J Mol Model (2012) 18:3267–3282
DOI 10.1007/s00894-011-1247-5

ORIGINAL PAPER

Pharmacophore modeling, molecular docking, and molecular dynamics simulation approaches for identifying new lead compounds for inhibiting aldose reductase 2

Sugunadevi Sakkiah · Sundarapandian Thangapandian · Keun Woo Lee

Received: 6 June 2011 / Accepted: 19 September 2011 / Published online: 18 January 2012

Abstract Aldose reductase 2 (ALR2), which catalyzes the reduction of glucose to sorbitol using NADP as a cofactor, has been implicated in the etiology of secondary complications of diabetes. A pharmacophore model, Hypo1, was built based on 26 compounds with known ALR2-inhibiting activity values. Hypo1 contains important chemical features required for an ALR2 inhibitor, and demonstrates good predictive ability by having a high correlation coefficient (0.95) as well as the highest cost difference (128.44) and the lowest RMS deviation (1.02) among the ten pharmacophore models examined. Hypo1 was further validated by Fisher's randomization method (95%), test set (r=0.91), and the decoy set shows the goodness of fit (0.70). Furthermore, during virtual screening, Hypo1 was used as a 3D query to screen the NCI database, and the hit leads were sorted by applying Lipinski's rule of five and ADME properties. The best-fitting leads were subjected to docking to identify a suitable orientation at the ALR2 active site. The molecule that showed the strongest interactions with the critical amino acids was used in molecular dynamics simulations to calculate its binding affinity to the candidate molecules. Thus, Hypo1 describes the key structure–activity relationship along with the estimated activities of ALR2 inhibitors. The hit molecules were searched against PubChem to find similar molecules with new scaffolds. Finally, four molecules were found to satisfy all of the chemical features and the geometric constraints of Hypo1, as well as to show good dock scores, PLPs and PMFs. Thus, we believe that Hypo1 facilitates the selection of novel scaffolds for ALR2, allowing new classes of ALR2 inhibitors to be designed.

Keywords Aldose reductase · Molecular dynamics simulation · Pharmacophore · Molecular docking · Virtual screening

Abbreviations

ADME	Absorption, distribution, metabolism, and excretion
ALR2	Aldose reductase 2
BBB	Blood–brain barrier
DS	Discovery Studio v.2.5
EF	Enrichment factor
GF	Goodness of fit
HBA	Hydrogen bond acceptor
HAli	Hydrophobic aliphatic
HAro	Hydrophobic aromatic
MD	Molecular dynamics
NI	Negative ionization
NADPH	Nicotinamide dinucleotide
53N	3-[5-(3-Nitrophenyl) thiophen-2-yl] propanoic acid
PME	Particle mesh Ewald
PLP	Piecewise linear potential
PMF	Potential of mean force
RA	Ring aromatic
RMS	Root mean square
RMSD	Root mean square deviation

S. Sakkiah · S. Thangapandian · K. W. Lee (✉)
Division of Applied Life Science (BK21 Program),
Systems and Synthetic Agrobiotech Center (SSAC),
Plant Molecular Biology and Biotechnology Research Center (PMBBRC)f, Research Institute of Natural Science (RINS),
Gyeongsang National University (GNU),
501 Jinju-daero, Gazha-dong,
Jinju 660-701, Republic of Korea
e-mail: kwlee@gnu.ac.kr

RMSF	Root mean square fluctuation
TIM	Triose phosphate isomerase
VMD	Visual molecular dynamics

Introduction

Diabetes is a chronic disease that occurs when the pancreas is not able to produce enough insulin or when the body cannot effectively utilize the insulin present [1]. There are two types of diabetes: type 1 diabetes (known as insulin-dependent or childhood-onset diabetes) is characterized by a lack of insulin secretion, and type 2 diabetes (known as non-insulin-dependent or adult-onset diabetes) is caused by the body's ineffective use of insulin. Diabetes often results from excess body weight and physical inactivity.

The polyol pathway has been implicated in the etiology of secondary complications of diabetes. Aldose reductase 2 (alditol/NADP + oxidoreductase, EC 1.1.1.21, ALR2) is the first enzyme in the polyol pathway; it reduces excess D-glucose to D-sorbitol, with the concomitant conversion of NADPH to $NADP^+$ [2–5]. Inhibiting ALR2 blocks the glucose flux and prevents or reverses functional deficits and structural abnormalities in the lens, retina, kidney, and peripheral nerves [6]. ALR2 is a 36 kD triosephosphate isomerase (TIM)-barrel-shaped monomeric protein of 315 amino acids, and mapped to chromosome region 7q35. Human aldose reductase, a cytosolic protein belonging to the aldo-keto reductase superfamily reductase [7], is a nicotinamide dinucleotide (NADPH)-dependent oxidoreductase that catalyzes the reduction of a broad range of aldehydes including glucose [8]. It also catalyzes the reduction of aldehydyes, xenobiotic aldehydes, ketones, and trioses using NADPH as a reducing cofactor [9]. ALR2 converts glucose into fructose using sorbitol dehydrogenase, and the accumulation of sorbitol in cells leads to diabetic complications. Under increased glucose flux, sorbitol can accumulate in cells with insulin-independent glucose uptake, leading to augmented osmotic pressure in the cells [10]. A variety of diabetic complications have been attributed to these biochemical phenomena caused by glucose processing through the polyol pathway. Thus, ALR2 has received considerable attention as a target for therapeutic intervention. Inhibiting ALR2 prevents the entry of glucose into the sorbitol pathway which reduces the damaging effects of late-onset diabetic disorders [11]. Therefore, inhibiting ALR2 is a potential therapeutic approach to curing the debilitating pathologies associated with chronic hyperglycemia.

The details of the binding pocket of ALR2 were identified by analyzing the various crystal structures of ALR2 [12–14]. The active site of ALR2 is located at the C-terminal face of the TIM barrel, which is well suited to participating in efficient interactions with $NADP^+$, a cofactor required for ALR2's reduction reactions. The binding pocket involves Ala299, Leu300, and Phe122 at the solvent-exposed face, and the side chain of Trp111, which orientated towards the center of the TIM barrel [15–18]. Extensive research over the past few decades has identified the involvement of ALR2 in the pathophysiology of diabetic complications, as it plays a pivotal role in glucose metabolism [19]. Inhibiting ALR2 may represent a direct treatment for diabetic complications, independent of controlling blood sugar levels, and research in this area has led to the development of a number of structurally diverse ALR2 inhibitors. The plethora of well-resolved crystal structures of ALR2 also mark it out as a promising target for rational drug design. The aim of ALR2 inhibitor therapy is to normalize the elevated fluxes of blood and sorbitol through the polyol pathway in the target tissue. However, although it undoubtedly shows great therapeutic potential for the treatment of diabetes, ALR2 is a particularly challenging molecule to use for drug design due to its remarkably broad substrate promiscuity.

Pharmacophore modeling is one of the strategies used to identify new ligands and develop novel inhibitors for ALR2. In the work reported here, we generated a chemical feature based pharmacophore hypothesis using the HypoGen algorithm implemented in Discovery Studio v.2.5 (DS) [20]. This approach provided structure–activity relationship data on a set of compounds, and the most potent inhibitors in various databases were retrieved by implementing the best hypothesis in a virtual screening process. The screened molecules were sorted based on Lipinski's rule of five as well as their absorption, distribution, metabolism, and excretion (ADME) properties, and the selected hit molecules were subjected to molecular docking studies. The docked molecules that showed good orientations at the active site of ALR2 based on the known interactions between ligands and the ALR2 binding pocket were selected for molecular dynamic (MD) simulations. The best molecule obtained from the MD simulations was used as a query to perform a similarity search against the PubChem database.

Methods and materials

Pharmacophore modeling is one of the most frequently applied and valuable methods of discovering novel scaffolds for various targets, so this approach was used in this work to find novel inhibitors of ALR2. Two different types of pharmacophore procedures can be used to discover the most potent leads: (i) ligand-based and (ii) structure-based. Here we used ligand-based pharmacophore generation, which depends completely on the reported activity value (IC_{50}) of the known antagonist of ALR2.

Preparation of the training and test sets

The 51 known inhibitors of ALR2 were identified from various reports [21–23] and divided into two different sets: (i) a training set to generate the hypothesis, and (ii) a test set to validate the generated hypothesis. Among the 51 compounds, 26 were picked as training compounds (these were selected based on certain rules, such as: the compounds should span over four orders of magnitude, and they should be structurally diverse), while the remaining 25 compounds were selected for the test set. The training set molecules were divided into three categories based on their activity values: highly active (IC_{50}<300 nM, +++), moderately active (300 nM ≤ IC_{50}<3000 nM), and low activity (IC_{50}≥3000 nM). The 2D forms of the training and test set molecules were sketched using ChemSketch [24] and converted into their corresponding 3D forms using DS. DS provides two types of conformational analysis: FAST and BEST quality analysis. FAST conformational analysis is the method of choice for generating a database of compounds, because the tolerances in a database query can be adjusted to minimize the effect of incomplete conformational coverage. Here, the BEST conformation analysis was applied for conformation generation, and the compounds were energy minimized to the closest local minimum by applying the CHARMm force field [25–27] for hypothesis generation. A maximum number of 255 diverse conformations were modeled for each compound using a Monte Carlo-like algorithm with an energy range of 20 kcal mol^{-1} together with a poling [28] algorithm.

Quantitative pharmacophore generation

Before performing the quantitative pharmacophore modeling, active compounds of the training set molecules were submitted to the Feature Mapping protocol to identify the chemical features that are crucial to potent inhibition of ALR2. The resulting chemical features were used to generate the quantitative hypothesis using the 3D QSAR Pharmacophore Generation protocol in DS, which correlates the observed biological activities for a series of compounds with their chemical structures. Based on the activity values presented by the training set compounds, the top ten hypotheses were identified and evaluated using the Debnath method. This method implies that the best pharmacophore model should have a high correlation coefficient, the lowest total cost, and the lowest root mean square (RMS) deviation, and the total cost should be close to the fixed cost and far from the null cost. The reliability of a pharmacophore model depends on the difference between the total cost of the generated hypothesis and the null hypothesis.

Methods used to validate the selected hypothesis

Once the best hypothesis had been selected from among the top ten hypotheses, it was validated by applying various potent methods. In this work, three potent methods were used to validate the selected hypothesis: (i) Fisher's randomization, (ii) the application of a test set, and (iii) the use of a decoy set. Fisher's randomization was performed at the same time the hypothesis was generated; it produces a number of random spreadsheets that depend on the selected significance level (90%, 95%, 98%, 99%) by shuffling the activity values present in the training set. Test and decoy sets were used to check whether the best hypothesis had the ability to determine the order of activity of molecules other than the training set compounds, and also to find out how well it differentiated the ALR2 inhibitors from other compounds, respectively. The test set contained a wide range of activity values, and the molecules in it were classified as either highly active, moderately active, or inactive, based on their activity values, just as in the training set. The decoy set consisted of 2,200 molecules, including ten known ALR2 inhibitors. The Ligand Pharmacophore Mapping protocol was used to calculate parameters such as the total number of compounds in the hit list (H_t), the percentage yield of actives (Y), the proportion of active molecules in the hit list (A), the enrichment factor (EF), and the goodness of fit (GF). The following equations were used to calculate the EF and GF [29]:

$$\mathrm{EF} = [(H_a \times D)/(H_t \times A)] \quad (1)$$

$$\mathrm{GF} = [(H_a/4H_tA)(3A + H_t) \times (1-((H_t - H_a)/(D - A))]. \quad (2)$$

Here, H_a is the total number of active molecules in the hit list, D is the total number of molecules in the decoy set, and A is the total number of actives in the decoy set. The values of EF and GF indicate the specificity and selectivity of the best pharmacophore model.

Virtual screening

Virtual screening technology is used to discover or identify novel potent compounds that can repress or trigger the activity of a particular target. In this work, the best hypothesis was selected based on the above validations and used as a 3D structural query in virtual screening to retrieve a novel scaffold for ALR2 inhibition from the NCI database (~200,000 compounds). The screened compounds were filtered by applying some conditions; for example, maximum fit values of greater than 10 were selected based

on the highest fit values of active compounds in the training set, the ADME properties of the compounds were examined, and Lipinski's rule of five was applied [30]. Mainly blood–brain barrier (BBB), solubility, and absorption criteria were the focus of the ADME investigation; if the molecule had values of 3, 3, and 0 for BBB, solubility, and absorption, respectively, it was considered that the molecule had good solubility, absorption, and BBB [31]. Lipinski's rule of five is a simple model that is used to forecast the absorption and intestinal permeability of a compound. According to the rule of five, compounds are well absorbed when they possess a log*P* of less than 5, a molecular weight of less than 500, fewer than five hydrogen-bond donors, fewer than ten hydrogen-bond acceptors, and fewer than ten rotatable bonds. Compounds that satisfied all of the above filters were selected for the subsequent molecular docking studies, which were performed to obtain the optimal orientations of the leads at the active site.

Molecular docking protocol

Molecular docking is one of the best filtering methods used in the drug design process. Hence, docking was carried out to find the most suitable orientation and interactions (hydrogen bonds and hydrophobic interactions) of each lead at the protein's active site. The LigandFit protocol in DS [26], which is one of the best docking techniques available, was used to dock the screened compounds. The protein complex (receptor) was selected from the Protein Data Bank [32, 33] (PDB, http://www.rcsb.org) to perform the molecular docking studies. Many ALR2 complexes have been reported; among these, PDB ID: 3DN5 [11] was selected as the receptor based on the resolution of the complex as well as the deposited date. The receptor was first prepared by removing the water molecules, and the CHARMm force field [25–27] was applied using the Receptor-Ligand Interaction protocol in DS. After preparing the receptor, the binding site was identified based on the volume occupied by the ligand in the ALR2 complex. To validate our docking parameters, the co-crystal from PDB was initially sketched and docked into the active site of ALR2. The docked pose was checked to see whether it was able to produce the hydrogen-bonding interactions with the critical amino acids. The RMSD between the docked pose and the co-crystal was calculated to determine if the docking parameters were able to reproduce a conformation comparable to that of the co-crystal at the active site of ALR2. Then the leads were docked using the same parameters as in the co-crystal docking. During the docking process, the top ten conformations for each ligand based on the dock score after energy minimization using the smart minimizer method (which begins with the steepest descent method and is followed by the conjugate gradient method) were assessed. The docked poses were validated based on the hydrogen-bonding interactions between the candidate molecules and the active site residues.

Molecular dynamics simulation

Based on the docking results, MD simulation was performed for the complexes of ALR2 obtained (with the ligands NCI0036494 and 53N) using the GROMACS 4.0.5 [34, 35] computational package. The Gromos96 force field [34] was applied to the two systems, placed in the center of the cubic box, and solvated by the water molecules. Topology files and other force field parameters for ligands were generated using the PRODRG program [36, 37]. Eight water molecules were replaced with Cl^- counterions to keep the system electrically neutral. A steepest-descent algorithm was used to minimize the energy and thus relax the water molecules in each system. All covalent bonds containing the hydrogen atoms were constrained using the SHAKE algorithm [38] with a tolerance of 10^{-7} Å. The Berendsen temperature and pressure coupling methods were applied to keep the system in a stable environment (300 K, 1 Bar); both of the coupling constants were set to 0.1. The particle mesh Ewald (PME) [39] method for long-range electrostatics, a 9 Å cutoff for Coulomb interactions, and a 10 Å cutoff for van der Waals interactions were selected. The LINCS [40] algorithm for bond constraints was used. Five-nanosecond MD simulations were performed at 300 K with an integration step of 1 fs, and periodic boundary conditions were used in all three dimensions. The three-dimensional structures and trajectories were visually inspected using PyMol and visual molecular dynamics (VMD), respectively. The frame that showed the smallest RMSD from the average structure obtained during the last 2 ns of MD simulation was selected as a representative structure. The representative structure was refined by performing 1000 steps of steepest decent followed by conjugated gradient energy minimization, and this refined structure was used for further analysis.

Similarity search

The best molecule from the MD simulation was used as a reference to find similar molecules in the PubChem database. The compounds that showed similarities of greater than 90% to this best molecule were selected for further refinement, such as the application of the rule of five and hypothesis screening using best conformation generation. These well-screened molecules were then subjected to molecular docking studies to find the optimal orientation and binding affinity of each molecule at the active site of ALR2. The LigandFit docking program in DS was used to dock the screened molecules using the same parameters employed for the hit compounds from the virtual screening process.

Results and discussion

HypoGen model for ALR2 inhibitors

The hypothesis was generated based on 26 compounds present in the training set (Fig. 1). These compounds were collected from the various literatures having different scaffolds. Hydrogen-bond acceptor (HBA), hydrophobic aliphatic (HAli), hydrophobic aromatic (HAro), negative ionization (NI), and ring aromatic (RA) groups were selected as required chemical features for ALR2 inhibitors using the Feature Mapping protocol, and were used as input to the 3D QSAR pharmacophore generation module. The top ten hypotheses generated and their statistical parameters were obtained based on the activity values of the training set compounds (Table 1). Among the five stated above, three chemical features (HBA, HAli, and HAro) were present in all of the hypotheses, indicating that these chemical groups are necessary for ALR2 inhibition. The ten hypotheses thus generated were categorized into three classes based on their chemical features. Three hypotheses (Hypo1, Hypo6, and Hypo7) were in the first class, which exhibited two HBA, two HAro, and one HAli feature; five hypotheses (Hypo2, Hypo3, Hypo4, Hypo8, and Hypo9) were in the second class, which presented one HBA, one HAli, two HAro, and one NI feature; the third class contained two hypotheses (Hypo5, Hypo10), which had one HBA, one HAli, one HAro, one RA, and one NI pharmacophoric feature. The hypothesis with the highest fit value was selected each of the three classes (Hypo1 in class I, Hypo2 in class II, and Hypo5 in class III) in order to elucidate the chemical groups that are crucial to ALR2 inhibition. Hypo5 had a maximum fit value of 7.61, but when the RA group was replaced by a HAro group, its maximum fit value increased to 9.75, as seen for Hypo2. Hypo1 had a maximum fit value (10.92) that was greater than that of Hypo2, and Hypo1 differed from Hypo2 in that it has another HBA feature rather than an NI feature. Hypo6 and Hypo7, which had the same chemical features, yielded maximum fit values of 9.4 and 10.15, respectively, but Hypo1, which also has the same features as Hypo6 and Hypo7, produced a maximum fit value of 10.92. The only differences between these three hypotheses were their geometric constraints. Based on this analysis, we inferred that the HAro and HBA chemical features are crucial to the inhibition of ALR2. Hypo1 showed the greatest ALR2-

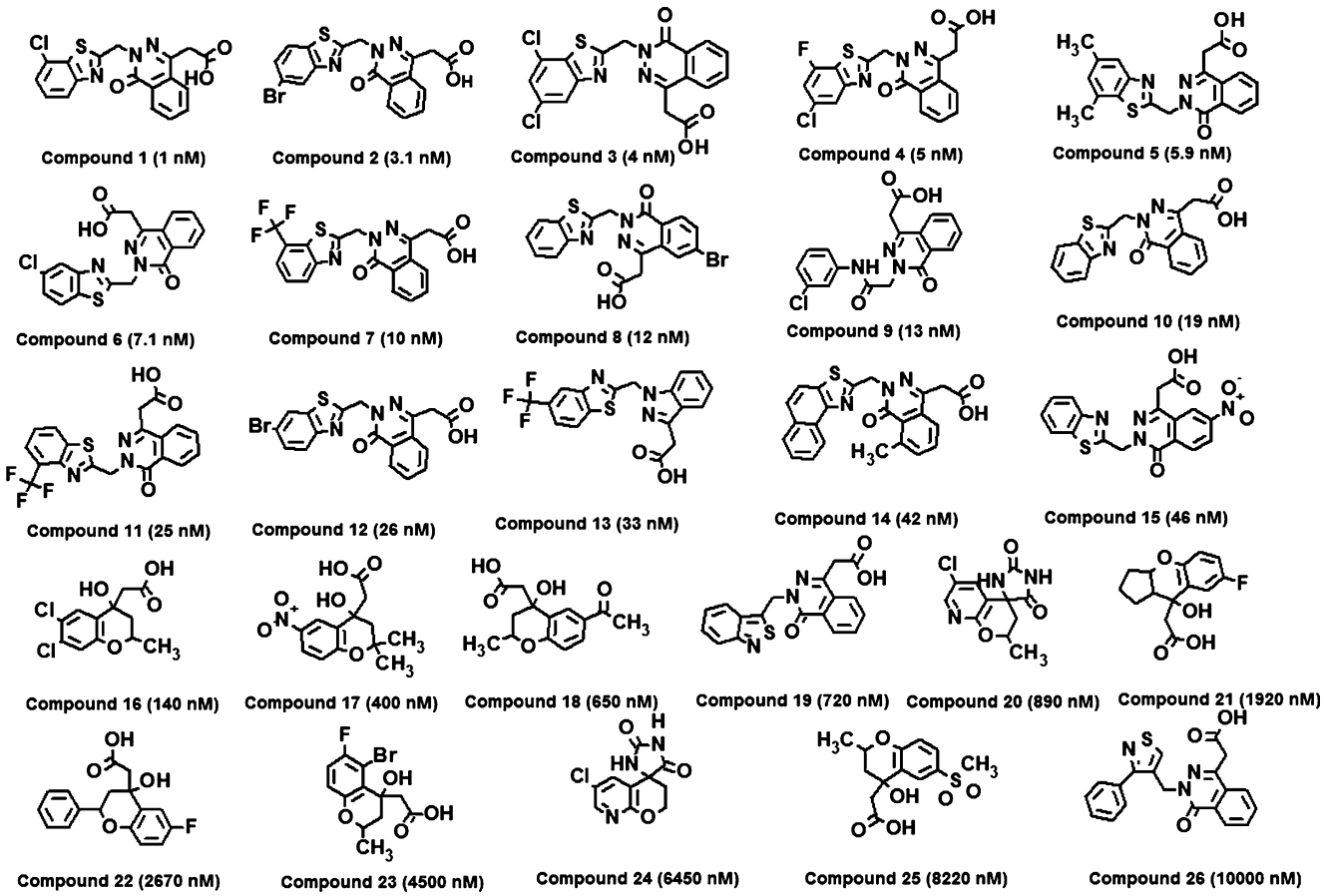

Fig. 1 2D chemical structures of the 26 training set molecules used for hypothesis generation; their experimental activity values (IC_{50}) are given in parentheses

Table 1 Statistical data and predictive powers (presented as the cost, measured in bits) of the top ten hypotheses resulting from automated 3D-QSAR pharmacophore generation

Hypothesis	Total cost	Difference in cost[a]	RMS[b]	Correlation	Features[c]	Max. fit value
Hypo1	108.56	128.44	1.02	0.96	HBA, HBA, HAli, HAro, HAro	10.92
Hypo2	113.84	123.16	1.20	0.93	HBA, HAli, HAro, HAro , NI	9.75
Hypo3	114.36	122.64	1.22	0.93	HBA , HAli, HAro, HAro, NI	9.75
Hypo4	118.03	118.97	1.33	0.92	HBA, HAli, HAro, HAro, NI	9.46
Hypo5	119.50	117.50	1.34	0.92	HBA, HAli, HAro, NI, RA	7.61
Hypo6	120.77	116.23	1.41	0.91	HBA, HBA, HAli, HAro, HAro	9.44
Hypo7	121.00	116.00	1.41	0.91	HBA, HBA, HAli, HAro, HAro	10.15
Hypo8	121.39	115.61	1.41	0.91	HBA, HAli, HAro, HAro, NI	8.29
Hypo9	121.56	115.44	1.42	0.91	HBA, HAli, HAro, HAro, NI	8.70
Hypo10	121.72	115.28	1.30	0.91	HBA, HAli, HAro, NI, RA	7.31

[a] Cost difference between the null and the total cost. The null cost, the fixed cost, and the configuration cost were 237.009, 94.8158, and 14.862, respectively

[b] *RMS* root mean square deviation: the deviation of the log(estimated activity) from the log(measured activity) normalized to log(uncertainty)

[c] *HBA* hydrogen-bond acceptor; *HAli* hydrophobic aliphatic; *HAro* hydrophobic aromatic, *NI* negative ionization, *RA* ring aromatic

inhibitory activity, and this contained five features: HAli, two HBA, and two HAro groups. Its 3D geometric constraints are shown in Fig. 2.

Debnath analysis states that the best pharmacophore model should have the lowest total cost value, the highest cost difference, the smallest RMS, and the best correlation coefficient. The predictive power of Hypo1 was confirmed using Debnath analysis. A hypothesis is considered to be valid when the overall cost of the hypothesis is far from the null cost and close to the fixed cost. The difference in cost is the difference between the null cost and the total cost of the hypothesis; a cost difference of 40–60 bits of cost difference leads to a predictive correlation probability of 75–90%, and if the difference is greater than 60 bits, the hypothesis is considered to have a correlation probability of of greater than 90% [31]. The cost difference bfor Hypo1 was found to be 128.44—more than 60 bits, indicating that this hypothesis has a >90% chance of being able to select an ALR2 antagonist. We also determined the correlation coefficients of the hypotheses from linear regressions derived from the geometric fit index values. All of the hypotheses had correlation coefficients of >0.90, but Hypo1 presented the highest correlation coefficient (0.96), which demonstrates the good predictive ability of Hypo1. The RMS factor (1.02 for Hypo1) represents the deviation of the estimated activity value from the experimental activity value, normalized to the uncertainty, and indicates the predictive quality for training set compounds. The fixed and total cost values for Hypo1 were 94.82 and 108.57, respectively. The highest cost difference and good correlation along with low RMS and minimum error values were observed for Hypo1 (Table 1) when compared with the other hypotheses. Hence, Hypo1 was selected as the best hypothesis and employed in further analyses.

To elucidate the predictive accuracy of Hypo1, the training set was classified into three subsets based on the activity values of the compounds in the set: highly active, IC_{50}<300 nM=+++; moderately active, 300 nM≥IC_{50}<3000 nM=++;

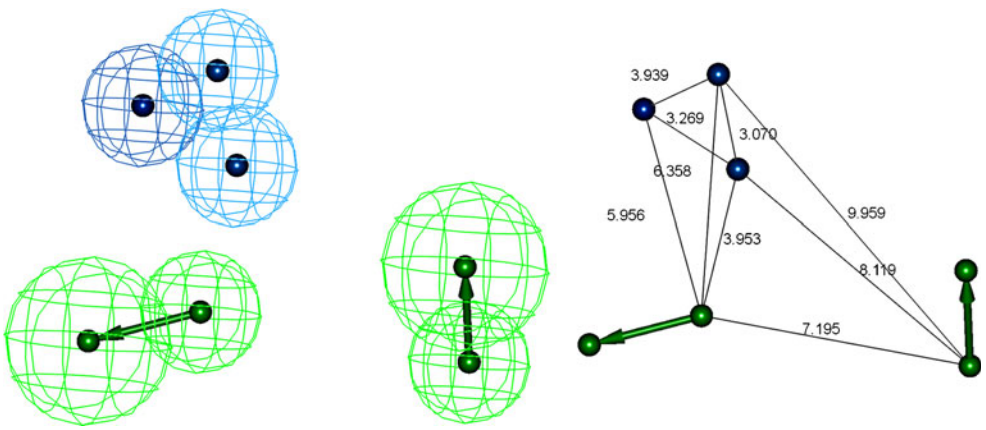

Fig. 2 Chemical features of the best pharmacophore (Hypo1), with its 3D spatial constraints. *Green* shows a hydrogen-bond acceptor (HBA), *blue* shows a hydrophobic aliphatic (HAli) feature, and *dark blue* represents a hydrophobic aromatic (HAro) feature

and low activity, $IC_{50} \leq 3000$ nM=+. The activity of each compound was estimated using regression analysis. One active and two moderately active compounds were underestimated as moderately active and low activity compounds, respectively. Three inactive compounds were overestimated as being moderately active, and the remaining compounds were classified correctly (Fig. 3). Based on the results of the above analysis, we can conclude that the activities of the compounds estimated by Hypo1 were close to the corresponding experimental IC_{50} values, and the error values defined as the ratio between the experimental and estimated activity values which demonstrated remarkable consistency between the estimated and experimental IC_{50} values. The experimental and predicted activities of Hypo1 for the training set compounds are shown in Table 2.

Validation of Hypo1

Fisher's randomization method, a test set, the goodness of fit, and the enrichment factor were used to validate and confirm the robustness of Hypo1.

Fisher's randomization method

Fisher's test was used to evaluate the statistical relevance of the best hypothesis by assigning a particular confidence level. Here we set a confidence level of 95%, so 19 random spreadsheets were created by shuffling the experimental activity values of the training set compounds, and a hypothesis was generated for each spreadsheet. The significance of the hypothesis was calculated using the following formula: $S = [1 - -(1 + X)/Y] \times 100$, where X is the total number of hypotheses with total costs that are lower than the original hypothesis, and Y is the total number of HypoGen runs (initial+random runs). Here, X=0 and Y=(1+19), hence $95\% = [1 - ((1 + 0)/(19 + 1)))] \times 100$. Two spreadsheets failed to produce the pharmacophore. The total costs of the other 17 pharmacophore models were compared with that of Hypo1, and it was found that the original hypothesis was far superior to the 17 random hypotheses, indicating a 95% confidence level for the Hypo1 model. Figure 4 clearly shows that the Hypo1 model was not generated by chance.

Test set

A good pharmacophore should have the ability to predict the activities of external compounds (i.e., compounds other than those in the training set). A test set consisting of 25 compounds was prepared using the same protocol as the training set, and the molecules were classified based on their activity values: highly active, $IC_{50} < 300$ nM=+++; moderately active, $300 \geq IC_{50} < 3000$ nM=++; low activity, $IC_{50} \leq 3000$ nM=+. Two active molecules were underestimated as being moderately active, and one moderately active molecule was overestimated as an active compound. The remaining compounds were classified correctly, indicating that Hypo1 was able to provide accurate estimates for the activities of compounds. The experimental and Hypo1-predicted activities of the compounds in the test set are shown in Table 3. The test set exhibits a correlation coefficient of 0.91 between the predicted and experimental values (Fig. 5). This result shows that Hypo1 can be used to gauge the activities of compounds beyond those in the training set, and it can be used in a predictive capacity.

Decoy set

Another two parameters that can be used to validate the hypothesis are the goodness of fit and the enrichment

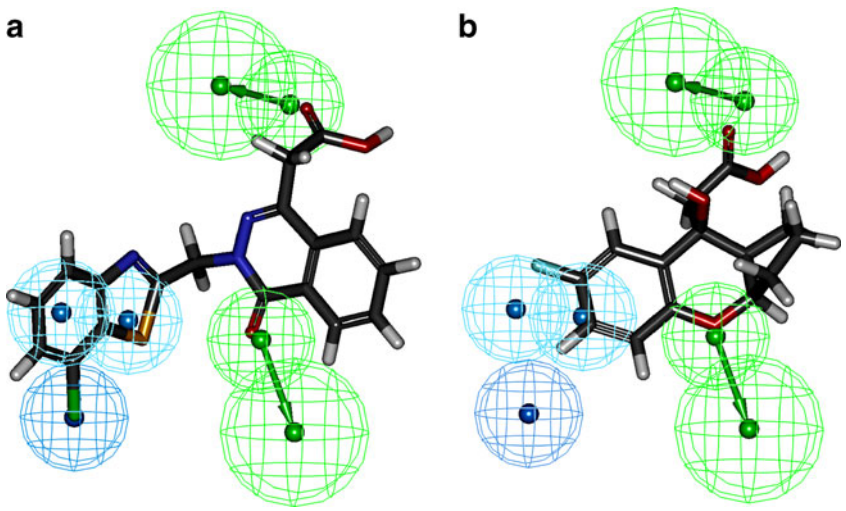

Fig. 3 **a**–**b** Mapping the best model pharmacophore for ALR2 antagonists (Hypo1) onto the training set compounds. **a** Hypo1 mapped onto the most active molecule (compound 1, IC_{50}: 1 nM). **b** Hypo1 mapped onto the least active molecule (compound 26, IC_{50}: 10000 nM). Features are color-coded: *green* hydrogen-bond acceptor (HBA), *light blue* hydrophobic aliphatic (HAli) feature, *dark blue* hydrophobic aromatic (HAro) feature

Table 2 Actual and estimated activities of the training set molecules, obtained using the pharmacophore model Hypo1

Compound no.	Fit value[a]	Exp. IC_{50} (nM)	Predicted IC_{50} (nM)	Error[b]	Experimental scale[c]	Predicted scale[c]
1	9.15	1	0.99	−1.0	+++	+++
2	8.37	3.1	6.00	1.9	+++	+++
3	8.61	4	3.40	−1.2	+++	+++
4	8.32	5	6.80	1.4	+++	+++
5	8.23	5.9	8.30	1.4	+++	+++
6	8.30	7.1	7.00	−1.0	+++	+++
7	8.00	10	14.00	1.4	+++	+++
8	7.84	12	20.00	1.7	+++	+++
9	7.66	13	31.00	2.4	+++	+++
10	7.93	19	17.00	−1.1	+++	+++
11	7.63	25	33.00	1.3	+++	+++
12	8.41	26	5.50	−4.7	+++	+++
13	7.76	33	24.00	−1.3	+++	+++
14	7.48	42	47.00	1.1	+++	+++
15	7.72	46	27.00	−1.7	+++	+++
16	6.47	140	480.00	3.4	+++	++
17	6.39	400	580.00	1.4	++	++
18	6.14	650	1000.00	1.6	++	++
19	6.22	720	840.00	1.2	++	++
20	6.08	890	1200.00	1.3	++	++
21	5.32	1900	6800.00	3.5	++	+
22	5.61	2700	3500.00	1.3	++	+
23	5.36	4500	6100.00	1.4	+	+
24	6.00	6400	1400.00	−4.5	+	++
25	5.88	8200	1900.00	−4.4	+	++
26	6.01	10000	1400.00	−7.3	+	++

[a] Fit value indicates how well the features in the pharmacophore overlap with the chemical features in the molecule. Fit=weight × [max (0, 1 –SSE)], where $SSE=(D/T)^2$, D=displacement of the feature from the center of the location constraints, and T=the radius of the location constraint sphere for the feature (i.e., the tolerance)

[b] Difference between the predicted and experimental values. "+" indicates that the predicted IC_{50} is higher than the experimental IC_{50}; "-" indicates that the predicted IC_{50} is lower than the experimental IC_{50}; a value of 1 indicates that the predicted IC_{50} is equal to the experimental IC_{50}

[c] Activity scale : IC_{50}<300 nM=+++ (highly active); 300 nM≥IC_{50}<3000 nM=++ (moderately active); IC_{50} ≥ 3000 nM=+ (low activity)

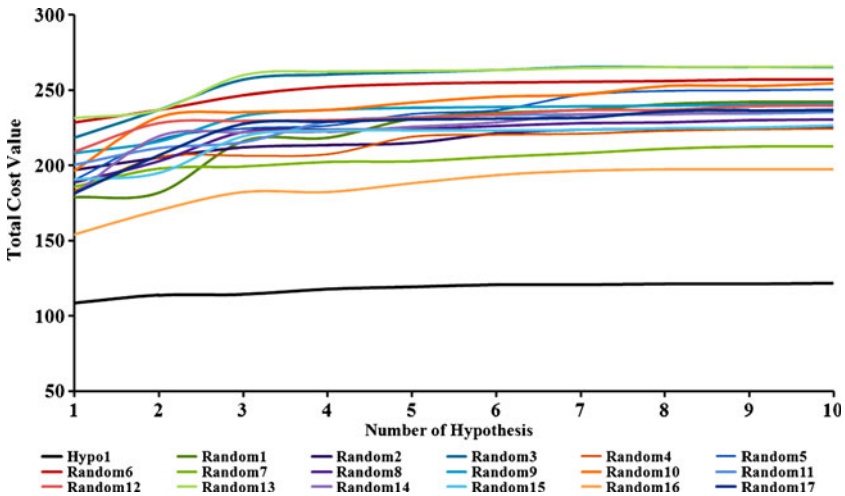

Fig. 4 Comparing the total cost of Hypo1 with the total costs of the 17 random hypotheses generated in the Fisher randomization run

Table 3 Experimental and Hypo1-predicted IC_{50} values of the 25 test set molecules

Compound no.	Exp. IC_{50} (nM)	Pred. IC_{50} (nM)	Error[a]	Exp. scale[b]	Pred. scale[b]
1	3.1	3.99	1.29	+++	+++
2	3.5	3.02	−0.86	+++	+++
3	6.2	7.29	1.18	+++	+++
4	15	5.37	−0.36	+++	+++
5	20	20.35	1.02	+++	+++
6	76	78.25	1.03	+++	+++
7	94	66.85	−0.71	+++	+++
8	100	735.27	7.35	+++	++
9	100	97.03	−0.97	+++	+++
10	150	152.95	1.02	+++	+++
11	170	162.50	−0.96	+++	+++
12	230	291.24	1.27	+++	+++
13	250	1629.34	6.52	+++	++
14	320	560.15	1.75	++	++
15	320	505.12	1.58	++	++
16	420	410.52	−0.98	++	++
17	660	859.31	1.30	++	++
18	700	532.02	−0.76	++	++
19	1,000	2364.14	2.36	++	++
20	1,400	80.12	−0.06	++	+++
21	1,800	1803.96	1.00	++	++
22	2,300	2237.28	−0.97	++	++
23	3,200	3034.31	−0.95	+	+
24	4,200	3996.00	−0.95	+	+
25	4,300	3424.47	−0.80	+	+

[a]Difference between the predicted and experimental values. "+" indicates that the predicted IC_{50} is higher than the experimental IC_{50}; "-" indicates that the predicted IC_{50} is lower than the experimental IC_{50}; a value of 1 indicates that the predicted IC_{50} is equal to the experimental IC_{50}

[b]Activity scale: IC_{50}<300 nM=+++ (highly active); 300 nM≥IC_{50}<3000 nM=++ (moderately active); $IC_{50} \geq 3000$ nM=+ (low activity)

factor. A decoy set containing 10 active and 2190 ALR2-inactive compounds was used to calculate these properties. Screening was performed using the Ligand Pharmacophore Mapping module in DS, and the values were calculated using the formula

$$EF = [(H_a \times D)/(H_t \times A)] \quad (3)$$

$$GF = [(H_a/4H_tA)(3A + H_t) \times (1 - ((H_t - H_a)/(D - A))]. \quad (4)$$

The number of hits (H_t) in the "hit list" was 14, the percentage yield of actives (Y) was 64.28%, the proportion of actives in the hit list (A) was 90%, and there was one false negative and five false positives. The best hypothesis should select the active compounds during the screening process, and a high percentage of the compounds that it picks should be active compounds; it should also be efficient at reducing true negatives and false positives. The EF (141.42) and GF (0.70) values are very good, indicating that Hypo1 screening is highly efficient (Table 4).

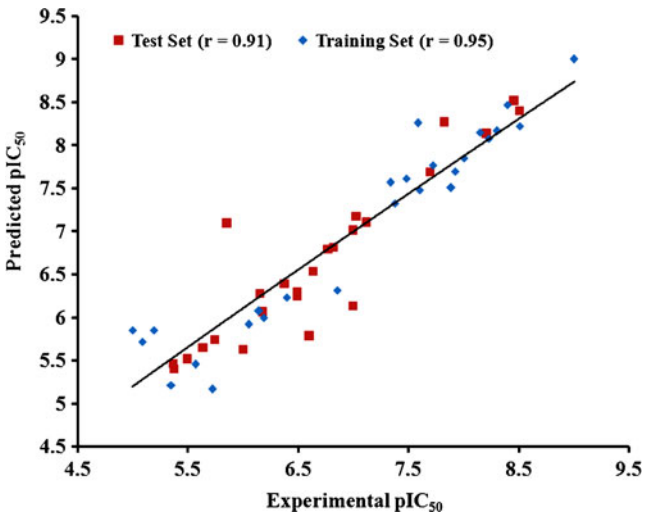

Fig. 5 Graph showing the correlation (r) between the experimental and Hypo1-predicted ALR2-inhibitory activities of 25 test set molecules along with 26 training set molecules

Database screening

Virtual screening of databases is an effective alternative to high-throughput screening methodologies, as well as a

Table 4 Statistical data associated with the screening of the molecules in the decoy set using Hypo1

Parameter	Value
Total number of molecules in the database (D)	2200
Total number of actives in database (A)	10
Total number of hit molecules from the database (H_t)	14
Total number of active molecules in the hit list (H_a)	9
% Yield of actives $[(H_a/H_t) \times 100]$	64.29
% Ratio of actives $[(H_a/H_A) \times 100]$	90.00
Enrichment factor (EF)	14.1
False negatives $[A - H_a]$	1
False positives $[H_t - H_a]$	5
Goodness of fit score[a] (GF)	0.70

[a] $[(H_a/4H_tA)(3A + H_t))^*(1 - ((H_t - H_a)/(D - A))]$; GH score above 0.6 indicates a very good model

sophisticated approach to the drug discovery process. Pharmacophore modeling, virtual screening, and molecular docking have become important tools in computer-aided drug design. We therefore used Hypo1 as a 3D structural query to screen the NCI database for potential inhibitors of ALR2. The NCI database consists of ~0.2 million compounds; among these, 2017 compounds had all of the chemical features present in Hypo1. These molecules were then further filtered down to 69 compounds by specifying that the maximum fit value must be above 10. However, even if a molecule has all of these required chemical features, it may not be active towards ALR2, so we tested the selected molecules for potential as a lead compound, based on their ADME properties and using Lipinski's rule of five. ADME and Lipinski's rule of five are the important criteria's to sort the small molecules based on drug-like properties. Therefore, we further sorted these molecules using ADME and the rule of five. Only three molecules (Fig. 6) passed the ADME and rule of five criteria, and these molecules were then subjected into a molecular docking program to study their critical interactions with the vital amino acids present in the active site of ALR2.

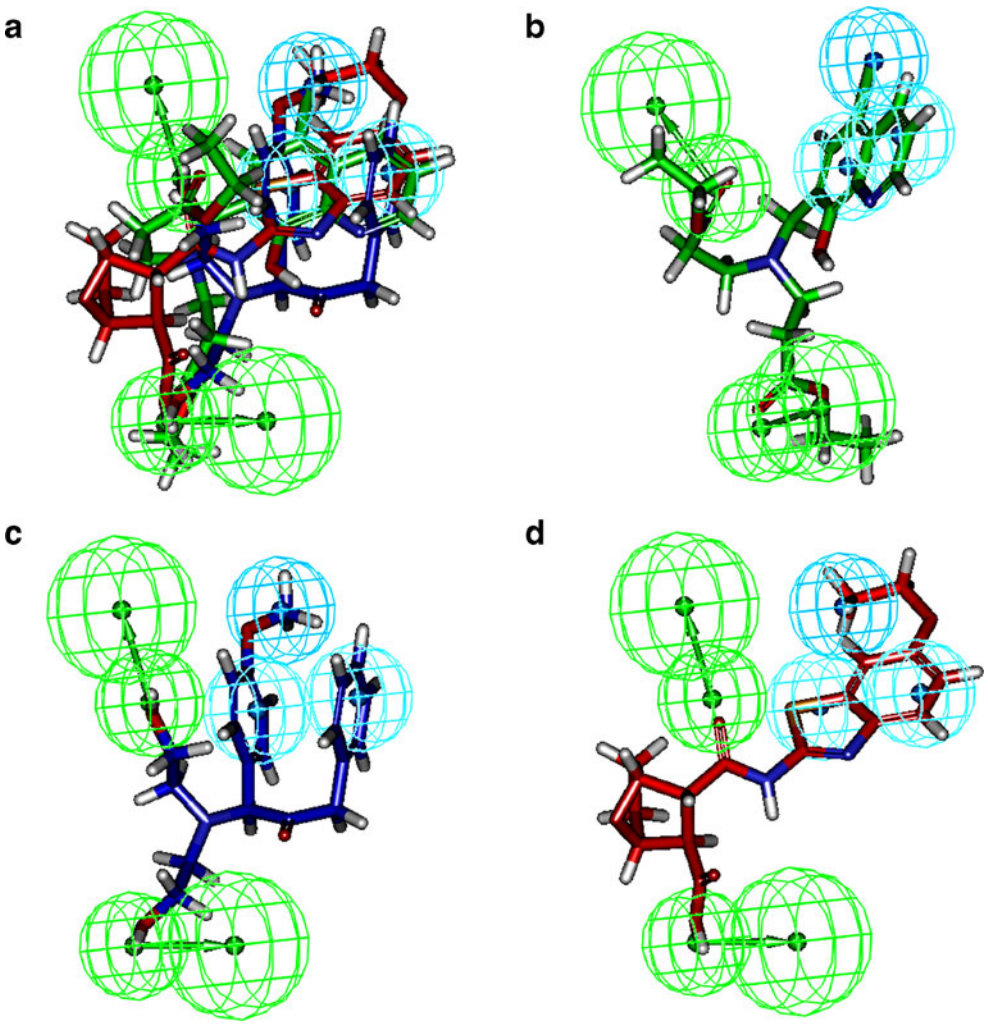

Fig. 6 **a**–**d** The pharmacophore model of ALR2 inhibitors Hypo1 was mapped onto compounds from the NCI database. **a** All three hit molecules (NCI0095667, NCI0036494, and NCI0019597) aligned with Hypo1. **b** NCI0095667 aligned with Hypo1. **c** NCI0036494 aligned with Hypo1. **d** NCI0019597 aligned with Hypo1. Features are color coded: *green* hydrogen-bond acceptor (HBA), *light blue* hydrophobic aliphatic (HAli), *dark blue* hydrophobic aromatic (HAro)

Molecular docking studies on ALR2

LigandFit in DS [41] was used to perform molecular docking in order to find the accurate orientations of the ligands at the active site of the protein. The candidate molecules from the NCI were docked into the active site of ALR2 in order to determine their optimal orientations and binding abilities. Initially the docking parameters were validated by docking the co-crystal molecule into the active site of ALR2. This validation process was performed to check whether the selected parameters were able to produce a suitable binding orientation of the ALR2 antagonist at the active site or not. Thus, the 3D structure of ALR2 combined with 3-[5-(3-nitrophenyl)thiophen-2-yl] propanoic acid (53N) was taken from the PDB (PDB ID: 3DN5), and 53N was sketched and docked into the active site of ALR2. The binding site contained two main subpockets: the catalytic anion binding site and the hydrophobic-specific pocket. The catalytic subpocket was deeply buried and consisted of residues that are presumably involved in the catalytic mechanism (Try48, Lys77, and His110). Additionally, the nicotinamide moiety of $NADP^+$ and Trp111 interact with the most of the ligands. Hydrophobic contacts can form with the side chains of Trp20, Val47, Trp79, and Trp219. 53N showed hydrogen bonds between its polar head group and the residues of the catalytic pocket (Tyr 48, His110, and Trp111), as well as short van der Waals interactions between the side chains of Trp111 and Leu300. An RMSD value of 0.598 Å (Fig. 7) between the best docked pose and the co-crystal of the ligand was obtained, revealing that the selected parameters are good for determining the orientation of ALR2 in the active site. Therefore, the same parameters were employed to dock the candidate compounds. All three molecules showed very strong interactions with critical residues like Leu300, Trp111, His110, and Tyr48. Among the three hit molecules, one (NCI0036494) showed a hydrogen bond as well as hydrophobic interactions with the active site residues, just as seen in the X-ray structure (Fig. 8). This compound was therefore selected for use in MD simulations aimed at probing the potential binding affinity and the adaptability of ALR2 towards the ligands.

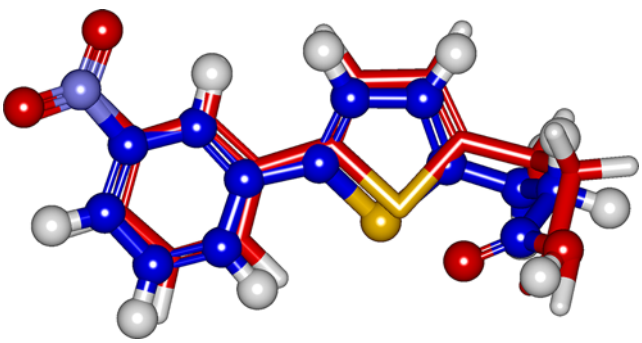

Fig. 7 Superposition of the co-crystal (53N) on its docked pose

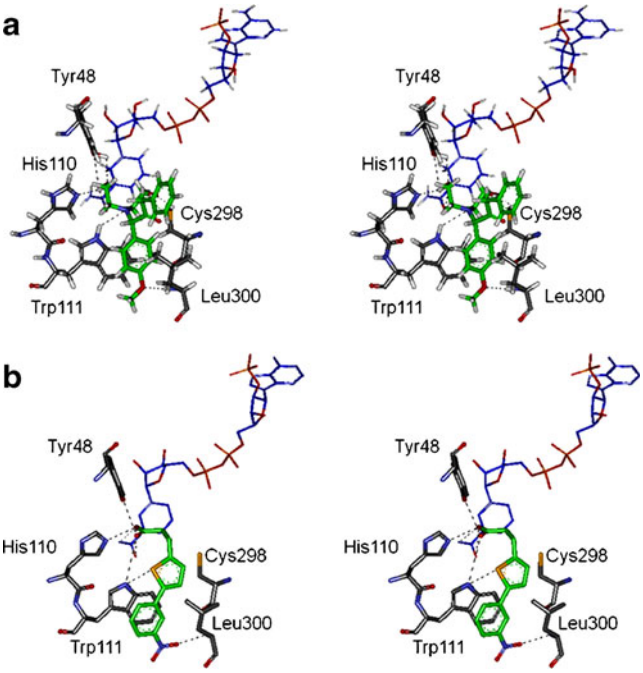

Fig. 8 **a**–**b** Stereo view of the binding modes of the compounds at the ALR2 active site, as suggested by molecular docking. **a** NCI0036494; **b** 53N. *Black lines* show hydrogen bonds

Molecular dynamics simulation

Based on the docking results, MD simulation was carried out for the ALR2–NCI0036494 complex and the complex structure from the PDB (PDB ID: 3DN5) using GROMACS. The behavior of the predicted complex was studied in a dynamic context, considering the flexibility of the protein. Superposing the coordinates of each complex structure allowed us to monitor how the RMS evolved. The drift in the root mean square deviation (RMSD) of the Cα atom of each protein was determined for both systems. The fluctuations tended to converge to around 1.5–2.5 Å after 2 ns of simulation (Fig. 9), which indicates that the whole system is stable and well equilibrated. The plot of the root mean square fluctuation (RMSF) shows that the residues present in the loop regions exhibit a great deal of movement, and the N- and C-terminal regions are also prone to significant fluctuations (Fig. 9). The potential energy plot showed a substantial decline in energy from the initial energy. A representative structure with an RMSD that was closest to the average structure from the last 2 ns of the MD simulations was chosen for use in the comparative analyses.

Upon analyzing the simulated complexes for multiple hydrogen-bond interactions, NCI0036494 and 53N were found to show similar interactions with the critical residues His110, Trp111, Tyr48, and Leu300, and these interactions were maintained throughout the simulation process. Multiple strong hydrogen-bond interactions were observed throughout

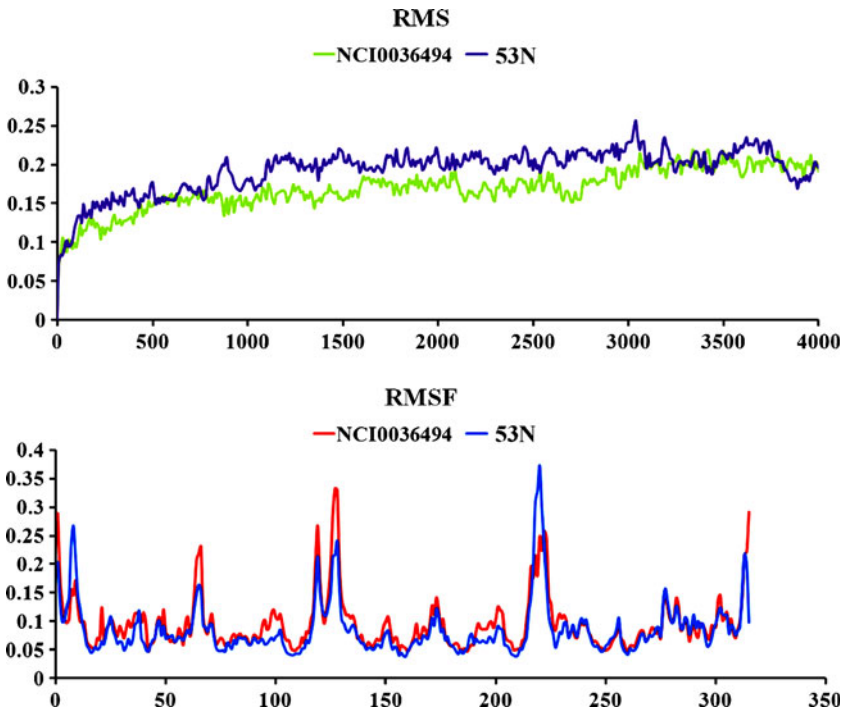

Fig. 9 RMS deviation profiles of the Cα atoms for the complexes of NCI0036494 and 53N with ALR2, and the local conformational changes as indicated by the RMSFs of the backbone atoms as a function of residue number

the simulation between the critical residues His110, Trp111, Tyr48, and Leu300 and the ligands (NCI0036494 and 53N). The distances between the lead molecule and the critical residues were similar to those between 53N and the residues, and the lead molecule also presented a similar orientation to 53N at the active site of ALR2. However, the residues His111 and Try 48 got closer to NCI0036494 than to 53N. His110 and Tyr48 function as H-bond acceptors by forming H-bonds with O25 and O1 of NCI0036494. The nitrobenzene group in the lead molecule exhibited a strong π–π interaction with Trp111, and a similar interaction was noted between the methoxy group of 53N and Trp111. Similar data were derived from the pharmacophore hypothesis and the docking studies, so this molecule could be a more potent antagonist of ALR2 than 53N (Fig. 10). Thus, when designing potent ALR2 inhibitors, it is important to ensure that the chemical features of Hypo1 are present in the structure of the potential inhibitor, as we have shown that hydrogen bonding plays an important role in the ligand–receptor interactions.

Similarity search and molecular docking studies

The NCI database compound NCI0036494 was selected as a potent ALR2-inhibiting compound based on various drug-like validations. We then performed a similarity search using this compound as the reference in the PubChem database. In the similarity search, a total of 41 molecules showed greater than 90% similarity to NCI0036494. These 41 compounds were tested for oral bioavailability by applying Lipinski's rule of five, and 38 compounds of those passed this test, indicating that these compounds are had good oral bioavailabilities. Next, these 38 molecules were minimized by applying the CHARMm force field [25–27], and the BEST conformational analysis was applied to generate a maximum number of 255 conformers using a Monte Carlo-like algorithm with an energy range of 20 kcal mol^{-1}. The generated conformers were screened using Hypo1 to check whether they satisfied the geometric constraints of the best hypothesis. In total, eight "hit" compounds (Fig. 11) from the PubChem database contained all of the critical chemical features present in Hypo1, and these molecules were subsequently subjected to molecular docking studies to determine whether these compounds adopted the appropriate orientation when bound at the active site of ALR2.

NCI0036494 and the eight hit compounds from the PubChem database showed better dock scores than the co-crystal (Table 5). The anion-binding site (a hydrophobic pocket) contained Trp111, His110, Tyr48, and the nicotinamide moiety of the bound cofactor NADPH/$NADP^+$. A hydrophobic-specific pocket also opened up next to the residues Thr113 and Leu300. Trp111 bordered the anion-binding pocket and exposed its π-face to the hydrophobic-specific pocket. The hit molecules showed a π–π interaction with Trp111, which was present at the anion-binding site (hydrophobic pocket), and the nicotinamide moiety of the bound cofactor NADPH/$NADP^+$. In addition, all of the hit molecules presented strong hydrogen-bond interactions with Trp111, His110, and Tyr48.

All of the critical interactions of the hit molecules with the active site of ALR2, and therefore the activities of the

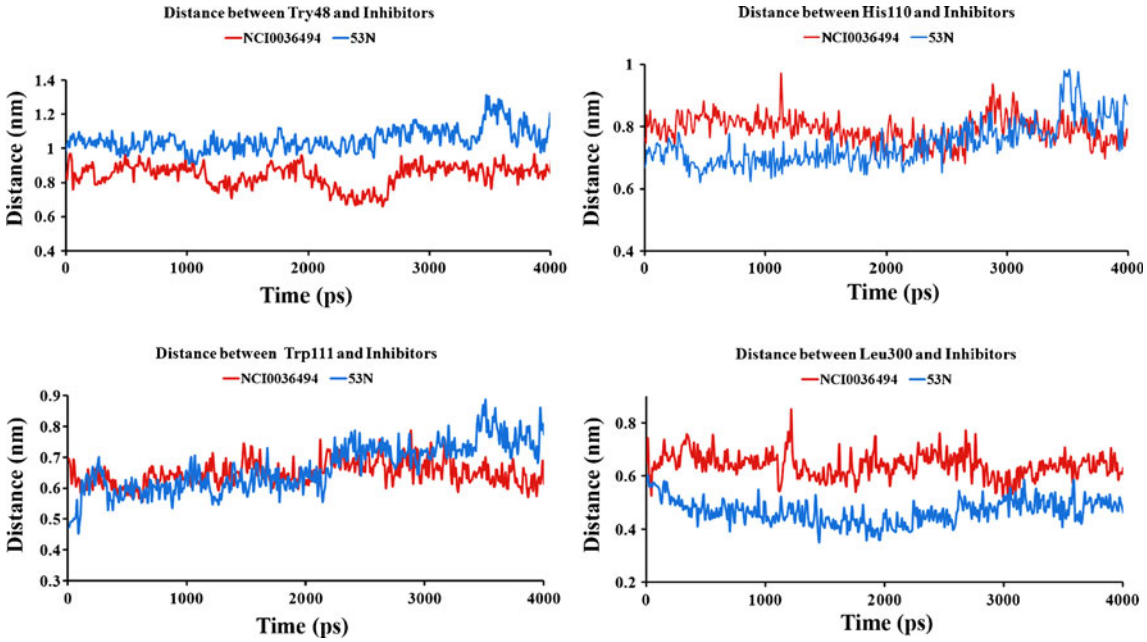

Fig. 10 The distances between each inhibitor and the Tyr48, His110, Trp111, and Leu300 groups in ALR2, plotted as a function of time

hit molecules towards ALR2, were confirmed by calculating various scoring functions such as the dock score, the Jain score, the potential of mean force (PMF), and the piecewise linear potential (PLP), and comparing the scores with those of 53N. All of the hit compounds presented higher scores than 53N. Among these nine molecules (one from NCI and eight from PubChem), five (411929, 110778826, 411600, 411165, and NCI36494) gave higher dock scores than 53N. The remaining four molecules (4386961, 39453612, 39453613, and 5135707) gave similar dock scores to 53N. The PLP scores were calculated based on the formation of hydrogen bonds between the protein and ligand. Compounds 411929, 411600, 411165, and 5135707 gave higher PLP scores than 53N, indicating that they bind strongly with ALR2. The PMF scores were obtained based on a statistical analysis of the three-dimensional structures of the protein–ligand complexes, and they correlate with the protein–ligand binding free energies. All of the hit compounds showed PMF values that were higher than that of 53N (157.4), indicating that our hit compounds had stronger receptor–ligand affinities (Table 6). The Jain score was predicted according to the sum of interactions such as the lipophilic

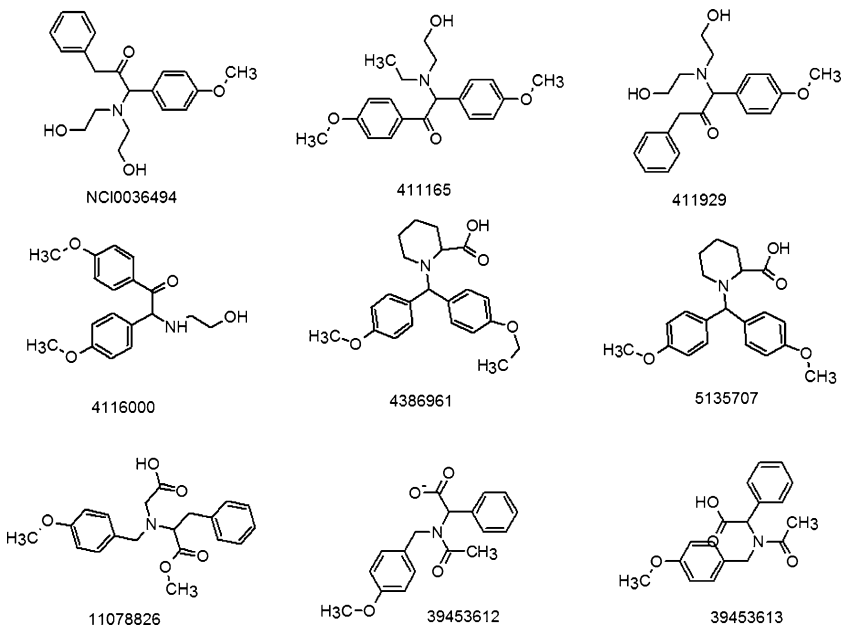

Fig. 11 The two-dimensional structures of the hit compounds

Table 5 Critical chemical interactions at the active site of ALR2

Name	π–π interaction	Hydrogen-bond Interaction			
	Trp111	Trp111	His110	Leu300	Tyr48
411929	Face to face	√	√	√	√
11078826	Face to face	√	√	√	√
411600	Face to face	√	×	√	√
411165	Face to face	√	×	√	√
4386961	Face to face	√	√	√	√
39453612	Face to face	√	√	√	√
6739453613	Face to face	√	√	√	√
39453613	Face to face	√	√	√	√
5135707	Face to face	√	√	√	√

interaction, the solvation of the protein and ligand, polar attractive interactions, polar repulsive interactions, as well as an entropy term for the ligand. The hydrophobic interactions, the entropy, and the energy changes were represented by the Jain and Ludi scores. The NCI compound gave a good dock score but a lower PLP score than 53N. All of these scores indicated that the compounds 411929, 4116000, 411165, and 5135707 present greater binding stabilities than 53N. Thus, we can conclude that these four lead compounds are good candidates for potent inhibitors of ALR2.

Conclusions

The purpose of this study was not only to construct a pharmacophore model that will predict the activities of the proposed ALR2-inhibitory compounds, but also to apply the best hypothesis in the virtual screening of a database to find novel scaffolds. Pharmacophore and molecular docking methods were applied to determine the structural features of inhibitors that are needed for them to bind to ALR2. MD simulation was also used to gain deep insight into the structural properties of the model and the binding details of the ALR2 complex. Various quantitative hypotheses were generated based on 26 known ALR2 antagonists, and the best ten hypotheses were retained for further evaluation. Ultimately, Hypo1 was selected as the best of these hypotheses based on Debnath analysis, Fisher's randomization, the use of a test set, and also by calculating GF and EF values. Hypo1 had the highest cost difference (128.44), the lowest RMSD (1.02), a good correlation coefficient (0.96), and a low total cost value (108.56) in comparison to the other hypotheses. Hypo1 produced good results during the validation process, so it was selected as a 3D structural query to screen the NCI database, and the hit molecules thus obtained were then filtered by applying a cutoff maximum fit value, studying their ADME properties, and using Lipinski's rule of five. Three of the hit molecules satisfied all of these conditions, and these were used in molecular docking studies to evaluate their optimal orientations and their interactions with the critical residues of ALR2. The three molecules showed strong hydrogen-bond interactions as well as hydrophobic contacts with critical residues such as His110, Trp111, Tyr48, and Leu300. The molecules NCI0036494 and 53N were then used in MD simulations to find the behavior of these molecules during 5 ns production runs. The hit molecule

Table 6 Comparison of the hit molecules with 53N using various scoring functions

Hit molecule	Ligscore1	Ligscore2	PLP1	PLP2	PMF	PMF04	Dock score	Ludi1	Ludi2	Ludi3	Jain
411929	1.5	2.43	97.86	106.85	201.39	128.14	108.543	383	374	1,182	5.94
11078826	3.38	5.24	95.66	97.22	180.12	121.2	107.47	425	377	717	6.13
411600	3.92	5.65	105.92	108.94	170.51	113.3	110.918	527	490	949	4.11
411165	2.28	4.57	103.64	104.01	192.13	135.23	108.21	576	507	1,035	4.63
4386961	2.64	4.29	90.96	97.36	167.76	106.44	98.429	427	413	754	4.92
39453612	2.85	4.91	94.22	91.39	185.63	112.4	98.775	469	400	868	5.65
39453613	1.58	3.86	96.47	98.37	179.49	113.22	100.459	479	446	1,149	7.49
5135707	2.69	4.7	100.6	103.79	180.16	111.59	95.114	442	427	882	6.3
NCI0036494	2.26	4.4	94.49	100.09	158.66	117.67	103.878	395	393	745	6.74
53N	5.66	6.17	102.81	104.66	157.4	128.4	104.832	509	463	941	5.84

NCI0036494 showed very similar interactions to 53N, as confirmed by analyzing the simulation trajectories (RMSD, RMSF, and the distance plots). NCI0036494 was then used as a query to search for similar compounds in the PubChem database. The 41 compounds found presented similarities to NCI0036494 of greater than 90%. After further analysis, eight of these compounds underwent docking studies, to check whether they complied with the geometric constraints of Hypo1. All of these molecules presented high values of properties such as the dock score, PMF, PLP, etc. Following a further round of analysis, we were able to conclude that four of these lead molecules are good candidate ALR2 inhibitors, and that Hypo1 is a good tool to select the novel candidate ALR2-inhibiting ligands from databases.

Acknowledgments This research was supported by the Basic Science Research Program (2009–0073267), the Pioneer Research Center Program (2009–0081539), and the Management of Climate Change Program (2010–0029084) through the National Research Foundation of Korea (NRF), as funded by the Ministry of Education, Science and Technology (MEST) of the Republic of Korea. This work was also supported by the Next-Generation BioGreen21 Program (PJ008038) from the Rural Development Administration (RDA) of the Republic of Korea.

References

1. National Institute of Diabetes and Digestive and Kidney Diseases (2011) National Diabetes Information Clearinghouse (NDIC). http://diabetes.niddk.nih.gov
2. Sato S, Kador PF (1990) Inhibition of aldehyde reductase by aldose reductase inhibitors. Biochem Pharmacol 40:1033–1042
3. Carper DA, Wistow G, Nishimura C, Graham C, Watanabe K, Fujii Y, Hayashi H, Hayaishi O (1989) A superfamily of NADPH-dependent reductases in eukaryotes and prokaryotes. Exp Eye Res 49:377–388
4. Costantino L, Rastelli G, Cignarella G, Vianello P, Barlocco D (1997) New aldose reductase inhibitors as potential agents for the prevention of long-term diabetic complications. Expert Opin Ther Patents 7:843–858
5. Singh SB, Malamas MS, Hohman TC, Nilakantan R, Carper DA, Kitchen D (2000) Molecular modeling of the aldose reductase–inhibitor complex based on the X-ray crystal structure and studies with single-site-directed mutants. J Med Chem 43:1062–1070
6. Nishimura C, Matsuura Y, Kokai Y, Akera T, Carper D, Morjana N, Lyons C, Flynn TG (1990) Cloning and expression of human aldose reductase. J Biol Chem 17:9788–9792
7. Sotriffer CA, Krämer O, Klebe G (2004) Probing flexibility and "induced-fit" phenomena in aldose reductase by comparative crystal structure analysis and molecular dynamics simulations. Proteins Struct Funct Bioinf 56:52–66
8. Fournier B, Bendeif EE, Bt G, Podjarny A, Lecomte C, Jelsch C (2009) Charge density and electrostatic interactions of fidarestat, an inhibitor of human aldose reductase. J Am Chem Soc 131:10929–10941. doi:10.1021/ja8095015
9. Gual P, Le Marchand-Brustel Y, Tanti JF (2003) Positive and negative regulation of glucose uptake by hyperosmotic stress. Diabetes Metab 29:566–575
10. Oates PJ (2002) Polyol pathway and diabetic peripheral neuropathy. Int Rev Neurobiol 50:325–328
11. Eisenmann M, Steuber H, Zentgraf M, Altenkamper M, Ortmann R, Perruchon J, Klebe G, Schlitzer M (2009) Structure-based optimization of aldose reductase inhibitors originating from virtual screening. Chem Med Chem 4:809–819
12. Mc VZ, Doan B, Dr S, Sredy J, Podjarny AD (2009) Discovery of [3-(4,5,7-trifluoro-benzothiazol-2-ylmethyl)-pyrrolo[2,3-b]pyridin-1-yl]acetic acids as highly potent and selective inhibitors of aldose reductase for treatment of chronic diabetic complications. Bioorg Med Chem Lett 7:2006–2008
13. Petrova T, Lunin VY, Ginell S, Hazemann I, Lazarski K, Mitschler A, Podjarny A, Joachimiak A (2009) X-ray-radiation-induced cooperative atomic movements in protein. J Mol Biol 387:1092–1105
14. El-Kabbani O, Darmanin C, Schneider TR, Hazemann I, Ruiz F, Oka M, Joachimiak A, Schulze-Briese C, Tomizaki T, Mitschler A, Podjarny A (2004) Ultrahigh resolution drug design II. Atomic resolution structures of human aldose reductase holoenzyme complexed with fidarestat and minalrestat: implications for the binding of cyclic imide inhibitors. Proteins Struct Funct Bioinf 55:805–813
15. El-Kabbani O, Dk W, Petrash M, Quiocho FA (1998) Structural features of the aldose reductase and aldehyde reductase inhibitor-binding sites. Mol Vision 4:19–25
16. Oka M, Matsumoto Y, Sugiyama S, Tsuruta N, Matsushima M (2000) A potent aldose reductase inhibitor, (2*S*,4*S*)-6-fluoro-2′,5′-dioxospiro[chroman-4,4′-imidazolidine]-2-carboxamide (fidarestat): its absolute configuration and interactions with the aldose reductase by X-ray crystallography. J Med Chem 43:2479–2483
17. Steuber H, Zentgraf M, Podjarny A, Heine A, Klebe G (2006) High-resolution crystal structure of aldose reductase complexed with the novel sulfonyl-pyridazinone inhibitor exhibiting an alternative active site anchoring group. J Mol Biol 356:45–56
18. Ramana KV, Srivastava SK (2010) Aldose reductase: a novel therapeutic target for inflammatory pathologies. Int J Biochem Cell Biol 42:17–20
19. Miyamoto S (2002) Molecular modeling and structure-based drug discovery studies of aldose reductase inhibitors. Chem-Bio Inf J 2:74–85
20. Accelrys Software, Inc. (2011) Corporate homepage. http://www.accelrys.com
21. Mylari BL, Larson ER, Beyer TA, Zembrowski WJ, Aldinger CE, Dee MF, Siegel TW, Singleton DH (1991) Novel, potent aldose reductase inhibitors: 3,4-dihydro-4-oxo-3-[[5-(trifluoromethyl)-2-benzothiazolyl]methyl]-1-phthalazineacetic acid (zopolrestat) and congeners. J Med Chem 34:108–122
22. Mylari BL, Zembrowski WJ, Beyer TA, Aldinger CE, Siegel TW (1992) Orally active aldose reductase inhibitors: indazoleacetic, oxopyridazineacetic, and oxopyridopyridazineacetic acid derivatives. J Med Chem 35:2155–2162
23. Lipinski CA, Aldinger CE, Beyer TA, Bordner J, Burdi DF, Bussolotti DL, Inskeep PB, Siegel TW (1992) Hydantoin bioisosteres. In vivo active spiro hydroxy acetic acid aldose reductase inhibitors. J Med Chem 35:2169–2177
24. Advanced Chemistry Development, Inc. (2010) Homepage. http://www.acdlabs.com
25. Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: the biomolecular simulation program. J Comput Chem 30:1545–1614
26. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4:187–217

27. MacKerell AD, Brooks CL, Nilsson L, Roux B, Won Y, Karplus M (1998) CHARMM: the energy function and its paramerterization with an overview of the program. In: Schleyer (ed) The encyclopedia of computational chemistry. Wiley, Chichester, pp 271–277
28. Smellie A, Teig SL, Towbin P (1995) Poling: promoting conformational variation. J Comput Chem 16:171–187
29. Sakkiah S, Thangapandian S, John S, Lee KW (2011) Pharmacophore based virtual screening, molecular docking studies to design potent heat shock protein 90 inhibitors. Eur J Med Chem 46:2937–2947
30. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46:3–26
31. Sakkiah S, Thangapandian S, John S, Kwon YJ, Lee KW (2010) 3D QSAR pharmacophore based virtual screening and molecular docking for identification of potential HSP90 inhibitors. Eur J Med Chem 45:2132–2140
32. Berman HM (2008) The Protein Data Bank: a historical perspective. Acta Crystallogr A 64:88–95
33. The PDB Team (2003) The Protein Data Bank. Methods Biochem Anal 44:181–198
34. Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Commun 91:43–56
35. Spoel DVD, Patriksson A, Seibert MM (2007) Protein folding properties from molecular dynamics simulations. In: Proc 8th Int Conf on Applied Parallel Computing: State of the Art in Scientific Computing (PARA 2006), Umeå, Sweden, 18–21 June 2006
36. Schuttelkopf AW, van Aalten DMF (2004) PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. Acta Crystallogr D 60:1355–1363
37. Aalten DMF, Bywater R, Findlay JBC, Hendlich M, Hooft RWW, Vriend G (1996) PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. J Comput Aided Mol Des 10:255–262
38. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J Comput Phys 23:327–341
39. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N. log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
40. Hess B, Bekker H, Berendsen H, Fraaije J (1997) LINCS: A linear constraint solver for molecular simulations. J Comput Chem 18:1463–1472
41. Venkatachalam CM, Jiang X, Oldfield T, Waldman M (2003) LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. J Mol Graph Model 21:289–307