Identification of selective Lyn inhibitors from the chemical databases through integrated molecular modelling approaches

V.V. Shetve[a*], S. Bhowmick[b*], S.A. Alissa[c], Z.A. Alothman[d], S.M. Wabaidur[d], F.A. Alasmary[d], H.M Alhajri[d] and M.A. Islam [e,f,g]

[a]Department of Bioinformatics, Rajiv Gandhi Institute of IT and Biotechnology, Bharati Vidyapeeth Deemed University, Pune, India;
[b]Department of Chemical Technology, University of Calcutta, Kolkata, India;
[c]Chemistry Department, College of Science, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia;
[d]Department of Chemistry, College of Science, King Saud University, Riyadh, Saudi Arabia; [e]Division of Pharmacy and Optometry, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK; [f]School of Health Sciences, University of Kwazulu-Natal, Durban, South Africa; [g]Department of Chemical Pathology, Faculty of Health Sciences, University of Pretoria and National Health Laboratory Service Tshwane Academic Division, Pretoria, South Africa

CONTACT M.A. Islam ataul.islam80@gmail.com
*These authors contributed equally to this work.

ABSTRACT

In the current study, the Asinex and ChEBI databases were virtually screened for the identification of potential Lyn protein inhibitors. Therefore, a multi-steps molecular docking study was carried out using the VSW utility tool embedded in Maestro user interface of the Schrödinger suite. On initial screening, molecules having a higher XP-docking score and binding free energy compared to Staurosporin were considered for further assessment. Based on in silico pharmacokinetic analysis and a common-feature pharmacophore mapping model developed from the Staurosporin, four molecules were proposed as promising Lyn inhibitors. The binding interactions of all proposed Lyn inhibitors revealed strong ligand efficiency in terms of energy score obtained in molecular modelling analyses. Furthermore, the dynamic behaviour of each molecule in association with the Lyn protein-bound state was assessed through an all-atoms molecular dynamics (MD) simulation study. MD simulation analyses were confirmed with notable intermolecular interactions and consistent stability for the Lyn protein-ligand complexes throughout the simulation. High negative binding free energy of identified four compounds calculated through MM-PBSA approach demonstrated a strong binding affinity towards the Lyn protein. Hence, the proposed compounds might be taken forward as potential next-generation Lyn kinase inhibitors for managing numerous Lyn associated diseases or health complications after experimental validation.

KEYWORDS: Lyn protein; virtual screening; molecular docking; molecular dynamics; MM-PBSA

Introduction

Lyn protein is an important member of the Src family and belongs to the intracellular membrane-associated tyrosine kinases. It is essentially known to be acting as signalling intermediaries to regulate various cellular outcomes like proliferation, differentiation, apoptosis, immune responses, adhesion, metabolism and migration [1]. It has been reported that Lyn protein can mediate both negative and positive signalling processes for regulating a number of important cellular/biological processes starting from cellular growth to trigger the cellular immune responses [2–4]. Most importantly, Lyn acts as a specific type of enzymes that can regulate information transfer in both ways by turning on and off kinase activity, depending on the types of cells it is working on and information input [1]. For example,

Lyn protein acts as a positive regulator of myofibroblast migration, proliferation, and collagen production, and therefore inhibition of Lyn kinase activity can be the potential target for preventing fibrosis by means of restricting the synthesis of procollagen and collagen [5]. Beside negative regulation or activation functions, Lyn protein also plays a crucial role in immune self-tolerance by acting on downstream of several immune receptors, including the B-cell receptor, toll-like receptors (e.g. TLR2 and TLR4) and many clusters of differentiation (e.g. CD79A, CD79B, CD5, CD19 and CD22) transmembrane proteins [6]. In particular, Lyn protein has a significant contribution in tightly regulating signalling pathways that are dysregulated in autoimmunity. Beyond above, Lyn is also greatly expressed in a variety of cells or tissues such as hemato-poietic, epidermoid, and neuronal cells, and helps in executing necessary signal transduction mechanism at the cytoplasmic side of the plasma membrane [7–10]. However, mutations in Lyn and overexpression of Lyn kinase activity trigger aggressive behaviour that warrants leading causes for at least more than 15 cancer types including lung cell carcinoma, colon, renal and ovarian, prostate, breast, liver cancer, etc. [11–15]. Moreover, the aberrant function of Lyn is also fundamentally associated with several serious pathophysiological conditions including various forms of other cancers like acute myeloid leukaemia (AML), chronic myeloid leukaemia (CML), melanoma and certain solid tumours [16–18]. Intriguingly, Lyn protein also plays several important key roles as intracellular signalling proteins or interactor protein to form a multi-protein complex for regulating cell proliferation and differentiation, and cell death [19,20]. Henceforth, as a key kinase protein, Lyn can be 'druggable' and might deliver a therapeutic opportunity for many diseases including cancer where Lyn protein expressions are explosively studied [21–23].

Based on pieces of evidence from a number of research outcomes, it is now quite clear that dysregulated Lyn kinase activity strongly associated with the ample number of human diseases or health complications. Hence, the Lyn protein kinase inhibition becomes an active area of drug development and scientific progress. Therefore, the present research objective specifically has focused on the identification of new small molecules directed towards the Lyn kinase protein inhibition for beneficial health allied therapeutic applications. A number of significant advancements have already been made for inhibition of Src-family kinases including Lyn kinase with small molecule inhibitors in the past few decades. However, effective therapeutic agents highly specific for inhibiting the Lyn kinase protein are still commercially absent. As of now, the major development of highly selective kinase inhibitors focused on targeting the conserved ATP-binding site of kinases to bind with small molecule inhibitors [24–26]. In that similar perspective, this study has also hypothesized to rationally identifying specific molecular determinants present in a chemical entity which specifically able to recognize catalytic domain of Lyn protein and could demonstrate exquisite selectivity for most likely to interact with an ATP-binding pocket.

Therefore, in the current work, an attempt was rigorously employed and analysed a state-of-the-art virtual screening scheme to find novel and potent chemical therapeutics that can be capable to interact with Lyn kinase domain, and thus Lyn kinase inhibition can be achieved to manage the several diseases or health complications such as cancers, autoimmune diseases, asthma and psoriasis, etc. [1]. Particularly adopted virtual screening strategy includes multi-step molecular docking analyses, molecular dynamics (MD) simulation studies and in silico pharmacokinetic assessment for the identified small molecules, and also analysed ligand-binding free energy estimations using Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) [27] method. The above-mentioned advanced and modern cheminformatic methodologies were applied to screen out a few potential drugs like small molecules from ~0.7 million compounds belonging to two highly known chemical library databases viz. Asinex [28] and Chemical Entities of Biological Interest (ChEBI) [29] chemical library. Moreover, a conventional state-of-the-art in silico approach was used to dig out the molecular mechanisms of action for finally selecting four compounds which demonstrated relatively stable and better binding interaction with Lyn kinase domain. Additionally, the bioactivity and ligand efficiency parameters of proposed Lyn

protein inhibitors were estimated using binding energy scores which also revealed acceptable and strong ligand potentiality for exerting biological modulation. The employed molecular docking-based virtual screening followed by the assessment of the ligands may give an insight for encompassing and investigating the developmental process to find new therapeutic measures for Lyn protein-associated therapeutic applications.

Materials and methods

In silico studies are being progressively used for the development of novel drugs or identification of drug-like potential candidate molecules for various disease states. Most importantly, virtual screenings (VS) of small molecular databases are getting continual attention for effectively and rationally identifying lead molecules. Herein, to identify potential small molecules that can be pharmacologically and biologically active and at the same time capable to interact with Lyn protein for modulating its role was extensively studied through applications of in silico VS techniques. Therefore, large virtual compound libraries were extensively filtered by employing several advanced levels of computational screening procedures such as multi-step molecular docking combined with MM-GBSA analyses, in silico pharmacokinetic analysis, and mapping on common pharmacophoric features of reference compound for hierarchically reducing the number of candidate molecules to a smaller sub-set of potential candidate molecules. Finally, molecular interaction stability was investigated through molecular dynamics simulation studies and MM-PBSA-based protein-ligand binding free energy calculations. More precisely, a total of 687,697 compounds were retrieved from Asinex and ChEBI screening library databases those used for virtual screening purposes. The Asinex database consisting of the natural product-like chemical compounds with mainly polar functional groups and suitable for exploration hit-to-lead as well as fragment-based drug design (FBDD), and structure-based drug design (SBDD), etc. On the other hand, ChEBI database contains separately distinguishable molecular entities either the product of nature or synthetic product. In this study, primarily molecules from both databases were screened out through a multi-step docking protocol using the 'Virtual Screening Workflow' (VSW) [30] in Maestro user interface of Schrödinger suite. On the other hand, a common feature pharmacophore model for standard ligand was generated using Discovery Studio [31], and then, it was used to fit all newly identified screened hit compounds in that model. More detailed methodologies of each employed technique are explained in the subsequent sections.

Preparation of ligand compounds and protein crystal structure

All downloaded compounds from both chemical library databases (Asinex and ChEBI) were carefully and correctly prepared before to be subjected to a stepwise docking study. Initially, using the Discovery Studio [31], all two-dimensional (2D) compounds in structural data format (.sdf) were checked to remove duplicate molecules, repaired the inappropriate valence compounds, and finally, the 3D coordinates were generated. Finally, from the above steps, 535,437 compounds were retained for further analysis. Additionally, one known potential Lyn kinase domain inhibitor i.e. Staurosporine [32] was used as a control compound in the current study for the assessment of outcomes. Finally, all cleaned compounds from both chemical library databases and Staurosporine were prepared using the 'LigPrep' [33] utility tool in Maestro interface following standard procedure which allowed to generate a maximum of 32 stereoisomers. Particularly, utilizing the Epik [34] tool, protonation states were generated at physiological pH of 7.4 for all compounds.

On the other hand, the crystal structure of the protein Lyn tyrosine kinase domain was retrieved from RCSB-Protein Data Bank (PDB) [35] (PDB ID: 3A4O [32]). The 'Protein Preparation Wizard' [36,37] tool of the Schrödinger suite was used to prepare the protein crystal structure for molecular docking study. All missing side and backbone chains were included during the protein preparation process. Hydrogen

atoms were added to the crystal protein. Water molecules, co-factors were removed from the typical PDB structure. Missing information on connectivity was corrected with the appropriate assignment of bond orders, formal charges and by capping the protein terminals. Loop refinement was carried out to rectify the missing and invalid residues on the protein structure. The protonation state of the protein was determined at close to the physiological pH by selecting PROPKA function in 'Protein Preparation Wizard'. In the final step, protein structure was minimized using OPLS3 [38] molecular mechanics force field to improve the steric clashes that might be present in the protein structure. The restrained minimization process was allowed until when the convergence of the heavy atoms reaches to the root-mean-square deviation (RMSD) of 0.30 Å. Thereafter, using the 'Receptor Grid Generation' panel of Glide (Grid-Based Ligand Docking with Energetics) module [39] embedded in Schrödinger's Maestro interface, the grid was generated for the prepared protein structure. The receptor grid box was generated around the surrounding active site residues by selecting information of co-crystallized ligand, i.e. Staurosporine, at the centre and thereby enclosing the specified space inside a rectangular box which usually contains receptor and binding site information. In detail, this study targeted the Lyn kinase ATP or catalytic binding site consisting of important amino acid residues such as Leu22, Gly23, Ala42, Lys44, Glu89, Met91, Ala92, Ser95, Ala140, Asn141, Val142, Leu143, Ala153, and Asp154 as described in a number of previous studies for inhibiting the role of Lyn kinase protein [32,40]. More specifically, Lyn-Staurosporine binding region and the substrate-binding groove/ATP-binding site in the hinge region between the N- and C-lobes in Lyn protein were selected and targeted for docking interaction analysis.

Virtual screening of large compound databases

Initially for screening out compounds with a strong binding affinity towards Lyn protein, an extensive hierarchical filter-based approach was utilized. Therefore, VSW utility tool embedded in Schrodinger suite was used for multi-step molecular docking. In particular, VSW is comprised of three sequential molecular docking methods, i.e. Glide-HTVS (high throughput virtual screening), Glide-SP (standard precision), and Glide-XP (extra precision) docking, and followed by MM-GBSA based binding free energy analysis for the best dock-scored ligand-protein complexes. These three levels of docking programs performed a systematic search for extracting the best conformational orientation for the docked ligand with Lyn protein in each step. In the CHPC server (https://www.chpc.ac.za/index.php/resources/lengau-cluster), the VSW protocol was executed under some specified parameters. In VSW panel, under 'Input' tab, browsing of .sdf compounds was done as the source of ligands. Under 'Filtering' tab, execution of QikProp module [41] was carried out for pre-filtering compounds by means of Lipinski's rule of five (RoF). Violation of such recommended rule kicked out the compounds from further proceeding with VSW. The RoF describes the four important rules for a molecule being a drug-like compound. These rules include, molecular weight and hydrophobicity (log *P*) should not be more than 500 kDa and 5, respectively. The number of hydrogen bond donors and acceptors should be less than or equals to 5 and 10, respectively. In 'Receptor' tab previously generated grid file was browsed. As the selected parameter 'Docking' tab, 10% of the best compounds were retained after docking in HTVS, SP and XP mode and considered for succeeding steps. To explore the binding interactions the XP docking outcomes were written into the file. In each of the molecular docking steps of VSW, the 'all good scoring states' were reserved. All other settings in VSW panel were maintained as default. Finally, the MM-GBSA approach of the Prime module was adopted to obtain ligand-binding free energies of the left-out molecules in Glide-XP docking. After the successful completion of VSW, top-ranked molecules were selected based on both Glide XP and MM-GBSA scores.

Pharmacophore model development

In order to screen out the molecules having a different pharmacophoric pattern in comparison to the Staurosporin, the Common Feature Pharmacophore model was developed from the Staurosporin using the Discovery Studio [31]. Before the pharmacophore model development, the conformations of the Staurosporin were generated. Out of two methods BEST/FAST, the BEST method was used which develops multiple acceptable conformations with the help of rigorous energy minimization and optimization through poling algorithm [42,43]. The chemical features are properly arranged rather than simply atom arrangements in the BEST algorithm [43]. The conformers having a range of 20 Kcal/mol energy value with respect to the global minimum were used to develop the pharmacophore models using the HipHop approach. The HipHop approach uses the conformers of active molecules only to generate the pharmacophore model. The input features were given as hydrogen bond (HB) acceptor (a), HB donor (d), hydrophobic (p) and ring aromatics (r).

The selected model was validated through the decoy set validation approach. The decoy set validation method checks the proficiency of the model in respect of select active molecules over the inactive compounds. A small dataset of active Lyn inhibitors having inhibition concentration (IC_{50}) less than 2.1 nM was mixed up with a large dataset of inactive Lyn inhibitors (IC_{50} greater than 10,000 nM). Both datasets were collected from the BindingDB database. The merged dataset was screened by the selected pharmacophore model to find outfitted active and inactive molecules. From the above data, a number of parameters were calculated included true positive (TP), true negative (TN), false positive (FP), false negative (FN), the enrichment factor (EF) from the top 1% hits, etc. To check the significance of the screened data the Boltzmann-enhanced discrimination of receiver operating characteristic (BEDROC) was calculated. The BEDROC is an inclusive form of receiver operating characteristic (ROC) that identifies difficulties in the screening method. After successful validation, the model was used to map the molecules left out following the molecular docking and pharmacokinetic analyses.

In silico ADME and drug-likeness prediction

The in silico pharmacokinetics (ADME) assessment is one of the most crucial and critical drug-likeness analyses of the molecules obtained from a large chemical database. Molecules retained in VSW are then allowed for ADME profile predictions by using freely available online SwissADME web tool [44]. A number of physiochemical, lipophilicity, water-solubility, pharmacokinetics, drug-like properties including RoF [45] and Veber's rule [46] were documented. Due to the fast predictive power and spontaneous, straightforward interpretation of the molecular design aspect, SwissADME became widely popular in the scientific community. Initially, all generated molecules were checked for RoF and Veber's rule. The RoF explains the drug-likeness characteristics of the molecules. The flexibility and surface area of promising molecules can be explained through Veber's rule. According to this rule, being a potential molecule, the total polar surface area (TPSA) and the number of rotatable bonds should not be more than 140 $Å^2$ and 10, respectively. Moreover, another two pharmacokinetic parameters, the human intestinal absorption (HIA) and blood-brain barrier (BBB), were assessed. Both parameters play an important role in facilitating the appropriate selection of good candidate drug-like molecules [47]. The intestine is normally the primary site for the absorption of a drug from an orally administered solution. The HIA parameter explains that for a given compound how much percentage will be absorbed through the human intestine [47]. The BBB parameter indicates the capability of the molecules to enter the brain cells. High HIA and low BBB explain low adverse effect and toxicities within the brain [47].

MD simulation

Proposed potential inhibitors complexed with Lyn kinase domain were considered for all-atoms MD simulation study for 100 ns of time span. The Gromacs 2018-2 software tool (http://www.gromacs.org/) available at the Lengau CHPC server (https://www.chpc.ac.za/index.php/resources/lengau-cluster) was used for the MD simulation study. In the simulation, the time step, constant temperature and constant pressure were considered as 2 fs, 300 K and 1 atm, respectively. To generate the ligand, topology an external opensource online server the SwissParam tool [48] was used. The CHARM36 all-atom force field [49] was applied and the system was solvated by the TIP3P water model [50]. The simulation system was built confining the protein-ligand complex within a cubic box with a diameter of 1 Å from the centre of the system. A total of five Na^+ ions were added to neutralize the system prior to the energy minimization and simulation. In order to overcome the close-contacts or overlaps between the atoms, the steepest descent algorithm was considered. In order to equal distribution of solvent and ions around the protein-ligand complex, the entire system was equilibrated with NVT (constant number of particles, volume, and temperature) followed by NPT (constant number of particles, pressure, and temperature) ensemble approaches. To address the long-range interaction parameters the van der Waals and electrostatic cut off were used to 0.9 and 1.4 nm, respectively. The trajectory information was updated in 1ps interval. Different parameters including root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF) and radius of Gyration (Rg) were calculated from the entire trajectory of MD simulation to observe and explore behaviour of molecules in dynamic states.

Binding free energy calculation through MM-PBSA method

The binding free energy (ΔG_{bind}) of the final proposed Lyn inhibitors was calculated through MM-PBSA approach using g_mmpbsa tool [51]. The MM-PBSA approach in binding free energy calculation is a widely used and important application to study biomolecular complexes [51]. It is important to note that MD simulation combined with MM-PBSA can also consider the conformational fluctuation and entropic contributions [51]. In particular, the compiled g_mmpbsa program is a standalone tool that does not need any dependency except it requires four input files. Those required input files were generated during MD simulation execution such as trajectory (i.e. trr or xtc) file, a topology-parameter (i.e. tpr) file, an index (i.e. ndx) file, and also the $G_{solvation}$ parameters (i.e. mdp) file. Herein, these all types of the file obtained for specific Lyn protein-ligand complexes which were used to calculate the ΔG_{bind} energy upon execution of g_mmpbsa tool. The ΔG_{bind} can be calculated using the following expression.

$$\Delta G_{bind} = G_{\mathrm{Complex}} - G_{\mathrm{protein}} - G_{\mathrm{ligand}} \quad (1)$$

The $G_{complex}$ is defined by the total free energy of the complex between protein and ligand. $G_{protein}$ and G_{ligand} describe the free energy of the protein and ligand, respectively, in the solvent. The free energy of individual complex (G), protein (G) and ligand (G) can be calculated as below.

$$G = \langle E_{\mathrm{MM}} \rangle - TS + \langle G_{solvation} \rangle \quad (2)$$

The average molecular mechanics (MM) potential energy denoted by E_{MM} in a vacuum. The T and S describe the temperature and entropy, respectively. The free energy of solvation denoted by $G_{solvation}$.

E_{MM} can be defined as follows:

$$E_{\mathrm{MM}} = E_{\mathrm{bonded}} + E_{\mathrm{nonbonded}} \quad (3)$$

E_{bonded} is the combination of bonded interactions such as bond length, bond angle and dihedral angle. The $E_{nonbonded}$ describes the nonbonding interactions included electrostatic and van der Waals interactions.

The $G_{solvation}$ is defined by the energy required to transfer a solute from a vacuum to the solvent. It can

be expressed as follows:

$$G_{solvation} = G_{polar} + G_{nonpolar} \quad (4)$$

The G_{polar} and $G_{nonpolar}$ are denoted by the electrostatic and non-electrostatic contribution to the solvation free energy, respectively.

Results and discussion

Virtual screening using multi-steps docking and MM-GBSA analysis

In the last few decades, an exponential increase in computationally employing virtual screening utility of large molecular databases has intensified as an effective way for drug candidates like hit-to-lead identification and optimization [52]. In that respect, structure-based virtual screening (SBVS) is an impressive computational approach which being implemented in recent years towards the selection of novel and selective potential molecules for a specific target [53]. Generally, a large chemical dataset of potentially bioactive compounds is screened through SBVS to find out some subset of novel and potent chemical compounds that can be further utilized for digging out more relevant pharmacologically active lead-like compounds and finally allowing them for experimental testing [52]. In the current study, about 0.7 million compounds belonging to the entire Asinex and the ChEBI databases were considered to screen against the Lyn protein through the multi-steps molecular docking such as Glide-HTVS, Glide-SP and Glide-XP followed by binding energy assessment by Prime-MM-GBSA approach. The flow diagram of the employed computational work is given in Figure 1. Prior to execute the VSW of both databases, the molecular protocol was validated using the self-docking approach. In this method, the co-crystal Staurosporine was re-drawn and docked in the same active site where it was bound. The best docked pose was extracted and superimposed with co-crystal Staurosporine conformer. The superimposed structure is given in Figure S1 (Supplementary file). The RMSD value of both superimposed conformers was found to be 0.275 Å. It is reported that RMSD less than 2 Å of co-crystal ligand and best docked pose of the same validate the docking protocol [54]. Therefore, the docking protocol considered in the current study was successfully validated.

Moreover, for comparison of the outcomes in the present study, the potent Lyn inhibitor, Staurosporine was chosen as a control or reference compound. In VSW method, sequential filtering and continuing with the best 10% compounds were implemented in each step of Glide-HTVS, Glide-SP and Glide-XP docking techniques. Final compound sets obtained in Glide-XP docking approach and Staurosporine were used to calculate the binding free energy using the Prime-MM-GBSA method. For the standard compound Staurosporine, the Glide-XP score and MM-GBSA score were found to be −6.812 and −65.68 Kcal/mol, respectively, and henceforth used those scores as the cut-off for filtering out the compounds in successive steps. From the molecular docking, the best ligand binding pose was considered on the basis of molecular binding interaction pattern and Glide-XP score. Based on the Glide-XP score and binding free energy analyses, finally, a total of 36 compounds were carefully sorted and then used for further assessment. To identify drug-like lead compounds it is necessary to check the

ADME profile of each compound. Therefore, all 36 screened compounds were subject to the ADME analyses and the obtained results are presented in Table S1 (Supplementary data). From Table S1, it was revealed that a total of 24 compounds showed acceptable pharmacokinetic profiles and again we have considered those compounds for further analyses. For simplicity, the above 24 molecules were named as Mol_1, Mol_2, Mol_3 and so on. In order to identify molecules having pharmacophoric features similar to the Staurosporine, a common feature ligand-based pharmacophore model was generated using the standard Lyn inhibitor Staurosporine, and subsequently, each compound mapped to the pharmacophore model by setting maximum omitted feature value as 0 (zero). All mapped compounds on the pharmacophore model are depicted in Figure S2 (Supplementary data). It was observed that out of 24 compounds, 17 compounds failed to fit with at least one

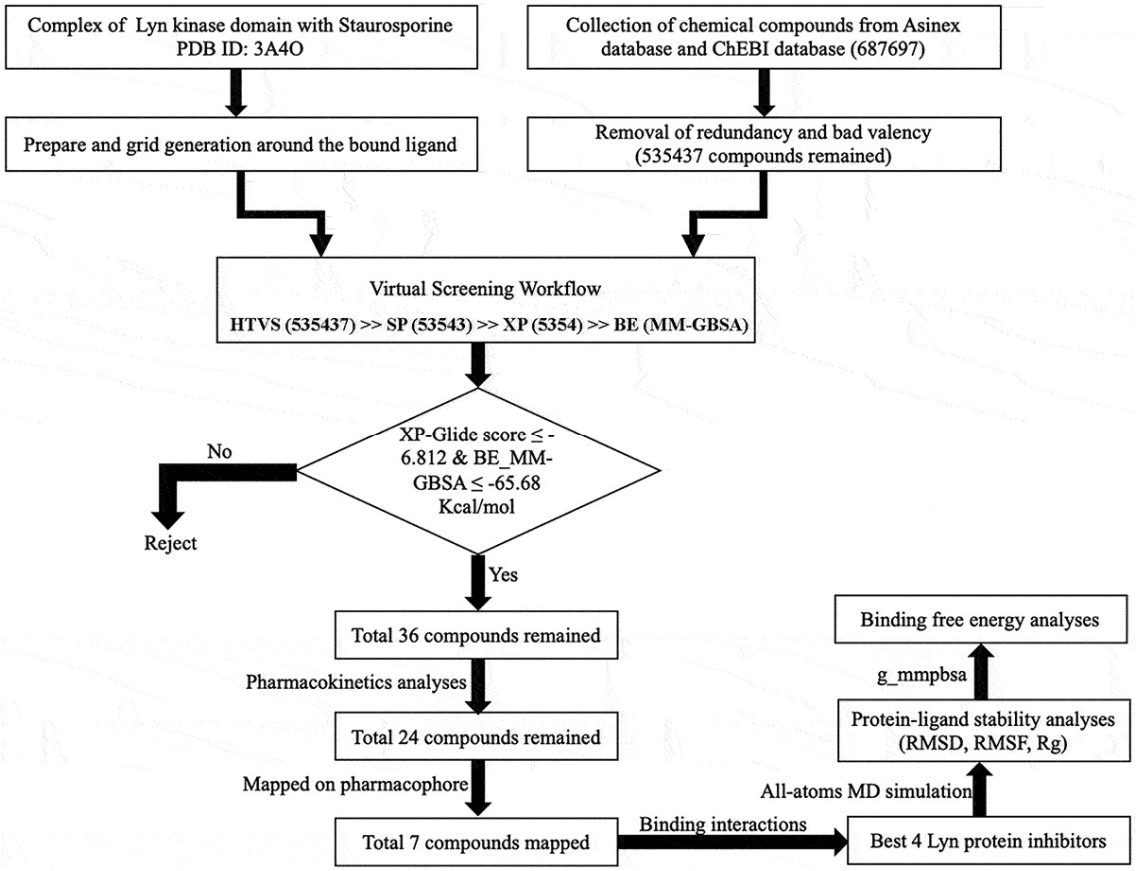

Figure 1. Schematic virtual screening workflow for identification of Lyn inhibitors

pharmacophore feature and hence they were removed from further processing. Finally, seven molecules (Mol_1, Mol_9, Mol_14, Mol_15, Mol_18, Mol_21 and Mol_24 in Figure S2 (Supplementary file)) those mapped to all pharmacophoric features were considered for a deeper level of molecular binding interaction analyses and comparisons with Staurosporine. The binding interaction profile of each of the above seven molecules along with Staurosporine is given in Figure S3 (Supplementary file). Based on binding interaction profiles, four molecules (Mol_1, Mol_9, Mol_21 and Mol_24) were finally selected as promising LynB inhibitors. Two-dimensional (2D) representations of selected four LynB inhibitors are given in Figure 2. It has been observed that all identified proposed Lyn inhibitors hold several different types of functional groups (e.g. hydroxyl, methyl, carbonyl, amino, thiol and fluorine)

which might help in the formation of potential chemical interactions with Lyn amino acid residues for exerting necessary biological effects. The presence of suitable functional groups in any chemical moiety enhances the chance of potential binding interaction formation with the counter functional groups of the amino acids of the target protein macromolecules. Hence, the identified molecules can create essential intermolecular interactions utilizing their functional groups with the active site residues of Lyn protein. It was observed that all molecules hold phenyl ring which was found to be common and that might be crucial to establish hydrophobic interactions with potential hydrophobic amino acid residues of Lyn protein. The presence of pyrimidine or pyridine ring in the chemical scaffold was indicated the chances of formation of hydrogen bonding (H-bond) interaction. The presence of '-oxo' groups in Mol_1, Mol_9, Mol_21 and Mol_24 was undoubtedly explained that all ligands can behave as hydrogen bond acceptor for participating in molecular interactions. Methyl group present in Mol_9 and Mol_21 can be a crucial component for hydrophobic interactions. The hydroxyl group attached to the pyrimidine ring in Mol_24 may be proved as important as either hydrogen bond acceptor or donor. Apart from the above-mentioned structural features, a number of heterocyclic rings were observed in all four inhibitors which can be participated in potential molecular interactions. Overall, all proposed Lyn inhibitors were consisting of a number of crucial functional groups, and those can be key components to stabilize the protein-ligand complex upon interaction formation.

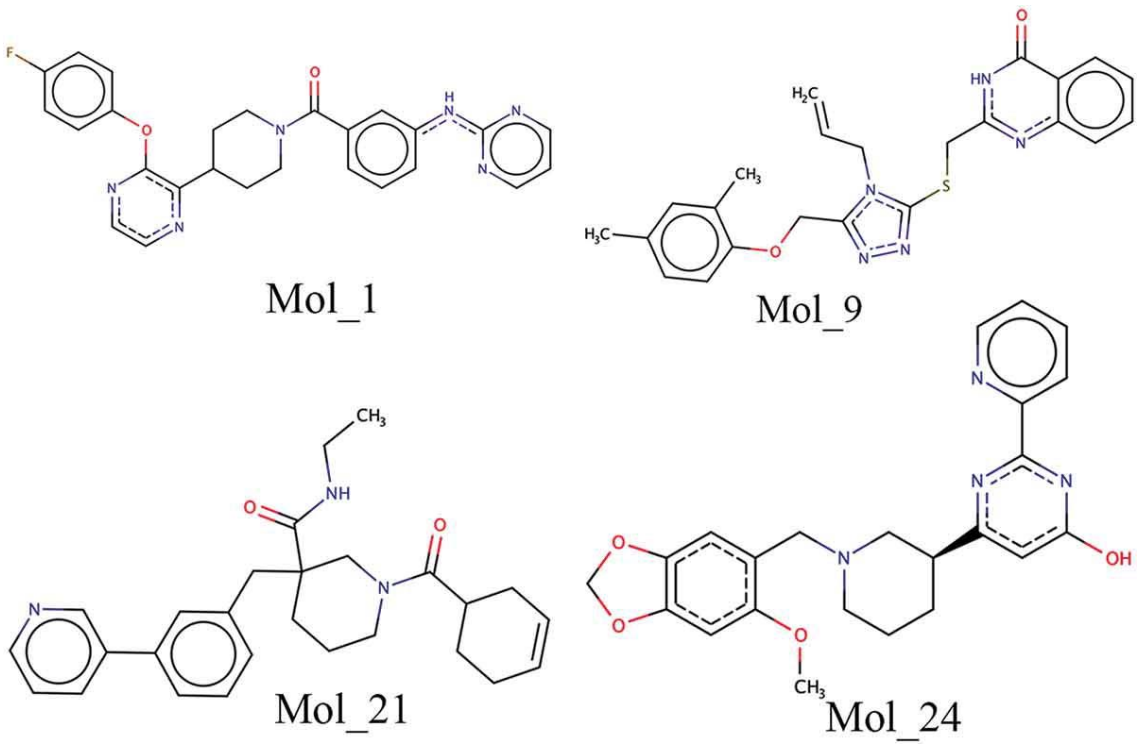

Figure 2. Two-dimensional representation of finally selected Lyn inhibitors

Analysis of molecular interactions between Lyn and identified inhibitors

Molecular interaction patterns of selected four compounds and standard Staurosporine bound to Lyn kinase domain were analysed critically using Protein–Ligand Interaction Profiler (PLIP) [55]. However, before going to analyse the docking generated molecular binding interaction profile of identified (Mol_1, Mol_9, Mol_21 and Mol_24) and reference Staurosporine compounds, the interaction profile of the

original co-crystallized Staurosporine has been extracted with the help of ligand explorer module in NGL viewer available in PDB to investigate the original molecular binding interaction profile for the standard compound Staurosporine (presented in Figure S4, supplementary file), which bound with the Lyn crystal structure. It was revealed that the Staurosporine tightly interacted at the Lyn kinase ATP-binding site or catalytic site. Close observation on it suggested that interactions were detected with several catalytic amino acid residues (such as Leu22, Gly23, Ala42, Glu89, Met91,

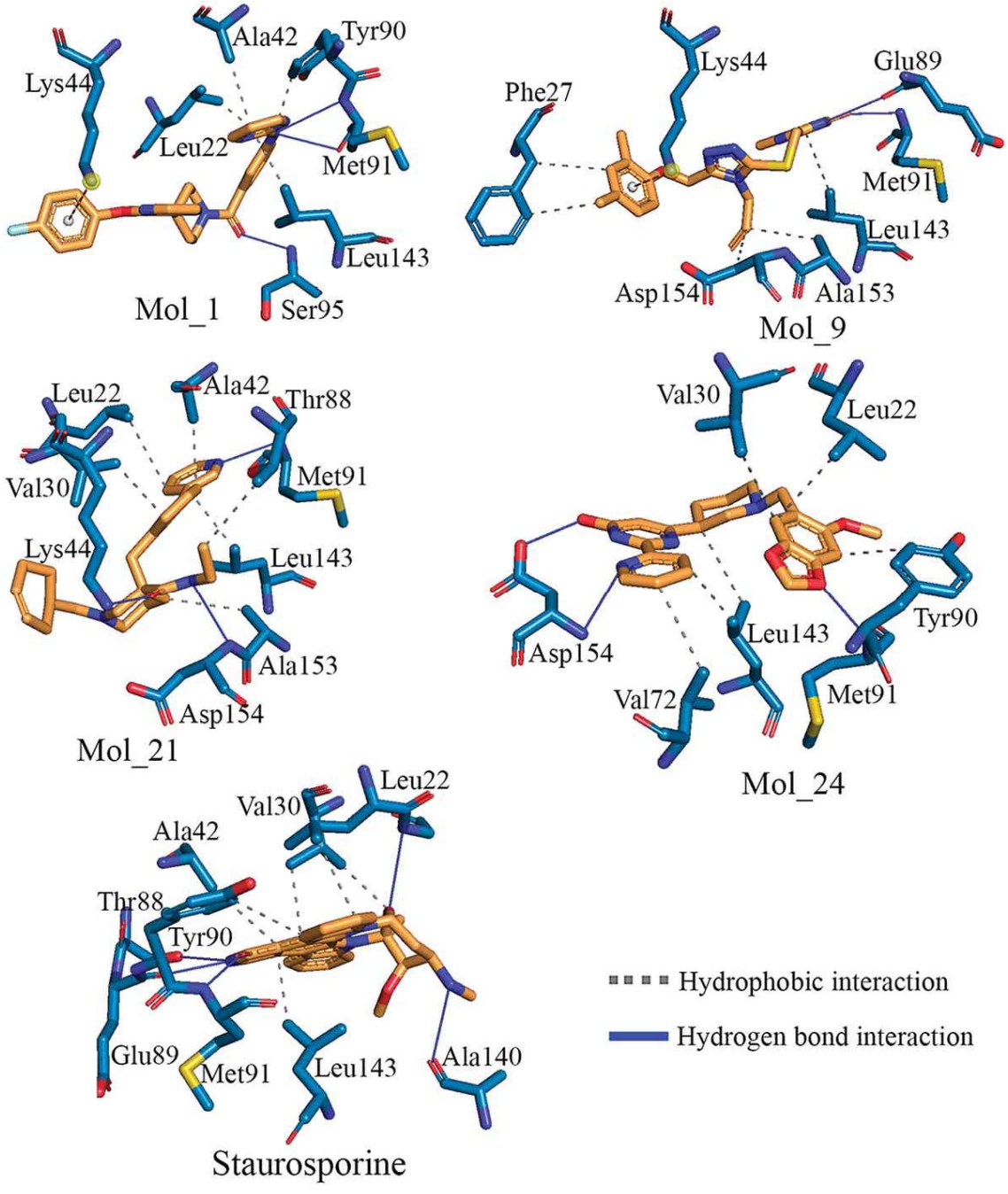

Figure 3. Binding interaction profile of Lyn inhibitors with Lyn protein

Ala92, Ser95, Ala140, Asn141, Val142, Leu143, Ala153 and Asp154) which present in the ATP or substrate-binding groove of Lyn protein for implicating potential inhibitory action biologically. Realizing such original molecular binding interaction profile for the reference compound Staurosporine helps in understanding the potentiality of newly identified compounds in terms of their inhibitory mechanism analysis. Molecular binding interaction profiles of all identified Lyn inhibitors are given in Figure 3. It can be noted that the -NH group present between phenyl and pyrimidine rings in Mol_1 was found to be essential for hydrogen bond interactions with key amino acid Met91. Another catalytic amino acid residue, Ser95 has also established hydrogen bond interaction with the -oxo group present between phenyl and pyridine rings of Mol_1. It was revealed that the '-oxo' group behaved as hydrogen bond acceptor and the amine group of Ser95 as hydrogen bond donor. Further, it was also observed that Mol_1 formed the hydrophobic interactions with residues Leu22, Ala42, Tyr90 and Leu143 with measured distances of 2.57, 2.59, 2.84, 2.35 Å, respectively. It was worth noting that the phenyl and pyrimidine rings were involved in the above hydrophobic interactions. There was a π-cation interaction observed with basic amino acid Lys44 at the bonding distance of 3.66 Å with fluorobenzene moiety present in compound Mol_1.

The -oxo group attached to the pyrimidine ring and nitrogen atom of the same were found to act as hydrogen bond acceptor and form potential hydrogen bond interaction with catalytic amino residues Met91 and Glu89, respectively. Dimethyl benzene ring in Mol_9 has established a π-cation interaction with Lys44 at a distance of 3.80 Å. A number of hydrophobic interactions were observed between Mol_9 and catalytic amino residues of Lyn protein. The phenyl ring fused with a pyrimidine, alkyl group attached with triazolidine ring and dimethyl benzene ring in Mol_9 participate in the hydrophobic interactions with Lyn protein. In particular, compound Mol_9 was found to form two hydrophobic contacts with Phe27 with a distance of 2.71 and 2.95 Å, while, each of Leu143, Ala153, and Asp154 residue establishes single hydrophobic contacts at the distances of 2.93, 2.76, and 2.82 Å, respectively. A number of functional groups present in Mol_21 included benzene, piperideine and pyridine rings along with alkyl group were found to be crucial to form hydrophobic interactions with Lyn protein. The catalytic amino residues viz. Leu22, Val30, Ala42, Thr88, Leu143 and Ala153 formed hydrophobic contacts with Mol_21 at the measured distances of 2.99, 2.57, 2.21, 2.81, 2.58 and 2.82 Å, respectively. The -oxo group present between piperideine ring and methylamine and the nitrogen atom of methylamine itself were found to be crucial to form hydrogen bond interactions with Lys44 and Asp154, respectively. Another catalytic amino residue, Met91 critically formed hydrogen bond interaction with the nitrogen atom of the pyridine ring. In the case of Mol_24, the nitrogen atom of pyridine and hydroxyl group attached in pyrimidine formed hydrogen bond interaction with Asp154. One of the oxygen atoms of dioxolane ring fused with benzene was also found to form hydrogen bond interaction with Met91. Several other amino acid residues Leu22, Val30, Val72, Tyr90, and Leu143 formed hydrophobic contacts with Mol_24 at a distance of 2.84, 2.99, 2.92, 2.63, 2.92, 2.63 Å, respectively. All hydrophobic contacts were found with benzene, pyridine and piperidein rings of Mol_24. In the case of docking-based molecular interaction analysis for standard compound Staurosporine, it was observed that Staurosporine formed several numbers of hydrophobic interactions with Leu22, Val30, Ala42, Tyr90, Leu143 and the distances measured as 2.74, 2.67, 2.85, 2.98, 2.73, 2.37 Å, respectively. Further, the catalytic amino acid residues Gly23, Thr88, Glu89, Met91 and Ala140 participated to interact with Staurosporine through hydrogen bonding. It is reported that the Staurosporine molecule binds to the ATP-binding pocket in Lyn protein. In the present study, a similar pattern of binding interactions was found with all identified compounds, i.e. Mol_1, Mol_9, Mol_21 and Mol_24 with Lyn protein. Moreover, a number of studies explored small molecule inhibitors for Lyn protein. For instance, experimentally proven that three small molecule inhibitors for Lyn protein were bound to the ATP-binding site in hinge region/substrate-binding groove present between N- and C-lobes in Lyn kinase. In Lyn kinase domain, a number of hydrophilic interactions and hydrogen bonds were observed between Glu89, Met91, Ala92, Asp154, Asn141, Gly23 and the identified inhibitors [40].

In the present study, the molecular interaction analysis of docked ligands was demonstrated alike hydrogen bond interactions with the amino acid residues Glu89, Met91, Asp154 and Gly23 which definitely correlate the significance of current observations. Moreover, a molecular surface view portrait was obtained which could possibly identify the unique binding orientation of each molecule in protein binding pockets in a surface curvature (Figure 4). It can be observed that each identified molecule along with Staurosporine was perfectly localized inside the receptor cavity of Lyn protein and facilitating important intermolecular contacts in three-dimensional geometry.

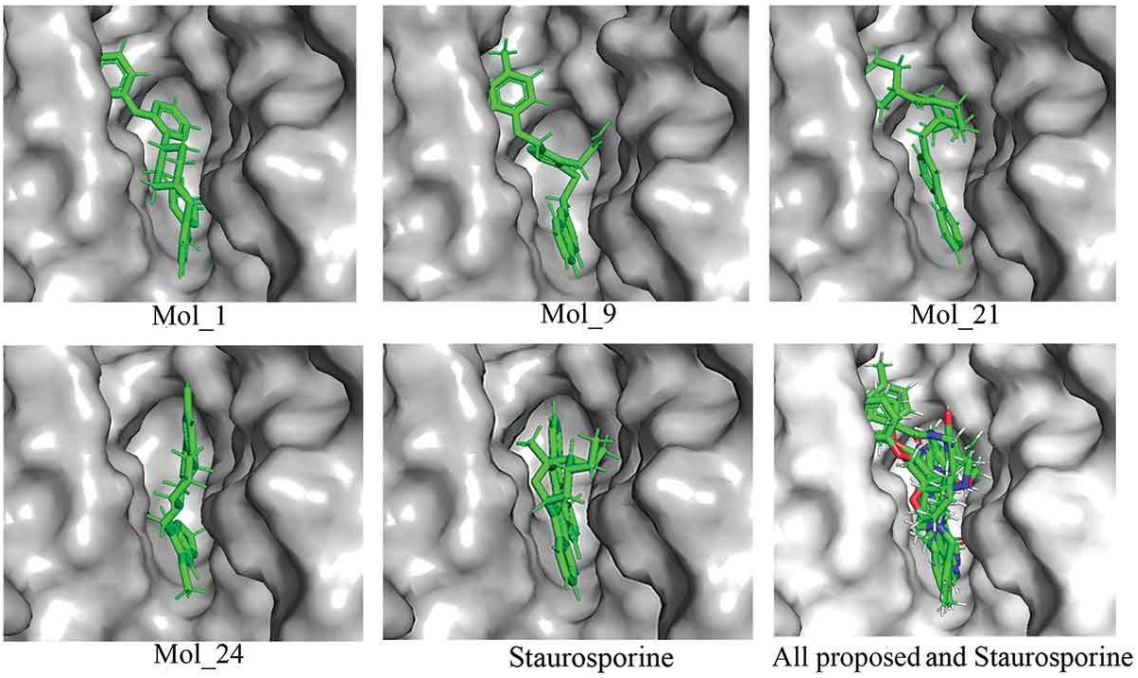

Figure 4. Binding mode of proposed inhibitors in the Lyn protein

Most interestingly, the present findings are highly in agreement with the previous study [40] outcomes in terms of molecular binding interaction analyses, where few flavonoid compounds (e.g. quercetin, apigenin and catechin) were studied for inhibition of Lyn protein kinase domain. In that study, important amino acid residues such as Glu89, Met91 and Ala92 of substrate-binding groove of Lyn kinase domain were found to participate in H-bond and hydrophobic interactions with apigenin and catechin for exhibiting inhibition of kinase activity at a significant level. Above all, it was interesting to note in many instances that almost all identified compounds (Mol_1, Mol_9, Mol_21 and Mol_24) show alike intermolecular binding interactions profile as originally bound reference compound Staurosporine (Figure S4 in supplementary file). Therefore, it can be postulated that computationally obtained conformational analysis of all identified compounds can also yield an equally probable inhibition mechanism as observed for staurosporine under in vivo state.

In silico pharmacokinetics study

Analyses of pharmacokinetics and physiological properties of chemical compounds are very crucial and also represent an important step in the drug discovery process. A number of parameters are taken into consideration to find out promising drug-like hit-to-lead compounds. To explore the pharmacokinetic analyses of proposed Lyn inhibitors the freely available SwissADME

(http://www.swissadme.ch/index.php) web tool was used. During analysing the results obtained from SwissADME, on very first observation, drug-likeness parameters were checked for all compounds following the Ghose [56] and Veber rules. More precisely, both rules imply the prediction of drug-like pharmacological properties of molecules applied to drug discovery research to prioritize the selection of active potential molecules with increased drug-like characteristics. Several pharmacokinetic and drug-likeness parameters were calculated and presented in Table 1. The molecular weight of all four identified molecules was found <500 g/mol, which suggested that the size of each molecule suitable to penetrate inside the cell membrane. The total polar surface area (TPSA) of all four molecules was found within the range of 62 to 111 $Å^2$ which also satisfied the drug-likeness criteria (<140 $Å^2$). Gastrointestinal absorption indicates the high absorbable nature of each molecule in the intestine. Solubility class property showed that all molecules were moderately soluble in nature. Synthetic accessibility of molecules Mol_1, Mol_9, Mol_21 and Mol_24 were found to be 3.93, 3.62, 4.29 and 3.94, respectively, which undoubtedly suggested that not a single molecule found difficult to synthesize.

Table 1. Physiochemical and ADME properties of selected Lyn protein inhibitors

Parameters	*Mol_1*	*Mol_9*	*Mol_21*	*Mol_24*
Formula	$C_{26}H_{23}FN_6O_2$	$C_{23}H_{23}N_5O_2S$	$C_{27}H_{33}N_3O_2$	$C_{23}H_{24}N_4O_4$
MW(g/mol)	470.50	433.53	431.57	420.46
NHA	35	31	32	31
NAHA	24	21	12	18
NRB	7	8	8	5
MR	132.25	123.46	131.79	118.03
TPSA ($Å^2$)	93.13	110.99	62.30	89.83
log *S*	−5.09	−4.86	−4.58	−4.25
SC	Moderately soluble	Moderately soluble	Moderately soluble	Moderately Soluble
GI	High	High	High	High
vROF	0	0	0	0
BS	0.55	0.55	0.55	0.55
SA	3.93	3.62	4.29	3.94
log *P*	3.41	2.92	3.60	3.90

MW: molecular weight; NHA: no. of heavy atoms; NAHA: no. of aromatic heavy atoms; NRB: no. of rotatable bonds; MR: molar refractivity; TPSA: topological polar surface area; S: solubility; SC: solubility class; GI: gastrointestinal absorption; vROF: violation if Lipinski's rule of five; BS: bioavailability score; SA: synthetic accessibility.

The BOILED-EGG representation was obtained from the SwissADME web tool to explore two important aspects of the molecules such as HIA (Human Intestinal Absorption) and BBB (Blood-Brain Barrier). Pictorial explanation of boiled-egg is given in Figure 5. The permeation ability of all four molecules in albumin (white) and yolk (yellow) area is explained in Figure 5. Molecules present in the albumin region are more prone to penetrate in the intestine, while, compounds found in the yolk area are considered to be favourable for BBB. It is also important to note that the yolk and white areas are not mutually exclusive. From Figure 5, it can be observed that Mol_1, Mol_9 and Mol_24 more favourable for the permeation in HIA, whereas, Mol_21 more favourable towards the BBB. Further, it was observed that Mol_1, Mol_21 and Mol_24 were effluated from the central nervous system (CNS) by P-Glycoprotein and denoted by the blue colour in Figure 5. On the other hand, Mol_9 was not effluated from CNS by P-Glycoprotein. From the above observation of pharmacokinetics and drug-likeness analyses, it can be postulated that all four proposed molecules possess strong characteristics to be potential Lyn inhibitors.

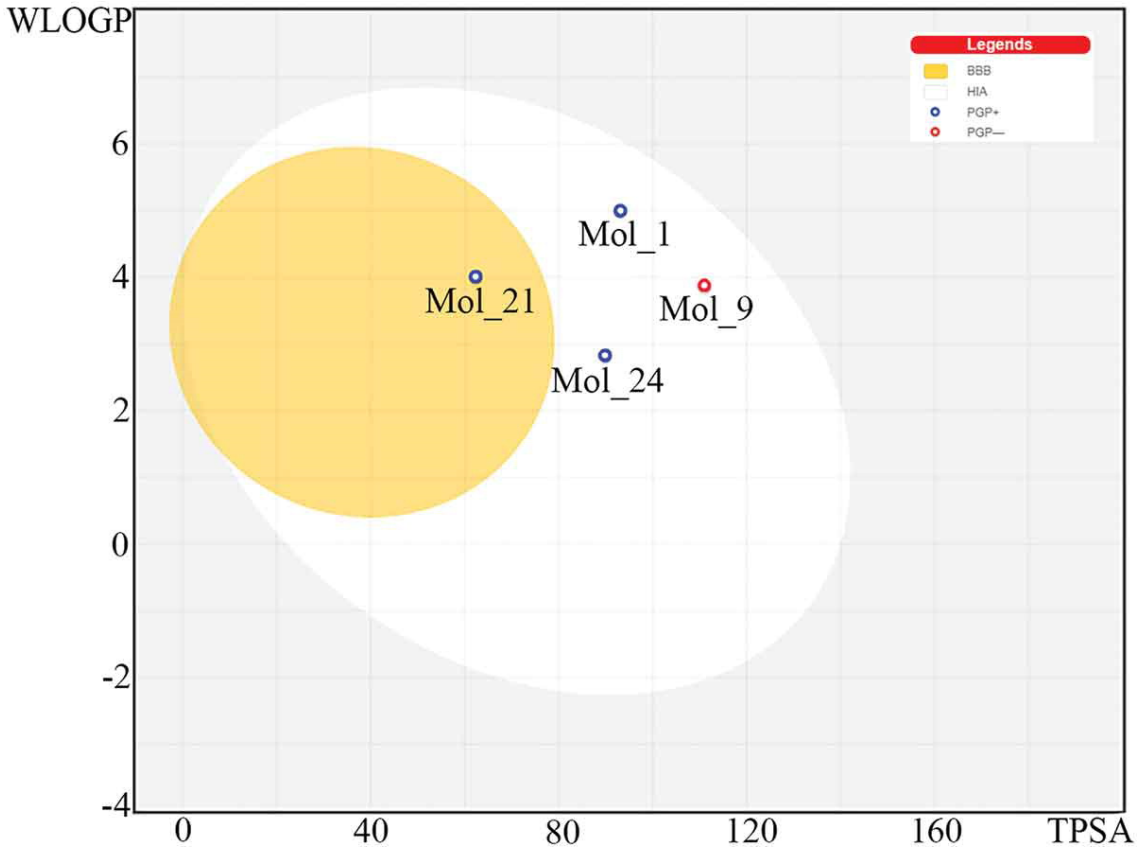

Figure 5. The EGG-BOILED model for the final screened Lyn inhibitors

Ligand efficiency assessment

A number of crucial parameters including Ligand Efficiency (LE), Ligand Efficiency Scale (LE_Scale), Fit Quality (FQ) and LE dependent Lipophilicity (LELP) were calculated to assess the quality and drug-likeness properties of the proposed Lyn inhibitors. All the above-mentioned parameters were calculated and are presented in Table 2.

Table 2. Bioactivity and efficiency parameters of proposed Lyn protein inhibitors

Molecule	*BE*	*[2]LE*	*[3]LE_Scale*	*[4]FQ*	*[5]LELP*
Mol_1	−10.110	0.288	0.288	1	11.840
Mol_9	−9.072	0.292	0.326	0.895	8.957
Mol_21	−9.750	0.304	0.316	0.962	11.392
Mol_24	−9.001	0.290	0.326	0.889	11.963

BE: binding energy; LE: ligand efficiency; LE_Scale: ligand efficiency scale; FQ: fit quality; LELP: ligand-efficiency dependent lipophilicity.

First of all, the LE is calculated by using Equation (5) which was proposed by Hopkins et al. [57]. The LE is basically the negative ratio between the binding energy (BE) obtained in molecular docking study and the number of heavy atoms (NHA). The recommended value of LE for a drug-like molecule is ≤0.4. For the proposed Lyn inhibitors, the LE was found to be 0.288, 0.292, 0.304, and 0.290 for Mol_1,

Mol_9, Mol_21 and Mol_24, respectively. Hence, the above values clearly show that all proposed Lyn inhibitors present drug-like characteristics.

$$LE = \frac{-BE}{NHA} \quad (5)$$

The LE_Scale was proposed by Reynolds et al. [58] and can be calculated using Equation (6). The size-dependent comparison of small molecules was portrayed by LE_Scale. LE_Scale of Mol_1, Mol_9,

Mol_21 and Mol_24 was calculated as 0.288, 0.326, 0.316 and 0.326, respectively (Table 2).

$$LE\_Scale = 0.873 \times e^{-0.026 \times NHA} - 0.064 \quad (6)$$

The good binding ability of the molecule can be checked through FQ and calculated from Equation (7). The FQ score is basically the ratio between LE and LE_Scale of the molecules. The FQ value of a molecule should be approximately 1 [59]. The FQ scores of Mol_1, Mol_9, Mol_21 and Mol_24 were 1.000, 0.895, 0.962, and 0.889, respectively, showing the good binding ability of molecules to Lyn protein.

$$FQ = \frac{LE}{LE\_Scale} \quad (7)$$

The LELP parameter can be calculated by Equation (8) which is the ratio between log *P* and LE and which was proposed by Keseru and Makara [60]. For a drug-like molecule, the LELP value should be more than 3. The LELP of Mol_1, Mol_9, Mol_21 and Mol_24 was found to be 11.840, 8.957, 11.392 and 11.963, respectively. Hence, the LELP parameter explains without any doubt that all proposed Lyn inhibitors show drug-like characteristics.

$$LELP = \frac{\log p}{LE} \quad (8)$$

Common pharmacophore-based assessment of identified compounds

The presence of essential pharmacophoric features in the small molecules usually contributes in determining the formation of molecular binding pattern and orientation inside the receptor cavity. In particular, a pharmacophore model represents a specific binding mode of small molecules/ligands within the macromolecular target. Therefore, it is undoubtedly interesting to have the same set of pharmacophoric features in the proposed molecules as of any standard established inhibitor. In this study, a common-feature pharmacophore model was developed from the standard compound Staurosporine, and the developed pharmacophore model was explained by the presence of two of each hydrogen bond acceptor (HBA), and hydrophobic (p) pharmacophoric features along with one hydrogen bond donor (HBD) pharmacophoric feature in 3D space. Further, the pharmacophore model was validated through the decoy set validation approach. In this method, a set of 323 molecules (D) having 23 actives (A) and the remaining 300 as inactive were collected from the BindingDB. The entire set was screened through the pharmacophore model. The total hit molecules (Ht) were found to be 190. A total of active hits were screened of 19. Therefore, the TP, TN, FP and FN were found to be 19, 129, 171 and 4, respectively. The enrichment factor was found to be 4.68 which implies that the selected model is efficient to pick active molecules 4.68 times higher from the database than expected by chance. The goodness of hits score (GH) was calculated and it was found to be 0.655. It is reported that the GH value of more than 0.5 is significant to select any pharmacophore model for database screening. The ROC curve is another crucial parameter to assess the quality of the model. This curve explains the

relationship between model sensitivity, i.e. capability to pick true positives, and specificity, i.e. capability to leave out the false positives [61]. The higher value of ROC score (represented by the area under ROC curve, i.e. AUC) of any model explains the proficiency to distinguish the good and bad molecules. The ROC curve of the pharmacophore model was developed and it is given in Figure S5 (Supplementary file). The ROC score was found to be 0.60 which implies that out of 10, six cases randomly selected Lyn inhibitors are ranked more than inactive inhibitors. The calculation procedure of all parameters can be found in Tai et al. [62]. Hence, the above parameters clearly validate the model and can be used in the virtual screening of the Lyn inhibitors.

All four proposed Lyn inhibitors were mapped on the developed pharmacophore model and presented in Figure 6. From Figure 6, it may be noticed that all identified four molecules perfectly map with all pharmacophoric features in the model obtained for standard compound Staurosporine. Hence, the proposed Lyn inhibitors consist of proper geometric and chemical compatibility with similar pharmacophoric profiles as of Staurosporine and such 3D pharmacophore matching similarity might indicate physically reasonable scoring terms for being promising drug candidate like molecules for exhibiting significant Lyn inhibitory activity.

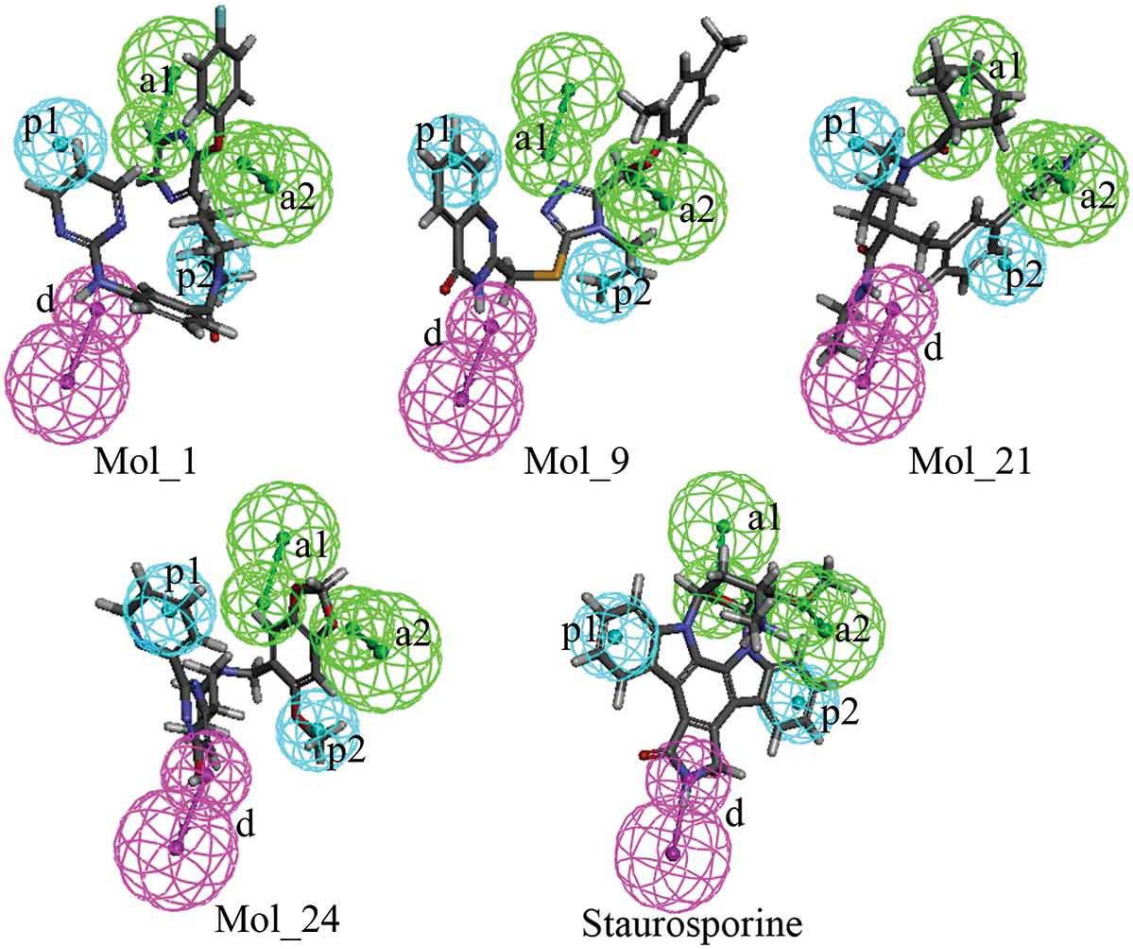

Figure 6. Pharmacophoric features of four identified Lyn inhibitors (Mol_1, Mol_9, Mol_21, Mol_24) and standard compound mapped on developed common feature pharmacophore model

Molecular dynamics simulation analyses Lyn protein-inhibitors complexes

In order to reproduce the nearly accurate or actual dynamic behaviour of protein-ligand complex, the MD simulation is an important and widely used approach implemented in recent days in drug discovery research for facilitating the understanding of energetic information about protein and ligand interactions in a time-affordable manner [63]. Herein, to explore the biophysical mechanism at the atomic level and elucidate the dynamic characteristics of the final proposed inhibitors inside the Lyn receptor cavity, all-atoms classical MD simulation for 100 ns time span of each complex was carried out. To evaluate the stability/flexibility of each protein-ligand complex, a number of parameters including RMSD, RMSF and *radius of gyration (Rg) were analysed from the MD simulation trajectory.*

Root-mean-square deviation (RMSD)

The RMSD value obtained from the MD simulation trajectory is one of the critical parameters that explains changes in structural conformation of the protein backbone over time upon system equilibration. To accurately analyse overall structural changes, the RMSD of backbone atoms of Lyn protein bound with compounds Mol_1, Mol_9, Mol_21, Mol_24 and Staurosporine was calculated and plotted against time of the simulation. The RMSD plot of Lyn protein bound with each inhibitor is depicted in Figure 7. In addition, the minimum, maximum and average RMSD of Lyn protein bound with Mol_1, Mol_9, Mol_21, Mol_24 and Staurosporine are given in Table 3. It was observed that Lyn backbone bound with proposed inhibitors was reached equilibration state and remained stable till the end of the simulation. Notice that the highest RMSD of Lyn protein was found to be 0.369 nm when bound with compound Mol_24. In the case of Lyn-Staurosporine complex, the RMSD value of backbone atoms was found to be fluctuated in a larger scale in comparison to the other compounds. Precisely, from 30 to 70 ns, the RMSD values oscillated much and afterwards it equilibrated around 0.3 nm till the simulation end period. Deviation from the native state of the crystal structure can be explained by the average values of RMSD of the Lyn backbone. The average RMSD of Lyn backbone bound with Mol_1, Mol_9, Mol_21, Mol_24 and Staurosporine was found to be 0.251, 0.226, 0.225, 0.255 and 0.331 nm, respectively. The above low RMSD values undoubtedly suggest the greater conformational stability of Lyn protein when bound with proposed potential inhibitors. Overall, the stability of the MD simulated systems in regards to the RMSD was clearly adjudged to the findings of a number of crucial binding interactions obtained in the molecular docking study. Moreover, there was no substantial variation in Lyn backbone RMSD found when bound with compounds Mol_1, Mol_9, Mol_21 and Mol_24. By observing the magnitude of deviation, taken together, it can be concluded that Lyn protein backbone bound with proposed inhibitors and was found to be more stable in comparison to the Lyn backbone bound with standard compound Staurosporine.

Table 3. Average, maximum and minimum RMSD, RMSF and Rg values of proposed Lyn protein bound with proposed inhibitors and Staurosporine

	Mol_1	*Mol_9*	*Mol_21*	*Mol_24*	*Staurosporine*
RMSD (nm)					
Min	0.0004	0.0004	0.0005	0.0005	0.0005
Max	0.318	0.329	0.295	0.369	0.526
Average	0.251	0.226	0.225	0.255	0.331
RMSF (nm)					
Min	0.054	0.041	0.041	0.044	0.048
Max	0.679	0.737	0.421	0.855	1.532
Average	0.162	0.143	0.128	0.151	0.169
Rg (nm)					
Min	1.903	1.908	1.914	1.932	1.914
Max	2.010	2.009	2.003	2.017	2.083
Average	1.949	1.964	1.961	1.968	2.011

Max: maximum; Min: minimum.

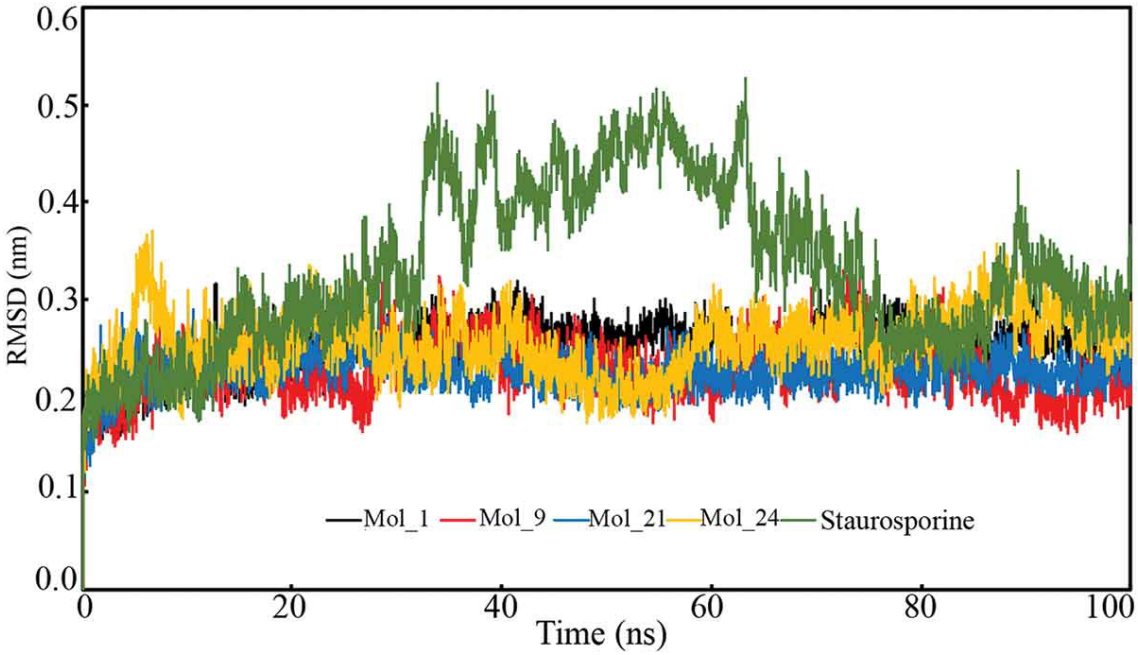

Figure 7. RMSD of Lyn backbone over time of simulation bound with Mol_1, Mol_9, Mol_21, Mol_24 and Staurosporine

Root-mean-square fluctuation (RMSF)

In protein-ligand stability, individual amino acid residue plays a significant role in protein stability specially bound with a small molecule. The amino acid residue fluctuation can be measured through the RMSF parameter calculated from the MD simulation trajectories. The analysed RMSF values plotted against each amino acid residue are presented in Figure 8. A similar pattern of RMSF variation was found for each amino residue of Lyn bound with Lyn inhibitors. Only both terminal residues (C-terminal and N-terminal) bit high, however, not a single amino acid was found to have an RMSF value of more than 0.8 nm. Due to binding interaction with the ligand, the RMSF of catalytic amino residues was found significantly low. Maximum, minimum and average RMSF values were calculated and are given in

. The fluctuation range of all complexes was found to be 0.041 to 1.532 nm. Difference between maximum and average, and, average and minimum can give an idea of fluctuation of the amino residues of Lyn. It was of 0.517 and 0.625 nm, 0.594 and 0.696 nm, 0.293 and 0.380 nm, 0.704 and 0.811 nm, and, 0.1.363 and 1.484 nm for the complex of Lyn with Mol_1, Mol_9, Mol_21, Mol_24 and Staurosporine, respectively. Such low values clearly explain that individual amino residue remained intact during the MD simulation.

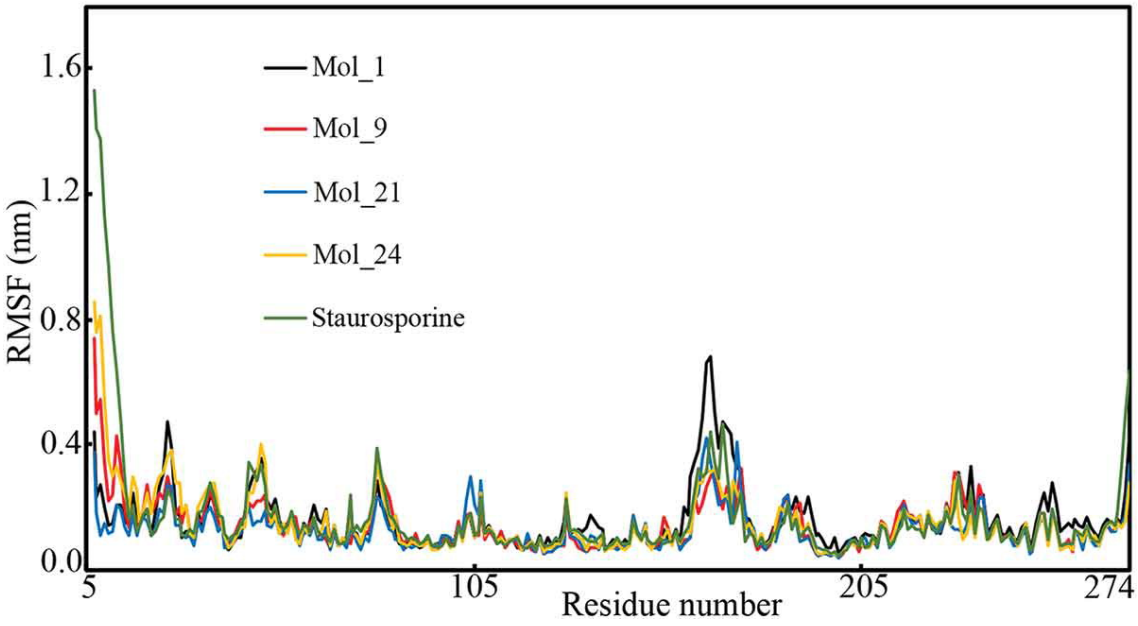

Figure 8. RMSF of individual amino residue of Lyn bound with proposed inhibitors and Staurosporine

Radius of gyration (Rg)

The compactness of the protein-bound with small molecules can be assessed through the Rg parameter calculated from the entire set of MD simulation trajectories. The relatively stable Rg signifies stably folding of protein during the MD simulation. In the case of unfolding the protein, the molecule gives a fluctuated Rg value over the time of the simulation. All Lyn-ligand complexes were used to calculate the Rg and are given in Figure 9. The Lyn protein that bound with Mol_1, Mol_9, Mol_21 and Mol_24 was found to be intact throughout the simulation. Rg values varied between 1.017 and 2.017 nm. The Lyn protein bound with Staurosporine was seen to fluctuate with higher Rg in comparison to others and finally equilibrated. Average Rg values of Lyn protein were found to be 1.949, 1.964, 1.961, 1.968 and 2.011 nm (Table 3) in case of bound with Mol_1, Mol_9, Mol_21, Mol_24 and Staurosporine, respectively. In the comparative observations, it was found that residual backbone and folding of the Lyn protein were consistently stable after binding with Mol_1, Mol_9, Mol_21 and Mol_24.

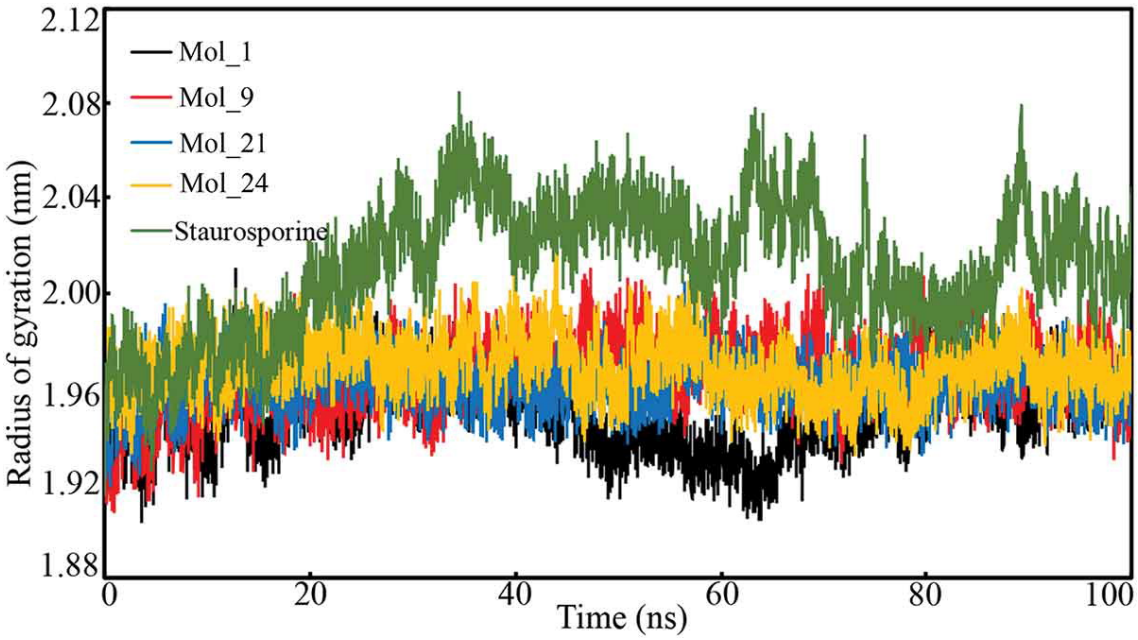

Figure 9. Radius of gyration Lyn protein-bound proposed inhibitor and Staurosporine

Binding free energy estimation using MM-PBSA method

It is essential to calculate the ΔG_{bind} of the small molecule in an accurate and trustworthy method. The scoring functions are approximated to calculate the binding energy in any molecular docking tool. The MM-PBSA approach which combines the molecular mechanics and continuum solvent models is considered to be accurate and trusted binding free energy calculation of small molecules. This approach used an ensemble of representative protein-ligand snapshots from the MD simulation trajectories and calculates the ΔG_{bind} of each snapshot. To explore the binding affinity of the proposed inhibitors towards the Lyn protein the binding free energy of each inhibitor calculated from the entire MD simulation trajectories using the MM-PBSA approach. The binding free energy of each frame was extracted and plotted against the simulation time (Figure 10). Maximum, minimum and average ΔG_{bind} are given in Table 4. Among proposed Lyn inhibitors it was observed that the highest average binding affinity was found to be −401.150 kJ/mol for Mol_24 followed by Mol_1, Mol_21 and Mol_9 with −272.070, −192.720 and 166.347 kJ/mol, respectively.

Table 4. Maximum, minimum and average values of binding free energy of proposed Lyn protein inhibitors and Staurosporine

MM-PBSA	*Mol_1*	*Mol_9*	*Mol_21*	*Mol_24*	*Staurosporine*
Binding free energy (ΔG_{bind}) kJ/mol					
Minimum	−415.480	−364.58	−353.52	−631.21	−930.688
Maximum	−134.470	−18.753	−83.753	−164.83	−431.77
Average	−272.070	−166.347	−192.72	−401.15	−671.507

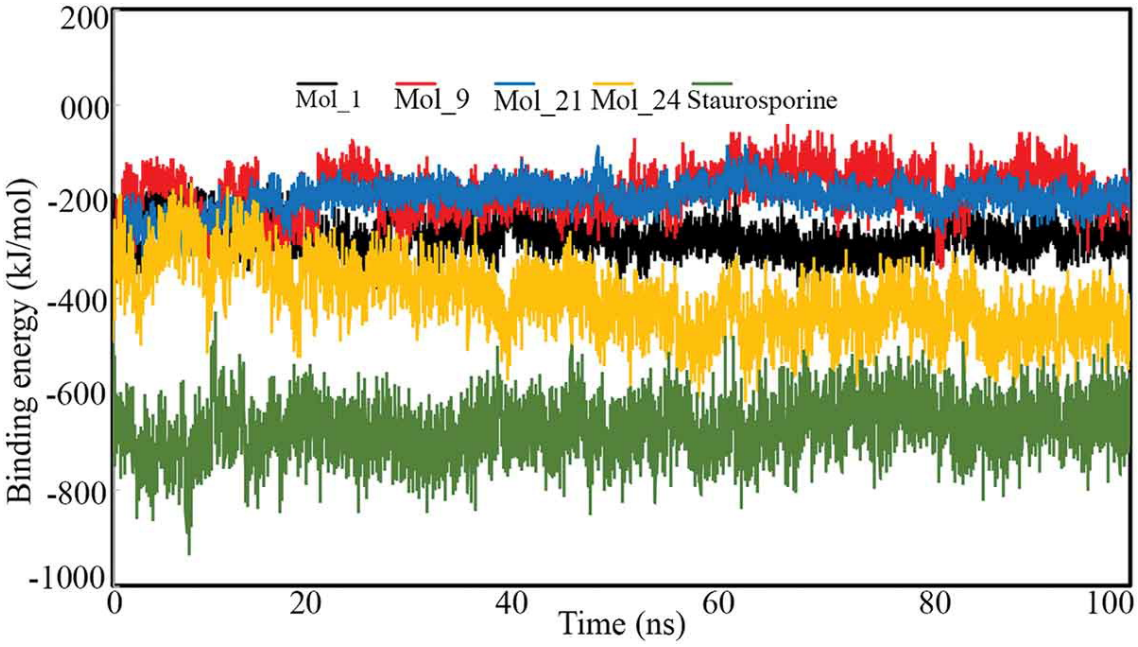

Figure 10. Binding free energy of proposed Lyn protein inhibitors and Staurosporine

From Figure 10, it can be seen that Staurosporine exhibits the highest ΔG_{bind} with an average value of −671.507 kJ/mol. Among the proposed Lyn inhibitors, Mol_24 was shown to have a comparable binding affinity with Staurosporine. The other three inhibitors (Mol_1, Mol_9 and Mol_21) show relatively less binding free energies compared to the reference compound Staurosporine. However, the molecules were found to have inconsistency in their binding free energies when dynamically interacted with Lyn protein. Hence, it was undoubtedly clear that all proposed Lyn inhibitors show a strong and good binding affinity towards the Lyn protein. Furthermore, it was seen that the Coulomb or electrostatic interaction ($\Delta G_{Coulomb}$) or van der Waals interaction energy (ΔG_{vdW}) mainly augmented to accomplish higher ΔG_{bind} value. Therefore, from the above results and discussions, it can be concluded that all the proposed Lyn protein inhibitors bind to the receptor and acquire a strong ability to inhibit the Lyn protein.

Conclusion

A multi-steps molecular docking-based virtual screening was performed to screen the Asinex and the ChEBI databases against the Lyn protein to identify potential chemical agents for the dysregulated Lyn associated therapeutic applications. Both prepared databases were used to VSW workflow for three consecutive molecular docking such as Glide-HTVS, Glide-SP and Glide-XP to remove low potential and inactive molecules. The binding energy of the remaining compounds and Staurosporine was calculated using the Prime-MM-GBSA approach. The Glide-XP score and Prime-MM-GBSA binding energy of Staurosporine were used as the threshold to consider the molecules for further analysis. Retained molecules were used to assess through the in silico pharmacokinetic analyses and removed molecules having poor pharmacokinetic profile. A common-feature pharmacophore model was developed from Staurosporine and remaining molecules from the above step were mapped. Molecules mapped with all pharmacophore features were considered for binding interaction analyses. Finally, four molecules were found to be potential against the Lyn protein. A number of strong binding interactions between proposed inhibitors and catalytic amino residues of Lyn were observed. The binding

interaction profile of the proposed molecules substantially matched with already published binding interactions observed in Lyn inhibitors. Drug-likeness characteristics were clearly explained that the proposed molecules were efficient enough to be Lyn inhibitors. The behaviour of the molecules in dynamic states was explored through all-atoms MD simulation study. A number of parameters were calculated from the MD simulation trajectories and found that protein-ligand complexes remained intact throughout the simulation. The MM-PBSA approach was used to calculate the binding free energy of the molecules and it was found that all proposed Lyn inhibitors showed a strong binding affinity towards the protein. Therefore, the proposed molecules might be crucial Lyn inhibitors and need an experimental validation.

Disclosure statement

No potential conflict of interest was reported by the authors.

Computational resource

The CHPC (www.chpc.ac.za), Cape Town, South Africa is thankfully acknowledged for computational resources and tools.

Funding

This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia through the Fast-track Research Funding Program. The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group No (RG-1441-430).

References

1. E. Ingley, *Functions of the Lyn tyrosine kinase in health and disease*, Cell. Commun. Signal. 10 (2012), pp. 21–31. doi:10.1186/1478-811X-10-21.
2. G. Manning, D.B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam, *The protein kinase complement of the human genome*, Science 298 (2002), pp. 1912–1934. doi:10.1126/science.1075762.
3. S.J. Parsons and J.T. Parsons, *Src family kinases, key regulators of signal transduction*, Oncogene 23 (2004), pp. 7906–7909. doi:10.1038/sj.onc.1208160
4. E. Ingley, *Src family kinases: Regulation of their activities, levels and identification of new pathways*, Biochim. Biophys. Acta 1784 (2008), pp. 56–65. doi:10.1016/j.bbapap.2007.08.012.
5. H. Pham, C. Birtolo, C. Chheda, W. Yang, M.D. Rodriguez, S.T. Liu, G. Gugliotta, M.S. Lewis, V. Cirulli, S.J. Pandol, and A. Ptasznik, *Essential role of Lyn in fibrosis*, Front. Physiol. 7 (2016), pp. 1–11. doi:10.3389/fphys.2016.00387.
6. E.J. Brodie, S. Infantino, M.S.Y. Low, and D.M. Tarlinton, *Lyn, lupus, and (B) lymphocytes, a lesson on the critical balance of kinase signaling in immunity*, Front. Immunol. 9 (2018), pp. 401–410. doi:10.3389/fimmu.2018.00401.
7. K. Kasahara, Y. Nakayama, K. Ikeda, Y. Fukushima, D. Matsuda, S. Horimoto, and N. Yamaguchi, *Trafficking of Lyn through the golgi caveolin involves the charged residues on alphaE and alphaI helices in the kinase domain*, J. Cell. Biol. 165 (2004), pp. 641–652. doi:10.1083/jcb.200403011.
8. M.L. Dykstra, A. Cherukuri, and S.K. Pierce, *Floating the raft hypothesis for immune receptors: Access to rafts controls receptor signaling and trafficking*, Traffic 2 (2001), pp. 160–166. doi:10.1034/j.1600-0854.2001.020302.x

9. T. Hayashi, H. Umemori, M. Mishina, and T. Yamamoto, *The AMPA receptor interacts with and signals through the protein tyrosine kinase Lyn*, Nature 397 (1999), pp. 72–76. doi:10.1038/16269.
10. A. Hirao, X.L. Huang, T. Suda, and N. Yamaguchi, *Overexpression of C-terminal Src kinase homologous kinase suppresses activation of Lyn tyrosine kinase required for VLA5-mediated Dami cell spreading*, J. Biol. Chem. 273 (1998), pp. 10004–10010. doi:10.1074/jbc.273.16.10004
11. H.M. Dingerdissen, J. Torcivia-Rodriguez, Y. Hu, T.C. Chang, R. Mazumder, and R. Kahsay, *BioMuta and BioXpress: Mutation and expression knowledgebases for cancer biomarker discovery*, Nucleic Acids Res. 46 (2018), pp. D1128–D1136. doi:10.1093/nar/gkx907.
12. E. Rupniewska, R. Roy, F.A. Mauri, X. Liu, M. Kaliszczak, G. Bellezza, L. Cagini, M. Barbareschi, S. Ferrero, A.M. Tommasi, E. Aboagye, M.J. Seckl, and O.E. Pardo, *Targeting autophagy sensitises lung cancer cells to Src family kinase inhibitors*, Oncotarget 9 (2018), pp. 27346–27362. doi:10.18632/oncotarget.25213.
13. A.K. Roseweir, T. Qayyum, Z. Lim, R. Hammond, A.I. MacDonald, S. Fraser, G.M. Oades, M. Aitchison, R.J. Jones, and J. Edwards, *Nuclear expression of Lyn, a Src family kinase member, is associated with poor prognosis in renal cancer patients*, BMC Cancer 16 (2016), pp. 1–10. doi:10.1186/s12885-016-2254-9.
14. D.R. Croucher, F. Hochgrafe, L. Zhang, L. Liu, R.J. Lyons, D. Rickwood, C.M. Tactacan, B.C. Browne, N. Ali, H. Chan, R. Shearer, D. Gallego-Ortega, D.N. Saunders, A. Swarbrick, and R.J. Daly, *Involvement of Lyn and the atypical kinase SgK269/PEAK1 in a basal breast cancer signaling pathway*, Cancer Res. 73 (2013), pp. 1969–1980. doi:10.1158/0008-5472.CAN-12-1472.
15. B. Elsberger, R. Fullerton, S. Zino, F. Jordan, T.J. Mitchell, V.G. Brunton, E.A. Mallon, P.G. Shiels, and J. Edwards, *Breast cancer patients' clinical outcome measures are associated with Src kinase family member expression*, Br. J. Cancer 103 (2010), pp. 899–909. doi:10.1038/sj.bjc.6605829.
16. N.K. Williams, I.S. Lucet, S.P. Klinken, E. Ingley, and J. Rossjohn, *Crystal structures of the Lyn protein tyrosine kinase domain in its Apo- and inhibitor-bound state*, J. Biol. Chem. 284 (2009), pp. 284–291. doi:10.1074/jbc.M807850200.
17. S. Berndt, V.V. Gurevich, and T.M. Iverson, *Crystal structure of the SH3 domain of human Lyn non-receptor tyrosine kinase*, PLoS One 14 (2019), pp. e0215140. doi:10.1371/journal.pone.0215140.
18. C. Dos Santos, C. Demur, V. Bardet, N. Prade-Houdellier, B. Payrastre, and C. Recher, *A critical role for Lyn in acute myeloid leukemia*, Blood 111 (2008), pp. 2269–2279. doi:10.1182/blood-2007-04-082099.
19. A.L. Samuels, S.P. Klinken, and E. Ingley, *Liar, a novel Lyn-binding nuclear/cytoplasmic shuttling protein that influences erythropoietin-induced differentiation*, Blood 113 (2009), pp. 3845–3856. doi:10.1182/blood-2008-04-153452.
20. K. Borzęcka-Solarz, J. Dembińska, A. Hromada-Judycka, G. Traczyk, A. Ciesielska, E. Ziemlińska, A. Świątkowska, and K. Kwiatkowska, *Association of Lyn kinase with membrane rafts determines its negative influence on LPS-induced signaling*, Mol. Biol. Cell 28 (2017), pp. 1147–1159. doi:10.1091/mbc.e16-09-0632.
21. P. Sutton, J.A. Borgia, P. Bonomi, and J.M.D. Plate, *Lyn, a Src family kinase, regulates activation of epidermal growth factor receptors in lung adenocarcinoma cells*, Mol. Cancer 12 (2013), pp. 76. doi:10.1186/1476-4598-12-76.
22. G. Tornillo, C. Knowlson, H. Kendrick, J. Cooke, H. Mirza, I. Aurrekoetxea-Rodriguez, M.D.M. Vivanco, N.E. Buckley, A. Grigoriadis, and M.J. Smalley, *Dual mechanisms of Lyn kinase dysregulation drive aggressive behavior in breast cancer cells*, Cell Rep. 25 (2018), pp. 3674–3692 e10. doi:10.1016/j.celrep.2018.11.103

23. L.J. Schwarz, E.M. Fox, J.M. Balko, J.T. Garrett, M.G. Kuba, M.V. Estrada, A.M. González-Angulo, G.B. Mills, M. Red-Brewer, I.A. Mayer, V. Abramson, M. Rizzo, M.C. Kelley, I.M. Meszoely, and C.L. Arteaga, *LYN-activating mutations mediate antiestrogen resistance in estrogen receptor–positive breast cancer*, J. Clin. Invest. 124 (2014), pp. 5490–5502. doi:10.1172/JCI72573.
24. S. Chakraborty, T. Inukai, L. Fang, M. Golkowski, and D.J. Maly, *Targeting dynamic ATP-binding site features allows discrimination between highly homologous protein kinases*, ACS Chem. Biol. 14 (2019), pp. 1249–1259. doi:10.1021/acschembio.9b00214
25. K.R. Brandvold, M.E. Steffey, C.C. Fox, and M.B. Soellner, *Development of a highly selective c-Src kinase inhibitor*, ACS Chem. Biol. 7 (2012), pp. 1393–1398. doi:10.1021/cb300172e.
26. M.I. Davis, J.P. Hunt, S. Herrgard, P. Ciceri, L.M. Wodicka, G. Pallares, M. Hocker, D.K. Treiber, and P.P. Zarrinkar, *Comprehensive analysis of kinase inhibitor selectivity*, Nat. Biotechnol. 29 (2011), pp. 1046–1051. doi:10.1038/nbt.1990.
27. S. Genheden and U. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities*, Expert. Opin. Drug Discov. 10 (2015), pp. 449–461. doi:10.1517/17460441.2015.1032936.
28. *ASINEX database*, ASINEX Corporation, Winston-Salem, USA, 2019.
29. P. de Matos, R. Alcantara, A. Dekker, M. Ennis, J. Hastings, K. Haug, I. Spiteri, S. Turner, and C. Steinbeck, *Chemical entities of biological interest: An update*, Nucleic Acids Res. 38 (2010), pp. D249–D254. doi:10.1093/nar/gkp886.
30. *Virtual screening workflow*, Schrödinger, LLC, New York, 2018; software available at https://www.schrodinger.com/.
31. *Discovery Studio*, Dassault Systèmes BIOVIA, San Diego, 2016; software available at https://www.3ds.com/products-services/biovia/.
32. N. Miyano, T. Kinoshita, R. Nakai, Y. Kirii, K. Yokota, and T. Tada, *Structural basis for the inhibitor recognition of human Lyn kinase domain*, Bioorg. Med. Chem. Lett. 19 (2009), pp. 6557–6560. doi:10.1016/j.bmcl.2009.10.038.
33. *LigPrep*, Schrödinger, LLC, New York, 2018; software available at https://www.schrodinger.com/.
34. J.C. Shelley, A. Cholleti, L.L. Frye, J.R. Greenwood, M.R. Timlin, and M. Uchimaya, *Epik: A software program for pKa prediction and protonation state generation for drug-like molecules*, J. Comput. Aided. Mol. Des. 21 (2007), pp. 681–691. doi:10.1007/s10822-007-9133-z.
35. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne, *The protein data bank*, Nucleic Acids Res. 28 (2000), pp. 235–242. doi:10.1093/nar/28.1.235.
36. *Protein Preparation Wizard*, Schrödinger, LLC, New York, 2018; software available at https://www.schrodinger.com/.
37. G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, and W. Sherman, *Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments*, J. Comput. Aided Mol. Des. 27 (2013), pp. 221–234. doi:10.1007/s10822-013-9644-8.
38. E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J.Y. Xiang, L. Wang, D. Lupyan, M.K. Dahlgren, J.L. Knight, J.W. Kaus, D.S. Cerutti, G. Krilov, W.L. Jorgensen, R. Abel, and R.A. Friesner, *OPLS3: A force field providing broad coverage of drug-like small molecules and proteins*, J. Chem. Theory. Comput. 12 (2016), pp. 281–296. doi:10.1021/acs.jctc.5b00864.
39. R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, M.P. Repasky, E.H. Knoll, M. Shelley, J.K. Perry, D.E. Shaw, P. Francis, and P.S. Shenkin, *Glide: A new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy*, J. Med. Chem. 47 (2004), pp. 1739–1749. doi:10.1021/jm0306430.

40. B. Wright, K.A. Watson, L.J. McGuffin, J.A. Lovegrove, and J.M. Gibbins, *GRID and docking analyses reveal a molecular basis for flavonoid inhibition of Src family kinase activity*, J. Nutr. Biochem. 26 (2015), pp. 1156–1165. doi:10.1016/j.jnutbio.2015.05.004.
41. *QikProp*, Schrödinger, LLC, New York, 2018; software available at
42. R. Kristam, V.J. Gillet, R.A. Lewis, and D. Thorner, *Comparison of conformational analysis techniques to generate pharmacophore hypotheses using catalyst*, J. Chem. Inf. Model. 45 (2005), pp. 461–476. doi:10.1021/ci049731z.
43. A. Smellie, S.L. Teig, and P. Towbin, *Poling: Promoting conformational variation*, J. Comp. Chem. 16 (1995), pp. 171–187. doi:10.1002/jcc.540160205.
44. A. Daina, O. Michielin, and V. Zoete, *SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules*, Sci. Rep. 7 (2017), pp. 42717. doi:10.1038/srep42717.
45. C.A. Lipinski, F. Lombardo, B.W. Dominy, and P.J. Feeney, *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings*, Adv. Drug Deliv. Rev. 46 (2001), pp. 3–26. doi:10.1016/S0169-409X(00)00129-0.
46. D.F. Veber, S.R. Johnson, H.Y. Cheng, B.R. Smith, K.W. Ward, and K.D. Kopple, *Molecular properties that influence the oral bioavailability of drug candidates*, J. Med. Chem. 45 (2002), pp. 2615–2623. doi:10.1021/jm020017n.
47. M. De Vrieze, P. Janssens, R. Szucs, J. Van der Eycken, and F. Lynen, *In vitro prediction of human intestinal absorption and blood-brain barrier partitioning: Development of a lipid analog for micellar liquid chromatography*, Anal. Bioanal. Chem. 407 (2015), pp. 7453–7466. doi:10.1007/s00216-015-8911-z.
48. V. Zoete, M.A. Cuendet, A. Grosdidier, and O. Michielin, *SwissParam: A fast force field generation tool for small organic molecules*, J. Comput. Chem. 32 (2011), pp. 235923–235968. doi:10.1002/jcc.21816.
49. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, H. Grubmuller, and A.D. MacKerell Jr., *CHARMM36m: An improved force field for folded and intrinsically disordered proteins*, Nat. Meth. 14 (2017), pp. 71–73. doi:10.1038/nmeth.4067.
50. P. Mark and L. Nilsson, *Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K*, Phys. Chem. A 105 (2001), pp. 9954–9960. doi:10.1021/jp003020w.
51. R. Kumari, R. Kumar, Open Source Drug Discovery, and A. Lynn, *g_mmpbsa – A GROMACS tool for high-throughput MM-PBSA calculations*, J. Chem. Inf. Model. 54 (2014), pp. 1951–1962. doi:10.1021/ci500020m
52. T. Zhu, S. Cao, P.C. Su, R. Patel, D. Shah, H.B. Chokshi, R. Szukala, M.E. Johnson, and K.E. Hevener, *Hit identification and optimization in virtual screening: Practical recommendations based on a critical literature analysis*, J. Med. Chem. 56 (2013), pp. 6560–6572. doi:10.1021/jm301916b.
53. Q. Li and S. Shah, *Structure-based virtual screening*, Meth. Mol. Biol. 1558 (2017), pp. 111–124.
54. M.A. Al-Sha'er and M.O. Taha, *Application of docking-based comparative intermolecular contacts analysis to validate Hsp90alpha docking studies and subsequent in silico screening for inhibitors*, J. Mol. Model. 18 (2012), pp. 4843–4863. doi:10.1007/s00894-012-1479-z.
55. S. Salentin, S. Schreiber, V.J. Haupt, M.F. Adasme, and M. Schroeder, *PLIP: Fully automated protein-ligand interaction profiler*, Nucleic Acids Res. 43 (2015), pp. W443–W447. doi:10.1093/nar/gkv315.
56. A.K. Ghose, V.N. Viswanadhan, and J.J. Wendoloski, *A knowledge-based approach in Designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases*, J. Comb. Chem. 1 (1999), pp. 55–68. doi:10.1021/cc9800071.

57. A.L. Hopkins, C.R. Groom, and A. Alex, *Ligand efficiency: A useful metric for lead selection*, Drug. Discov. Today 9 (2004), pp. 430–431. doi:10.1016/S1359-6446(04)03069-7.
58. C.H. Reynolds, S.D. Bembenek, and B.A. Tounge, *The role of molecular size in ligand efficiency*, Bioorg. Med. Chem. Lett. 17 (2007), pp. 4258–4261. doi:10.1016/j.bmcl.2007.05.038.
59. X. Xue, G. Bao, H.Q. Zhang, N.Y. Zhao, Y. Sun, Y. Zhang, and X.L. Wang, *An application of fit quality to screen MDM2/p53 protein-protein interaction inhibitors*, Molecules 23 (2018), pp. 3174. doi:10.3390/molecules23123174.
60. G.M. Keseru and G.M. Makara, *The influence of lead discovery strategies on the properties of drug candidates*, Nat. Rev. Drug. Discov. 8 (2009), pp. 203–212. doi:10.1038/nrd2796.
61. X.M. Chen, T. Lu, S. Lu, H.F. Li, H.L. Yuan, T. Ran, H.C. Liu, and Y.D. Chen, *Structure-based and shape-complemented pharmacophore modeling for the discovery of novel checkpoint kinase 1 inhibitors*, J. Mol. Model. 16 (2010), pp. 1195–1204. doi:10.1007/s00894-009-0630-y.
62. W. Tai, T. Lu, H. Yuan, F. Wang, H. Liu, S. Lu, Y. Leng, W. Zhang, Y. Jiang, and Y. Chen, *Pharmacophore modeling and virtual screening studies to identify new c-Met inhibitors*, J. Mol. Model. 18 (2012), pp. 3087–3100. doi:10.1007/s00894-011-1328-5.
63. X. Liu, D. Shi, S. Zhou, H. Liu, H. Liu, and X. Yao, *Molecular dynamics simulations and novel drug discovery*, Expert. Opin. Drug. Discov. 13 (2018), pp. 23–37. doi:10.1080/17460441.2018.1403419.