Screening and Docking Studies of 266 Compounds from 7 Plant Sources as Antihypertensive Agents

**Enireddy Vamsidhar[4]*, Galla Venkata Swamy[3], Sashikanth Chitti[1],
P Ajay Babu[1], Galla Venkatasatyanarayana[2] and Adidela Daveedu Raju[3]**

[1]Translational Research Institute of Molecular Sciences (TRIMS) Sai Narasimha Towers,
MVP Colony, Sector-10, Visakhapatnam-530017; India
[2]Prism Degree & PG College, Dwarakanagar, Visakhapatnam, India
[3]St. Theressa Institute of Engineering and Technology, Garividi,Vizianagaram, India
[4]GITAM University, Rushikonda, Visakhapatnam, India

Abstract

Various proteins play important roles in hypertension and a number of plants have been tested for their efficacy in modulating hypertension. Angiotensin 1-converting enzyme, renin and extracellular regulated kinase 2(ERK2) proteins, respectively, has major role in hypertension and therefore protein - ligand interaction studies were performed on 266 compounds from different parts of 7 plants (Allium sativum, Coriandrum sativum, Dacus carota, Murrayya koneigii, Eucalyptus globus, Calendula officinalis and Lycopersicon esculentum). Analysis was conducted using GOLD (Genetic Optimisation for Ligand Docking) software as docking program and the molecules drawn in ISIS Draw software are energy minimized using cosmic - optimize 3D module of Tsar (Tools for structure activity relationships) software. Before docking plant compounds, software validation was performed and found that all co-crystallized ligands are within 2.0 A°. Further, docking and re-scoring of 266 compounds with GOLD, Molegro and eHiTS followed by rank-sum technique revealed high binding affinity of compound 27, from Allium sativum, with Angiotensin converting enzyme, 1UZE and Renin, 2IKO. The docked pose of compound 27 (Phytic acid) exactly fits into the active site region and the ligand formed more number of H-bond interactions than the co-crystallized ligand. The best compound that exhibited high binding affinity with 3ERK was molecule 23 (Stigmasterol) from Lycopersicon esculentum.

Introduction

Hypertension is a highly prevalent cardiovascular risk factor (Zoccali et al., 2002) and an increase in blood pressure (BP) increases the risk of developing heart disease, obesity (Chow et al., 2000), kidney disease (Johnson et al., 2007), eye damage, and stroke (Stokes et al., 1987). Excessive salt intake has been suspected as a cause of high prevalence of hypertension. A high dietary salt intake contributes to the risk of hypertension, which further antagonizes the BP-lowering effect of most anti-hypertensive drugs (Joseph et al., 2003).

It has been reported that a combination of diet rich in fruits, vegetables, and low-fat dairy products could substantially lower blood pressure levels (Appel et al., 1997). In another study, re-searchers evaluated the association between dietary habits and levels of blood pressure (BP) in 4,304 women and men aged 18 to 30 observed that whole grains have been consistently associated with lower blood pressure (Steffen et al., 2005) Furthermore, the researchers suggested that high intakes of plant foods and low intakes of meat products may help high blood pressure treatment and these benefits can be linked to the presence of specific compounds in plants. Various plants have been tested for their efficacy in modulating hypertension, however, when literature was searched for computer-aided docking studies on compounds from plants vs. proteins that mediate hypertension, none of the reports were found to contain the required information. Also, many virtual screening studies have been reported in literature stating the importance of dataset, algorithms and scoring functions, whereas none of the works contain screening compounds from plants. This provided us the rationale to screen plant based compounds using GOLD (Genetic Optimisation for Ligand Docking) software. Hence, in this paper we report screening various compounds from seven plant sources (*Allium sativum, Coriandrum sativum, Dacus carota, Murrayya koneigii, Eucalyptus, Calendula officinalis and Lycopersicon esculentum*) against proteins, Angiotensin 1-converting enzyme 1UZE (PDB id) (Natesh et al., 2004), renin 2IKO (White et al., 2000) and extracellular regulated kinase 2(ERK2) 3ERK, (Xu et al., 1996) extracted from Protein Data Bank (PDB), respectively.

Materials and Methods

Virtual screening

Virtual screening utilizes docking and scoring of each compound from a dataset and the technique employed is based on the prediction of binding modes and binding affinities of each compound in the dataset by means of docking to an X-ray crystallographic structure (Jalaie and Shanmugasundaram, 2006). Some recent studies (Warren et al., 2006) have focused on certain crucial factors such as the size and diversity of the ligand dataset, wide range of targets and the evaluation of docking pro-

grams. Taking these aspects into consideration, diverse compounds from seven plants and three protein targets are evaluated.

However, in general, it is important to visualize the docked poses of high-scoring compounds because many ligands are docked in different orientations and may often miss interactions that are known to be important for the target receptor. This sort of study becomes more difficult as the size of the dataset increases. Therefore, an alternative approach is to eliminate unpromising compounds before docking by restricting the dataset to drug-like compounds; by filtering the dataset based on appropriate property and sub-structural features and by performing diversity analysis (Waszkowycz, 2008).

Dataset

Chemical compound names from each plant were obtained from Dukes Ethnobotany (http://www.ars-grin.gov/duke/) and the respective structures are searched in various literature databases. This resulted in 266 compounds, selected based on the property and sub-structural features, from seven plants (*Allium sativum, Coriandrum sativum, Dacus carota, Murrayya koneigii, Eucalyptus, Calendula officinalis and Lycopersicon esculentum*) were drawn using ISISDraw software (www.mdli.com). The 2D structures are converted into 3D structures by using corina 3D analysis tool in Tsar (Tools for structure activity relationships) software (www.accelrys.com). The geometries of these compounds were optimized using *cosmic optimize* 3D module and the charges were added. All molecules were written as mol2 files.

Receptor X-ray structure

The X-ray crystal structure of Angiotensin 1-converting enzyme, 1UZE, in complex with inhibitor enalprilat (Natesh et al., 2004) was recovered from the Protein Data Bank. Although residues of the enzyme were missing from 435 to 438 in the structure, they are too far away from the binding site and hence repairing was not done. Similarly, renin enzyme, 2IKO, in complex with inhibitor 7IG (Sarver et al., 2007) and 3ERK in complex with SB4 (Wang et al., 1998) were extracted from PDB. We used the molecular docking program GOLD (Genetic Optimisation for Ligand Docking) for virtual ligand screening based on docking, and a consensus scoring and ranking was employed to generate classes using GOLD score, eHiTS_Score (electronic High Throughput Screening), MolDock score of Molegro software respectively.

GOLD docking

Water molecules were discarded from the pdb file, added hydrogens and missing side chains were reconstructed using the option '*prepare file for docking programs*' available at the WHAT-IF web interface (http://swift.cmbi.ru.nl/servers/html/index.html). The structure was then eventually converted to *mol2* format using Mercury (v. 1.4.2; Cambridge Crystallographic Data Centre (CCDC)) (http://www.ccdc.cam.ac.uk/products/mercury).

Automated docking studies were then performed using the genetic algorithm GOLD (Jones et al., 1997) (v. 3.1; CCDC, Cambridge, UK). The algorithm had been previously validated and successfully tested on a data set of over 300 complexes extracted from the PDB (Nissink et al., 2002).

The GOLD program uses a genetic algorithm (GA) to explore the full range of ligand conformational flexibility and the rotational flexibility of selected receptor hydrogens. The mechanism for ligand placement is based on addition of fitting points to hydrogen-bonding groups on the protein and ligand and maps acceptor points in the ligand on donor points in the protein and vice versa. Additionally, GOLD generates hydrophobic fitting points in the protein cavity onto which ligand CH groups are mapped. The genetic algorithm optimizes flexible ligand dihedrals, ligand ring geometries, dihedrals of protein OH and NH3 groups, and the mappings of the fitting points. The docking poses are ranked based on a molecular mechanics-like scoring function, GoldScore (Verdonk et al., 2003). In the present work, the binding site was defined as a spherical region which encompasses all protein atoms within 10.0 A$^\circ$ of each crystallographic ligand atom. Default settings were used for all calculations.

For each of the 10 independent GA runs, a maximum number of 25000 GA operations were performed on a single population of 50 individuals. Operator weights for crossover, mutation, and migration were set to defaults. To further speed up the calculation, the GA docking was stopped when the top three solutions were within 1.5 A$^\circ$ RMSD (Root Mean Square Deviation) of each other.

Before screening 266 plant compounds, the docking protocol was validated. 1UZE, 2IKO and 3ERK proteins with bound ligand were docked individually into their corresponding binding pockets to obtain the docked pose and the RMSD of all atoms between these two conformations in each case are 0.89, 1.15 and 1.04 A$^\circ$ indicating that the parameters for docking simulation are good in reproducing the X-ray crystal structure.

Consensus scoring and ranking

Generally, docking programs have the ability to predict protein-ligand complex structures with reasonable accuracy and speed. The ability to predict the probable binding mode of a ligand to differentiate correct poses from incorrect ones is based on reliable scoring functions. However, combinations of various scoring functions would reduce the errors in single scoring schemes and improve the probability of identifying true hits (Kitchen et al., 2004). Thus, it has been demonstrated that consensus scoring is generally more effective than single scoring for molecular docking (Wang et al., 2003; Clark et al., 2002) and represented an effective way in getting improved hit rates in various virtual database screening studies (Charifson et al., 1999).

In our study, we tested three different scoring functions such as GOLD score, dock score implemented in eHiTS (electronic High Throughput Screening) and MolDock score of Molegro software respectively. Docking program GOLD was used to dock compounds to generate an ensemble of docked conformations and each scoring function is applied to generate classes based on the obtained dock scores followed by ranking the best conformations. During ranking, signs of some scoring functions are changed to make certain that a lower score always indicates a higher affinity.

Results and Discussion

Dock runs of all 266 compounds on proteins 1UZE, 2IKO and 3ERK using GOLD resulted in few best compounds that were evaluated based on the binding compatibility [docked energy (kcal/mol)] with the receptor. The software generated 10

conformers for each docked molecule in about 3-5 minutes. In each case, binding energies greater than the co-crystallized ligand are selected.

Dock scores of co-crystallized ligands run in triplicates are within 66.29-67.69 kcal/mol range for 1UZE, 41.32-42.25 for 2IKO and 53.55-54.70 kcal/mol for 3ERK proteins, respectively, and hence any molecule from the dataset that result in scores higher than the range are considered more appropriate. Therefore, in the first step, virtual screening with docking and scoring resulted in few best hits (Table 1). In the second step, consensus scoring was applied to generate different scores for these compounds. Likewise, re-scoring docking poses with independent functions is another valuable approach which gained prominence in recent studies. Therefore, re-scoring of best docked poses based on their interaction energies with respective protein active site residues was done using GoldScore scoring function.

Further re-scoring was carried out using eHiTS scoring function (http://www.simbiosys.ca) and each molecule was optimized using eHiTS_Score optimization routine to remove steric clashes, if any, or other artifacts that may influence the score (Zsoldos et al., 2007). Post-scoring results are evaluated for RMSD (Root Mean Square Deviation) and found to be within 2A°.

Similarly, the next scoring function MolDock score of Molegro

software was applied to remove any steric clashes and the RMSD was found to be less than 2A°. This scoring function was derived from PLP (Piecewise Linear Potential) scoring functions originally proposed by Gehlhaar et al. (1995) and later extended by Yang et al (Yang and Chen, 2004). The scoring function was further improved to include new hydrogen bonding term and new charge schemes (Thomsen and Christensen, 2006).

In all the above cases, ranking was done individually by clustering best scored compounds into equally split four classes using Tsar software, of which compounds in Class4 represents the highest class or top rank. Classes were generated for all scoring functions and instead of taking an average, rank-sum technique (Clark et al., 2002) was employed to retrieve best compounds. The ranks obtained from each of the individual scoring functions were added to give a rank-sum (Table 1).

The advantage of a sum over an average is that the contribution from each individual score can more easily be split out for illustrative purposes in the former instance. Finally, from top rank-sum classes, m27_3 and m27_9 (Phytic acid) compound conformers from *Allium sativum* are considered as potential ligands against Angiotensin converting enzyme, 1UZE and Renin, 2IKO proteins and m23_2 (Stigmasterol) conformer from *Lycopersicon esculentum* as potential ligand against extracellular regulated kinase, 3ERK protein, respectively.
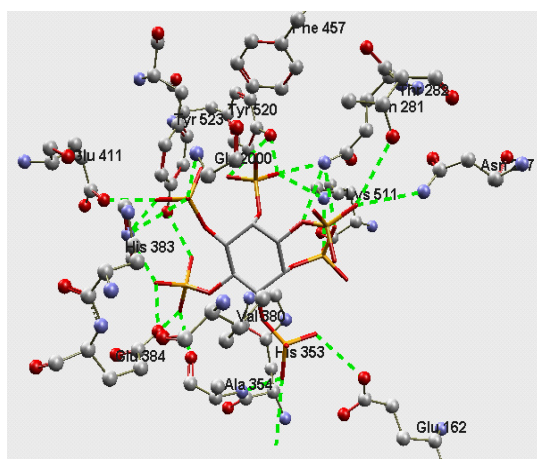
Protein, PDB id	Plant	Conformer (Molecule name)	GOLD score (kcal/mol)	MolDock score (kcal/mol)	eHiTS dockscore (kcal/mol)	Classes			
						GOLD score	MolDock score	eHiTS dock score	Sum
Angiotensin converting enzyme, 1UZE	Allium sativum	m27_3 (Phytic acid)	80.18	150.99	6.37	4	4	4	12
Renin, 2IKO	Allium sativum	m16_9 (Glutathione)	54.64	107.21	6.88	1	2	4	7
	Allium sativum	m27_9 (Phytic acid)	81.94	141.65	2.62	4	4	1	9
	Corinadrum Sativum	m40_6 (Roboflavine)	50.97	99.89	3.35	1	1	1	3
	Calendula Officinalis	m14_4 (Kaempferitin)	47.95	132.17	5.2	1	4	3	8
	Lycopersicon esculentum	m25_1 (Thiamin)	52.03	106.86	5.26	1	2	3	6
Extracellular regulated Kinase,3ERK	Allium sativum	m27_9 (Phytic acid)	75.6	105.69	3.55	4	3	1	8
	Lycopersicon Esculentum	M23_2 (Stigmasterol)`	70.13	124.94	7.77	4	4	4	12

Table 1: Optimized dock scores of best compounds that scored higher than respective original ligands against 1UZE, 2IKO and 3ERK proteins and their respective classes.
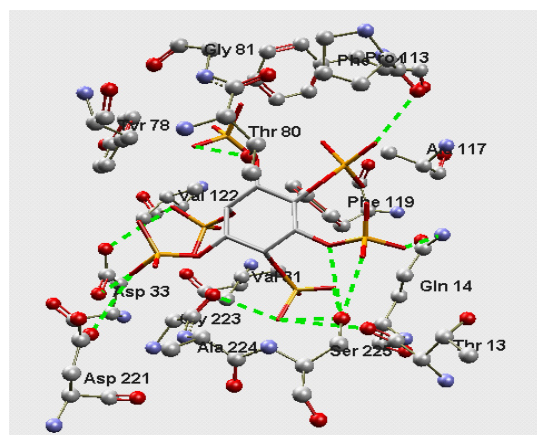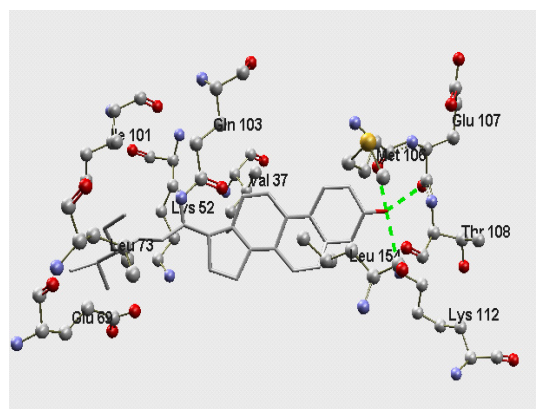
Binding modes

Active site of 1UZE offers different binding modes for these compounds as they are strongly dependent on the attached substituent. 1UZE bound ligand was docked deeply within the binding pocket region forming interactions with His353, Ala354, Lys511, His513, Tyr520 and Tyr523, respectively. Whereas molecule m27_3 conformer from *Allium sativum* formed more number of H-bond interactions with Glu162, Asn277, Gln281, Thr282, His353, Ala354, His383, Glu384, Glu411, Lys511, Tyr520, Tyr523 and Gly2000, respectively. It is probably be-

cause of interactions with these residues, the scores are higher in all softwares under study. Similar observation was made for 2IKO protein docked with molecule 27 from *Allium sativum* formed major H-bond interactions with Thr13, Gln14, Asp33, Thr80, Pro113, Ala117, Asp221, Gly223, Ala224 and Ser225 and the co-crystallized ligand of 2IKO formed Asp33, Ser79, Thr80 and Asp221 residue interactions respectively. Original bound ligand of 3ERK showed three H-bond interactions with Asp104, Met106 and Asp109 residues whereas the best compound m23_2 conformer from *Lycopersicon esculentum* identified against extracellular regulated kinase showed Met106, Thr108 and Lys112 residue interactions with ligand, respectively.

From our analysis, it is evident that plant based compounds may exhibit anti-hypertension and the best compounds from docking and scoring runs resulted in few hits, the major interacting residues in each case are given in Figure 1. Interestingly, compound 27, phytic acid from *Allium sativum* represented the best compound with 1UZE and 2IKO proteins, whereas the same compound reported to be moderate than molecule m23, stigmasterol of *Lycopersicon esculentum*.

Conclusion

Screening methods are routinely and extensively used to reduce cost and time of drug discovery. It has been clearly demonstrated that the approach utilized in this study is successful in finding novel anti-hypertensive inhibitors from plants. Compound 27, in particular, from *Allium sativum* showed high binding affinity against Angiotensin converting enzyme, 1UZE and Renin, 2IKO. The docked pose of compound 27 exactly fits into the active site region and the ligand formed more number of H-bond interactions than the co-crystallized ligand. Though the number of H-bond interactions is less, the best compound observed from *Lycopersicon esculentum* was molecule 23, yet this compound exhibited high binding affinity with 3ERK. Therefore, this study states the importance of small molecules from various plant sources and their use to enhance protein-ligand interaction studies, *in silico*. This approach to screen compounds from plants depends on various parameters such as size and shape of the compound and pharmacophoric groups attached on the compounds, among others. Further, work can be extended to study the receptor-ligand interactions experimentally and evaluation of their biological activity would help in specifying compounds against hypertension based on screening, docking and consensus scoring techniques.

a) Angiotensin converting enzyme, 1UZE

b) Renin, 2IKO

c) molecule 23_2

Figure 1: Images of molecule m27 conformers from Allium sativum bound within active site region of a) Angiotensin converting enzyme, 1UZE b) Renin, 2IKO and c) molecule 23_2 conformer of Lycopersicon esculentum bound to extracellular regulated kinase, 3ERK. Hydrogens are removed for clarity.

References

1. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, et al. (1997) A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. N Engl J Med 336: 1117-1124. » CrossRef » PubMed » Google Scholar

2. Charifson PS, Corkery JJ, Murcko MA, Walters WP (1999) Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J Med Chem 42: 5100-5109. » CrossRef » PubMed » Google Scholar

3. Chow WH, Gridley G, Jr Fraumeni JF, Järvholm B (2000) Obesity, hypertension, and the risk of kidney cancer in men. N Engl J Med 343: 1305-1311. » CrossRef » PubMed » Google Scholar

4. Clark RD, Strizhev A, Leonard JM, Blake JF, Matthew JB (2002) Consensus scoring for ligand/protein interactions. J Mol Graph Model 20: 281-295. » CrossRef » PubMed » Google Scholar

5. Gehlhaar DK, Verkhivker G, Rejto PA, Fogel DB, Fogel LJ, et al. (1995) Docking Conformationally Flexible Small Molecules into a Protein Binding Site through Evolutionary Programming. Proceedings of the Fourth International Conference on Evolutionary Programming 615-627. » CrossRef » PubMed » Google Scholar

6. Izzo JL, Black HR, Goodfriend TL (2003) Hypertension Primer: The Essentials of High Blood Pressure. Lippincott Williams & Wilkins. » CrossRef » PubMed » Google Scholar

7. Jalaie M, Shanmugasundaram V (2006) Virtual screening: are we there yet? Mini Rev Med Chem 6: 1159-1167. » CrossRef » PubMed » Google Scholar

8. Johnson RJ, Segal MS, Sautin Y, Nakagawa T, Feig DI, et al. (2007) Potential role of sugar (fructose) in the epidemic of hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease. Am J Clin Nutr 86: 899-906. » CrossRef » PubMed » Google Scholar

9. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267: 727-748. » CrossRef » PubMed » Google Scholar

10. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 3: 935-949. » CrossRef » PubMed » Google Scholar

11. Natesh R, Schwager SL, Evans HR, Sturrock ED, Acharya KR (2004) Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme. Biochemistry 43: 8718-8724. » CrossRef » PubMed » Google Scholar

12. Nissink JW, Murray C, Hartshorn M, Verdonk ML, Cole JC, et al. (2002) A new test set for validating predictions of protein-ligand interaction. Proteins 4: 457-471. » CrossRef » PubMed » Google Scholar

13. Sarver RW, Peevers J, Cody WL, Ciske FL, Dyer J, et al. (2007) Binding thermodynamics of substituted diaminopyrimidine renin inhibitors. Anal Biochem 360: 30-40. » CrossRef » PubMed » Google Scholar

14. Steffen LM, Kroenke CH, Yu X, Pereira MA, Slattery ML, et al. (2005) Associations of plant food, dairy product, and meat intakes with 15-y incidence of elevated blood pressure in young black and white adults: the Coronary Artery Risk Development in Young Adults (CARDIA) Study. Am J Clin Nutr 82: 1169-1177. » CrossRef » PubMed » Google Scholar

15. Stokes J III, Kannel WB, Wolf PA, Cupples LA, D'Agostino RB (1987) The relative importance of selected risk factors for various manifestations of cardiovascular disease among men and women from 35 to 64 years old: 30 years of follow-up in the Framingham Study. Circulation 75: 65-73. » CrossRef » PubMed » Google Scholar

16. Thomsen R, Christensen MH (2006) MolDock: a new technique for high-accuracy molecular docking. J Med Chem 49: 3315-3321. » CrossRef » PubMed » Google Scholar

17. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD (2003) Improved protein-ligand docking using GOLD. Proteins 52: 609-623. » CrossRef » PubMed » Google Scholar

18. Wang R, Lu Y, Wang S (2003) Comparative evaluation of 11 scoring functions for molecular docking. J Med Chem 46: 2287-2303. » CrossRef » PubMed » Google Scholar

19. Wang Z, Canagarajah BJ, Boehm JC, Kassisa S, Cobb MH, et al. (1998) Structural basis of inhibitor selectivity in MAP kinases. Structure 6: 1117-1128. » CrossRef » PubMed » Google Scholar

20. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, et al. (2006) A critical assessment of docking programs and scoring functions. J Med Chem 49: 5912-5931. » CrossRef » PubMed » Google Scholar

21. Waszkowycz B (2008) Towards improving compound selection in structure-based virtual screening. Drug Discov Today 13: 219-226. » CrossRef » PubMed » Google Scholar

22. White WB, Prisant LM, Wright JT Jr (2000) Management of patients with hypertension and diabetes mellitus: advances in the evidence for intensive treatment. Am J Med 108: 238-245. » CrossRef » PubMed » Google Scholar

23. Xu Q, Liu Y, Gorospe M, Udelsman R, Holbrook NJ (1996) Acute hypertension activates mitogen-activated protein kinases in arterial wall. J Clin Invest 97: 508-514. » CrossRef » PubMed » Google Scholar

24. Yang JM, Chen CC (2004) GEMDOCK: A Generic Evolutionary Method for Molecular Docking. Proteins 55: 288-304. » CrossRef » PubMed » Google Scholar

25. Zoccali C, Mallamaci F, Tripepi G (2002) Hypertension as a cardiovascular risk factor in end-stage renal failure. Curr Hypertens Rep 4: 381-386. » CrossRef » PubMed » Google Scholar

26. Zsoldos Z, Reid D, Simon A, Sadjad SB, Johnson AP (2007) eHiTS: a new fast, exhaustive flexible ligand docking system. J Mol Graph Model 26: 198-212. » CrossRef » PubMed » Google Scholar