# Computational analysis on Cuminum cyminum compounds against aldose reductase as anti-diabetic agents

**Naresh Babu Muppalaneni[1], Allam Appa Rao[2]**

[1] Associate Professor, Department of Computer Science & Engineering, Avanthi's Research & Technological Academy

[2] Vice-Chancellor, Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh, INDIA

**Abstract**

*Various proteins play important roles in diabetes and a number of plants have been tested for their efficacy in modulating diabetes. Of all the proteins, we selected aloes reductase enzyme to analyze few plant compounds computationally for their efficacy towards protein inhibition. A total of 85 compounds from different parts of a plant, Cuminum cyminum were studied. Analysis was conducted using Molegro Virtual Docker 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 the co-crystallized ligand was within 2.0 A°. Further, docking and re-scoring of top ten compounds with GOLD, AutoDock vina, eHiTS, PatchDock and MEDock followed by rank-sum technique revealed high binding affinity of compound Apigetrin.*

**Keywords**—Computer Science, Computer Application, Computer Aided Drug Design, type 2 Diabetes, Docking, GOLD, Molegro, aldose reductase

## 1. Introduction

Human body gets energy by making glucose from foods like bread, rice, potatoes etc., To use this glucose human body needs insulin. Insulin is hormone that helps the body control the level of glucose in the body. Type 2 diabetes is disease in which pancreas does not produce enough insulin or body may not utilize insulin produced. Diabetes mellitus is a group of metabolic diseases characterized by hyperglycemia resulting from defects in insulin secretion, insulin action, or both. The chronic hyperglycemia of diabetes is associated with long-term damage, dysfunction, and failure of various organs, especially the eyes, kidneys, nerves, heart, and blood vessels. [1].

Furthermore, the researchers suggested that high intakes of plant foods and low intakes of meat products may help high blood pressure treatment and proper insulin levels and hence these benefits can be linked to the presence of specific compounds in plants. Various plants have been tested for their efficacy in modulating diabetes, however, when literature was searched for computer-aided docking studies on compounds from plants vs proteins that mediate diabetes, very few 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 Molegro Virtual Docker software. Hence, in this paper we report screening various compounds from Cuminum cyminum against Aldose reductase extracted from Protein Data Bank (PDB).

## 2. Materials And Methods

### 2.1 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 [2]. Some recent studies [3] 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 programs. 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 [4].

## 2.2 Data Set

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 85 compounds, selected based on the property and sub-structural features, from Cuminum cyminum 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.

## 2.3 Receptor X-ray structure

The X-ray crystal structure of Aldose reductase, 1AH3, in complex with inhibitor was recovered from Protein Data Bank. We used the molecular docking program Molegro Virtual Docker (MVD) for virtual ligand screening based on docking, and a consensus scoring and ranking was employed to generate classes using MolDock score of Molegro software respectively.

## 2.4 Molegro Docking

Water molecules were discarded from the pdb file, added hydrogens and missing side chains were reconstructed. Automated docking studies were then performed using the genetic algorithm to explore the full range of ligand conformational flexibility and the rotational flexibility of selected receptor hydrogens. The docking poses are ranked based on a scoring function, MolDock score. The scoring scheme was derived from PLP [Piecewise Linear Potential] scoring functions originally proposed by Gehlhaar et al [5] and later extended by Yang et al [6]. In the present work, the binding site was defined as a spherical region which encompasses all protein atoms within 15.0 Å of each crystallographic ligand atom. Default settings were used for all calculations.

Before screening plant compounds, the docking protocol was validated. 1AH3 with bound ligand was docked individually into its corresponding binding pocket to obtain the docked pose and the RMSD of all atoms between these two conformations was 0.87 A° (Table 1) indicating that the parameters for docking simulation are good in reproducing the X-ray crystal structure.

Table1: Table showing the RMSD values of 1AH3 in three runs.

| SlNo | PDB ID | Run1 | Run2 | Run3 |
|---|---|---|---|---|
| 1 | 1AH3 | 0.8736 | 0.8721 | 0.670 |

## 2.5 Consensus Scoring and Ranking

Generally, docking programs have the ability to predict the experimental orientations of protein-ligand complexes. The ability to predict the ideal binding mode of a ligand and to differentiate correct poses from incorrect ones is based on reliable scoring functions. However, it has been reported that various combinations of scoring functions would reduce errors when compared to single scoring scheme which improves the probability of identifying true hits [7]. Thus, it has been demonstrated that consensus scoring is generally more effective than single scoring for molecular docking [8,9] and represented an effective way in getting improved hit rates in various virtual database screening studies [10]

In our study, we tested three different scoring functions such as GOLD score of GOLD docking routine, 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.

# 3. Results

Dock runs of 85 compounds on protein 1AH3 using MVD resulted in few best compounds that were evaluated based on the binding compatibility [docked energy (kcal/mol)] with the receptor. The software generated 5 conformers for each docked molecule and in each case, binding energies greater than the co-crystallized ligand were only selected.

Dock scores of co-crystallized ligand of 1AH3 run in triplicates are within -105.52 to -107.01 kcal/mol, 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-2]. 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 MolDock score scoring function.

Table 2: Table showing the dock scores of best compounds from Cuminum cyminum

| S.No. | Compound | Affinity (kcal/mol) |
|---|---|---|
| 1 | Riboflavin | -133.388 |
| 2 | Apigenin-5-o-glucoside | -131.705 |
| 3 | Apigetrin | -130.833 |
| 4 | Apiin | -127.982 |
| 5 | Benzyl Cinnamate | -117.458 |

| | | |
|---|---|---|
| 6 | Luteolin | -116.643 |
| 7 | Stigmasterol | -116.379 |
| 8 | Cosmosin | -115.701 |
| 9 | Luteolin-7-o-glucoside | -112.54 |
| 10 | Cynaroside | -111.372 |

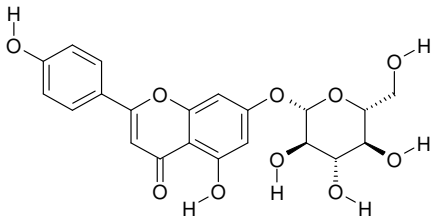

**Figure 1: 2-dimensional structure of Apigetrin**

## 4. Discussion

In our study, we tested seven different scoring functions such as GOLD, Molegro, AutoDock vina (Windows platform), e-HiTS (Linux platform) and PathDock, MEDock (docking servers). Re-scoring was carried out using all the above scoring functions and each molecule was optimized using optimization routine. Post-scoring results are evaluated for RMSD (Root Mean Square Deviation) and found to be within 2A°. 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 [8] was employed to retrieve best compounds. The ranks obtained from each of the individual scoring functions were added to give a rank-sum [Table-3]. 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, Riboflavin, Apigenin-5-o-glucoside and Apigetrin compound conformers are considered as potential ligands against Aldose reductase. The docking scores of the above best compounds in the seven different softwares, generated classes using Tsar software and the sum of the classes for each ligand are shown in Table 3 and Table 4.(Appendix)

From our analysis, it is evident that plant based compound Apigetrin exhibited anti-diabetic activity as it obtained best rank among other compounds and the the major interacting residues are reported in Table-5 and the 2-D image of apigetrin in Figure-1.

Table 5: Number of H-bond interactions and the corresponding interacting residues of apigetrin with active site amino acid residues of aldose reductase.

| Compound | MolDock Score | No. of Interactions | Interacting residues |
|---|---|---|---|
| Apigetrin | -133.388 | 4 | OG - Ser302 |
| | | | NE1 - Trp20 |
| | | | NE2 - Gln49 |
| | | | O - Tyr48 |

## 5. 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-diabetic inhibitors from plants. Compound Apigetirn, in particular, from Cuminum cyminum showed high binding affinity against Aldose reductase, 1AH3. The docked pose of the compound exactly fits into the active site region and the ligand formed more number of H-bond interactions than the co-crystallized ligand. Therefore, this study states the importance of small molecules from various plant sources and their use to enhance protein-ligand interaction studies, in silico.

## 6. References


[1] Expert committee on diagnosis and Classification of Diabetes Mellitus, Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, DIABETES CARE, VOLUME 26, SUPPLEMENT 1, JANUARY 2003

[2] Jalaie M, V Shanmugasundaram 2006 Virtual screening: are we there yet? Mini Rev Med Chem 6:1159-1167

[3] Warren GL, CW Andrews, AM Capelli, B Clarke, J LaLonde, MH Lambert, M Lindvall, N Nevins, SF Semus, S Senger, G Tedesco, ID Wall, JM Woolven, CE Peishoff, MS Head 2006 A critical assessment of docking programs and scoring functions. J Med Chem 49:5912-5931

[4] Waszkowycz B 2008 Towards improving compound selection in structure-based virtual screening. Drug Discov Today 13:219-226

[5] D. K. Gehlhaar, G. Verkhivker, P. A. Rejto, D. B. Fogel, L. J. Fogel, S. T. Freer, Proceedings of the Fourth Annual Conference on Evolutionary Programming. 1995, 615-627

[6] J. M. Yang, C. C. Chen, Proteins. 2004, 55, 288-304

[7] Kitchen DB, H Decornez, JR Furr, J Bajorath 2004 Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 3:935-949

[8] Clark RD, A Strizhev, JM Leonard, JF Blake, JB Matthew 2002 Consensus scoring for ligand/protein interactions. J Mol Graph Model 20:281-295

[9] Wang R, Y Lu, S Wang 2003 Comparative evaluation of 11 scoring functions for molecular docking. J Med Chem 46:2287-2303;
[10] Charifson PS, JJ Corkery, MA Murcko, WP Walters 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

# APPENDIX

Table 3*:* Scores of the top 10 *Cuminum cyminum* compounds obtained from different docking softwares. All values are in kcal/mol

| S.No. | Cuminum compounds | Molegro (kcal/mol) | Ehits (kcal/mol) | Vina (kcal/mol) | Gold (kcal/mol) | MEDock (kcal/mol) | Patchdock (kcal/mol) |
|---|---|---|---|---|---|---|---|
| 1 | Riboflavin | -133.388 | -7.4034 | -7.9 | 29.15 | -9.24 | 4670 |
| 2 | Apigenin-5-o-glucoside | -131.705 | -4.7679 | -8.4 | 58.05 | -7.51 | 4792 |
| 3 | Apigetrin | -130.833 | -6.5608 | -9.3 | 56.89 | -11.61 | 5114 |
| 4 | Apiin | -127.982 | -5.471 | -8.7 | 50.47 | -8.14 | 5748 |
| 5 | Benzyl Cinnamate | -117.458 | -4.7007 | -7.8 | 55.91 | -9.21 | 4476 |
| 6 | Luteolin | -116.643 | -5.5682 | -8.3 | 49.89 | -12.6 | 3864 |
| 7 | Stigmasterol | -116.379 | -2.4912 | -9.2 | 21.18 | -6.13 | 5436 |
| 8 | Cosmosin | -115.701 | -6.1684 | -9.1 | 46.83 | -11.76 | 5090 |
| 9 | Luteolin-7-o-glucoside | -112.54 | -4.8228 | -8.7 | 51.73 | -10.77 | 4946 |
| 10 | Cynaroside | -111.372 | -5.9918 | -8.8 | 52.7 | -11.95 | 4932 |

Table 4: Classes generated using Tsar software.

| S.No. | Compound | Molegro | Ehits | Vina | Gold | MEDock | Patchdock | Sum |
|---|---|---|---|---|---|---|---|---|
| 1 | Riboflavin | 4 | 4 | 1 | 1 | 2 | 2 | 14 |
| 2 | Apigenin-5-o-glucoside | 4 | 2 | 2 | 4 | 1 | 2 | 15 |
| 3 | Apigetrin | 4 | 4 | 4 | 4 | 4 | 3 | 23 |
| 4 | Apiin | 4 | 3 | 3 | 4 | 2 | 4 | 20 |
| 5 | Benzyl Cinnamate | 2 | 2 | 1 | 4 | 2 | 2 | 13 |
| 6 | Luteolin | 1 | 3 | 2 | 4 | 4 | 1 | 15 |
| 7 | Stigmasterol | 1 | 1 | 4 | 1 | 1 | 4 | 12 |
| 8 | Cosmosin | 1 | 3 | 4 | 3 | 4 | 3 | 18 |
| 9 | Luteolin-7-o-glucoside | 1 | 2 | 3 | 4 | 3 | 3 | 16 |
| 10 | Cynaroside | 1 | 3 | 3 | 4 | 4 | 3 | 18 |