**International Journal for Computational Biology (IJCB)**
**Vol.3, No.2, August 2014, pp. 10-17**
**ISSN: 2278-8115** 

# Screening of Bioactive Compounds against Nonreceptor Fyn Kinase: Virtual Screening and Network Approach

Sudharsana Sundarrajan*, Thabitha Amalraj, Sweta Kumari, Sajitha Lulu, MohanaPriya Arumugam

Department of Bioinformatics, School of Biosciences and Technology, VIT University, Vellore 632014, India

**Article Info**

***Article history:***

Received May 5th, 2014
Revised Jul 20th, 2014
Accepted Jul 24th, 2014

***Keyword:***

Fyn kinases
Src kinase
Oncogenesis
Ligand efficiency
Reverse virtual screening
Feature score
Drug-target network

**ABSTRACT**

Tyrosine phosphorylation is a key controlling mechanism in signal transduction and enzyme activity regulation. Dysfunction of Fyn kinase, a unique member of non-receptor Src kinase family is implicated in oncogenesis, T-cell mediated diseases and neuronal disorders. Fyn kinase has been recognized as an important target for anti-cancer therapeutics. An insilico virtual screening of open and closed states of Fyn with seventy phytochemicals used in cancer treatment was carried out. Molecular properties and bioactive spectrum analysis further improved the screening process by forming a data set containing seven potential hits. Ligand efficiency score which combines biological and chemical space together identified three secondary metabolites apigenin, genistein and quercetin as efficient inhibitors of Fyn kinase. A reverse virtual screening approach validated the target selection by identifying Src kinase family members as potential drug targets. Drug-target interaction network based on feature scores of 22 phytochemicals which survived the initial screening process further validated our findings. A conceptual optimization may be required to reduce attrition and increase the activity of the lead.



***Corresponding Author:***

Sudharsana Sundarrajan,
Department of Bioinformatics,
School of Biosciences and
Technology, VIT University,
Vellore. India.
Email: s.sudharsana@vit.ac.in

***How to Cite:***

Sudharsana Sundarrajan *et. al.* Screening of Bioactive Compounds against Nonreceptor Fyn Kinase: Virtual screening and Network approach. IJCB. 2014; Volume 3 (Issue 2): Page 10-17.

## 1. INTRODUCTION

Tyrosine kinases (TK) are crucial mediators in several signaling cascades acting as molecular switches [1]. Tyrosine kinase signaling pathways normally prevent deregulated proliferation or contribute to sensitivity towards apoptotic stimuli. These signaling pathways are often genetically or epigenetically altered in cancer cells to impart a selection advantage to the cancer cells [2].

The dysregulation of TKs occur in many different ways in cancer cells. The first mechanism is by fusion of TKs with a partner protein, a consequence of chromosomal translocation. Second mechanism of impairment is by mutations degrading the kinase auto-regulation process. Other mechanisms include over-expression of their ligands and decrease in factors that limit their activity. All these aberrant actions may lead to increase angiogenesis, invasiveness and drug resistance in malignant cells [1]. Two major classes of TKs have been reported [3]. Receptor tyrosine kinases (EC 2.7.10.1) are transmembrane proteins with ligand-binding domain and catalytic domain. The catalytic activity is triggered by ligand binding, after which the receptor oligomerizes, breaks the cytoplasmic juxtamembrane interactions and autophosphorylates a regulatory tyrosine residue within the kinase activation domain. Upon activation, auto-phosphorylation generates binding sites for signaling proteins, recruiting them to the membrane and activating multiple signaling cascades. Similarly non-receptor

kinases (NTKs) (EC 2.7.10.2) upon activation, cleaves the signal inhibitors, oligomerizes and autophosphorylates a regulating tyrosine residue leading to signal transduction. In humans, about 20 families of receptor tyrosine kinases and 10 families of non- receptor tyrosine kinases have been reported [4].

Fyn is a member of the nonreceptor Src-family tyrosine kinases. Other members of Src-kinase family include Lck, Src, Yes, Lyn, Hck, Blk, Fgr and Yrk tyrosine kinase [5]. Fyn kinase actively participates in T-cell signal transduction. It is activated by the stimulation of T-cell antigen receptors. ITAM sequences (immunoreceptor tyrosine activation motif) of the CD3 subunit and ζ-chain of T-cell receptors are then phosphorylated by Fyn. This is the prerequisite step for ZAP-70 kinase binding to ITAM and phosphorylation of LAT18 and SLP-76 adaptor proteins that enable further downstream signaling pathways. This signaling process results in the proliferation and differentiation of T-cells. Fyn resides in specialized plasma membrane micro domains, termed as lipid rafts. Genetic evidence demonstrates that Fyn activation is strictly connected with T-cell receptor-induced translocation of Lck. Activated Fyn plays a critical role in skin cancer development. In keratinocyte cell-cell adhesion Rho/PKR2 signaling, Fyn is reported as a downstream mediator [4].

Crystallographic and biochemical studies reveal that kinase activity is regulated through reversible phosphorylation of two key tyrosine residues. In closed conformation Fyn kinase remains inactive by two week intra-molecular interactions formed between SH2 domain and phosphorylated C-terminal tyrosine (TYR531) and SH3 domain and the linker region. Dephosphorylation of the C-terminal tyrosine mediated by $CD_{45}$ results in open conformation. The catalytic activity is conferred by phosphorylation of regulatory tyrosine (TYR420) in the activation loop [6].

Our study aims at identifying inhibitors of plant origin against Fyn kinase by reverse virtual screening approach. Insilico docking based virtual screening aids the discovery of potent hits. Lipinski filter and bioactive spectrum predictions are employed to identify more potent leads. A Drug-target network is constructed to identify compounds interacting with common partners. A more stringent screening strategy adopted by employing ligand efficiency score identified the final potential lead compounds.

## 2. RESEARCH METHOD

### 2.1 Non receptor kinase Fyn as a therapeutic target

Discovery of small molecular drug candidates typically begins with the identification and validation of the target. In this study we consider Fyn kinase as a potential target because of its greater impact in mediating diseases such as cancer, allergy and Alzheimer's. Fyn is a proto-oncogene tyrosine-protein kinase ($p59^{fyn}$) classified as a phosphotransferase. Fyn kinase sequence was retrieved from UniProt database (UniProt ID: P06241). Closed and open structures of Fyn kinase were modeled by homology modeling using modeller 9.7 [7]. In order to identify homologous sequences with known 3D structures, a protein-protein BLAST was performed against protein data bank (PDB) [8]. Models were built with phosphorylated side chains using 1Y57, 2DQ7 and 2H8H as templates. **Table 1** represents the detailed information ranging from template identification to model validation. The backbone dihedral angle distributions of all amino acids were verified using Procheck Ramachandran plot [9]. The fold quality was inspected by ProSA web server [10] and the predicted model quality was verified by QMEAN score [11].

**Table 1.** Template and structure validation statistics for the modeled protein

| Protein | Template PDB ID | Query coverage (%) | Identity (%) | Procheck | | | | ProSA Z score[5] | QMEAN score[6] |
|---|---|---|---|---|---|---|---|---|---|
| | | | | A[1] | B[2] | C[3] | D[4] | | |
| Fyn open | 1Y57 | 99 | 77 | 91.6 | -7.11 | 6.1 | 1.5 | -9.77 | 0.736 |
| | 2DQ7 | 61 | 99 | | | | | | |
| Fyn Close | 2H8H | 99 | 77 | 92.3 | 5.9 | 1.3 | 0.5 | -10.07 | 0.749 |

[1]Percentage of residues in most favored region, [2] Percentage of residues in additionally allowed region, [3]Percentage of residues in generously allowed region, [4]Percentage of residues in disallowed region, [5]Overall model quality and [6]Model quality estimation.

### 2.2 Identification of bioactive compounds

Phytocompounds offer wide scope of cure for many inherent diseases. The high-throughput experimental findings of anti-cancer bioactive compounds were manually curated from three different databases: PubChem [12], DIACAN [13] and cancer plants database (http://cancerplantsdatabase.com). The selection of phytochemicals of our work is restricted to compounds whose activity has been proved against various cancer cell lines. To give an emphasis on Indian medicinal plants, seventy compounds of Indian origin were selected for the study. The structures of all the compounds were retrieved from PubChem database. **S.Table. 1** presents the source of phytochemical compounds of our work.

### 2.3 Virtual screening

Virtual screening offers computational prediction of binding affinity for diverse set of compounds [14]. A high throughput docking study was carried out using AutoDock Vina tool [15]. AutoDock 4.2 [16] was used to prepare the initial files. Hydrogen atoms were added, non-polar contacts were merged and Gastegier charges were computed for the proteins. Open and closed conformations of Fyn were processed separately. The ligand files were prepared using AutoDock Raccoon tool [17]. The search space of was set for the protein allowing the ligands to choose their preferential binding site.

### 2.4 Rule of five and bioactive spectrum prediction

Lipinski rule of five (RO5) aids in identifying chemical entities with favorable physiochemical properties. The compounds were subjected to RO5 analysis using MolInspiration server (http://www.molinspiration.com). Bioactive spectrum of the compounds determines their potential roles based on their signature pharmacophore. The probable activity score towards kinase inhibition was calculated using PASS server [18].

### 2.5 Network Building

Network analysis interprets the complex relationship existing between the drug and the target. The drug-target network was established by connecting the potential compounds and their potential target. PharmMapper server [19] employs principles of Pharmacophore mapping to spot potential molecular target candidates for a known small bioactive molecule. The database was searched for human targets with a threshold of 1000 hits. The final list was scrutinized to obtain only NTKs from the predicted results using Perl script. A drug – target network was constructed and visualized using Cytoscape 2.7 [20]. In the graphic network, molecular species (compounds and proteins) are represented as nodes and inter-molecular connections (drug-target) are indicated as edges between the nodes.

## 3. RESULTS AND ANALYSIS

### 3.1 Fyn in open and closed conformations – Structural and functional uniqueness

The three-dimensional structure of Fyn tyrosine kinase has not been reported previously, with the exception of few crystal structures of Fyn SH3 or SH2 domains. Human Fyn kinase contains 537 amino acids. It has a unique structure comprising an N-terminal region anchoring to the membrane, SH4 domain, SH3 and SH2 domains mediating interactions with protein partners, SH1 kinase domain and a C-terminal tail involved in negative regulatory function [21]. The functional uniqueness is maintained by acting as a molecular switch turning itself on and off through phosphorylation of either Tyr420 or Tyr531 residue. The phosphorylation status of Tyr531 in the C-terminal tail determines the open and closed conformation of Fyn. In closed state Tyr531 is phosphorylated and conversely unphosphorylation keeps Fyn in open state and it can be activated by phosphorylation at Tyr420. The open state of Fyn forms the basis for kinase activity. Hence Fyn open structure with phosphorylated Tyr420 and closed structure with phosphorylated Tyr531 have been modeled.

### 3.2 Model building and validation

The BLASTP search identified several homologous sequences with known 3D protein structures. Based on the query coverage and sequence identity, 2 templates for open conformation of Fyn (PDB ID: 1Y57, 2DQ7) and a single template (PDB ID: 2H8H) for closed conformation were chosen. In open state Tyr420 and in closed state Tyr531 were phosphorylated. The SH3 Kinase domain and C-terminal tail plays a major role in the enzyme activation and hence the first 85 residues corresponding to N terminus region were excluded from model building (**Figure. 1**). The secondary structural profile and accessibility were estimated by Ramachandran plot using Procheck server. The open conformation has 95.5% of residues in the allowed region and the closed conformation has 99.2% of in the allowed region. This shows that the models generated have good stereochemical features. The fold quality assessed by ProSA web server produced a Z score of -9.77 for open and -10.07 for close conformation validating the model produced. The QMEAN score of the closed and open Fyn models are 0.75 and 0.737 as predicted by the Swissmodel server. The QMEAN computes a composite scoring from global (i.e. for the entire structure) and local (i.e. per residue) error estimates on the basis of a single model. The QMEAN score falls within the interval of 0 to 1 and hence the high accuracy of the models is proven.

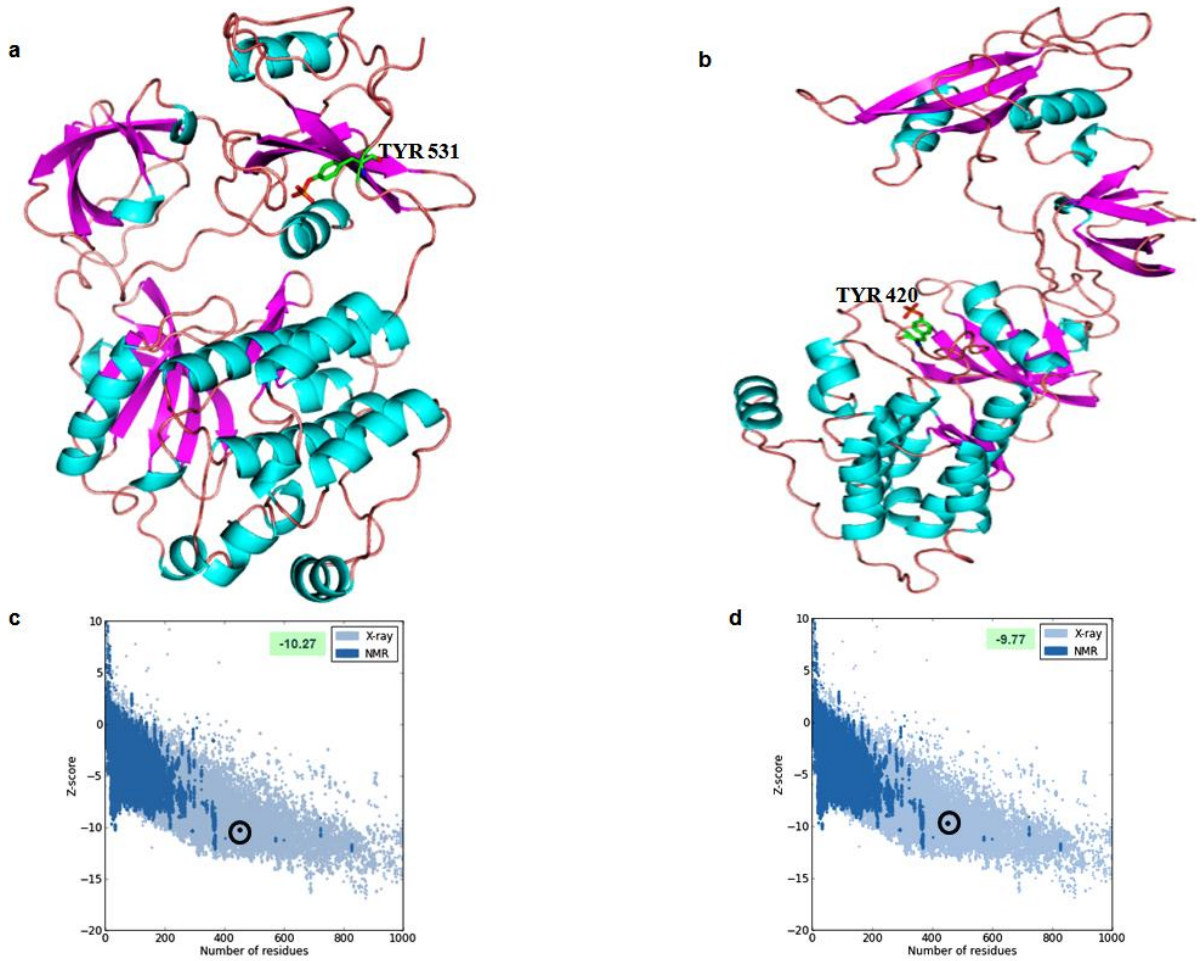


**Figure. 1** Fyn kinase modeled structures and their Z-score plots **(a)** Cartoon representation of Fyn kinase in close state with phosphorylated Tyr531 shown as sticks (**b)** Cartoon representation of Fyn kinase in open state with phosphorylated Tyr420 shown as sticks **(c)** Z-score plot of Fyn closed conformation (**d)** Z-Score plot of Fyn open conformation. The circle indicates the position of predicted model.

### 3.3 Hit identification - Virtual screening, physiochemical properties and bioactive spectrum filters

Hit identification ideally results in set of compounds that interact with the target. We have applied high throughput virtual screening method to identify compounds with favorable binding affinity towards open and closed conformation of Fyn kinase. The ligands with high binding energy ranging from -10 to -8 Kcal/mol were selected for next step of processing. 35 compounds for open conformation and 25 compounds for closed conformation were selected based on their binding energy values. 22 compounds which showed high preference in both close and open states formed the dataset for next step of screening process.

The physiochemical properties identify compounds which have futuristic lead like potential [22]. Optimizing the binding potency of a ligand independent of the size of the ligand to improve large scale screening experiments is of adverse need. One such measure is ligand efficiency (LE) [23] which is calculated using the formula described in equation (1).

**LE = affinity /N ------- (1)**

The affinity corresponds to compounds binding energy and N is the number of non-hydrogen atoms. Molecular weight of the compounds plays a major governing role in ligand efficiency. Further to the extension of screening strategy kinase inhibition scores predicted by PASS server eliminated compounds with minimum (< 0.4) kinase inhibitory scores. **Table 2** presents the parameter scores of aloe-Emodin, apigenin, brazilin, daidzein, genistein, rhein and quercetin identified as fruitful hits from the above screening process.

**Table 2** Physiochemical properties, biospectrum and ligand efficiency scores of final seven compounds

| Sl.No | Phytochemicals | milogP | Molecular Weight | nON[1] | nOHNH[2] | Kinase inhibition P[a] score[3] | Ligand efficiency |
|---|---|---|---|---|---|---|---|
| 1. | Aloe-Emodin | 2.424 | 270.24 | 5 | 3 | 0.441 | -0.46 |
| 2. | Apigenin | 2.463 | 270.24 | 5 | 3 | 0.941 | -0.45 |
| 3. | Brazilin | 1.285 | 286.283 | 5 | 4 | 0.507 | -0.42 |
| 4. | Daidzein | 2.559 | 254.241 | 4 | 2 | 0.722 | -0.47 |
| 5. | Genistein | 2.268 | 270.24 | 5 | 3 | 0.844 | -0.45 |
| 6. | Quercetin | 1.683 | 302.238 | 7 | 5 | 0.957 | -0.4 |
| 7. | Rhein | 2.997 | 284.223 | 6 | 3 | 0.468 | -0.41 |

[1]Hydrogen bond acceptor, [2]Hydrogen bond donor, [3]Probability to be active

### 3.4 Lead identification

Biological activity spectrum reflects the interaction of chemical entity with biological entity. The predicted Pa falling above 0.7 indicates high probability of the substance being analogue of a known pharmaceutical agent. The compounds apigenin (Pa - 0.941), genistein (Pa - 0.844) and quercetin (Pa - 0.957) were observed to show greater predicted kinase inhibition activity. Size related efficiency indices are currently used for selecting leads for fragment based drug designing. Ligand Efficiency of -0.40 was set as minimum gain for the compounds to evolve from 'lead-like' to 'drug-like'. The ligand efficiency score for apigenin (-0.45), genistein (-0.45) and quercetin (-0.40) were observed to fall above the desired range. **Figure. 2** depicts a distribution plot of key properties playing a major role in lead identification.

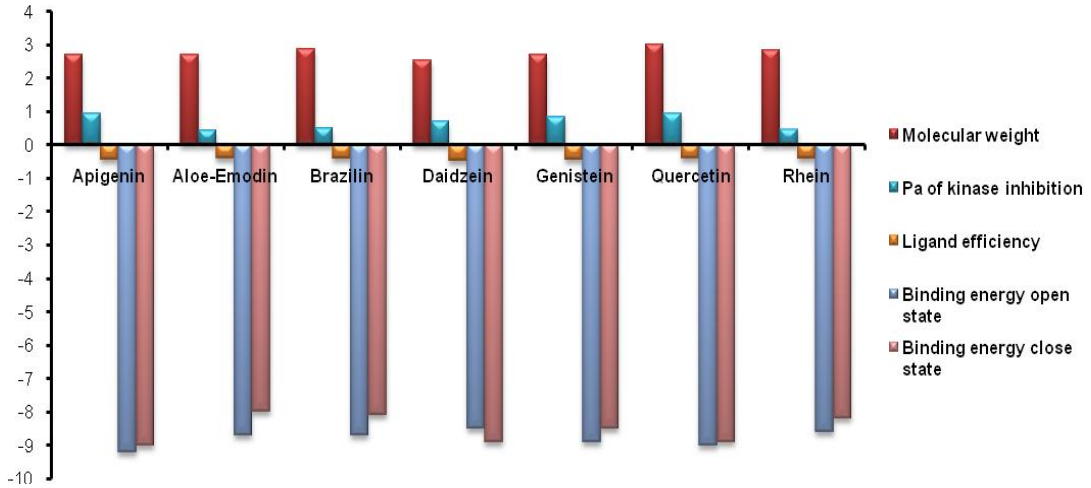


**Figure. 2** Screening parameter values of final seven hits. (Molecular weight is scaled by multiplying $10^2$).

### 3.5 Lead validation – A networks approach

The study further validated the reliability of the candidate target and compounds using network approach. A detailed analysis about the connections between the compounds and targets were carried out using network pharmacology approach. The drug-target network [24] computed based on fit score and matching pharmacophore features validated the above results. Successors of initial screening process were reported to interact with many members of NTKs. Upon close observation of the final network revels that apigenin, genistein and quercetin interact with all the members of SRC kinase family. By homology we confirm that these compounds can act as potential inhibitors of Fyn kinase. **Figure. 3** portrays the drug-target network of 22 compounds identified from initial screening and also their interaction with NTKs.

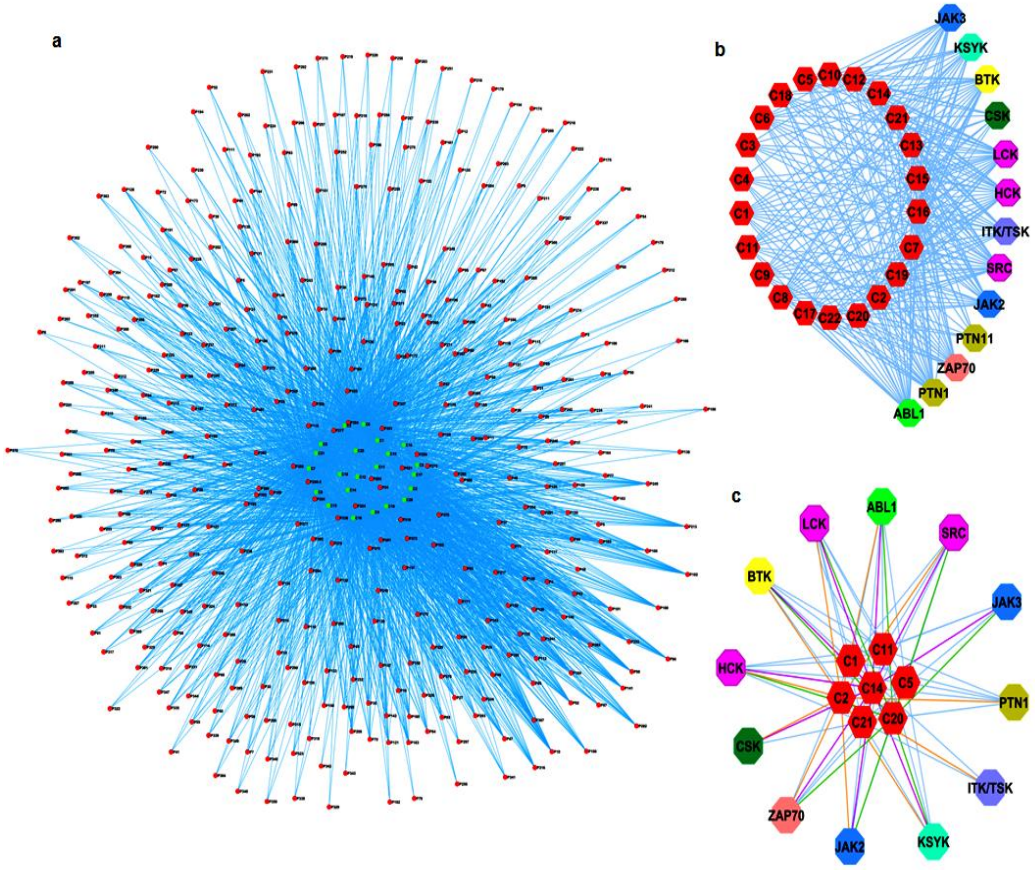


**Figure. 3** Drug- Target Networks predicted using Cytoscape 2.7 **(a)** Drug target network of proteins with 22 compounds. Red circles indicate the proteins and green circles indicate the compounds. **(b**) Interaction network of NTKs with the initial survivors. **(c)** Interaction network of final seven hits with NTKs. (see **S.Table 2** for compound names).

### 3.6 Binding mode analysis of lead compounds

Several groups reported the crystal structure of SH3 and SH2 domain of Fyn kinase. Based on their observation an ATP binding domain is observed opposite to SH2 and SH3 domains. A 28 residue long (408····436) segment forms an activation center of Fyn kinase domain. Asp408 residue chelates magnesium ion and positions the phosphate for phosphor-transfer. An ATP binding site (280····420) is also included in this activation loop. Signature residues like Glu343, Met345 and Asp408 which were reported to show interaction with the drug staurosporine and ADP (Dubravko Jelic et al., 2007) were also observed to interact with our top three leads. The final successful leads interact with residues in kinase activation loop thereby promising efficient inhibition. The target-ligand interaction details are summarized in **Table 3**.

**Table 3** Binding energy and interacting residues of top three lead compounds

| Phytochemical | Binding Energy Kcal/mol | | Interacting residues in open conformation | Interacting residues in closed conformation |
|---|---|---|---|---|
| | Open state | Close state | | |
| Apigenin | -9.2 | -8.6 | Lys299, Glu343, Met345 | Lys295, Tyr330, Asn400, Lys405 (+-π) |
| Quercetin | -9.0 | -8.0 | Glu314, Thr342(σ-π), Tyr344(π-π), Met345, Asp408 | Tyr330, Glu343, Asn400 |
| Genistein | -8.9 | -8.5 | Lys299,Tyr344(π-π), Met345,Asp408 | Lys299, Thr330, Arg392, Asp408 |

The binding groove and interaction pattern of apigenin with Fyn kinase is illustrated in **Figure. 4**.

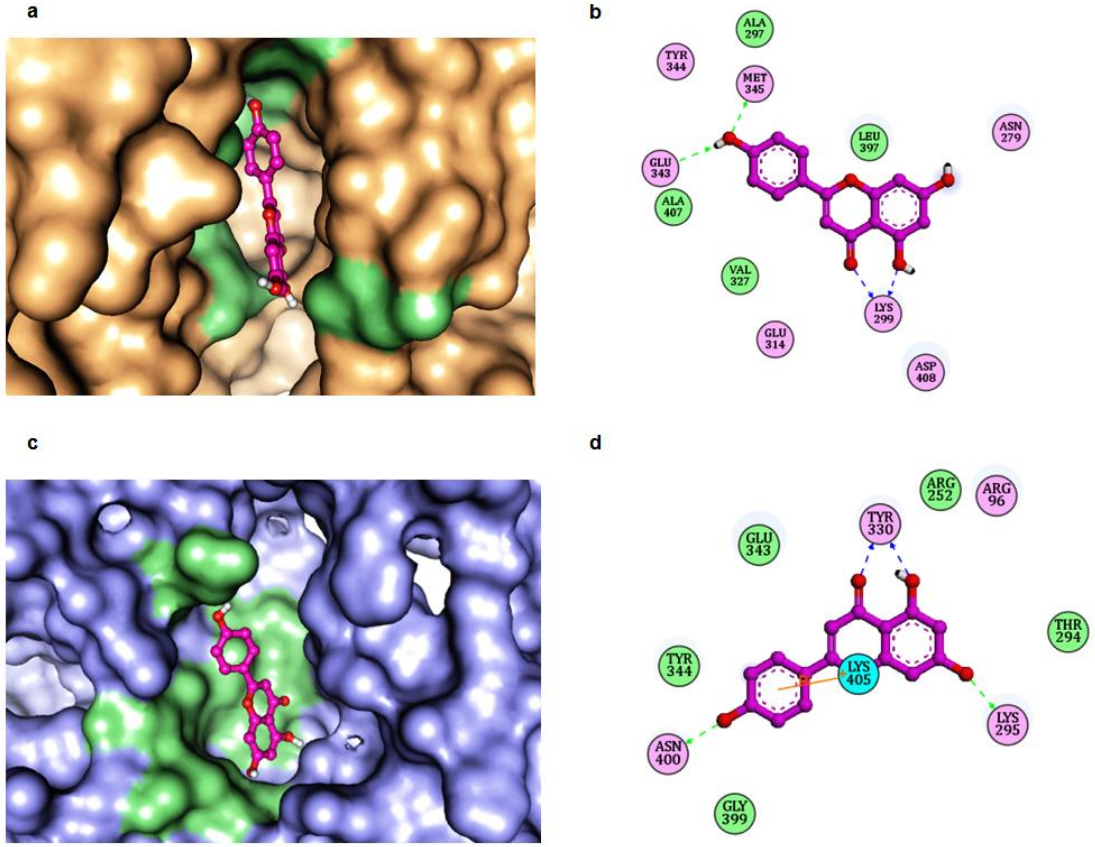


**Figure. 4** Binding pose of Fyn kinase with apigenin in open and close state **(a)** The open state Fyn is represented as surface and the bound apigenin is represented as ball and sticks. The crucial binding site residues are coloured green. Binding of apigenin represented in ball and stick with open Fyn represented in surface. The binding site residues are coloured green. **(b**) 2D ligand plot showing the bonding patterns and interacting residues of apigenin with open Fyn. **(c)** The closed state Fyn is represented as surface and the bound apigenin is represented as ball and sticks. The crucial binding site residues are coloured green. Binding of apigenin represented in ball and stick with open Fyn represented in surface. The binding site residues are coloured green. **(d)** 2D ligand plot depicting interacting residues of closed Fyn with apigenin. The green and blue arrows represent hydrogen bondings and the orange line indicates ca-pi interaction.

## 4. CONCLUSION

The current analysis provides a brief scenario of Fyn kinase structure and function. The virtual screening process in search of potential leads against Fyn kinase identified compounds of high potency. The filters such as physiochemical properties, bio-active spectrum and ligand efficiency provided a numerical framework combining the potency towards the biological space (target) and chemical space (ligands). These numerical systems will facilitate drug discovery process by minimizing the cost and time to narrow down more potent and efficient leads.

## ACKNOWLEDGEMENTS

We acknowledge VIT University for computing facilities.

## SUPPLEMENTARY FILES

1. Tables showing the data used in this study

## REFERENCES


[1] Daniela SK, Richard AV., Tyrosine Kinases as Targets for Cancer Therapy, N ENGL J Med, 2005; 353: 172-187

[2] Manash KP, Anup KM., Tyrosine kinases – Role and Significance in Cancer. Int. J. Med. Sci, 2004; 1: 101-115

[3] Dan RR, Yi-Mi WU, Su-Fang L., The Protein Tyrosine Kinase Family of the Human Genome. Oncogene, 2000; 19: 5548-5557

[4] Marilyn DR., Fyn, A SRC Family Tyrosine Kinase. Int J Biochem Cell Biol, 1998; 30: 1159-1162

[5] Evan I., Src Family Kinases: Regulation of their activities, levels and identification of new pathways. Biochim Biophys Acta, 2008; 1784: 56–65

[6] Dominik F, Behrouz M, Alessandra F, Kirishanthy K, Jenny Z, Ondrej B, Dominique D, Andre V, Michael J., Lck-dependent Fyn Activation Requires C Terminus-dependent Targeting of Kinase-active Lck to Lipid Rafts. J Biol Chem, 2008; 283: 26409–26422

[7] Sali A, Blundell TL., Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol, 1993; 234: 779-815.

[8] Stephen FA, Warren G, Webb M, Eugene WM, David JL., Basic local alignment search tool. J. Mol. Biol, 1990; 215: 403-410

[9] Laskowski RA, MacArthur MW, Moss DS, Thornton JM., PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst, 1993; 26: 283-291

[10] Markus W, Manfred JS., ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res, 2007; 35: W407–W410

[11] Pascal B, Michael K, Torsten S., QMEAN server for protein model quality estimation. Nucleic Acids Res, 2009; 1: 510-W514

[12] Bolton E, Wang Y, Thiessen PA, Bryant SH., PubChem: Integrated Platform of Small Molecules and Biological Activities. Chapter 12 IN Annual Reports in Computational Chemistry, vol 4, American Chemical Society, Washington, DC

[13] Priyanka J, Vipin AM, Silpa S, Priya AP, Parvathi S, Raghunath K, Puthiyaveetil AN., DIACAN: Integrated Database for Antidiabetic and Anticancer medicinal Plants. Bioinformation, 2013; 9: 941-943

[14] Ingo M, Scott O.,Advances in virtual screening. Drug Discov Today: Insilico techniques, 2006; 3: 405-411

[15] Trott O, Olson AJ., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem, 2010; 31: 455-461

[16] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ., Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. J. Comput Chem, 2009; 16: 2785-2791.

[17] Stefano F., Raccoon|AutoDock VS: an automated tool for preparing AutoDock virtual screenings. http://autodock.scripps.edu/resources/raccoon. 2010

[18] Filimonov D, Poroikov V., Bioactive compound design: Possibilities for industrial use. PASS: Computerized prediction of biological activity spectra for chemical substances. BIOS Scientific Publishers, 1996;44-56

[19] Liu X, Ouyang S, Yu B, Liu Y, Huang K, Gong J, Zheng S, Li Z, Li H, Jiang H., PharmMapper server: a web server for potential drug target identification using pharmacophore mapping approach. Nucleic Acids Res, 2010; 1: W609-W614

[20] Paul S, Andrew M, Owen O, Nitin SB, Jonathan TW, Daniel R, Nada A, Benno S, Trey I., Cytoscape: a software for integrated models of biomolecular interaction networks. Genome Res, 2001; 13: 2498-2504

[21] Dubravko J, Boris M, Sanja K, Krunoslav N, Donatalla V, Ognjen C, Roberto A, Wolfganag B., Homology modeling of human Fyn Kinase structure: Discovery of Rosmarinic acid as a new Fyn Kinase inhibitor and in silico study of its possible binding modes. J. Med. Chem, 2007; 50: 1090-1100

[22] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev, 2001; 46: 3–26

[23] Cele AZ., Ligand efficiency indicies for effective drug discovery. Expert Opin. Drug Discov, 2007; 2:469-488

[24] Peter C, Tamas K, Huba J.M, Gabor L, Ruth N., Structure and dynamics of molecular networks: A novel paradigm of drug discovery a comprehensive review. Pharmacol Ther, 2013; 138:333–408.