmembrane domains 7 and 8. The deletion of *Drosophila* Ptch L2 peptides, a homolog of human Ptch, results in an interruption of the Shh signaling, and this result describes the important role of the L2 peptides in the Shh signaling.[35,36]

From the multiple sequence alignment, the L2 part of Hhip exhibited sequence similarity with the conserved Ptch L2 peptides. We investigated the interactions between the Shh and Hhip L2 peptides on the grounds that the part of the Hhip L2 peptides has sequence similarity with the L2 of the human Ptch. A part of the Hhip L2 peptide (M373-D387) has displayed strong interactions with the critical residues of the Shh pseudo-active site in the Shh-Hhip complex X-ray crystal structure (Figs. 5(a) and 5(b)). In this interaction, the D383 of Hhip coordinated the zinc ion and E381 of Hhip was hydrogen bonded with the K87 of the Shh and it was

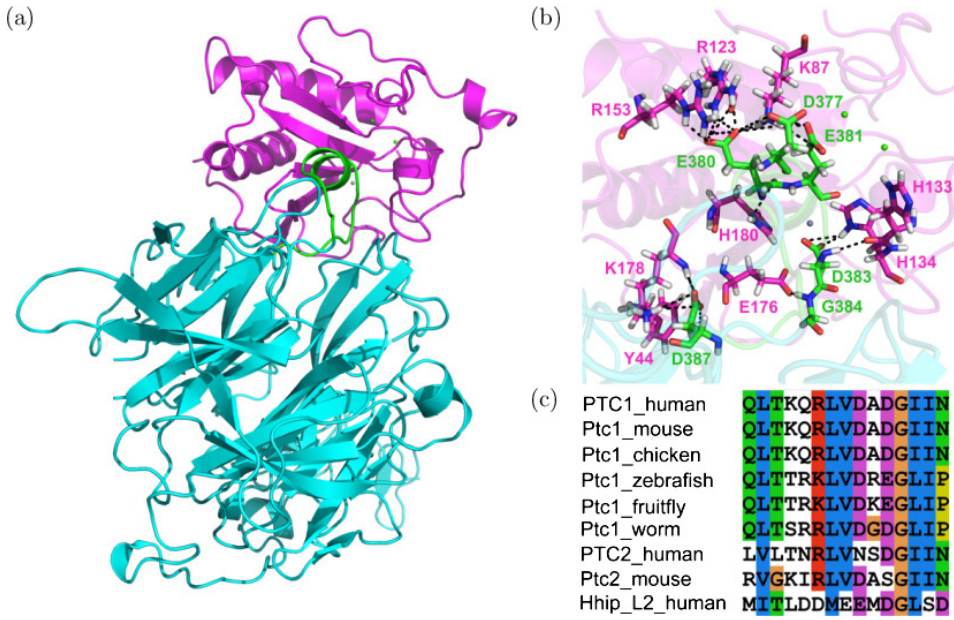


Fig. 5. Assessment of binding components at the pseudo-active site of ShhN and sequence analysis between the L2 peptide of Hhip and Ptchs present in various species. (a) The X-ray crystal structure of ShhN-Hhip complex. The structure of the ShhN, Hhip and the L2 peptide of Hhip were shown in magenta, cyan and green colors, respectively. (b) The hydrogen bond interactions in the ShhN-Hhip complex structure are indicated by black dashed lines. The zinc and two calcium ions depicted as bluish gray and green spheres, respectively. Pseudo-active site solvent molecules in this structure were omitted for clarity. (c) The result of multiple sequence alignment between Hhip L2 peptide and various species of Ptchs. We reproduced the result of sequence alignment that previously reported.[33]

also presented close to one of the calcium ions bound in the Shh. The necessity of the metal ions in the activation of the Shh signaling pathway has already been reported through many experiments.[17,37] Because the E381 and the D383 of Hhip are conserved in the L2 regions in various species of Ptch, we deduced that if a hit compound can coordinate or strongly bind the regions of zinc and calcium ion instead of the Ptch, it could be a promising drug candidate for interrupting the Shh signaling pathway. In order to identify the interactions between protein and ligand, robotnikinin and the hit compounds resulted from the drug-likeness filtration were docked into to the pseudo-active site of the Shh containing zinc and calcium ions.

The *LigandFit* docking simulations require a specified binding site and this site was defined based on the shape of the pseudo-active site of the Shh structure. In order to obtain a clearly defined binding site, a known inhibitor, robotnikinin, was docked into the binding site of the Shh protein. The best conformation of robotnikinin has shown interactions with the metal ions and important pseudo-active site residues (data not shown). From the result of robotnikinin docking, the

Table 3. The docking scores of robotnikinin and the hit compounds computed using 11 different scoring functions.

	Robotnikinin	GK 03795	HTS 06264	HTS 06766	HTS 07674	SEW 05450
LigScore1	3.93	5.54	4.02	4.83	4.73	5.54
LigScore2	5.15	5.21	4.27	4.68	4.21	5.68
-PLP1	98.99	98.58	98.25	91.33	89	104.72
-PLP2	94.03	97.85	82.22	80.88	82.49	97.13
Jain	3.07	3.79	4.18	5.64	3.38	2.42
-PMF	192.03	166.53	212.41	205.5	186.5	177.94
Ludi_1	497	716	575	610	554	564
Ludi_2	402	619	477	525	468	489
Ludi_3	521	785	820	644	718	733
-PMF04	130.24	136.49	153.94	158.67	119.89	142.07
DOCK_ SCORE	95.079	90.98	101.74	92.667	95.535	94.742
Consensus[a]	9	11	11	11	11	11

[a]The consensus scores represent a comprehensive score for each ligand.

location of the specified binding site was obtained and it was used to dock other hit compounds in the pseudo-active site of the Shh structure. Subsequently, the 495 hit compounds obtained from the database screening were docked into the specified binding site. Structurally diverse conformations of each of the hit compounds were evaluated by a variety of scoring functions (11 properties) that used in calculating the fitness at the specified binding site (Table 3). Consensus scoring function was employed to consider all the scoring functions resulted from *LigandFit* docking simulations. We selected the 12 diverse hit compounds with the best consensus score of 11 that is higher than the consensus score of robotnikinin (9). Finally, five compounds that have shown strong molecular interactions or bound close to the metal ions and the residue K87 were selected and their protein–ligand complexes were subjected to MD simulations (Fig. 6).

3.4. *Molecular dynamics simulation*

In order to elucidate the motions of individual particle of molecules,[38] the Shh-robotnikinin and the Shh-hit compound complexes (Shh with GK 03795, HTS 06264, HTS 06766, HTS 07674, SEW 05450) were subjected to 4 ns MD simulations. The root mean square deviation (RMSD) values were calculated by differences in the coordinates between backbone carbon atoms of initial complex structure and that of the individual snapshots generated as a function of time (in ps). In the result of RMSD plot, we observed that all the systems have converged and sustained in the native state (Fig. 7(a)).

In addition to the calculation of RMSD, we analyzed the number of hydrogen bonds, the representative structures and changes of distance between specific atoms, the interaction energy between protein and ligands by calculating the Lennard-Jones and the Coulombic energy parameters. The criterion for hydrogen bond formation between protein and ligands was 0.35 nm. In this result, HTS 06766 had

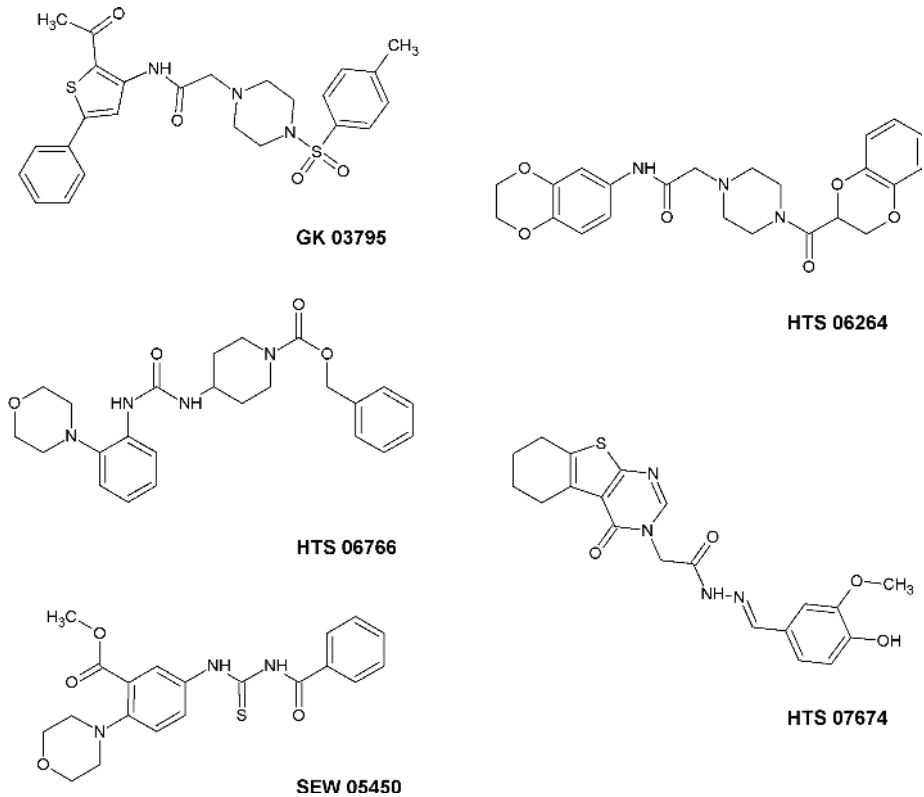


Fig. 6. 2D chemical structures of the identified hit compounds.

shown strong hydrogen bond interactions with the Shh before about 3300 ps, but it was dramatically decreased in the later stage. In case of GK 03795 complex, the formed hydrogen bonds were persistently maintained with the Shh and the average number of hydrogen bonds during the meantime of the MD simulations was as high as the Shh-robotnikinin complex (Fig. 7(b)). In Lennard-Jones short-range (LJ-SR) and Coulombic short-range (Coul-SR) potential energy calculations between protein and ligands, the Shh-GK 03795, Shh-SEW 05450 complexes displayed relatively low energy pattern compared to that of other systems including robotnikinin complex (Figs. 7(c) and 7(d)).

We also calculated the representative structures of the protein–ligand complexes to compare the binding modes of robotnikinin and the hit compounds against the Shh (Fig. 8).

In order to find out the representative structures from structurally adjusted conformations of the complexes, we chose stabilized time interval from the results of the each of RMSD plots (2 ns to 4 ns), and the covariance matrices of them were obtained and diagonalized. The RMSD values were calculated by comparison between these covariance matrices and each snapshot that is generated from

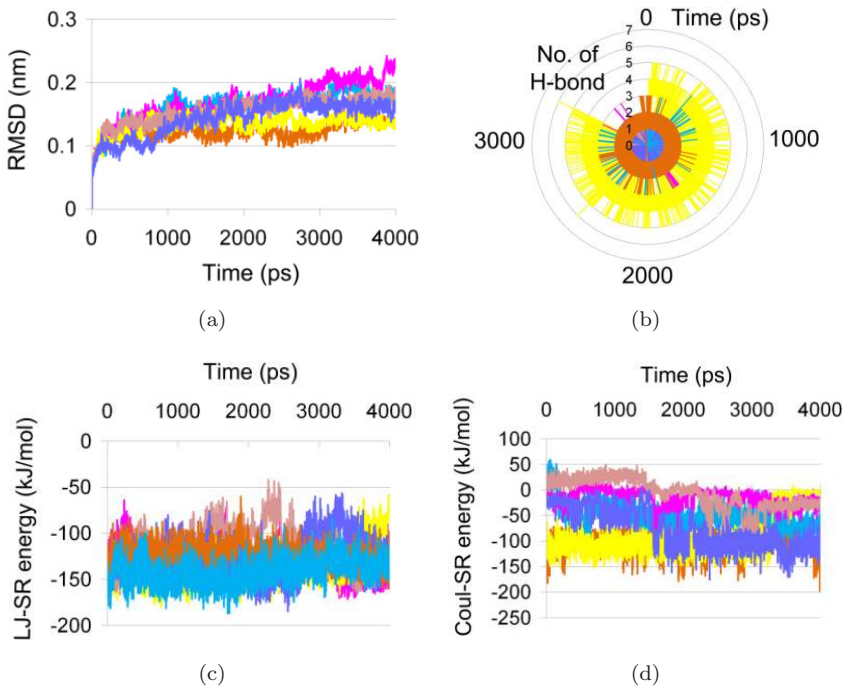


Fig. 7. The results of MD simulations. (a) The RMSD of complex structures, (b) The number of hydrogen bonds between Shh and ligands, (c) Lennard-Jones potential energy between Shh and ligands in short range (LJ-SR energy), (d) Coulomb energy between Shh and ligands in short range (Coul-SR energy). Complex structures are color coded as follows : Shh-robotnikinin, orange; Shh-GK 03795, cyan; Shh-HTS 06264, magenta; Shh-HTS 06766, yellow; Shh-HTS 07674, pink; SEW 05450, slate blue.

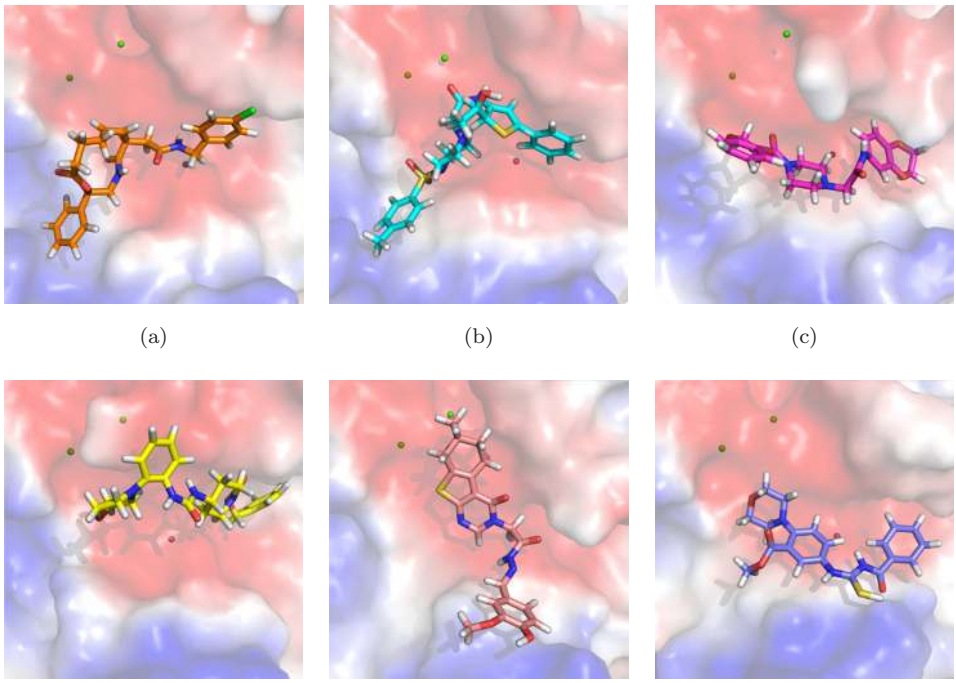
a structurally adjusted time interval (2 ns to 4 ns). Finally, the snapshot that is corresponding to RMSD minima was selected as a representative structure of each complex.

As a result of comparing the binding modes, all of the hit compounds were observed to surround the zinc ion, whereas GK 03795 showed a similar binding orientation compared to robotnikinin. A significant point of these comparisons is that GK 03795 is located within the coordinating distance to one of the calcium ions, which was not observed in the binding mode of robotnikinin (Fig. 9).

Thus we calculated the variation in distance between the calcium ion and closest oxygen atom of the amide carbonyl group of GK 03795 during the 4 ns MD simulation. Surprisingly, it was restrained within the average distance value of 0.23 nm throughout the simulation (Fig. 10).

This result has given an explanation that GK 03795 has high specificity for the Shh when considering biological meaning of the calcium ions that have the essential role in activating the Shh signaling pathway.[17,37]

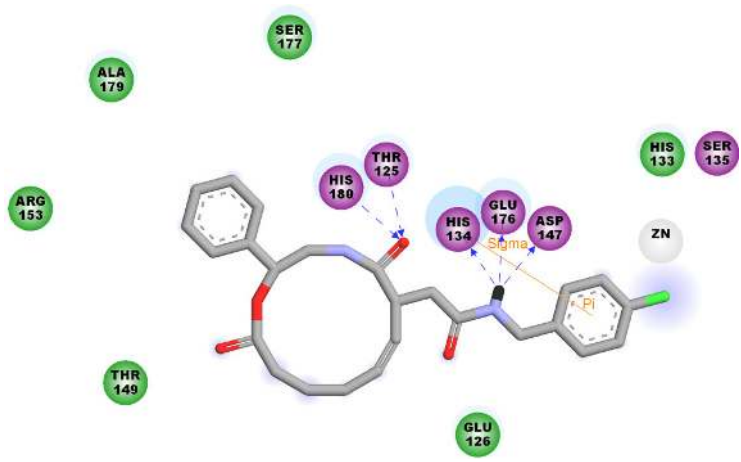
From the results of *in silico* simulations, we can infer that GK 03795 may have a better inhibitory effect than robotnikinin by blocking not only the zinc ion but also



(a) (b) (c)

(d) (e) (f)

Fig. 8. The binding modes of representative structures are shown in same orientations. The protein surface is colored based on electrostatic potential and bound compounds are shown in stick form. The two calcium ions and a zinc ion are represented as green and violet spheres. (a) Shh-robotnikinin, (b) Shh-GK 03795, (c) Shh-HTS 06264, (d) Shh-HTS 06766, (e) Shh-HTS 07674, (f) Shh-SEW 05450.



(a)

Fig. 9. 2D diagrams displaying the interactions between the residues of the pseudo-active site of Shh and (a) robotnikinin, (b) GK 03795. Blue dashed line indicates hydrogen bonding interaction between amino acid side-chain with electron donor.

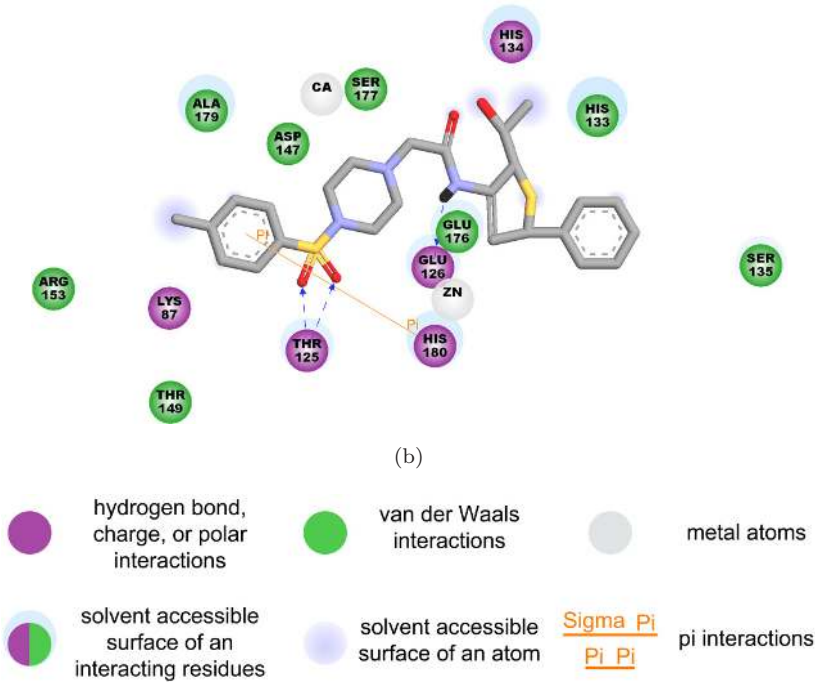


Fig. 9. (*Continued*)

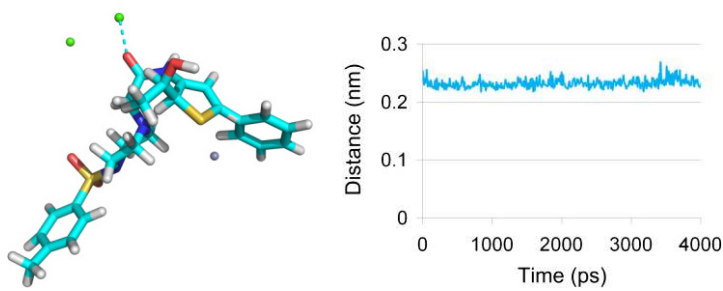


Fig. 10. Calculation of distance between one of the calcium ions and the functional group of inhibitor. (a) The distance between the calcium ion and the amide carbonyl group of GK 03795 is represented in cyan dashed line, and (b) the variation of distance between the calcium ion and the amide carbonyl group of GK 03795 as a function of time.

one of the calcium ions which is important in activating the Shh signaling pathway. This hit compound may keep the Ptch L2 peptides from binding to the Shh. Thus, GK 03795 could be a potential candidate inhibitor for disrupting the Shh signaling pathway by forming a spatial interruption in the pseudo-active site of the Shh.

4. Conclusion

The Shh signaling pathway has been studied steadily because it is involved in various cancers. Our computer-based studies have progressed from the 3D structural information of robotnikinin and its analogs to the discovery of new set of chemical scaffolds with Shh inhibitory profile. Robotnikinin containing 12-membered macrocycle moiety plays a role in interrupting the Shh-Ptch binding, which is the starting point in the activation of the Shh signaling pathway. Our first aim of this study was to develop a pharmacophore model that can elucidate common chemical features within various spatial conformations generated from 3D structures of robotnikinin and its analogs. The best pharmacophore model could discriminate robotnikinin and its analogs corresponding to biological activities. And we used this model to retrieve the hit compounds that map the identified chemical features. The chemical databases *Maybridge* and *Chembridge* containing diverse chemical compounds were used in a series of virtual screening and drug-like filtering processes. In order to test the fitness of robotnikinin and the hit compounds at the Shh structure, the *LigandFit* docking simulations and the 4 ns MD simulations were performed. In the results of docking simulations, we could predict the binding mode of robotnikinin in the pseudo-active site of the Shh, and the final five hit compounds (GK 03795, HTS 06264, HTS 06766, HTS 07674, SEW 05450) were selected based on the consensus scoring functions. In the analysis of the MD simulations of the Shh-robotnikinin and Shh-hit complexes, GK 03795 was selected as a potential Shh signaling pathway inhibitor based on the degree of energy, number of hydrogen bonds, and similar binding mode to that of robotnikinin complex. In particular, GK 03795 was the only compound among the five hit compounds that could coordinate with one of the calcium ions constantly. Combining these results, we concluded that GK 03795 can disrupt the interaction between the Shh and Ptch by blocking the metal ions that are present in the Shh and inducing spatial hindrance to the binding of Ptch L2 peptides in the pseudo-active site of the Shh. The chemical structure of GK 03795 is novel in the history of the Shh signaling pathway inhibitors and thereby can be utilized in designing future and potent inhibitors. This hit compound may receive more attention as this is a small compound when compared to the known inhibitors which are macrocycles and thereby provides space for further optimization.

Acknowledgments

This research was supported by 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. Shahi MH, Sinha S, Afzal M, Castresana JS, Role of sonic hedgehog signaling pathway in neuroblastoma development, *Biol and Med* **1**:1–6, 2009.
2. Scales SJ, de Sauvage FJ, Mechanisms of hedgehog pathway activation in cancer and implications for therapy, *Trends Pharmacol Sci* **30**:303–312, 2009.
3. Hardcastle Z, Mo R, Hui C, Sharpe PT, The Shh signaling pathway in tooth development: Defects in *Gli2* and *Gli3* mutants, *Development* **125**:2803–2811, 1998.
4. Lavine KJ, Kovacs A, Ornitz DM, Hedgehog signaling is critical for maintenance of the adult coronary vasculature in mice, *J Clin Invest* **118**:2404–2414, 2008.
5. Traiffort E, Angot E, Ruat M, Sonic hedgehog signaling in the mammalian brain, *J Neurochem* **113**:576–590, 2010.
6. Romer J, Curran T, Targeting medulloblastoma: Small–molecule inhibitors of the sonic hedgehog pathway as potential cancer therapeutics, *Canc Res* **65**:4975–4978, 2005.
7. Sanchez P, Hernández AM, Stecca B, Kahler AJ, DeGueme AM, Barrett A, Beyna M, Datta MW, Datta S, i Altaba AR, Inhibition of prostate cancer proliferation by interference with Sonic Hedgehog–Gli1 signaling, *Proc Natl Acad Sci USA* **101**:12561–12566, 2004.
8. Bale AE, Yu K, The hedgehog pathway and basal cell carcinomas, *Hum Mol Genet* **10**:757–762, 2001.
9. Zhao C, Chen A, Jamieson CH, Fereshteh M, Abrahamsson A, Blum J, Kwon HY, Kim J, Chute JP, Rizzieri D, Munchhof M, VanArsdale T, Beachy PA, Reya T, Hedgehog signaling is essential for maintenance of cancer stem cells in myeloid leukaemia, *Nature* **458**:776–779, 2009.
10. Jeong J, McMahon AP, Cholesterol modification of hedgehog family proteins, *J Clin Invest* **110**:591–596, 2002.
11. Rubin LL, de Sauvage FJ, Targeting the hedgehog pathway in cancer, *Nat Rev Drug Discov* **5**:1026–1033, 2006.
12. Dessaud E, McMahon AP, Briscoe J, Pattern formation in the vertebrate neural tube: A sonic hedgehog morphogen-regulated transcriptional network, *Development* **135**:2489–2503, 2008.
13. Bumcrot DA, Takada R, Mcmahon AP, Proteolytic processing yields two secreted forms of sonic hedgehog, *Mol Cell Biol* **15**:2294–2303, 1995.
14. Porter JA, Young KE, Beachy PA, Cholesterol modification of hedgehog signaling proteins in animal development, *Science* **274**:255–259, 1996.
15. Kavran JM, Ward MD, Oladosu OO, Mulepati S, Leahy DJ, All mammalian hedgehog proteins interact with cell adhesion molecule, down-regulated by oncogenes (CDO) and brother of CDO (BOC) in a conserved manner, *J Biol Chem* **285**:24584–24590, 2010.
16. Bishop B, Aricescu AR, Harlos K, O'Callaghan CA, Jones EY, Siebold C, Structural insights into hedgehog ligand sequestration by the human hedgehog-interacting protein HHIP, *Nat Struct Mol Biol* **16**:698–703, 2009.
17. Maun HR, Wen X, Lingel A, de Sauvage FJ, Lazarus RA, Scales SJ, Hymowitz SG, Hedgehog pathway antagonist 5E1 binds hedgehog at the pseudo-active site, *J Biol Chem* **285**:26570–26580, 2010.
18. Mahindroo N, Punchihewa C, Fujii N, Hedgehog-Gli signaling pathway inhibitors as anticancer agents, *J Med Chem* **52**:3829–3845, 2009.
19. Fuse N, Maiti T, Wang B, Porter JA, Hall TM, Leahy DJ, Beachy PA, Sonic hedgehog protein signals not as a hydrolytic enzyme but as an apparent ligand for Patched, *Proc Natl Acad Sci USA* **96**:10992–10999, 1999.

20. Peng LF, Stanton BZ, Maloof N, Wang X, Schreiber SL, Syntheses of aminoalcohol-derived macrocycles leading to a small-molecule binder to and inhibitor of Sonic hedgehog, *Bioorg Med Chem Lett* **19**:6319–6325, 2009.
21. Advanced Chemistry Development Inc., ACD/CNMR Predictor, V.3.0. 133, Richmond St, Suite 605, Toronto, Ontario M5H 2L3, Canada.
22. Smellie A, Teig SL, Towbin P, Poling: Promoting conformational variation, *J Comput Chem* **16**:171–187, 1995.
23. Yang S, Pharmacophore modeling and applications in drug discovery: Challenges and recent advances, *Drug Discov Today* **15**:444–450, 2010.
24. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv Drug Deliv Rev* **23**:3–25, 1997.
25. Ekins S, Waller CL, Swaan PW, Cruciani G, Wrighton SA, Wikel JH, Progress in predicting human ADME parameters *in silico*, *J Pharm Tox Methods* **44**:251–272, 2000.
26. Susnow RG, Dixon SL, Use of robust classification techniques for the prediction of human cytochrome P450 2D6 inhibition, *J Chem Inf Comput Sci* **43**:1308–1315, 2003.
27. Cheng A, Dixon SL, *In silico* models for the prediction of dose-dependent human hepatotoxicity, *J Comput Aided Mol Des* **17**:811–823, 2003.
28. Berendsen HJC, van der Spoel D, van Drunen R, GROMACS 4: A message-passing parallel molecular dynamics implementation, *Comput Phys Commun* **91**:43–56, 1995.
29. Hess B, Kutzner C, van der Spoel D, Lindahl E, GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation, *J Chem Theory Comput* **4**:435–447, 2008.
30. van Aalten DMF, Bywater R, Findlay JBC, Hendlich M, Hooft RWW, Vriend G, PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules, *J Comput Aided Mol Des* **10**:255–262, 1996.
31. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM, LINCS: A linear constraint solver for molecular simulations, *J Comput Chem* **18**:1463–1472, 1997.
32. Tom D, Darrin Y, Lee P, Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems, *J Chem Phys* **98**:10089–10092, 1993.
33. Bosanac I, Maun HR, Scales SJ, Wen X, Lingel A, Bazan JF, de Sauvage FJ, Hymowitz SG, Lazarus RA, The structure of SHH in complex with HHIP reveals a recognition role for the Shh pseudo active site in signaling, *Nat Struct Mol Biol* **16**:691–698, 2009.
34. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, Clustal W and Clustal X version 2.0, *Bioinformatics* **23**:2947–2948, 2007.
35. Briscoe J, Chen Y, Jessell TM, Struhl G, A hedgehog-insensitive form of patched provides evidence for direct long-range morphogen activity of sonic hedgehog in the neural tube, *Mol Cell* **7**:1279–1291, 2001.
36. Taipale J, Cooper MK, Maiti T, Beachy PA, Patched acts catalytically to suppress the activity of Smoothened, *Nature* **418**:892–896, 2002.
37. McLellan JS, Zheng X, Hauk G, Ghirlando R, Beachy PA, Leahy DJ, The mode of hedgehog binding to Ihog homologs is not conserved across different phyla, *Nature* **455**:979–983, 2008.
38. Karplus M, McCammon JA, Molecular dynamics simulations of biomolecules, *Nat Struct Biol* **9**:646–788, 2002.

Swan Hwang is a M.Sc. candidate at Division of Applied Life Science from Gyeongsang National University (GNU), Jinju, Korea. He holds a B.Sc. degree in Microbiology from GNU. His research interests include molecular modeling, structural bioinformatics, and systems biology. He has worked on computer-aided drug design for developing cancer drug candidates and redesigning new functional proteins.

Sundarapandian Thangapandian is a Ph.D. candidate in the Computational Biology and Bioinformatics Lab (CBBL), Gyeongsang National University, Jinju, Republic of Korea, since August 2008. He completed his Masters in Technology specializing in Pharmaceutical Chemistry. His interests lie in the area of computer-aided drug discovery and computer simulations addressing various biophysical problems related to various therapeutic areas.

Yuno Lee received his B.S. from the Department of Biochemistry from Gyeongsang National University (GNU), Jinju, Korea, and M.S. from the Division of Applied Life Science from Gyeongsang National University (GNU), Jinju, Korea, in 2007 and 2009, respectively. Since March 2009, he is in course of his Ph.D. degree at Division of Applied Life Science, Gyeongsang National University (GNU), Jinju, Korea.

Sugunadevi Sakkiah received her M.Sc. degree in Bioinformatics from Bharathiar University, India. From 2004 to 2005, she worked as a Research Associate in Genome Bioscience Research Institute, India. She then joined as a Project Assistant in Madurai Kamaraj University, India, from 2005–2006 and then got the opportunity to explore the field of computer-aided drug design in "Orchid Chemicals and Pharmaceutical" as a Research Executive from 2006–2008. Presently, she is a Ph.D. candidate in Computer Aided Drug Design, Molecular Modeling, and Molecular Dynamics, under the guidance of Prof. Keun Woo Lee, Gyeongsang National University, South Korea, and is expecting to complete her doctoral studies in November 2011.

Shalini John hails from South India and is working with Prof. Keun Woo Lee since February 2009 for her doctoral degree. She received her M.Sc. degree from Vellore Institute of Technology, India, with a specialization in organic chemistry. Her technical interests include computer-aided designing of organic molecules of therapeutic interest and engineering the biological molecules of therapeutic interests. In addition, she is also interested in synthesizing computationally designed molecules for biological evaluation.

Keun Woo Lee received his B.S., M.S., and Ph.D. at the Department of Chemistry from Seoul National University (SNU), Seoul, Korea, in 1988, 1990, and 1997, respectively. From 1997 to 2004, he pursued his postdoctoral career in USA at the Department of Chemistry, University of California, Berkeley (1997–1998) and Department of Biology and Biochemistry, University of Houston (UH) (from 1998–2002 as a Research Associate and from 2002–2004 as a Research Professor). Since 2004, he is an Assistant/Associate Professor at the Department of Biochemistry, Gyeongsang National University (GNU), Jinju, Korea. At present, he is an active member of the Korean Chemical Society (KCS) and Korean Society of Bioinformatics and Systems Biology (KSBSB).