BIOINFORMATION
Discovery at the interface of physical and biological sciences

BIOMEDICAL INFORMATICS

www.bioinformation.net
Volume 19(2)

Research Article

Declaration on Publication Ethics:
The author's state that they adhere with COPE guidelines on publishing ethics as described elsewhere at https://publicationethics.org/. The authors also undertake that they are not associated with any other third party (governmental or non-governmental agencies) linking with any form of unethical issues connecting to this publication. The authors also declare that they are not withholding any information that is misleading to the publisher in regard to this article.

Declaration on official E-mail:
The corresponding author declares that lifetime official e-mail from their institution is not available for all authors

Comments from readers:
Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article without open access charges. Comments should be concise, coherent and critical in less than 1000 words.

Edited by P Kangueane

Citation: Odhar *et al.* Bioinformation 19(2): 160-166 (2023)

Molecular docking and dynamics simulation analysis of the human FXIIa with compounds from the Mcule database

Hasanain Abdulhameed Odhar*, Ahmed Fadhil Hashim, Salam Waheed Ahjel & Suhad Sami Humadi**

Department of pharmacy, Al-Zahrawi University College, Karbala, Iraq; *Corresponding author:

Institution website: https://alzahu.edu.iq/en
Institution email: info@alzahu.edu.iq

Author contacts:
Hasanain Abdulhameed Odhar – E-mail: hodhar3@gmail.com & hassaninathar@g.alzahu.edu.iq
Ahmed Fadhil Hashim – E-mail: ahmedfadhil@g.alzahu.edu.iq
Salam Waheed Ahjel – E-mail: salam.waheed@g.alzahu.edu.iq
Suhad Sami Humadi – E-mail: suhad@g.alzahu.edu.iq

Abstract:
The human factor XIIa is a serine protease enzyme that is implicated in the pathological thrombosis. This coagulation factor represents an interesting molecular target to design safer antithrombotic agents without adversely influencing physiological hemostasis. Therefore, it is of interest to virtually screen the human factor XIIa crystal with millions of compounds in Mcule database in order to identify potential inhibitors. For this purpose, both molecular docking and dynamics simulation were employed to identify potential hits. Also, various predictive approaches were utilized to estimate chemical, pharmacokinetics and toxicological features for the top hits. As such, we report here that compound 4 (1-(4-benzylpiperazin-1-yl)-2-**[**5-(3,5-dimethylpyrazol-1-yl)-1,2,3,4-tetrazol-2-yl**]**ethanone) may be a potential ligand against the human factor XIIa for further consideration in the design and development of novel antithrombotic agents.

Keywords: Human plasma β-FXIIa, docking, molecular dynamics simulation, structure-based virtual screening, Mcule database.

Background:
The human factor XIIa (FXIIa) is a serine protease enzyme that is essential for the launch of the intrinsic coagulation pathway in a contact activation reaction involving both plasma prekallikrein (PPK) and high molecular weight kininogen (HMWK) **[1]**. In this activation reaction, the plasma FXII zymogen will undergo a proteolytic transformation into the active protease FXIIa. This proteolytic conversion process occurs on a negatively charged surface where a small quantity of activated FXII is generated **[2]**. In the same time, HMWK attached to the same surface will deliver PPK to FXIIa for activation. Then, this generated active plasma kallikrein (PK) will reciprocally activate extra quantity of FXIIa **[3]**. And during the later steps of intrinsic coagulation cascade, FXIIa will cleave its substrate FXI to produce FXIa, which in turn converts FIX to FIXa **[4,5]**. Finally, this cascade of events will drive the generation of thrombin and fibrin formation. It has been demonstrated that FXIIa activation of FXI is not substantial for hemostasis therefore it is believed that the deficiency of FXII is not usually associated with excessive bleeding whether in humans or animals **[3,6]**. In addition to the procoagulant activity, the FXIIa mediated contact system has a proinflammatory effect through the kallikrein-kinin pathway, which can release bradykinin from HMWK by PK **[7]**. As such, FXIIa represents an attractive molecular target to design and introduce new drug candidates that can hinder pathological thrombus formation and without adversely influencing physiological hemostasis. The use of available antithrombotic drugs like warfarin, heparin and antiplatelet agents can be related to a high risk of severe bleeding because these drugs can target components of the blood clotting cascade like thrombin, FVIIa, FIXa, FXa, FXIa **[3]**. Structure-based virtual screening is a computational approach that can identify potential hits against a molecular target with a known three-dimensional structure **[8]**. In this in-silico study, both molecular docking and dynamics simulation were employed to carry out structure-based virtual screening against human plasma β-FXIIa crystal in order to identify potential inhibitors. For this purpose, we have used the human plasma β-FXIIa complexed with benzamidine as a target to perform structure-based virtual screening **[3]**. A cartoon representation for the three-dimensional structure of the human plasma β-FXIIa co-crystalized with benzamidine can be seen in **Figure 1**.

Methodology:
Setting up a plan for virtual screening study:
The major steps for this in-silico screening study can be summarized in **Figure 2**. As seen in this figure, more than 40 million compounds were filtered through different screening steps. Concisely, the database of these chemical compounds was first subjected to diversity-based filtration to select the most dissimilar molecules. Then, REOS (rapid elimination of swill) filter was used to remove any frequent and non-selective compounds in the database. After that, molecular docking was applied to virtually assess binding of these compounds to the active site of the human plasma β-FXIIa. Moreover, prediction of toxicity potential, drug-likeness and pharmacokinetics were utilized to further refine the screening of these compounds. Finally, two rounds of molecular dynamics (MD) simulation for 20 and 50 nanoseconds were applied to the filtered compounds.

Structure-based virtual screening:
For this stage, we have employed the structure-based virtual screening pipeline available through the online drug discovery website Mcule.com **[9]**. The methodology and options of this structure-based virtual screening are identical to what we had used in our previously published studies **[10–13]**. In this stage, the full Mcule database with more than 40 million compounds was used for this screening study. At first, the diversity-based filter was applied in order to choose the most dissimilar compounds. Then, the REOS (rapid elimination of swill) filter was utilized to eliminate non-selective hitters. After that, molecular docking was employed to assess the binding of these filtered compounds against the human plasma β-FXIIa. Both AutoDock tools and AutoDock Vina, embedded in Mcule.com website, were used to carry out molecular docking **[14,15]**. For the docking step, the human plasma β-FXIIa crystal (PBD: 6B74) was used as a target **[3]**. The employed docking coordinates were X: 62, Y: -1, Z: 15 while the dimensions for the used grid box were 22*22*22 Angstrom. To validate this docking protocol, the non-covalent inhibitor benzamidine was redocked to the plasma β-FXIIa crystal. Then, the Root Mean Square Deviation (RMSD) was calculated by comparing the conformations of the docked and the co-crystalized benzamidine **[16]**. Finally, the hits in the docking output were ranked based on their least binding energy and only the top ten compounds were selected for further validation. Both PyMOL v2.4.1 and Discovery Studio Visualizer v21.1.0 were used to inspect the docking orientation of the least binding energy pose for each compound against the plasma β-FXIIa crystal **[17,18]**. And by examination of docking interaction, we have excluded any compound with unfavorable interaction against target active site.

Prediction of toxicity potential, drug-likeness score and pharmacokinetics properties:
Then, the refined compounds were subjected to prediction of median lethal dose (LD_{50}) by using ProTox-II webserver [19]. Also, the AMES mutagenic potential for each hit was anticipated by using pkCSM sever [20]. Regarding pharmacokinetics features for these hits, both pkCSM and SwissADME servers were utilized to predict volume of distribution and water solubility [20,21]. And lastly, the drug-likeness score was predicted for each hit by using the molsoft website [22].

Molecular dynamics (MD) simulation study:
The aim of this MD simulation study is to calculate ligand proximity to the plasma β-FXIIa active site and to estimate Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) binding energy for the refined hits during simulation period. To accomplish this objective, The YASARA Dynamics v20.12.24 was used to carry out two rounds of MD simulation for 20 and 50 nanoseconds [23]. Initially, the MD simulation was executed for each refined hit for 20 nanoseconds only. Then, only those hits with an average RMSD of ligand movement that is less than 4 Angstrom

were subjected to another round of MD simulation for 50 nanoseconds. The detailed procedure used for this MD simulation is like to what we had employed in our previously published articles [11–13]. In summary, 0.9% of NaCl was used for this MD simulation study and an excess concentration of either Na^+ or Cl^- was added to ensure neutralization of ligand-target complex. Both steepest descent and simulated annealing minimizations were employed to abolish clashes likelihood during simulation. throughout this simulation, the following force fields were employed: AMBER14 for the solute, TIP3P for water molecules, AM1BCC and GAFF2 for the ligands [24–26]. GraphPad Prism v8.0.2 was used to plot RMSD values of ligand movement during the simulation period of 50 nanoseconds [27]. For the estimation of MM-PBSA binding energy, the YASARA Dynamics utilizes AMBER14 force field where the more positive binding energy indicates better interactions between ligand and target [28]. The YASARA Dynamics software employed the following equation to estimate MM-PBSA binding energy:

$$Binding\ Energy = EpotRecept + EsolvRecept + EpotLigand + EsolvLigand - EpotComplex - EsolvComplex$$

Table 1: Chemical characteristics for the top hits that were generated by virtual screening of Mcule database library against the crystal of human plasma beta-factor XIIa (β-FXIIa). These best compounds were ranked based on their minimum binding energy to the β-FXIIa crystal

Compound No	Molecular Formula	M.W. (g/ mol)	Log P	TPSA (Å²)	H-bond acceptors	H-bond donors
1	$C_{21}H_{15}F_2N_5O_3$	423.37	4.60	94.55	8	1
2	$C_{19}H_{19}ClN_6O_2S$	430.91	3.64	121.17	8	1
3	$C_{21}H_{19}ClN_4O_5S$	442.92	5.88	107.61	7	1
4	$C_{19}H_{24}N_8O$	380.45	0.70	84.97	9	0
5	$C_{21}H_{14}N_4OS$	370.43	4.75	106.91	5	1
6	$C_{20}H_{23}FN_4O_3$	386.42	2.79	78.68	7	1
7	$C_{19}H_{21}N_5O_2S$	383.47	4.03	119.19	7	2
8	$C_{17}H_{14}ClFN_4$	328.77	4.58	50.70	4	1
9	$C_{20}H_{23}N_7OS$	409.51	3.63	113.27	8	2
10	$C_{17}H_{14}ClF_3N_4O_2$	398.77	3.28	69.04	6	1

M.W.: molecular weight; **Log P:** logarithm of partition coefficient; **TPSA:** topological polar surface area.

Table 2: A summary for docking results, predicted pharmacokinetic characteristics and toxicity potential for the top hits generated by virtual screening of Mcule library against β-FXIIa crystal. These chemical candidates were ranked according to their minimum docking energy to β-FXIIa crystal.

No	Docking results		Drug-likeness score	Pharmacokinetics characteristics		Toxicity potential	
	Binding energy (Kcal/ mol)	Unfavorable interactions		Water solubility (mg/ ml)	VD_ss (L/Kg)	AMES toxicity	LD_50 (mg/ Kg)
1	-9.3	No	0.15	8.09e-04 Moderate	0.70	No	300
2	-9.0	No	2.15	8.80e-03 Moderate	1.38	No	1000
3	-8.7	No	-0.14	1.46e-04 Poor	0.65	No	1600
4	-8.7	No	0.92	1.01e-01 Soluble	1.15	No	489
5	-8.4	No	0.17	2.99e-04 Poor	0.76	Yes	1000
6	-8.4	Yes	0.92	5.82e-01 Soluble	1.60	No	235
7	-8.3	No	-0.28	6.68e-04 Moderate	0.52	Yes	1800
8	-8.3	No	0.24	9.07e-03 Moderate	0.64	No	750
9	-8.3	No	0.99	1.52e-02 Moderate	1.05	No	800
10	-8.3	No	0.12	1.28e-03 Moderate	0.38	No	800

VD_ss: steady state volume of distribution; **LD_50:** median lethal dose.

Table 3: A summary for molecular dynamics (MD) simulation for the filtered top hits

Hit no.	MD simulation for 20 nanoseconds			MD simulation for 50 nanoseconds			
	Ligand movement RMSD (Å)			Ligand movement RMSD (Å)			Average MM-PBSA binding energy (Kcal/ mol)
	Mean	Minimum	Maximum	Mean	Minimum	Maximum	
2	6.84	0.74	12.67	-	-	-	-
4	3.63	0.76	4.90	3.78	0.76	5.63	9.82
9	2.59	0.72	3.48	5.81	0.72	10.05	-17.2

MD: Molecular dynamics; **MM-PBSA:** Molecular Mechanics Poisson-Boltzmann Surface Area; **RMSD:** Root-Mean-Square Deviation; **Å;** Angstrom.
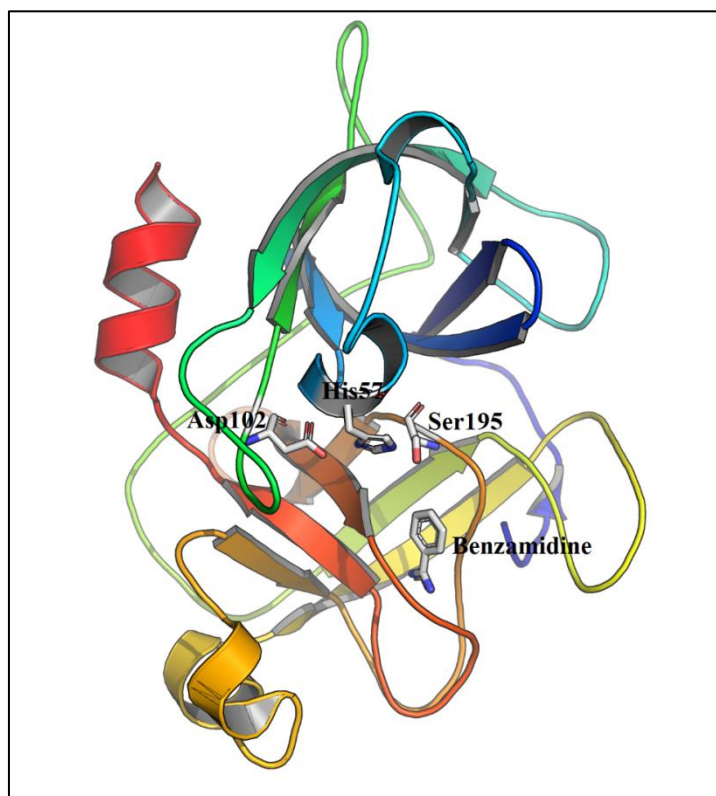

Figure 1: A three-dimensional cartoon illustration for the human plasma β-FXIIa in complex with the non-covalent inhibitor benzamidine (PBD: 6B74). Residues of the catalytic triad are represented as labelled sticks.

Results and Discussion:
At first, the accuracy of docking protocol was validated by using pose selection (redocking) approach. In this approach, the co-crystalized ligand is removed from target crystal and redocked into the same active site. Then, RMSD value for the difference between docked conformation and native conformation for that ligand is calculated. It is believed that RMSD value for the conformations difference between 1.5 and 2 Angstrom refers to a reliable docking protocol **[29]**. In this regard, our redocking result for the co-crystalized benzamidine indicates that the followed docking protocol is genuine. The RMSD value for the conformations difference between docked and co-crystalized ligand was 1.983 Å as seen in **Figure 3**. According to the chemical characteristics presented in **Table 1** for the top 10 hits screened against the plasma β-FXIIa, almost all of these hits are predicted to have good oral bioavailability **[30,31]**. The only exception is hit number 3 with a

logarithm of octanol/ water partition coefficient (Log P) that is greater than 5. In addition, the chemical structure of these top hits can be seen in **Figure 4**. Then, a tabular summary can be seen in **Table 2** for the docking results, pharmacokinetics features and toxicity potential of these top hits. As noted in this table, the only compound that showed unfavorable interactions with target crystal in docking study is hit number 6. And based on the prediction of drug-likeness score for each hit in **Table 2**, we can notice that four hits have a high score and these are compounds 2, 4, 6 and 9. While prediction of water solubility indicated that hits 3 and 5 may have poor solubility. On the other hand, volume of distribution (VD) anticipation referred to the possibility that compounds 1, 3, 5, 7, 8 and 10 may have low VD values. Finally, evaluation of toxicity potential prediction in **Table 2** pointed out to the mutagenic probability of both compounds 5 and 7. Also, both compounds 1 and 6 have a relatively low median lethal dose (LD_{50}) as predicted in **Table 2**. Based on prediction results in **Table 1** and **Table 2**, it is evident that compounds 2, 4 and 9 may represent potential candidates that can be further evaluated in molecular dynamics (MD) study. The overall results of this in-silico methodology can be summarized in **Figure 5**.
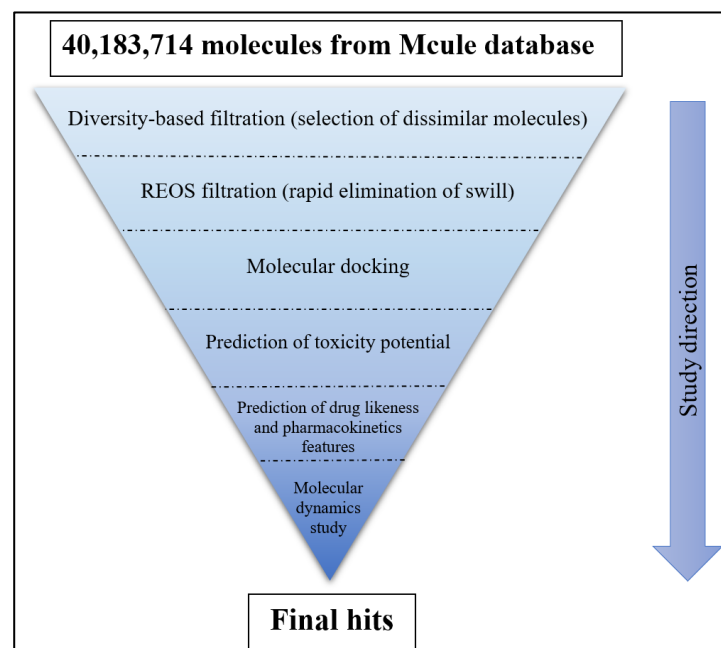

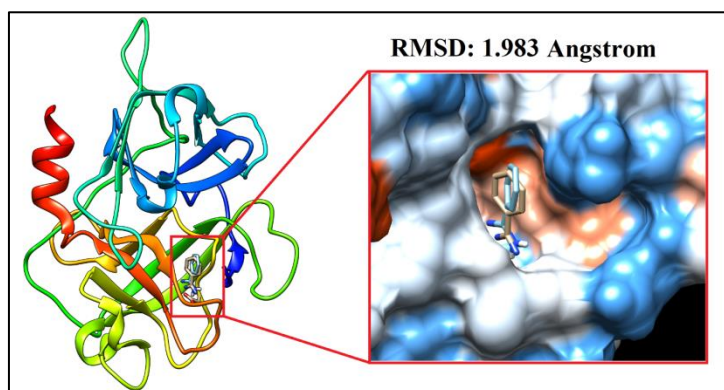**Figure 2:** A schematic representation for the main steps of virtual screening study.

©Biomedical Informatics (2023)

Figure 3: Redocking of the non-covalent inhibitor benzamidine into the β-FXIIa.
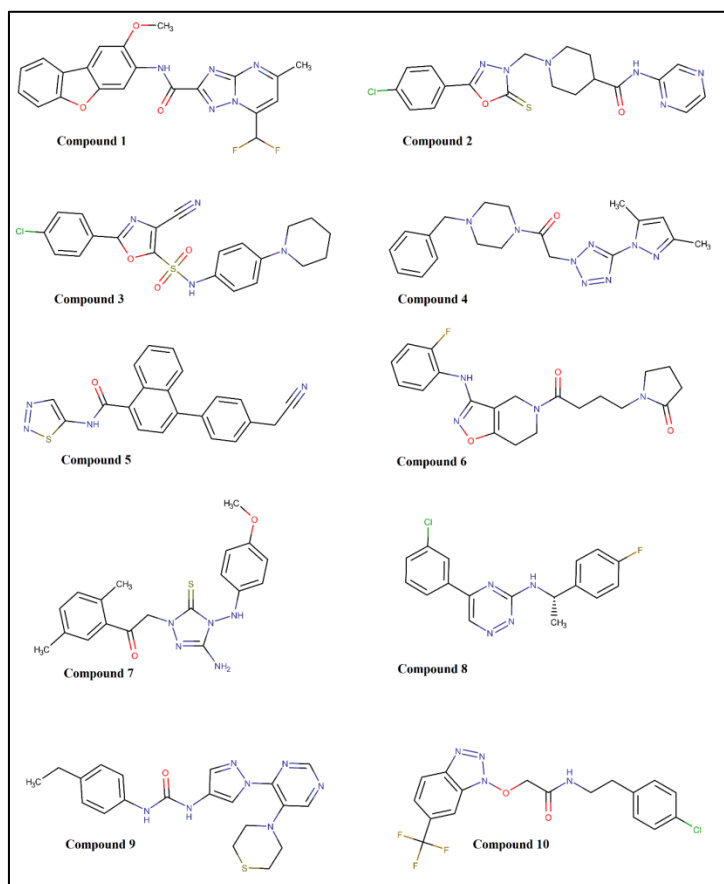

Figure 4: Chemical structures for the top ten hits that were virtually screened against the plasma β-FXIIa crystal.
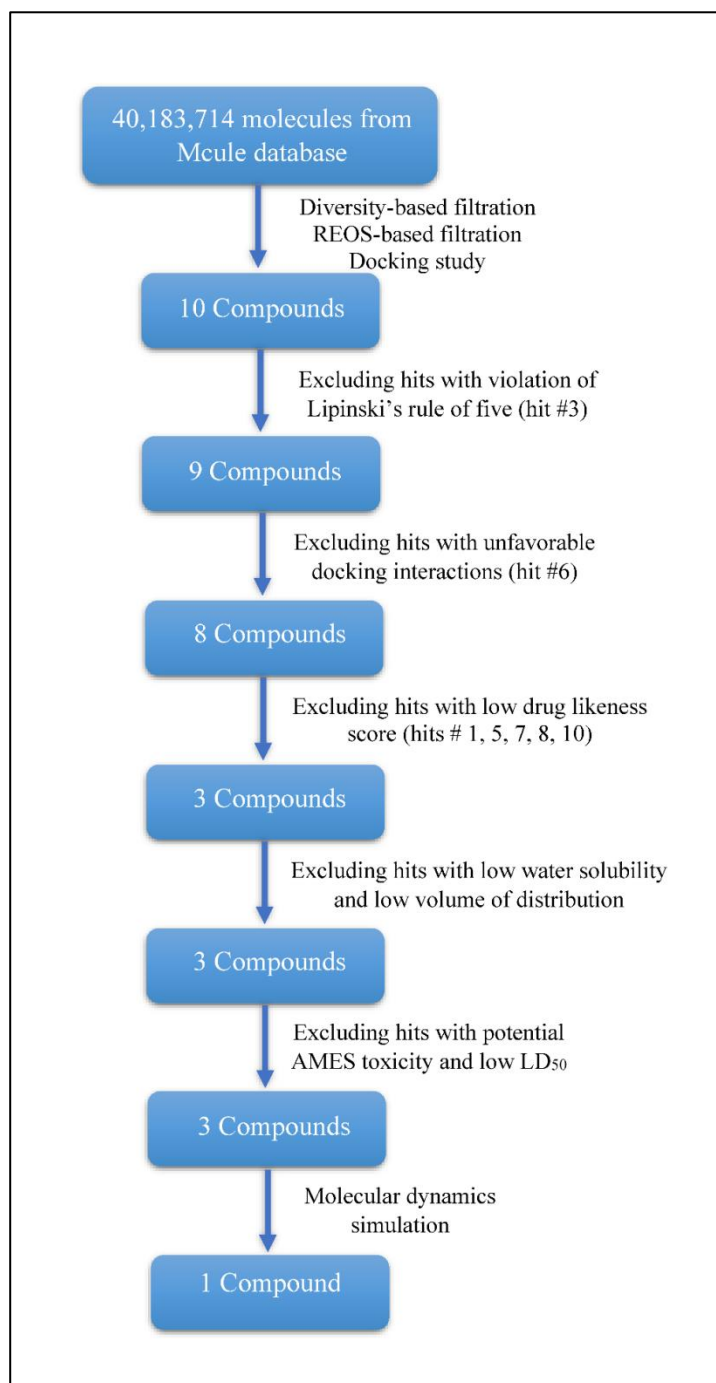

Figure 5: A concise overview for the stages of this structure-based virtual screening together with results for each filtration stage.
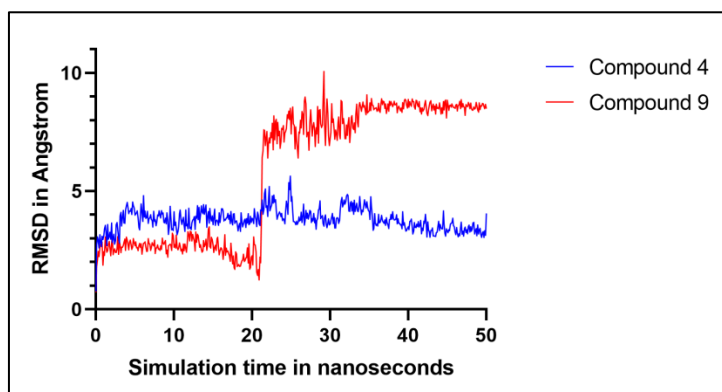
Figure 6: Ligand movement RMSD for compounds 4 and 9 against the human plasma β-FXIIa throughout simulation interval. This plot was generated by superposing ligand-target complex on its reference structure.
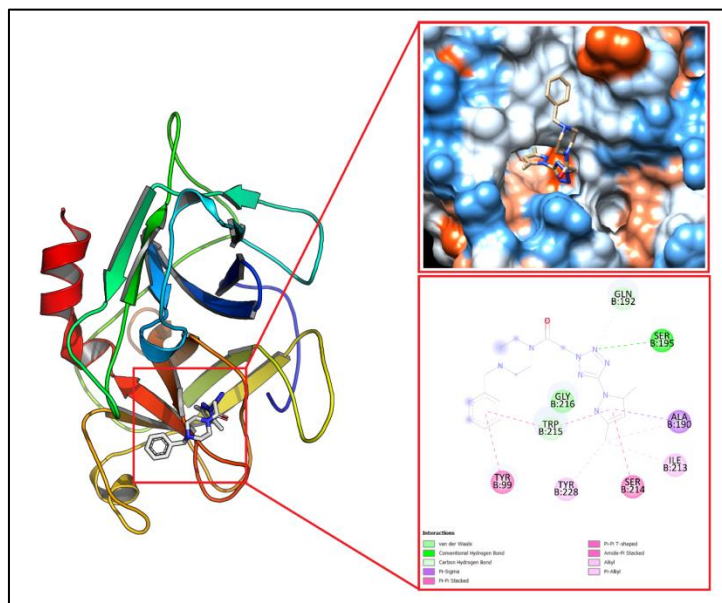

Figure 7: A cartoon representation for the docking of compound 4 into the human plasma β-FXIIa crystal.

As molecular dynamics (MD) simulation is considered a time-consuming process that requires extensive computational resources. Therefore, compounds 2, 4 and 9 were subjected at first to MD simulation for only 20 nanoseconds. It is well-known that ligand proximity to the target active site during simulation usually refers to stronger ligand binding to its target. This ligand proximity to target active site can be estimated by calculating ligand movement Root Mean Square Deviation (RMSD) throughout simulation **[32,33]**. As such, we have evaluated the mean RMSD value for ligand movement from target active site during simulation and subsequently eliminated those hits with mean RMSD value greater than 4 Angstrom. A tabular summary for MD simulation study can be seen in Table 3 for these three hits. According to **Table 3**, it is evident that compound 2 has a mean RMSD value of 6.84 Å during ligand movement analysis. As a

result, only compounds 4 and 9 were then subjected to an extended MD simulation for 50 nanoseconds period. Analysis of this extended simulation indicated that compound 4 was able to maintain a close proximity to target crystal with a mean RMSD value of 3.78 Å. While the mean RMSD of ligand movement for compound 9 was 5.81 Å throughout 50 nanoseconds. These results of ligand movement are in agreement with the calculated average MM-PBSA binding energy for compounds 4 and 9 which was 9.82 and -17.2 Kcal/mol respectively. As mentioned in YASARA Dynamics guideline, the more positive MM-PBSA binding energy points out to better interactions between ligand and target crystal **[28]**. Based on these findings in **Table 3**, it seems that compound 4 may be a better ligand to the human plasma β-FXIIa as compared to compound 9. This is because compound 4 was able to maintain a lower RMSD value of ligand movement throughout simulation. Also, the average MM-PBSA binding energy for compound 4 was more positive unlike compound 9. The detailed analysis of ligand movement RMSD for both compounds 4 and 9 throughout 50 nanoseconds period of simulation can be seen in **Figure 6**. Lastly, a cartoon illustration for the docking of compound 4 into the human plasma β-FXIIa crystal can be seen in **Figure 7**. It is evident from this figure that compound 4 is involved in hydrogen bonding with Serine 195, a key residue of the catalytic triad.

Conclusion:
Here, we report that compound 4 (1-(4-benzylpiperazin-1-yl)-2-**[**5-(3,5-dimethylpyrazol- 1-yl)-1,2,3,4-tetrazol-2-yl**]** ethanone) may represents a novel ligand for the human plasma β-FXIIa crystal. Prediction analysis carried out in this in-silico study showed that compound 4 obeys Lipinski's rule of five for oral activity. It also has a good docking orientation, a high drug-likeness score and acceptable pharmacokinetics and safety profiles. Finally, analysis of ligand movement and binding energy during molecular dynamics simulation indicated that compound 4 may be a potential ligand for β-FXIIa crystal; however these computational results must be validated both *in vitro* and *in vivo*.

References:
[1] Colman RW *et al. Blood*. 1997 **90**:3819 **[PMID: 9354649]**.
[2] Samuel M *et al. J Biol Chem*. 1992 **267**:19691 **[PMID: 1527088]**.
[3] Dementiev A *et al. Blood Adv*. 2018 **2**:549 **[PMID: 29519898]**.
[4] Gailani D *et al. Arterioscler Thromb Vasc Biol*. 2007 **27**:2507 **[PMID: 17916770]**.
[5] Müller F *et al. Cell*. 2009 **139**:1143 **[PMID: 20005807]**.
[6] Lammle B *et al. Thromb Haemost*. 1991 **65**:117 **[PMID: 1905067]**.
[7] Müller F *et al. Curr Opin Hematol*. 2008 **15**:516 **[PMID: 18695377]**.
[8] Rifaioglu AS *et al. Brief Bioinform*. 2019 **20**:1878 **[PMID: 30084866]**.
[9] https://mcule.com/
[10] Odhar HA *et al. Bioinformation*. 2019 **15**:627 **[PMID: 31787811]**.
[11] Odhar HA *et al. Bioinformation*. 2020 **16**:236 **[PMID: 32308266]**.

[12] Odhar HA *et al. Bioinformation*. 2022 **18**:304 **[**PMID: 36518132**]**.

[13] Hashim AF *et al. Bioinformation*. 2022 **18**:442. https://doi.org/10.6026/97320630018442

[14] Morris GM *et al. J Comput Chem*. 2009 **30**:2785 **[**PMID: 19399780**]**.

[15] Eberhardt J *et al. J Chem Inf Model*. 2021 **61**:3891 **[**PMID: 34278794**]**.

[16] Carbone J *et al.* Molecules, 2022 27:6884.

[17] https://pymol.org/2/

[18] https://discover.3ds.com/discovery-studio-visualizer-download

[19] https://tox-new.charite.de/protox_II/

[20] https://biosig.lab.uq.edu.au/pkcsm/prediction

[21] http://www.swissadme.ch/

[22] https://molsoft.com/mprop/

[23] Krieger E *et al. Bioinformatics*. 2014 **30**:2981 **[**PMID: 24996895**]**.

[24] Wang J *et al. J Comput Chem*. 2004 **25**:1157 **[**PMID: 15116359**]**.

[25] Maier JA *et al. J Chem Theory Comput*. 2015 **11**:3696 **[**PMID: 26574453**]**.

[26] Jakalian A *et al. J Comput Chem*. 2002 **23**:1623 **[**PMID: 12395429**]**.

[27] https://www.graphpad.com/

[28] http://www.yasara.org/macros.htm

[29] Hevener KE *et al. J Chem Inf Model*. 2009 **49**:444 **[**PMID: 19434845**]**.

[30] Lipinski CA *et al. Adv Drug Deliv Rev*. 2001 **46**:3 **[**PMID: 11259830**]**.

[31] Veber DF *et al. J Med Chem*. 2002 **45**:2615 **[**PMID: 12036371**]**.

[32] Odhar HA *et al. J Pharm Res Int*. 2021 **33**:308. https://doi.org/10.9734/jpri/2021/v33i39B32208

[33] Odhar HA *et al. J Pharm Res Int*. 2021 **33**:177. https://doi.org/10.9734/jpri/2021/v33i51B33528