



OPEN

Structure based pharmacophore modeling, virtual screening, molecular docking and ADMET approaches for identification of natural anti-cancer agents targeting XIAP protein

Firoz A. Dain Md Opo^{1,2}, Mohammed M. Rahman^{3✉}, Foyzal Ahammad⁴, Istiak Ahmed⁵, Mohiuddin Ahmed Bhuiyan² & Abdullah M. Asiri³

X-linked inhibitor of apoptosis protein (XIAP) is a member of inhibitor of apoptosis protein (IAP) family responsible for neutralizing the caspases-3, caspases-7, and caspases-9. Overexpression of the protein decreased the apoptosis process in the cell and resulting development of cancer. Different types of XIAP antagonists are generally used to repair the defective apoptosis process that can eliminate carcinoma from living bodies. The chemically synthesis compounds discovered till now as XIAP inhibitors exhibiting side effects, which is making difficulties during the treatment of chemotherapy. So, the study has design to identifying new natural compounds that are able to induce apoptosis by freeing up caspases and will be low toxic. To identify natural compound, a structure-based pharmacophore model to the protein active site cavity was generated following by virtual screening, molecular docking and molecular dynamics (MD) simulation. Initially, seven hit compounds were retrieved and based on molecular docking approach four compounds has chosen for further evaluation. To confirm stability of the selected drug candidate to the target protein the MD simulation approach were employed, which confirmed stability of the three compounds. Based on the finding, three newly obtained compounds namely Caucasicoside A (ZINC77257307), Polygalaxanthone III (ZINC247950187), and MCULE-9896837409 (ZINC107434573) may serve as lead compounds to fight against the treatment of XIAP related cancer, although further evaluation through wet lab is necessary to measure the efficacy of the compounds.

Hepatocellular carcinoma (HCC) is one of the most predominant type of primary liver cancer that has ranked fourth most common cause of cancer-related death worldwide^{1,2}. Different factors are associated with the development of liver cancer including fatty liver disease, consumption of alcohol, hepatitis B and hepatitis C virus³. Early detection rate of the cancer is low and when the symptoms appear, cancer began to spread and be difficult to treat^{4,5}. Liver transplantation, surgery, radiotherapy, chemotherapy are the most common approach to treat HCC, but in advance stage of cancer treatment failure rate are frequent⁶. To increase the survival rate early detection of liver cancer and better compounds, which are responsible to inhibit the growth of HCC is important to identify to reduce the HCC related disease.

Apoptosis is a type of programmed cell death that help to eliminate unwanted cells from multicellular organisms⁷. Apoptosis deficiency in the cells is an important cause of cancer that occur due to the IAPs. XIAP

¹Department of Biomedical Science. College of Natural Sciences, Chosun University, Chosun, South Korea. ²Department of Pharmacy, University of Asia Pacific, 74/A, Green Road, Farmgate, Dhaka 1215, Bangladesh. ³Department of Chemistry, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia. ⁴Department of Genetic Engineering and Biotechnology, Faculty of Biological Science and Technology, Jashore University of Science and Technology, Jashore 7408, Bangladesh. ⁵Department of Chemistry, Jahangirnagar University, Savar Upazila, Dhaka 1342, Bangladesh. ✉email: mmrahman@kau.edu.sa

is one of the main anti-apoptotic proteins in the IAP family capable of neutralizing caspase-9 via BIR3 domain, where the effector caspases-3 and 7 is neutralized by BIR2 domain⁸. Cancer cells can escape from drug-induced death due to defects in pro-apoptotic death regulators or the presence of highly expressed pro-survival proteins, which is one of the main reasons for failure of chemotherapy. Oncogene neutralization has shown the ability to reduce the lengthy process of chemotherapy and that is helpful to reduce the quantity and dose of drugs during cancer treatment. Therefore, XIAP based targeted would be an excellent treatment options for different cancer diseases including hepatocellular carcinoma⁹. Clinically antisense technology, SMAC-mimetics, siRNA has tried to decrease the overexpression of XIAP, but due to the neurotoxicity antisense based treatment (Example: AEG35156) were terminated in Phase-I clinical trial. The antisense based treatment process work based on reducing the XIAP mRNA level and increasing apoptotic cell death of stem cells. SMAC-mimetics are the most important compounds to neutralize the XIAP also work over the IAP families. Three amino acids present in the N terminal region of proline at the N-terminus of SMAC/Diablo are able to work over the binding groove of XIAP-BIR2 and XIAP-BIR3 domain at the same time¹⁰. Few important side effects have been observed in the case of developing of these compounds, SMAC/Diablo protein able to bind all BIR domains of several IAP most of are with cIAP1 and cIAP2 lead to the toxic effect or adverse effects due to the higher affinity of binding. To minimize adverse effects caused for higher binding affinity, antagonist molecules having the lower micromolar affinity to the XIAP-BIR2 domain is urgently need to develop^{11,12}.

In-silico drug design consist of theoretical and computational approaches can be used to identify novel hits or leads against selected biological active macromolecules¹³. Nowadays, computer aided drug design (CADD) approach like pharmacophore modeling, virtual screening, molecular docking and dynamic simulation approaches are widely used to discover, develop, and analyze drugs and similar biologically active molecules¹⁴. In CADD approach, structure and ligand-based pharmacophore model can able to identify similar active molecules against specific target protein, where binding affinity of a large scale compound with target macromolecule can be evaluate easily by in-silico molecular docking process¹⁵. The biological activity of a compound can be evaluating, whenever the compound binds with targeted macromolecule and trigger a specific response. Calculation of binding capacity of a compound was time consuming and costly in conventional drug development due to require a large-scale in-vitro and in-vivo experiment¹⁶, in that case molecular docking approach make it easier within a short time. Pharmacokinetics and pharmacology properties like absorption, distribution, metabolism, and excretion (ADME) even toxicity of a compound can predict by using computer aided drug design process¹⁷. So, this study focused mainly on computer aided drug design process like structure-based pharmacophore modeling, virtual screening, ADMET, molecular docking and dynamic simulation approaches to identify the possible natural antagonist against XIAP protein to treat the cancer.

Result and discussion

Structure-based pharmacophore modeling and virtual screening. *Pharmacophore model generation.* XIAP is a nonredundant modulator of tumor necrosis factor-related apoptosis and best-defined anti-apoptotic IAP family member that directly neutralizes caspase-9 via its BIR3 domain. Over-expression of the protein is responsible for developing different cancers⁷. Chemical that has developed targeting XIAP are mostly toxic and has adverse effect. Therefore, natural compounds identical to the previously originated antagonist can be developed as a drug instead of chemically synthesis compound. Ten (10) chemically synthesis active antagonist of XIAP (Table 1) were collected through ChEMBL and advance literature search, which were docked with XIAP protein. The best binding score found for the antagonist Hydroxythio Acetildenafil (PubChem CID: 46781908) was -6.8 kcal/mol, binding energy of other 9 molecules has shown in Table 1. Also, the interaction between XIAP protein and antagonist has provided in Table S1.

For drug design, protein 3D structure determination is necessary and nowadays the most validated structure of protein can be mining from several protein data banks or homology modeling. To identify antagonist against desire protein crystal x-ray structure of XIAP protein (PDB: 5OQW) in complex with compound 46781908 recovered and a structure-based pharmacophore model to the enzymatic cavity was generated. Ligands binding capacity to the selected XIAP protein are determined experimentally and validated through x-ray diffraction method having IC_{50} value 40.0 nM¹⁸. The overall expression can be regulated by binding of the inhibitor to the active site of XIAP protein. Sometimes proper efficacy of inhibitor against any protein might not be reliable due to the improper binding. So, the determination of active series of inhibitors should be examined for sufficient interaction to get more biological activity compared to the existing one. LigandScout4.3 essential advance molecular design software was used to generate the key chemical features based on pharmacophore model.

The different chemical features were determined, and total number was 14. Among them four were hydrophobics, one positive ionizable bond, three H bond acceptor, 5 H bond donor, 15 exclusion volume features were presented as a protein ligand complex interaction (Fig. 1). To maintenance optimum pharmacophore features, some features have omitted during the time of pharmacophore model generation.

Obtained pharmacophore features generated from the protein-ligand complex is depicted that hydrophobic interactions are predominant formed with the amino acid residues of the selected protein. HBD features have been found in several interactions with the protein, whereas nitrogen atoms in the benzene ring interacted with the THR308, ASP309, GLU314 amino acid (Fig. 2). One HBD was formed with the oxygen atom of the side chain of the amino benzene HOH523 number position. On the other site, HBD formation after ligand-protein interactions was shown as a red mark in the position of THR308, HOH556. HOH565. THR308, HOH556, which were bind to the oxygen atom, and nitrogen atom of the benzene ring bind to the HOH565. Partially two interactions such as HBD, HBA were formed also with the oxygen atom of HOH565. Positive ionizable pharmacophore features have also been found to be formed in GLU314 number position from the complex protein ligand structure.

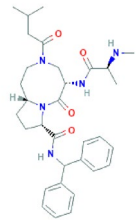
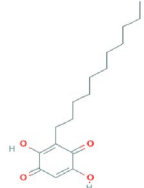
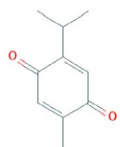
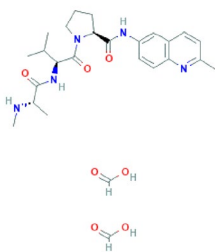
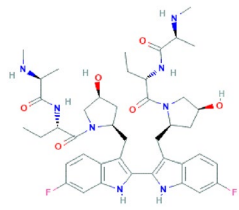
Molecule PubChem CID	Chemical Name	Molecular Formula	Chemical Structure	IC ₅₀ (nM)	Binding Affinity (kcal/mo)
25022340	AT-406	C ₃₂ H ₄₃ N ₅ O ₄		225	-6.1
3218	Embelin	C ₁₇ H ₂₆ O ₄		4100	-5
10281	Thymoquinone	C ₁₀ H ₁₂ O ₂		200	-4.4
44182275	MLS-0391037.0001	C ₂₆ H ₃₇ N ₅ O ₇		9400	-2.1
49836020	Birinapant	C ₄₂ H ₅₆ F ₂ N ₈ O ₆		126	-5.9

Table 1. (continued)

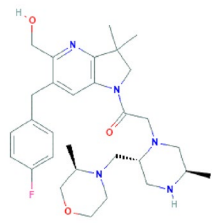
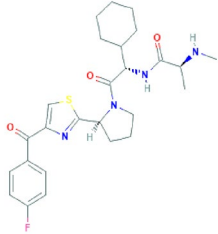
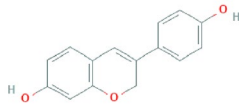
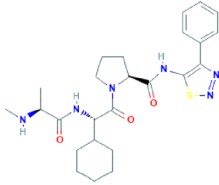
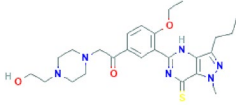
118169620	ASTX660	$C_{30}H_{42}FN_5O_3$		40	-6.5
24737642	LCL161	$C_{26}H_{33}FN_4O_3S$		1050	-6.4
219100	Idronoxil	$C_{15}H_{12}O_3$		4010	-6.1
46940575	GDC-0152	$C_{25}H_{34}N_6O_3S$		28	-6.2
46781908	Hydroxythio Acetildenafil	$C_{25}H_{34}N_6O_3S$		23	-6.8

Table 1. List of 10 known active antagonist of XIAP protein and their binding affinity towards the protein generated through molecular docking method.

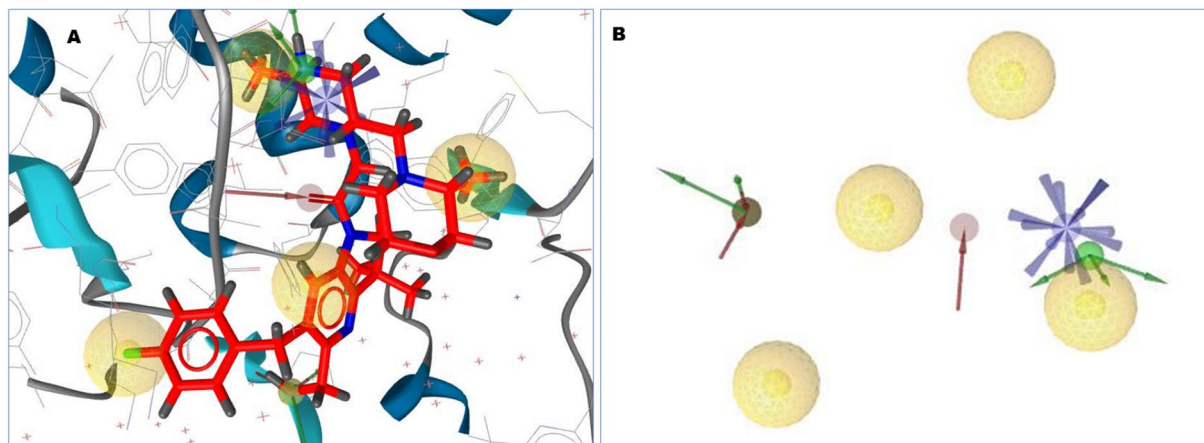


Figure 1. (A) The 3D structure-based pharmacophore model of XIAP protein in complex with 46781908 (CID) ligands derived from the X-ray derived crystal structure of XIAP protein (PDB ID: 5OQW). (B) Several pharmacophore features were generated after complex interaction, four yellow spherical shapes indicating hydrophobic interaction, one blue star shape depicting the positive ionizable with tolerance 2, three red color arrow and spherical shapes indicating H bond acceptor having tolerance 1.5, five hydrogen bond donors represented by green spherical or arrow shape have been identified within the protein–ligand complex interaction. 15 exclusion volume, which were generated during pharmacophore modeling has not showed in this schematic presentation.

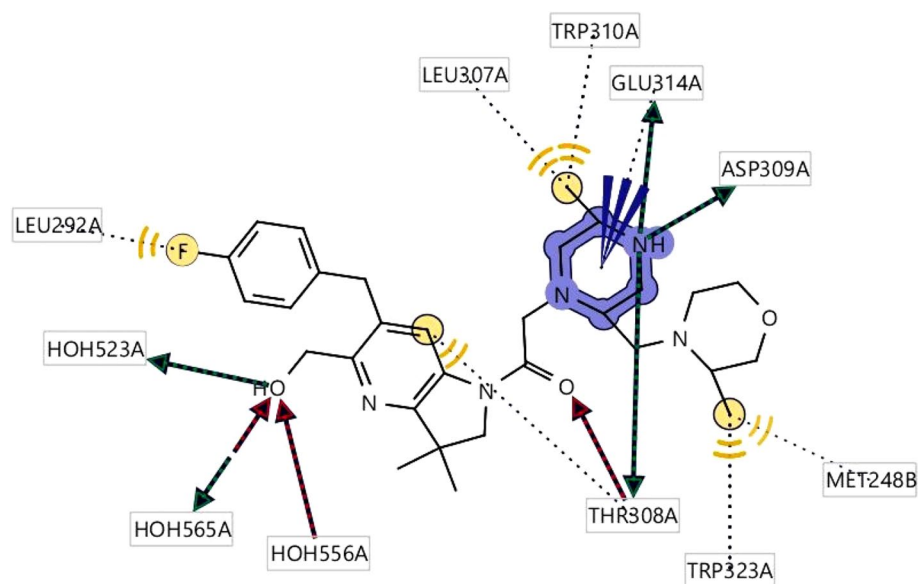


Figure 2. 2D structure obtained during the pharmacophore modeling showing the hydrophobic interaction depicting yellow color and the interaction with the amino acid residues in our selected XIAP protein. Hydrogen bond donor (HBD) features most frequently participated in ligand–protein interaction were shown in green color, whereas the red color showing the interaction of Hydrogen bond acceptors (HBA) to the oxygen, nitrogen atom of the benzene ring and its different side chains. Hydrogen atoms as well as restricted area maintain the shape and position of the binding pocket did not mention in the figure.

Pharmacophore model validation. Validation is necessary to get the authentic pharmacophore analysis as well as to evaluate the quality of the molecular model¹⁹. Structure-based pharmacophore model generated in this study was validated before database screening to evaluate whether or not our models are capable to distinguish the active compounds from decoy set. The pharmacophore model was validated by using 10 actives (Table 1) known XIAP antagonists with correspondence 5199 decoy compound (Supplementary file) obtained from the enhanced Database of Useful Decoys (DUDe). The active test set with inhibitor constant IC_{50} values were merged with the decoy compounds and an initial screening was run to validate to model. The performance of a classification model like the AUC value and EF value of the compounds was estimated from the receiver operating characteristic curve (ROC). In general, ROC is a probability graph express the performance of a classification model that can give an idea about degree of separability, where AUC is used to describe the summary of

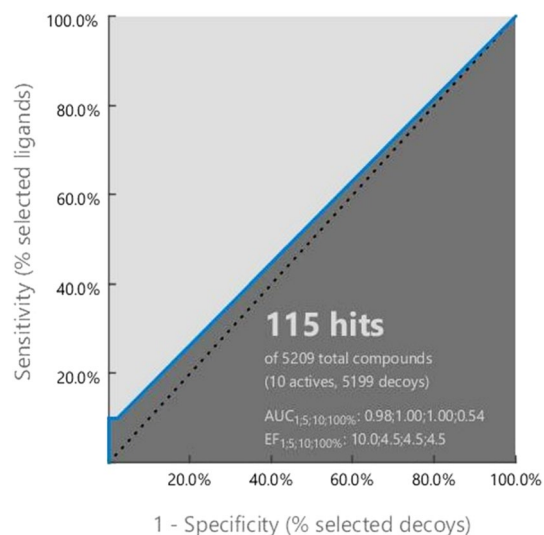


Figure 3. Receiver operating characteristic (ROC) curve generated based on the recognize ability of the active to decoy compounds of the structure-based pharmacophore model. The pharmacophore model was validated using a set of 10 XIAP active and 5199 decoy Compounds.

the model performance. A model with higher AUC value should have better predictability²⁰. The AUC value is ranging between 0 and 1, so the model whose prediction rate is 100% correct has an AUC value 1. In our validation process, the early enrichment factor (EF1%) was 10.0 with an excellent AUC (area under the ROC curve) value in 1% threshold was 0.98 (Fig. 3), which proved that our model has ability to distinguish true actives from decoy compounds.

Dataset generation for pharmacophore-based screening. Database generation is an important part for identification the best lead molecule during screening process. ZINC database is a curated collection of commercially available chemical compounds from which we can get the information regarding compound's molecular weight, chemical structure, physical and chemical properties against biological active macromolecules. It contains more than 230 million purchasable compounds in 3D format to the freely accessible website, which is ready to dock²¹. It also provides information regarding various compound from different vendor like Ambinter as a natural compound database library. To create the database for pharmacophore based virtual screening, the previously obtained pharmacophore model generated for each active compound was submitted to ZINCPharmer²². Initially, it searches hits from the ZINC database of "ZINC natural products and ZINC natural derivatives" consist millions of Drug-like, Natural Products and FDA approved drugs. A maximum of 0.5 Å RMSD from sphere centers were used as input parameters for ZINCPharmer and a total 11,000 compounds was retrieved for further screening. The database of hit compounds from the ZINCPharmer were saved and downloaded for further screening.

Pharmacophore-based virtual screening. Pharmacophore interaction features generated from the protein-ligand complex was applied to the 11,000 natural compounds. During the screening process relative pharmacophore was used as a scoring function, where all query features were used as a screening mode and maximum four (4) features have omitted. It is difficult to match all of the query features during screening process that's why some features have been omitted to increase the pharmacophore fit score. A higher score is desirable for optimum fitting with the desire environment than the compound will show better activity against the targeted macromolecules. A total seven hit compound with a fit score ranging from 94.75 to 105.31 were generated that matches all of the pharmacophore features. Usually, the pharmacophore fit value is shown the geometric fit of features to the 3D-structure-based pharmacophore model. The molecule with maximum fit score to the validated pharmacophore model should show activity against our desire XIAP protein. The compound, which remarked as hit was retrieved and saved for further evaluation.

Molecular docking based virtual screening

Binding site identification and receptor grid generation. Based on the crystal structure, the desire XIAP (PDB ID: 5OQW) protein was bounded with three ligands. So, the protein pockets have different attachment site as well as shape for binding the favorable ligand. The binding position of the complex structure was retrieved so that the binding site can further utilize during molecular docking simulation. Analysis of the protein-ligand interaction revealed a salt bridge at the position of GLU314, four pi-alkyl bonds formed to formed at the benzene ring and side chain of benzene ring by interacting with TRP323, TYR324, LYS297. Four conventional hydrogen bonds were formed between the ASP309 and THR308 position. One halogen bond has observed at VAL298 with the fluorine attached to the benzene ring. At the same time, two carbon-hydrogen bonds were formed at GLN319 with the two sides of the same benzene ring (Fig. 4).

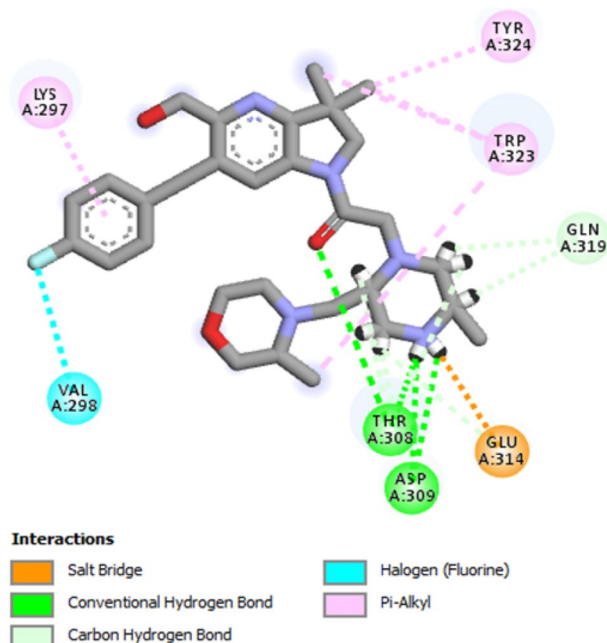


Figure 4. Depicted figure showing the binding site of the protein with the selected ligand complex (PDB ID: 5OQW). Several types or interaction were shown by color with different amino acids residues.

ZINC ID	Compound name	Docking Score (kcal/mol)	Pharm-Fit Score	Source of the compounds
ZINC77257307	Caucasicoside A	-8.0	95.53	Helleborus caucasicus
ZINC1070004335	Venturicidin B	-7.8	94.75	Streptomyces aureofaciens
ZINC247950187	POLYGALAXANTHONE III	-7.6	95.70	Polygala tenuifolia
ZINC107434573	MCULE-9896837409	-6.9	95.72	Unknown

Table 2. Docking score with XIAP protein, pharmacophore fit score and the source of the top four selected compounds.

Molecular docking. Molecular docking is an important part in drug design process, which is carried out in the study to evaluate the binding ability of the hits compounds to the target XIAP protein. XIAP monomeric proteins has decorated with two active chains such as A, B attached with three ligands (A4E, NA, ZN) with the protein. The protein was prepared and a receptor grid with box dimeter X = 30.06, Y = -4.19 and Z = -22.94 was generated depend on the previously obtained binding site to the chain A.

The specific number of drugs like hit (7) compounds were docked with XIAP by utilize PyRx tools Autodock vina to evaluate their binding capacity, which satisfied the characteristics of the pharmacophore model²³. Among them, four compounds ZINC77257307, ZINC1070004335, ZINC247950187 and ZINC107434573 with binding affinity -8.0 kcal/mol, -7.8 kcal/mol, -7.6 kcal/mol, and -6.9 kcal/mol (Table 2), respectively shown better binding affinity than the XIAP antagonist CID: 46781908 (-6.8 kcal/mol), that has used during main pharmacophore model generation. The binding affinity for all of hit shown in Table S2. Interestingly, compound which have higher pharmacophore fit score (Table 2) found higher docking score, and higher docking score indicates the better binding to the desire protein.

Interpretation of protein-ligands interactions. Here, it is observed that compound that have better pharmacophore fit score gained better binding affinity ZINC77257307 (-8.0 kcal/mol), ZINC1070004335 (-7.8 kcal/mol), ZINC247950187 (-7.6 kcal/mol) and ZINC107434573 in contrast to the compound CID: 46781908 (-6.8 kcal/mol) showed in Table 2. ZINC77257307 formed nine van der Waals interactions with LYS297, LEU292, GLY304, TYR305, TYR324, GLY306, LEU307, GLN319, LYS311, three conventional hydrogen bonds with ASP309, THR308, GLU314, one pi donor hydrogen bond with TRP310, one alkyl bond with LYS299, and two Pi-alkyl bonds with TRP323 with desire XIAP protein (Figs. 5 and 6).

In the case of ZINC1070004335 van der Waals bonds were predominantly formed with LYS334, TYR329, ALA263, PHE301, PRO257, ASN255, ARG268, ASP264, TYR265. The position of TRP317, GLN333, GLU332 acquired three conventional hydrogen bonds, one carbon-hydrogen bond at HIS302 and one alkyl bond with ARG286 position (Figs. 5 and 6).

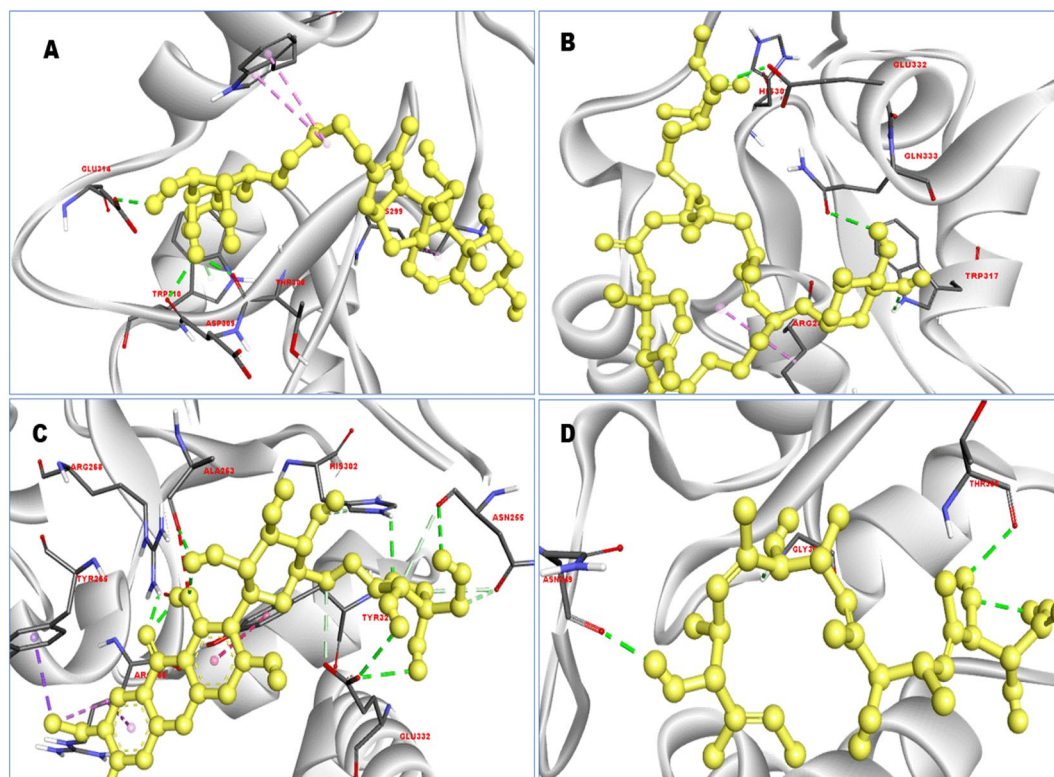


Figure 5. 3D interaction between the protein–ligand complex. Here, figure (A) ZINC77257307, (B) ZINC1070004335, (C) ZINC247950187, and (D) ZINC107434573 showing the ligand contact with the protein XIAP after molecular docking.

For the compound ZINC247950187, the number of van der Waals interaction has decreased but increased other types of bond such as 5 van der Waals interaction at position ASP264, GLN333, PRO257, PHE301, LYS328 has found to formed, which are less than the previous two compounds. Nine conventional hydrogen bonds with ALA263, ARG268, HIS302, GLU332, ASN255 position with four hydrogen atom, one pyridine, and one benzene ring with ARG268 and ALA263 position has found to formed (Figs. 5 and 6). For, ZINC107434573 it has observed to formed five carbon-hydrogen bonding in the position of ASN255, HIS302, GLU332, one pi-sigma with TYR265, one pi-pi t shaped bond with TYR329, one alkyl and pi alkyl bond with ARG286. The compounds also produced 11 van der Waals bond with GLY305, PRO251, TYR324, PHE250, TRP323, LYS322, GLN319, GLU314, TRP310, ASP309, LEU307, and 3 conventional hydrogen bonds with ASN249, GLY306, THR308 position. The protein–ligand interaction mode for all the four compounds has listed in Table 3.

Pharmacophore features analysis. Pharmacophore is a group of steric and electronic features that confirm optimal supramolecular interactions during virtual screening on large scale compound databases. It is powerful and more efficient method than molecular docking that can find molecules against specific target to induce or inhibit the macromolecular activity²⁴. The compound which has similar or relevant properties should show the same or better activity like to the query compound. In this study, the top four compounds (based on docking score) ZINC77257307, ZINC1070004335, ZINC247950187 and ZINC107434573 pharmacophore features was analyzed and compared with the features of the antagonist 46781908. All of the compounds have better pharmacophore properties than the antagonist CID: 46781908, so these compounds should be effective to our target protein. The pharmacophore feature found of the four compounds has shown in Fig. 7.

Absorption, distribution, metabolism and excretion (ADME) and toxicity test Analysis of ADME properties.

After administration of the drug through any route to the human body or in the animal model, it undergoes the absorption, distribution, metabolism, excretion resulting active or passive transport to the target site²⁵. Interaction with the target biological macromolecules might produce desirable or undesirable pharmacological effect. Drug design is a step-by-step evaluation process and lacking the evaluation my reason for rejection of the drug, which is costly for any companies. The bioavailability of a drug depends on the safety and efficacy, lack of safety and efficacy are the main cause of drug failure, which are mainly depend on the ADME properties. Here, we evaluate the ADME properties of the selected four compounds by using in silico SwissADME server to see the pharmacokinetic properties such as lipophilicity, water-solubility, drug-likeness, medicinal chemistry of the compounds²⁶. The lipophilicity of compounds means they can easily diffuse through the cell membrane; hence the oral preparation is not suitable. Moreover, an injectable dosage form may be a

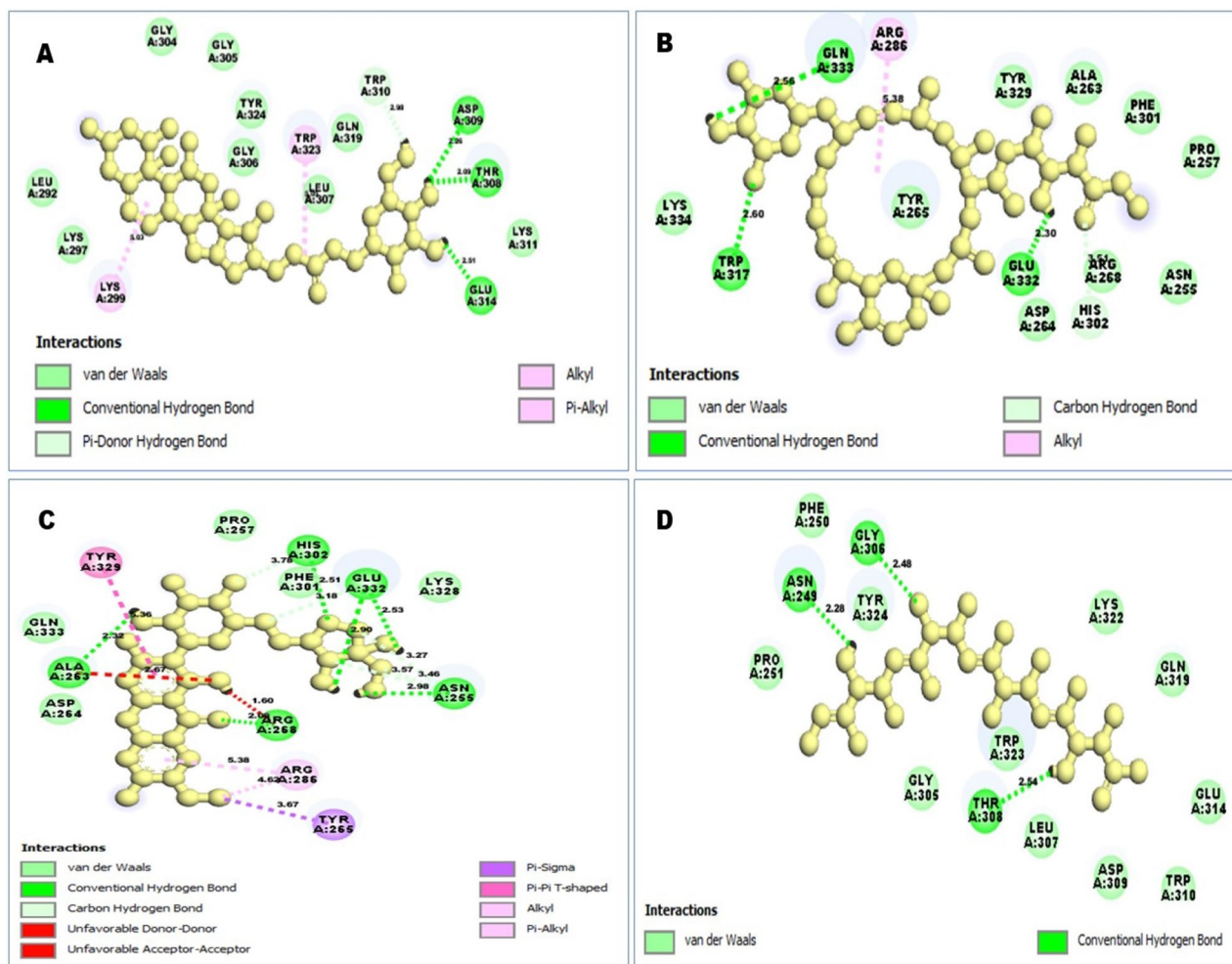


Figure 6. 2D interaction between the protein–ligand complex. Here, figure (A) ZINC77257307, (B) ZINC1070004335, (C) ZINC247950187, and (D) ZINC107434573 showing the ligand contact with the protein XIAP after molecular docking. Different bonds types were described by different colors such as blue, red, purple, light pink, deep pink, green and sky blue.

Compounds ID	Conventional hydrogen bonds	Pi donor hydrogen bond	Alkyl bond	Carbon hydrogen bond	Unfavorable donor-donor bond	Unfavorable acceptor-acceptor	Pi-alkyl bond
ZINC77257307	ASP309 (2.26 Å), THR308 (2.09 Å), GLU314 (2.51 Å)	TRP310 (2.98 Å)	LYS299 (5.03 Å)				TRP323 (4.36 Å, 5.03 Å)
ZINC1070004335	TRP317 (2.6 Å), GLN333 (2.56 Å), GLU332 (2.30 Å)		ARG286 (5.38 Å)	HIS302 (3.51 Å)			
ZINC247950187	ALA263 (2.32 Å), ARG268 (2.08 Å), HIS302 (2.51 Å), GLU332 (2.53 Å, 2.90 Å), ASN255 (3.46 Å, 3.23 Å, 2.98 Å, 3.57 Å)	Pi-sigma TYR265 (3.67 Å)	Pi-pi T shaped TYR329 (5.36 Å)	ASN255, HIS302 (3.78 Å), GLU332 (3.18 Å)	ARG268 (1.60 Å)	ALA263 (2.67 Å)	ARG286 (4.62 Å, 5.38 Å)
ZINC107434573	ASN249 (2.28 Å), GLY306 (2.48 Å), THR308 (2.54 Å)						

Table 3. Interaction result between the selected 4 ligands in complex with the protein XIAP.

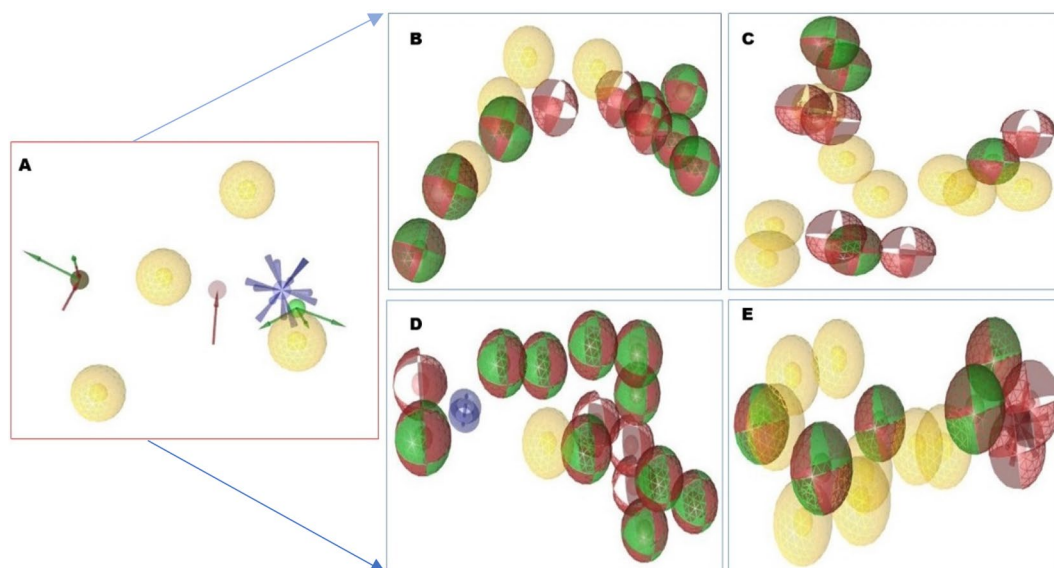


Figure 7. Pharmacophore features generated from the four selected compounds attach to the desire XIAP protein. Ligands attach to the protein (A) 46781908 had four hydrophobic (yellow), one positive ionizable (blue star), three H bond acceptor (red), and five hydrogen bond donors (green) pharmacophore features. Comparing to this most of our selected compounds (B) ZINC77257307, (C) ZINC1070004335, (D) ZINC247950187 and (E) ZINC107434573 have better pharmacophore features than antagonist 46781908.

Properties	Parameters	ZINC77257307	ZINC1070004335	ZINC247950187	ZINC107434573
Physico-chemical properties	MW (g/mol)	606.74 g/mol	706.95	568.48	454.60
	Heavy atoms	43	50	40	32
	Arom. heavy atoms	0	0	14	0
	Rotatable bonds	7	9	6	12
	H-bond acceptors	10	10	15	7
	H-bond donors	7	4	9	6
	Molar Refractivity	157.51	196.06	131.62	127.97
Lipophilicity	Log P _{o/w}	1.58	4.60	3.53	2.87
Water solubility	Log S (ESOL)	Soluble	Poorly soluble	Soluble	Soluble
Pharmacokinetics	GI absorption	Low	Low	Low	Low
Drug likeness	Lipinski, Violation	2	1	2	1
Medi. chemistry	Synth. accessibility	7.82	9.66	5.95	5.96

Table 4. List of pharmacokinetic properties (physico-chemical, lipophilicity, water solubility, drug likeness, and medicinal chemistry) of the selected 4 compounds.

better option to get a rapid onset of action as the gastrointestinal absorption is low. The ADME properties of the selected four compounds has shown in Table 4.

Analysis of toxicity. For better lead compound selection, in-silico toxicity measurement is an important procedure before drug candidate undergo clinical trial²⁵. Computational based in-silico toxicity measurement has been widely used due to their accuracy, rapidity, accessibility, which can provide information about any synthesis or natural compounds. To identify the toxicity and adverse effect of the selected four compounds, we used both freely accessible TEST tool²⁷, and ProTox-II server²⁸. Each software was used to evaluate several toxicological parameters such as acute toxicity, hepatotoxicity, cytotoxicity, carcinogenicity, mutagenicity, immunotoxicity, and the result was achieved based on predicted median lethal dose (LD₅₀) in mg/kg weight (Table 5). According to the ProTox-II server compound ZINC107434573 and ZINC247950187 both were belonging to class 4, LD₅₀ range from 300 to 2000 mg/kg, these would be harmful in case of oral delivery. For ZINC1070004335, the LD₅₀ value was less than < 50 mg/kg, so oral intake might be toxic or fetal, which were belongs to the toxicity class two in ProTox-II. ZINC77257307 was in the toxicity class 6 and the LD₅₀ value was also more than 5000 mg/kg and it is also nontoxic but having some immunotoxicity.

Endpoint	Target	ZINC77257307	ZINC1070004335	ZINC247950187	ZINC107434573
Organ toxicity	Hepatotoxicity	Inactive	Inactive	Inactive	Inactive
Toxicity endpoints	Carcinogenicity	Inactive	Inactive	Inactive	Inactive
	Immunotoxicity	Active	Active	Active	Inactive
	Mutagenicity	Inactive	Inactive	Inactive	Inactive
	Cytotoxicity	Active	Inactive	Inactive	Inactive
	LD ₅₀ (mg/kg)	6000	50	1469	665
	Toxicity class	6	2	4	4
Tox21-Nuclear receptor signaling pathways	Androgen Receptor (AR)	Inactive	Inactive	Inactive	Inactive
	Aryl hydrocarbon Receptor (AhR)	Inactive	Inactive	Inactive	Inactive
Tox21-Stress response pathway	Heat shock factor response element	Inactive	Inactive	Inactive	Inactive
Fathead minnow LC ₅₀ (96 h)	mg/L	0.81	N/A	N/A	9.95E-02
48-h <i>Daphnia magna</i> LC ₅₀	mg/L	9.11	11.61	63.28	10.76
Developmental toxicity	value	0.47	0.53	0.39	0.36
Oral rat LD ₅₀	mg/kg	53.32	682.02	727.46	1127.01
Bioaccumulation factor	Log10	1.57	N/A	1.20	-0.33

Table 5. List of toxicity properties (Organ Toxicity, Toxicity Endpoints, Tox21-Nuclear receptor signaling pathways, Tox21-Stress response pathway, Fathead minnow LC₅₀ (96 h), Developmental toxicity, Oral rat LD₅₀, Bioaccumulation factor), of the selected 4 compounds.

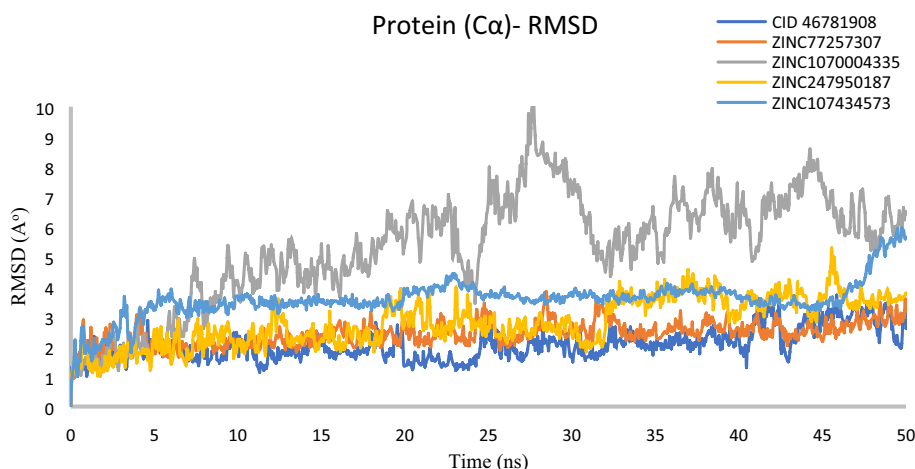


Figure 8. RMSD values extracted from Ca of the protein–ligand docked complexes, viz CID: 46781908 (Gray), ZINC77257307 (orange), ZINC1070004335, ZINC247950187 (Gold), ZINC107434573 (Blue), with respect to 50 ns simulation time.

Molecular dynamics (MD) simulation

MD simulation is used to explore the binding stability of protein–ligand docking complexes²⁹. The MD simulation also provide information regarding intermolecular interaction within a reference time. Herein, the complexes docking file of selected four natural compounds and one reference antagonist bind with XIAP protein were analyzed by utilize MD simulation approaches to confirm the stability and intermolecular interactions between protein and molecules against 50 ns time interval. Trajectories of MD were extracted by utilize SID in Maestro-Desmond interface and the simulation result has described based on RMSD, RMSF and Protein–Ligand (P–L) interaction mapping.

RMSD analysis. Root mean square deviation (RMSD) in MD simulation is used to measure the average distance generated by displacement of a selected atoms for a specific time frame with respect to a reference time frame¹⁴. Initially, RMSD value of specific protein structure such as Ca, backbone, sidechain and heavy atoms are computed, after that RMSD of the protein fit ligand from all the time frames during the reference time (in our case 50 ns) is calculated. RMSD for frame x can be calculated from the following equation (Eq. 1).

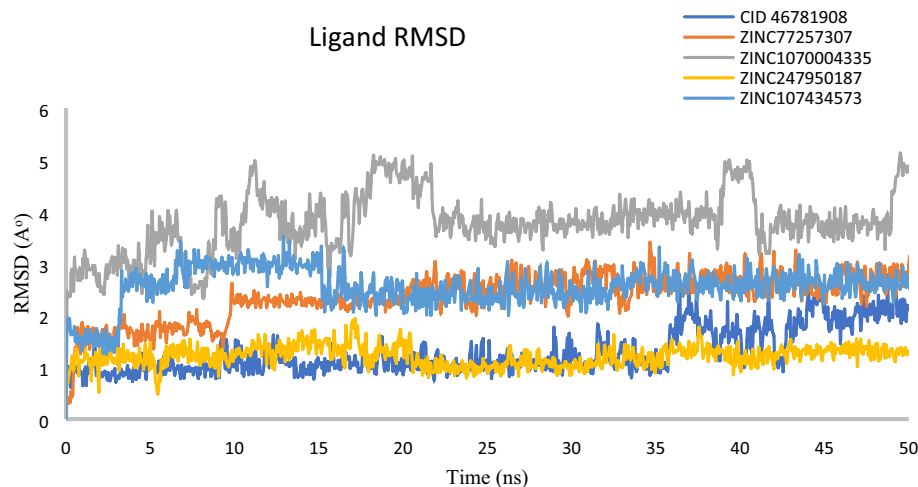


Figure 9. RMSD values extracted from protein fit ligand of the protein–ligand docked complexes, viz CID: 46781908 (Gray), ZINC77257307 (orange), ZINC1070004335, ZINC247950187 (Gold), ZINC107434573 (Blue), with respect to 50 ns simulation time.

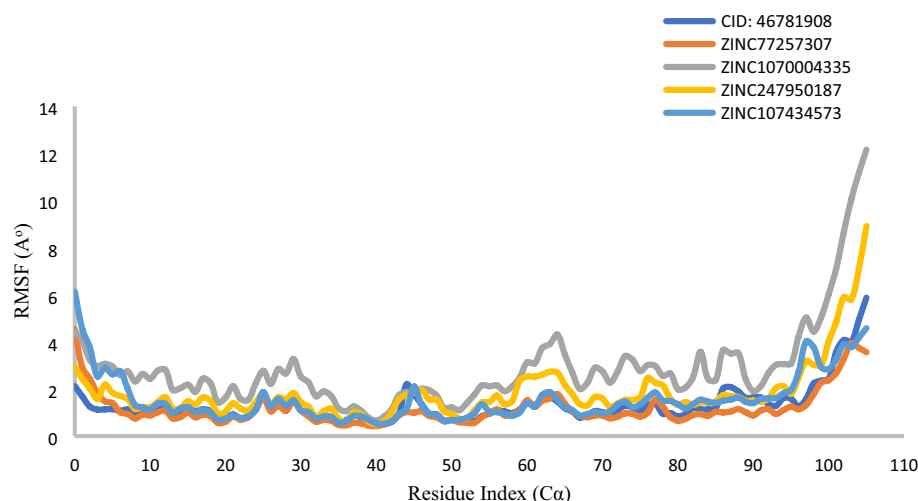


Figure 10. RMSF values extracted from protein Cα atoms, viz CID: 46781908 (Gray), ZINC77257307 (orange), ZINC1070004335, ZINC247950187 (Gold), ZINC107434573 (Blue), with respect to 50 ns simulation time.

$$RMSD_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (r'_i(t_x) - r_i(t_{ref}))^2} \quad (1)$$

Here, N can define as the number of atoms in the atom selection; t_{ref} is the reference time, and r' define the location of the selected atoms within the frame x after superimposing on the reference frame, t_x expressed the recoding intervals of x .

RMSD of protein. Based on RMSD result, it can be determined that simulation has equilibrated or not. Fluctuations between 1–3 Å within a reference protein structure is perfectly acceptable, where much larger value indicate large conformational change of the protein and the system is not stable. Analysis of our four protein–ligand docking complex found Cα atoms of XIAP showed acceptable fluctuations < 3 Å, except in XIAP–ZINC1070004335 complex. The compound ZINC1070004335 exhibited an extended variation 5.1 Å and a maximum fluctuation 8.81 Å (between 26 and 28 ns) observed during 50 ns simulation run (Fig. 8). From the data we can assume that XIAP undergo protein conformation changes influence by the binding of ZINC1070004335.

RMSD of ligand. Furthermore, analysis of RMSD from the data obtained from protein fit ligands showed minimum variations (< 3 Å), except for ZINC1070004335 complex (> 4.84 Å) at the end of 50 ns simulation interval (Fig. 9).

However, it is necessary to mention that three natural compound ZINC77257307, ZINC247950187, ZINC107434573 and the chemical antagonist CID: 46781908 docked with XIAP exhibited equilibrium within 10 to 15 ns, with an exception for XIAP- ZINC1070004335 complex, which has tried to exhibited state of equilibrium after 25 ns.

RMSF analysis. The Root Mean Square Fluctuation (RMSF) is necessarily important for characterization and determination the local conformational change in the protein chain and the compounds utilized as ligand¹⁴. The RMSF of the residue i can be calculated by the following equation (Eq. 2).

$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t=1}^T \langle (r'_i(t) - (r_i(t_{ref})))^2 \rangle} \quad (2)$$

Here, T can define as the trajectory time; t_{ref} is the reference time, and r' define the location of the selected atoms within the residue i after superimposing on the reference frame, and $\langle \rangle$ expressed the average of the square distance taken over residue b .

The local structural fluctuations of XIAP protein in complex with natural compound were calculated by using the deviations contributed by residues index α . Interestingly, residues for all protein have found a minimum RMSF values, except in N-terminal minimum 4.18 Å to maximum 12.16 Å (Fig. 10). So, analysis of RMSF and RMSD value for all protein–ligand complex supported the combined screened potential compounds except compound ZINC1070004335 against XIAP protein.

Protein–ligand interaction analysis. Protein–ligand contact occur through hydrogen bonding, ionic bonding, water bridges and hydrophobic bonding play an important role to design a molecule into effective drug. Protein–ligand contact occurred to XIAP protein and selected four natural compounds were analyzed from the MD trajectories by using default parameters of Desmond module. All the selected natural compounds ZINC77257307, ZINC247950187, ZINC107434573 and ZINC1070004335 showed tangible contact with most of the protein residues (Fig. 11), i.e., ASN 249 and LYS 299 except compound ZINC247950187 but maintain optimum binding with other residue like GLU 332 of the XIAP protein. Interestingly, the residue GLU332 were also found in the respective ligand–protein complex docking structure of the selected compounds (Table 3). Moreover, the four compounds screened through different filtering process exhibited considerable intermolecular interaction.

Conclusions

In this in-silico approach, three new natural compound ZINC77257307, ZINC247950187, and ZINC107434573 have discovered that may be able to induce apoptosis through freeing up caspases. These selected compounds have a higher binding affinity ranging between -6.9 and -8.0 kcal/mol with desire XIAP protein. Based on the in-silico toxicity test, they have found a lower toxicity, and ADME analysis determined the easily absorbability to the tissue site, which is readily fat soluble. Initially, a structure-based pharmacophore model was developed following by virtual screening, molecular docking, ADMET analysis and MD simulation. Four compounds were reached at the last step until MD simulation, but stability of the compound ZINC1070004335 was unfavorable to the protein XIAP in MD simulation, which has rejected. The top three natural compound that exist during the A-to-Z virtual screening process may serve as lead molecules to fight against cancer.

Material and methods

Structure-based pharmacophore modeling and virtual screening. *Structure-based pharmacophore modeling.* The active antagonists of X-linked inhibitor of apoptosis protein (XIAP) were generated by collecting all available target annotations from ChEMBL (on the basis of high-confidence activity data) and extensive literature search³⁰. In order to generate a structure-based pharmacophore models, the 10 active antagonists (Table 1) obtained from ChEMBL (<https://www.ebi.ac.uk/chembl/>) and literature search have docked with the XIAP (PDB ID: 5OQW) protein by utilizing the PyRx AutoDock Vina option based on scoring functions. The best compound with highest binding affinity (kcal/mol) was selected for structure-based pharmacophore modeling. Top scoring compound in complex with XIAP protein was used to interact with the natural compounds resulting retrieval of hits. LigandScout 4.3 advance software was used to produce a structure-based pharmacophore model³¹. This advanced software works by making the interaction between inhibitors and critical amino acids of the active sites in our target protein. This software interprets ligand–receptor interaction with different pharmacophore features such as hydrogen bond donor, charge transfer, hydrophilic and hydrophobic regions, and hydrogen bond acceptors. We have detected other features using stepwise algorithms such as the number of aromatic rings, hybridization state, the pattern of binding, the distance of receptor molecules³². For identify better and optimum compound structure, we deleted hydrophilic properties from the protein by using ligand scout, excluded or included features to the active site necessary to maintain sterically circumference of the macromolecule.

Pharmacophore model validation. Pharmacophore validation helps to evaluate the potential features of active and inactive compounds usually can be obtained from specific protein–ligand interaction³³. Pharmacophore model generated from the protein–ligand complex was validated for its performance to distinguish active compounds from decoys by screening a set of 10 known actives and correspondence 5199 decoy correspondence obtained from DUD-E decoys database³⁴. The ten active antagonists obtained from ChEMBL (<https://www.ebi.ac.uk/chembl/>) was remarked as “active” against XIAP, which was experimentally validated that’s why the compounds have chosen for further experiment. The database from DUD-E was converted in the .ldb format before

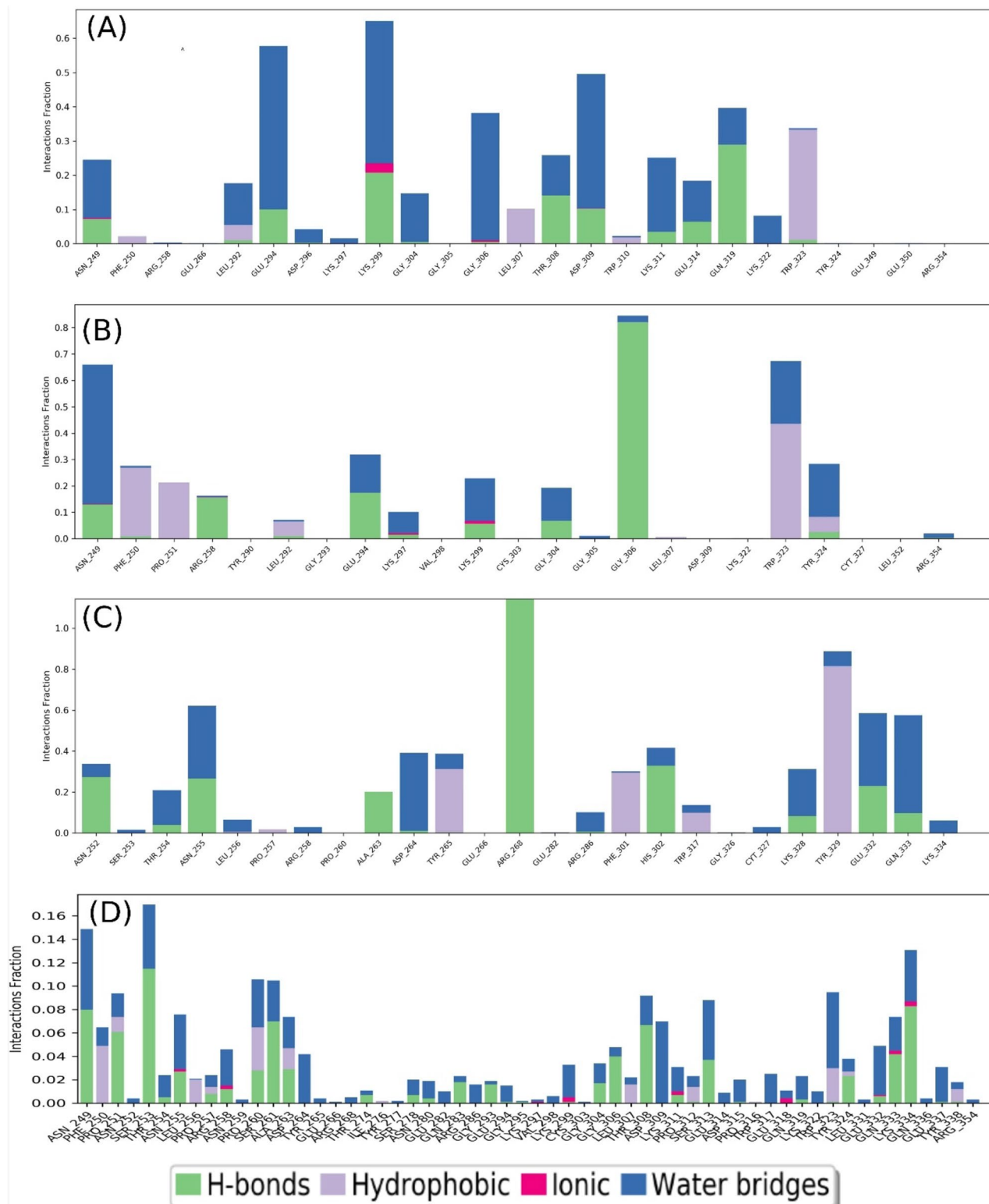


Figure 11. Protein–ligand contact mapping for XIAP with potential natural compounds, i.e. (A) ZINC77257307, (B) ZINC1070004335, (C) ZINC247950187, and (D) ZINC107434573 extracted from 50 ns MD simulations.

screening by the “create screening database” menu of LigandScout 4.3³¹. Here, it has been assessed the quality of the structure-based model based on the GH score and early enrichment factor (EF).

Dataset Generation for pharmacophore-based screening. It can be identified the structurally novel and active molecules by the completion of virtual screening depending on the generated pharmacophore model³⁵. ZINC (<https://zinc.docking.org/>) is a freely available chemical database, which is being utilized to identify the potential lead compounds²¹. Compounds from the database can be searched depending on the structure, name of the compound, or using the chemical smile ID. Physical and chemical properties such as 2D and 3D structure determination, the boiling point, the melting point were analyzed. Molecular weight, crystal structure, and biological application information can be also obtained of the desired compound. In the case of the desired compound, it has given priority, the compound having the most similar features matches the required pharmacophore features and can easily interact with our target protein. It has been chosen the possible hit compounds whose maximum features were matched to query pharmacophore. For the study, initially we have screened ZINC natural product library by using ZINCPharmer (<http://zincpharmer.csb.pitt.edu/pharmer.html>) server for target XIAP based on pharmacophore features²².

Pharmacophore-based virtual screening. Database generated from ZINCPharmer was subject to screen against the validated structural based pharmacophores features. LigandScout 4.3 advanced help to make and getting a 3D model in case of protein–ligand interaction and worked to change compounds into the specific (idb) file format. These compounds were passed directly into the database list for quick pharmacophore features based virtual screening. The securitizing has been completed by selecting relative pharmacophore-fit as a getting few functions based on omitting some features more than 2. Fitted hit compounds were arranged based on the pharmacophore fit score and subject to further validation.

Molecular docking based virtual screening

Protein and ligand preparation. Protein preparation in computational biology is a process by which macromolecular structure are converted into more suitable form for a computational experiment³⁶. Prior to docking crystal structures of protein is needed to prepare which are not part of the x-ray crystal structure refinement process such as addition and optimization hydrogen bonds, remove atomic clashes, and perform other operations. For the study, desired 3D structure of XIAP protein was obtained from the protein data bank (PDB ID: 5OQW), which was determined experimentally and validated through X-ray diffraction method having resolution 2.31 Å and R-value free score 0.246 that is significantly less than standard value 0.25. The X-ray crystallography structure of our desire protein was prepared by the following steps (i) water, metal ion and cofactors were removed, (ii) polar hydrogen bond was added and non-polar H was merged and (iii) gasteiger charges were calculated by using AutoDockTools (ADT) 1.5.6^{36,37}. Selected hit compounds retrieved from LigandScout was prepared and the energy was minimized and bond angle was optimized by default of the Universal Force Field (UFF) for each ligand.

Active site identification and grid generation. Binding of ligand or drug molecules to the specific site of protein is the key strategy to treat a particular disease. Improper attachment of ligand may show several side effects in the body with higher possibilities of toxicities also³⁸. These bininding affinities depend on several features H bond donars, hydrophobic or hydrophilic interaction, ionization, chelation of zinc compound. In the study, we used BIOVA Discovery Studio Visualizer Tool 16.1.0 to find the binding site of our desire protein. Moreover PrankWeb (<https://prankweb.cz/>) server was used to analyze the all probable binding site of the desire protein structure³⁹. The server utilizes novel machine learning-based method for prediction of ligand binding sites from protein structure. Receptor grid was generated after selection of the active site of protein by using the PyRx software.

Molecular docking. Selected hits compounds obtained by pharmacophore screening were subject to molecular docking, which was carried out by PyRx virtual screening software. In computational biology, PyRx is utilize as a virtual screening software that has identified many potential drug candidate against several diseases⁴⁰. The software includes both AutoDock and AutoDock Vina with the Lamarckian genetic algorithm (LGA) as scoring function. This study used PyRx tools AutoDock Vina to proceed the molecular docking interactions. Resultant docked compound with better binding affinity (kcal/mol) were retrieved and visualized by using BIOVA Discovery Studio Visualizer Tool 16.1.0.⁴¹

Absorption, distribution, metabolism and excretion (ADME) and toxicity test

Absorption, distribution, metabolism and excretion (ADME). Evaluation of Absorption, Distribution, Metabolism and Excretion (ADME) properties is one of the major criteria before developing molecule into a drug^{29,42}. Previously many drug candidate could not fit the clinical trial demand and so the computer-based prediction is important for the early stage of prediction. Physicochemical properties, hydrophobicity, lipophilicity, gastrointestinal environment, blood brain barrier are directly affected by the ADME profile before the excretion of drug from body through urine and feces⁴³. The freely accessible Swiss-ADME (<http://www.swissadme.ch/>) server was used to evaluate the ADME properties such as solubility profile, GIT absorption, bioavailability profile of the selected compounds²⁶.

Toxicity test. Computational based approach has made it possible to measure the toxicity by in silico methods for accessing the safety profile of the desired compounds. Otherwise, these compounds may show the harmful effect on human and animals. Toxicity profile can evaluate and determine the mutagenicity, carcinogenicity, LD₅₀ value, immunotoxicity in both quantitatively and qualitatively⁴⁴. Toxicity Estimation Software Tool (TEST) is freely accessible software, which was used in this study to estimate toxicity of our compounds without requiring any other external programs²⁷. The toxicity estimator TEST tools are being used for selective molecules that is work based on Quantitative Structure–Activity Relationships (QSARs) methodologies. Additionally, ProTox-II (http://tox.charite.de/protox_II) server was used to determine the toxic effect of the selected four compounds²⁸. Different toxicological pathways including nuclear receptor signaling pathways, stress response pathways can be obtained from this site.

Molecular dynamics (MD) simulation

Explicit-solvent MD simulation. In order to further evaluate of protein toward the binding mode of our candidate molecules, the best poses obtained from re-docking studies were evaluated through 50 ns molecular dynamics (MD) simulations for measuring the complex stability. The MD simulations was carried out by using Desmond module in Schrödinger Release 2020-3 (Academic version) suite under Linux environment^{1,45}. The complex protein–ligand interaction was first solvated with simple point charge (SPC) water model with boundary condition orthorhombic box shape. Buffer box calculation method with box distance 15 Å (a = 5 Å, b = 5 Å, and c = 5 Å) on both side for all of the complex's atoms has assigned. The system was neutralized by adding Na⁺ and Cl⁻ with a salt concentration 0.15 M. NPT ensemble was performed at constant pressure (1.01325 bar) and temperature (300 K) with recoding intervals 50 ps with energy 1.2, where OPLS-2005 force field was utilized to carried out the MD simulation.

Post-dynamics trajectory analysis. The trajectories generated after completing the MD simulation were further analyzed by using Simulation Interaction Diagram (SID) of Desmond module in Schrödinger package. Based on the simulation's trajectories, the stability of the ligand–protein complexes was determined according root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF) and Protein–Ligand contacts.

Received: 11 December 2020; Accepted: 5 February 2021

Published online: 18 February 2021

References

- Jiang, Y., Han, Q. J. & Zhang, J. Hepatocellular carcinoma: Mechanisms of progression and immunotherapy. *World J. Gastroenterol.* **25**, 3151–3167 (2019).
- Koduru, S. V. *et al.* Non-coding RNAs in various stages of liver disease leading to hepatocellular carcinoma: differential expression of miRNAs, piRNAs, lncRNAs, circRNAs, and sno. *Sci. Rep.* **8**, 7967 (2018).
- Waller, L. P., Deshpande, V. & Prysopoulos, N. Hepatocellular carcinoma: a comprehensive review. *World J. Hepatol.* **7**, 2648–2663 (2015).
- Samad, A. *et al.* Computational assessment of MCM2 transcriptional expression and identification of the prognostic biomarker for human breast cancer. *Heliyon* **6**, e05087 (2020).
- Rafi, J. H. *et al.* High expression of bone morphogenetic protein 1 (BMP1) is associated with a poor survival rate in human gastric cancer, a dataset approaches. *Genomics* **113**, 1141–1154 (2020).
- T. R., S. J., M., M. F., C. & I. L., W. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 6859–6887 (2001).
- Gao, C. *et al.* X-linked inhibitor of apoptosis inhibits apoptosis and preserves the blood-brain barrier after experimental subarachnoid hemorrhage. *Sci. Rep.* **7**, 1–12 (2017).
- Obexer, P. & Auserlechner, M. X-linked inhibitor of apoptosis (XIAP)—a critical death-resistance regulator and therapeutic target for personalized cancer therapy. *Front. Oncol.* <https://doi.org/10.3389/fonc.2014.00197> (2014).
- Fadó, R. *et al.* X-linked Inhibitor of Apoptosis Protein negatively regulates neuronal differentiation through interaction with cRAF and Trk. *Sci. Rep.* **3**, 1–11 (2013).
- Lukacs, C. *et al.* The structure of XIAP BIR2: understanding the selectivity of the BIR domains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1717–1725 (2013).
- Ardecky, R. J. *et al.* Design, synthesis and evaluation of inhibitor of apoptosis protein (IAP) antagonists that are highly selective for the BIR2 domain of XIAP. *Bioorg. Med. Chem. Lett.* **23**, 4253–4257 (2013).
- Kester, R. F. *et al.* Optimization of benzodiazepinones as selective inhibitors of the X-linked inhibitor of apoptosis protein (XIAP) second baculovirus IAP repeat (BIR2) domain. *J. Med. Chem.* **56**, 7788–7803 (2013).
- Larvol, B. L. & Wilkerson, L. J. In silico drug discovery: tools for bridging the NCE gap. *Nat. Biotechnol.* **16**(Suppl), 33–34 (1998).
- Samad, A. *et al.* Designing a multi-epitope vaccine against SARS-CoV-2: an immunoinformatics approach. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1792347> (2020).
- Meng, X.-Y., Zhang, H.-X., Mezei, M. & Cui, M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.* **7**, 146–157 (2011).
- Rahman, S. M. M. *et al.* Anti-inflammatory, antinociceptive and antidiarrhoeal activities of methanol and ethyl acetate extract of *Hemigraphis alternata* leaves in mice. *Clin. Phytosci.* **5**, 16 (2019).
- Kalyaanamoorthy, S. & Chen, Y. P. P. Structure-based drug design to augment hit discovery. *Drug Discov. Today* **16**, 831–839 (2011).
- Ward, G. A. *et al.* ASTX660, a novel non-peptidomimetic antagonist of cIAP1/2 and XIAP, potently induces TNF α -dependent apoptosis in cancer cell lines and inhibits tumor growth. *Mol. Cancer Ther.* **17**, 1381–1391 (2018).
- Temml, V. *et al.* Discovery of the first dual inhibitor of the 5-lipoxygenase-activating protein and soluble epoxide hydrolase using pharmacophore-based virtual screening. *Sci. Rep.* **7**, 1–8 (2017).
- Tai, W. *et al.* Pharmacophore modeling and virtual screening studies to identify new c-Met inhibitors. *J. Mol. Model.* **18**, 3087–3100 (2012).
- Sterling, T. & Irwin, J. J. ZINC 15—ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337. <https://doi.org/10.1021/acs.jcim.5b00559> (2015).
- Koes, D. R. & Camacho, C. J. ZINCPharmer: Pharmacophore search of the ZINC database. *Nucleic Acids Res.* **40**, W409 (2012).

23. Dallakyan, S. & Olson, A. J. Small-molecule library screening by docking with PyRx. *Methods Mol. Biol.* **1263**, 243–250 (2015).
24. Rohilla, A., Khare, G. & Tyagi, A. K. Virtual Screening, pharmacophore development and structure based similarity search to identify inhibitors against IdeR, a transcription factor of *Mycobacterium tuberculosis*. *Sci. Rep.* **7**, 1–14 (2017).
25. Han, Y. *et al.* In silico ADME and toxicity prediction of ceftazidime and its impurities. *Front. Pharmacol.* **10**, 434 (2019).
26. Daina, A., Michielin, O. & Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **7**, 42717 (2017).
27. Martin, T. M., Harten, P., Venkatapathy, R., Das, S. & Young, D. M. A hierarchical clustering methodology for the estimation of toxicity. *Toxicol. Mech. Methods* **18**, 251–266 (2008).
28. Banerjee, P., Eckert, A. O., Schrey, A. K. & Preissner, R. ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.* **46**, W257–W263 (2018).
29. Pradeepkiran, J. A., Kumar, K. K., Kumar, Y. N. & Bhaskar, M. Modeling, molecular dynamics, and docking assessment of transcription factor rho: a potential drug target in *brucella melitensis* 16M. *Drug Des. Dev. Ther.* **9**, 1897–1912 (2015).
30. Mendez, D. *et al.* ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Res.* **47**, D930–D940 (2019).
31. Wolber, G. & Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* **45**, 160–169 (2005).
32. He, G. *et al.* An improved receptor-based pharmacophore generation algorithm guided by atomic chemical characteristics and hybridization types. *Front. Pharmacol.* <https://doi.org/10.3389/fphar.2018.01463> (2018).
33. Khedkar, S., Malde, A., Coutinho, E. & Srivastava, S. Pharmacophore modeling in drug discovery and development: an overview. *Med. Chem. (Los Angeles)* **3**, 187–197 (2007).
34. Mysinger, M. M., Carchia, M., Irwin, J. J. & Shoichet, B. K. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* **55**, 6582–6594 (2012).
35. Luo, Y. *et al.* Synthesis and biological evaluation of new homocamptothecin analogs. *Eur. J. Med. Chem.* **54**, 281–286 (2012).
36. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* **27**, 221–234 (2013).
37. Valasani, K. R., Vangavaragu, J. R., Day, V. W. & Yan, S. S. Structure based design, synthesis, pharmacophore modeling, virtual screening, and molecular docking studies for identification of novel cyclophilin D inhibitors. *J. Chem. Inf. Model.* **54**, 902–912 (2014).
38. Brylinski, M. Local alignment of ligand binding sites in proteins for polypharmacology and drug repositioning. in *Methods in Molecular Biology* vol. 1611 109–122 (Humana Press Inc., 2017).
39. Tian, W., Chen, C., Lei, X., Zhao, J. & Liang, J. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* **46**, W363–W367 (2018).
40. Shaker, B. *et al.* User guide for the discovery of potential drugs via protein structure prediction and ligand docking simulation. *J. Microbiol.* **58**, 235–244 (2020).
41. Vyas, V., Jain, A., Jain, A. & Gupta, A. Virtual screening: a fast tool for drug design. *Sci. Pharm.* **76**, 333–360 (2008).
42. Hadni, H. & Elhallaoui, M. 2D and 3D-QSAR, molecular docking and ADMET properties: In silico studies of azaaurones as antimalarial agents. *New J. Chem.* **44**, 6553–6565 (2020).
43. Yamashita, F. & Hashida, M. In silico approaches for predicting ADME properties of drugs. *Drug Metab. Pharmacokin.* **19**, 327–338 (2004).
44. Raies, A. B. & Bajic, V. B. In silico toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **6**, 147–172 (2016).
45. Bowers, K. J. *et al.* Scalable algorithms for molecular dynamics simulations on commodity clusters. in *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, SC'06 84* (ACM Press, 2006). <https://doi.org/10.1145/1188455.1188544>.

Acknowledgements

This project was funded by the National Plan for Science, Technology and Innovation (MAARIFAH) – King Abdulaziz City for Science and Technology – the Kingdom of Saudi Arabia – award number (14-BIO1978-03). The authors also, acknowledge with thanks Science and Technology Unit, King Abdulaziz University for technical support. We thank also Abdus Samad and Rahat Alam, BioSol Centre (www.biosolcentre.org) for their support in molecular dynamics simulation.

Author contributions

F.A.D.M.O. Performed the experiment and wrote the initial draft. M.M.R. Investigated and wrote the draft. F.A. Performed the investigation. I.A. Analysis the data M.A.B. Edited the draft A.M.A. Supervised the project. All authors are reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-83626-x>.

Correspondence and requests for materials should be addressed to M.M.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021, corrected publication 2021