bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Identification of Natural Antiviral Drug Candidates Against Tilapia Lake Virus: Computational Drug Design Approaches

Md Afsar Ahmed Sumon[1], Amer H. Asseri[2], Mohammad Habibur Rahman Molla[3,*] Mohammed Othman Aljahdali[3], Md. Rifat Hasan[4], M. Aminur Rahman[5], Md. Tawheed Hasan[6], Tofael Ahmed Sumon[7], Mohamed Hosny Gabr[1], Md. Shafiqul Islam[8], Burhan Fakhurji[9], Mohammed Moulay[10], Earl Larson[11], Christopher L. Brown[12*]

[1]Department of Marine Biology, Faculty of Marine Sciences, King Abdulaziz University, Jeddah-21589, Saudi Arabia; Email: afsar.sumon@gmail.com; mgabr@kau.edu.sa

[2]Department of Biochemistry, Faculty of Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia ahasseri@kau.edu.sa

[3]Department of Biology, King Abdulaziz University, Jeddah-21589, Saudi Arabia; mrahmanmolla@stu.kau.edu.sa; moaljahdali@kau.edu.sa

[4]Department of Applied Mathematics, Faculty of Science, Noakhali Science and Technology University, Noakhali 3814, Bangladesh. rifatmathdu@gmail.com

[5]Department of Fisheries and Marine Bioscience, Faculty of Biological Science and Technology, Jashore University of Science and Technology, Jashore 7408, Bangladesh. Email: aminur1963@gmail.com

[6]Department of Aquaculture, Sylhet Agricultural University, Sylhet 3100, Bangladesh. Email: tawheed7788@yahoo.com

[7]Department of Fish Health Management, Sylhet Agricultural University, Sylhet-3100, Bangladesh. Email: tofael.fhm@sau.ac.bd

[8]Institute of Marine Sciences, University of Chittagong, Chittagong-4431, Bangladesh; prof.shafiqims@cu.ac.bd

[9]iGene Medical Training and Molecular Research Center, Jeddah 21589, Saudi Arabia; dr.burhan@igene-sa.com

[10]Embryonic Stem Cell Research Unit, King Fahd Medical Research Center, King Abdulaziz University,Jeddah, 21589, Saudi Arabia; moulay.a.m3@gmail.com

[11]Department of Microbiology, St Johns River State College, Orange Park, FL, USA. Email: e.larson@workmail.com

[12]FAO World Fisheries University Pilot Programme, Pukyong National University, Busan 47340, South Korea. Email: brownchristopher38@gmail.com

Correspondent Authors Email: mrahmanmolla@stu.kau.edu.sa; brownchristopher38@gmail.com

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Abstract

Tilapia Lake Virus (TiLV) is a disease that affects tilapia fish, causing a high rate of sudden death at any stage of their growth. Unfortunately, there are currently no effective antiviral drugs or vaccines to control the progression of this disease. Researchers have discovered that the CRM1 protein plays a critical function in the development and spreading of animal viruses. By inhibiting CRM1, the virus's spread in commercial fish farms can be suppressed. With this in mind, this study intended to identify potential antiviral drugs from two different tropical mangrove plants from tropical regions: *Heritiera fomes* and *Ceriops candolleana*. To identify promising compounds that target the CRM1 protein, a computer-aided drug discovery approach is employed containing molecular docking, ADME (absorption, distribution, metabolism and excretion) analysis, toxicity assessment as well as molecular dynamics (MD) simulation. To estimate binding affinities of all phytochemicals, molecular docking is used and the top three candidate compounds with the highest docking scores were selected, which are CID107876 (-8.3), CID12795736 (-8.2), and CID12303662 (-7.9). We also evaluated the ADME and toxicity properties of these compounds. Finally, molecular dynamics simulation was conducted to analyze the stability of the protein-ligand complex structures and confirm the suitability of these compounds. The computational study demonstrated that the phytochemicals found in *H. fomes* and *C. candolleana* could potentially serve as important inhibitors of TiLV, offering practical utility. However, further *in vivo* investigations are necessary to investigate the effectiveness of these compounds as antiviral drugs against the virus TiLV.

Keywords: CRM1; TiLV; *Heritiera fomes*; *Ceriops candolleana*; virtual screening; homology modeling; ADMET; MD simulation

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

1. Introduction

Tilapia is one of the most widely consumed, rapidly growing, and profitable species in worldwide aquaculture industries. Tilapia has desirable attributes including fast growth, tolerance of environmental changes, stress and disease resistance, ease of captive breeding, high protein content, and palatability [1]. However, aquatic viruses greatly impact the commercial aquaculture industry by causing sudden mortality and serious economic losses. Tilapia Lake Virus (TiLV) disease is one of the emerging and transboundary diseases of tilapia culture, causing up to 90% of sudden mortality in any culture stage [2]. Several tilapias, including the Nile (*Oreochromis niloticus*), Red (*Oreochromis* sp.), and hybrid tilapia derived from *O. niloticus* and *O. aureus* have been infected with the TiLV [3]. With the increase in surveillance and diagnostic techniques in recent years, it is established that TiLV disease has spread across Asia, Africa, and North and South America [4]. TiLV is caused by a single-stranded (ss)-RNA virus, and its infection is characterized by various symptoms including lethargy, loss of appetite, respiratory distress, skin erosion, and brain lesions. Unfortunately, there are currently no effective therapeutic drugs available to control the spread of this virus [5].

The development of drugs that effectively target host proteins essential to viral life cycle completion is a promising strategy for antiviral therapy development. Chromosome Region Maintenance1 (CRM1) is a nuclear transporter protein that is understood to have a vital function in various virus families such as coronaviruses, retroviruses, flaviviruses, orthomyxoviruses, rhabdoviruses, paramyxoviruses, and herpesviruses [6]. CRM1 is among the most well characterized Exportin-1 compounds (XPO1); CRM1-mediated export is used by a variety of viral species at various stages of their lifecycles to ensure proper protein localization [7]. When CRM1, a protein responsible for cellular transport is blocked, the levels of Protein 4 in the

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

cytoplasm and nuclei generally show increases. This reflects the crucial role of CRM1 in exporting TiLV RNA from the nucleus. Inhibiting CRM1-mediated export has multiple effects on the virus. It leads to alterations in the expression of viral proteins, disrupts viral replication, and hinders the complete assembly of viral particles. As a result, the infectivity of the virus is reduced. Additionally, inhibiting CRM1-mediated export triggers a stronger antiviral immune response in the host, enhancing its ability to fight against the virus [8]. CRM1 binds to nuclear export sequences (NES)-cargoes in the nucleus via a cluster of leucine-rich or hydrophobic amino acids in the presence of RanGTP. The tricomplex (RanGTP) is hydrolyzed by RanGAP (a GTPase) resulting to RanGDP after translocation into the cytoplasm. After being released, CRM1 protein returns to the nucleus, and this cycle is repeated to take advantage of or regulate CRM1-mediated nuclear export [9]. A CRM1-targeting therapeutics drug has the potential to be precisely effective against TiLV, thereby inhibiting damage from the infection.

It is imperative to discover viable phytochemicals and small bioactive molecules for drug discovery against infectious viral diseases in aquaculture through investigation of novel compounds interactions with the targeted viral protein [10]. Computational or *in silico* methods are assisting in making decisions and simulating virtually every aspect of drug discovery and development, which is a easy and rapid method for predicting liable compounds against certain drug targets [11]. Moreover, those compounds can be selected, documented and scored for their interactions based on molecular docking. The ADMET properties, which include absorption, distribution, metabolism, excretion, and toxicity, play a decisive role in determining the efficacy and safety of compounds in pharmacokinetics. Computer-aided methods can be employed for straightforward prediction of these properties. Additionally, MD simulation can be utilized to confirm potential drug candidate stability when interacting with the targeted protein [12].

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Several marine and terrestrial plants have been discovered to contain bioactive compounds possessing antiviral drug properties [13]. Natural phytocompounds are increasingly being used as antimicrobial drugs in aquaculture instead of synthetic antibiotics [14]. Although it is challenging to develop novel drugs, advancing investigations for the discovery of bioactive compounds that target specific protein of infectious aquatic disease etiologies is urgently needed [15]. Previously, mangrove plants and medicinal plant garlic (*Allium sativum*) derived natural bioactive ligand was used against Nervous Necrosis Virus (NNV) capsid protein and *in-silico* study proved that ligand potential to inhibit this viral replication in Asian Seabass, *Lates calcarifer* [16]. Similarly, amphene is the best screened phytochemical with the lowest binding energy with MX protein complex that can inhibit NNV infectivity in same fish species. Furthermore, a computational drug development approach revealed that phytocompounds, namely friedlein, phytosterols and 1-Triacontanol from *Avicennia alba* could bind to viral gag polyprotein and inhibit infectious Walleye Dermal Sarcoma Virus (WDSV) replication [15].

In recent decades, mangrove plants have been documented as a source of new chemical products with a rich assortment of naturally bioactive compounds. Mangrove plants, *Heritiera fomes* and *Ceriops candolleana* have a long history of use for medicinal and other purposes (making pickles, nuts, boats etc.) by people in coastal populations. Several studies have revealed that root, bark and leaves of both plants possess significant antioxidant, antinociceptive, antihyperglycemic, antimicrobial (such as antibacterial, antiviral and antifungal), and anticancer properties [17]. Compounds such as 10-undecenoic acid, methyl ester hexadecane dibutyl phthalate hexadecanoic acid, methyl ester undecylenic acid 9,12-octadecadienoic acid (Z,Z)-, methyl ester palmitic acid 2,5-anhydrogluconic acid aminopyrine procyanidins, phytosterols, 9-epiquinine, AC1L1UKB bis (acetic acid), tannins, cholesterol, delta 7-avenasterol, and squalene

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

have been isolated from different parts of these two plants including root, bark and leaves, with impressive potential for a wide range of pharmaceutical products as replication inhibitors for various types of viruses, including TiLV [18].

To evaluate the potential of a new drug candidate against the TiLV virus, this research employed various readily-available *in-silico* methodologies. These approaches included virtual screening, molecular docking, homology modeling, QM calculation, ADMET analysis, and MD simulation. The objective was to investigate the effectiveness of the drug candidate in targeting either CRM1 or XPO1 proteins associated with the TiLV virus.

2. Materials and Methods

2.1. Protein Structure Prediction through Homology Modeling

To obtain the TiLV structural CRM1 fasta sequence (NC_031978.2), NCBI database (https://www.ncbi.nlm.nih.gov/gene/100703095, 30 April 2022) was exploited and submitted to the popular online web portal (https://zhanglab.dcmb.med.umich.edu/I-TASSER/, on 30 April 2022), viz. Iterative Threading Assembly Refinement (I-TASSER) for confirmation or prediction of the 3D structure of anticipated protein. The C-score, TM-score value, and root mean square deviation (RMSD) of the top five models of that structure were then provided by the I-TASSER [19].

2.2. Protein Structure Refinement and Validation

The 3D protein structure was refined by uploading it in online to GalaxyRefine (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE, 7 May 2022), from which the data about the value of RMSD, energy score, and final structure quality was obtained. To select the

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

refined structure, several parameters including RMSD, average distance among atoms and energy score were used. The refined structure was pictured using PyMol v2.3.4 software. For validation of the refine model, the Ramachandran plot scoring function was used [20].

2.3. Preparation of Protein and Ligand

Protein structures need to be refined and further validated before docking can take place. Thus, the selected 3D structure of CRM1 was generated using AutoDockTools (ADT). Nonpolar hydrogen atoms were combined, polar hydrogen atoms were introduced, and the Gasteiger charges of the protein were computed [21]. Additionally, it served as a cofactor and metal ion discarded from its protein. An extensive literature review and the database (https://cb.imsc.res.in/imppat/home) of Indian Medicinal Plants, Phytochemistry, and Therapeutics (IMPPAT) were used to identify the selected plants compounds. Seventeen compounds were retrieved, among which 7 compounds were identified from *C. candolleana* and 10 compounds from *H. fomes*. To set and minimize energy for compounds selected from the database, The Universal Force Field (UFF) was used which is designed for each ligand.

2.4. Identification of Binding Site and Grid Box Generation

The refined protein structure of CRM1 was submitted to the CASTp 3.0 server (http://sts.bioe.uic.edu/, 10 May 2022) to predict the active site residues. Then, the server identified different active pockets, and the first of these was selected based on the pocket surface area (SA) and volume. The active pocket and corresponding AA residues were retrieved to visualize the binding pocket of the protein through BIOVA Discovery Studio Visualizer Tool 16.1.0. The residues of the binding sites were then generated by BIOVA, used to select the grid box for molecular docking simulations [22].

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

2.5. Molecular Docking Simulation

PyRx software was used to run virtual screenings of selected compounds using a molecular docking process. The virtual screening software had identified several hypothetical drugs to treat several diseases. In addition to Lamarckian genetic algorithm (LGA) for scoring, AutoDock and AutoDock Vina for docking were provided. This study employed PyRx tools AutoDock Vina to perform docking interactions between molecules. For the analysis of complex binding poses, the BIOVA Discovery Studio Visualizer Tools 16.1.0.41 was used [23].

2.6. ADME Analysis

During the preliminary drug development phase, an ADME assessment is required to determine a drug's safety and efficacy in living systems. ADME properties depict pharmacokinetic behavior to predict the movement of drugs through and out of the body. This study utilized the SwissADME (http://www.swissadme.ch, 11 May 2022) to interpret the response of selected drug candidates regarding pharmacodynamics [24].

2.7. Toxicity Test

It is economically advantageous and vitally important to conduct toxicology tests during the drug development process in order to determine whether the compound has toxic properties, as well as for the assessment of dosage requirements. To assess the toxicity of chosen chemicals, *in-silico* computational approaches were utilized to evaluate their safety profiles. For early-stage toxicity evaluation, ProTox-II was used (http://tox.charite.de/protox_II, 11 May 2022), which simulated characteristics associated with the compounds' acute toxicity, carcinogenicity, hepatotoxicity, mutagenicity, cytotoxicity, and immunotoxicity [25].

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

2.8. Molecular Dynamic (MD) Simulation

The thermodynamic stability of the receptor-ligand system was evaluated using the Desmond v3.6 Program, which facilitated the automated simulation and free energy perturbation (FEP) calculations at different temperatures to predict the equation of state (EOS). Orthorhombic periodic boundary conditions were applied to simulate larger complex structures. To stabilize the system and achieve energetic neutralization, sodium ions were introduced and randomly placed [26]. Molecular dynamics simulations were conducted using the buffer box calculation method with a size of 10 Å and a specific box shape. Periodic boundary conditions and the NPT ensemble were employed during the simulations. The retrieval and analysis of the resulting trajectory from the simulations has been done via the Simulation Interaction Diagram (SID) in Schrödinger Release 2020-2.

3. Results

3.1. Protein 3D Structure, Refinement and Validation

Among the models of CRM1 provided by I-TASSER, the best 3D protein structure was selected based on the lowest C-score (-0.01). Protein model-2 was picked following refinement, which had a 3D refined score of 37759.1, a GDTHA score of 0.9666, an RMSD value of 0.355, and no MolProbity issues. Prior to refinement, the Ramachandran plot analysis of CRM1 revealed that 97.629% of residues were in the favorable region, 2.062% were in the allowed region, and 0.309% were in the disallowed region. After refinement, the refined CRM1 model showed improved results, with 97.984% of residues in the favorable region, 2.016% in the allowed region, and 0.000% in the disallowed region.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

3.2. Phytochemical and Protein Preparation

The phytochemicals from the *C. candolleana* were sourced from the IMPPAT database, while those from *H. fomes* were obtained from published literature. In total, 17 compounds from *C. candolleana* and *H. fomes* were collected and stored in a 2D (SDF) file format. To prepare ligands for further analysis, the phytochemicals were optimized and converted into pdbqt files. Similarly for the docking process, optimization and preparation of the proteins via AutoDock tool has been done, and its corresponding file was saved in the pdbqt format.

3.3. Identification of Active Site and Receptor Grid Generation

The study first identified CRM1's active side (AS) from CASTpi, then obtained its combined binding location (Fig. 1). The protein binding site residue was identified through active pocket analysis (Fig. 1). Four active pockets were selected from a total of 197 active pockets based on their surface area. The chosen active pockets and their corresponding amino acid residues are represented as balls in Figure 1. During the simulation procedure of molecular docking, the server-identified binding sites were used to build a receptor grid with grid box dimensions of X = 77.1430, Y = 81.9808, and Z = 93.4587 in angstroms (Å).

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

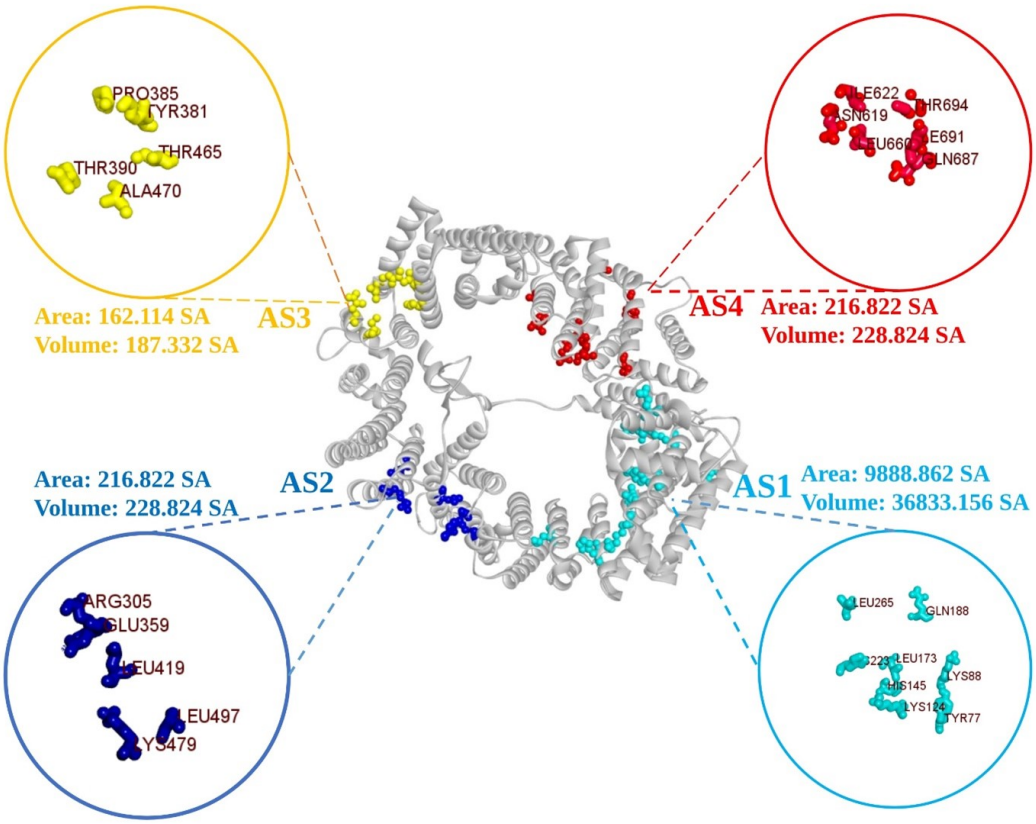


Fig. 1. The surface area of the active pockets of CRM (selected four) was calculated using the CASTp server. All the active sites i.e., AS1, AS2, AS3 & AS4 with their corresponding amino acids are depicted respectively in light blue, deep blue, yellow and red.

3.4. Molecular Docking

After molecular docking of phytochemicals compounds, the binding affinity was determined to have varied between −3.2 and −8.7. The top 20% of the 17 phytochemicals with the highest binding affinity were selected. The top three compounds such as procyanidins (CID107876), delta7-Avenasterol (CID 12795736), and phytosterols (CID12303662), were chosen according to their docking scores respectively of −8.3, −8.2 and −7.9 kcal/mol, then examined using various screening techniques. The three best compounds scores are enumerated in Table 1, while docking scores are recorded in Table S1 for all the phytochemicals. Validation of those scores

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

for the three selected compounds was confirmed by re-docking. The single binding poses of three preselected phytochemicals were obtained and re-docked at the same binding site as before. The RMSD values for the upper and lower ranges were determined, and it was found that the binding affinities of the phytochemicals remained similar to the earlier scores. The phytochemicals i.e., CID107876, CID12795736, and CID12303662 with docking scores respectively of -8.3, -8.2 and -7.9 kcal/mol and zero (0) RMSD value were chosen.

Table 1. The best three phytochemicals with their name, formula and CID.

PubChem ID	Compounds	Molecular Formula	Molecular Weight	Docking Score (kcal/mol)
107876	Procyanidin	C30H26O13	594.5	-8.3
12795736	delta7-Avenasterol	C29H48O	412.7	-8.2
12303662	Phytosterols	C29H50O	414.7	-7.9

3.6. Protein-Ligand Interaction Analysis

Compounds that have the highest binding scores were chosen to see how they could interact with CRM1. Several hydrogen and hydrophobic bonds were reported to form between the compound CID107876 and the desired CRM1 (Fig 2). The hydrogen bonds were discovered to form at PRO427 and LYS590 position, while the hydrophobic bonds from at the location TYR454, VAL465, PRO1062, PRO1062, PRO1062, GLU459, ARG458, LYS455, and LYS455 position is illustrated in Figure 3 and the bond types in Table 2.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

254 **Table 2. Compilation of bonding interactions among three chosen phytochemicals and**

255 **CRM1.**

PubChem ID	Residue	Distance	Category	Type
CID 107876	PRO427	2.58	Hydrogen Bond (HB)	Conv-H-Bond
	LYS590	3	HB	Conv-H-Bond
	TYR454	5.6	Hydrophobic (HP)	Pi-Pi- T-Shaped
	VAL465	4.82	HP	Pi-Alkyl
	PRO1062	4.63	HP	Carbon-H-Bond
	PRO1062	3.57	HP	Carbon-H-Bond
	PRO1062	5.32	HP	Carbon-H-Bond
	GLU459	4.59	HP	Pi-Cation
	ARG458	3.87	HP	Pi-Cation
	LYS455	4.86	HP	Pi-Cation
	LYS455	5.13	HP	Pi-Cation
CID 12795736	SER162	2.6	HB	Conv-H-Bond
	VAL120	4.64	HP	Alkyl
	ILE127	4.82	HP	Alkyl
	LYS124	5.09	HP	Alkyl
	VAL169	5.18	HP	Alkyl
CID 12303662	VAL169	4.6	HP	Alkyl
	LYS124	5.07	HP	Alkyl
	ILE127	4.7	HP	Alkyl
	VAL120	4.78	HP	Alkyl

256

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

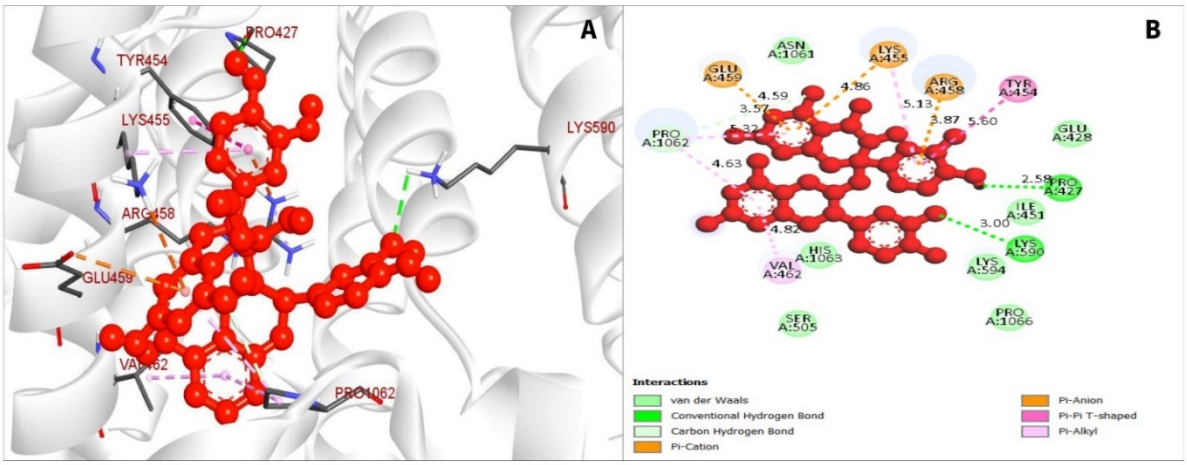


Fig. 2. The interaction between the chemical CID107876 and CRM1 (A). The 3D interaction is shown on the left of the protein ligands, while 2D interaction is shown on the right (B).

The compound CID12795736 has formed several bonds in the VAL120, ILE127, LYS124, and VAL169 residual position which were hydrophobic. Another two conventional hydrogen bonds were formed at the position of SER162 AA position (Fig. 4) and are listed in Fig. 3.

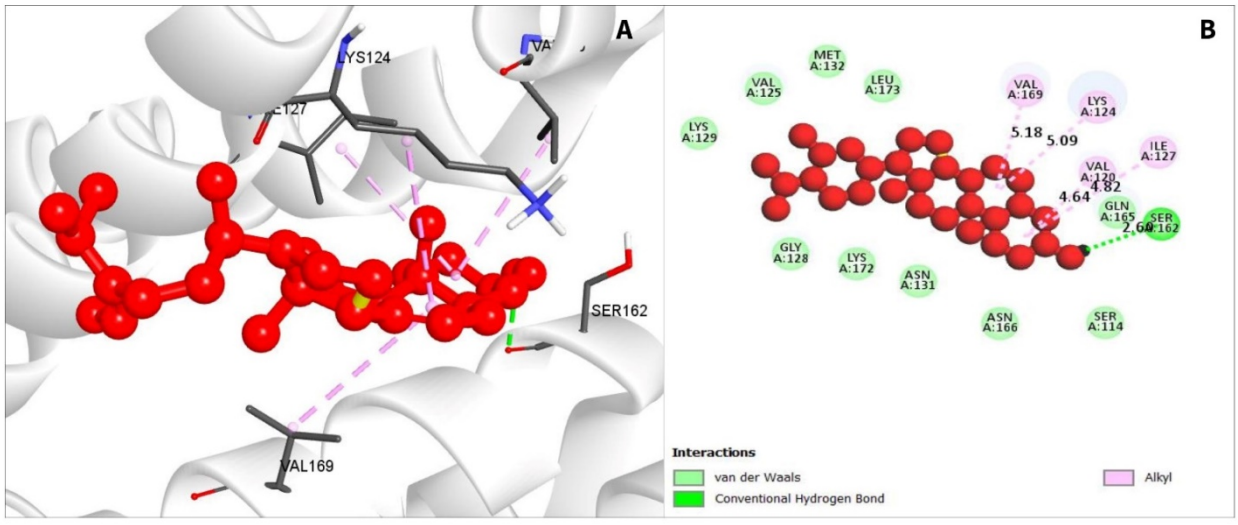


Fig. 3. The interaction among chemical CID12795736 and CRM1 (A). The 3D interaction is shown on the left of the protein ligands, while 2D interaction is shown on the right (B).

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

For compound CID12303662, also various hydrophobic bonds occurred with desired phytochemicals, with alkyl bonds in VAL169, LYS124, ILE127, and VAL120 residual positions as depicted in Fig. 4 (listed in Table 2).

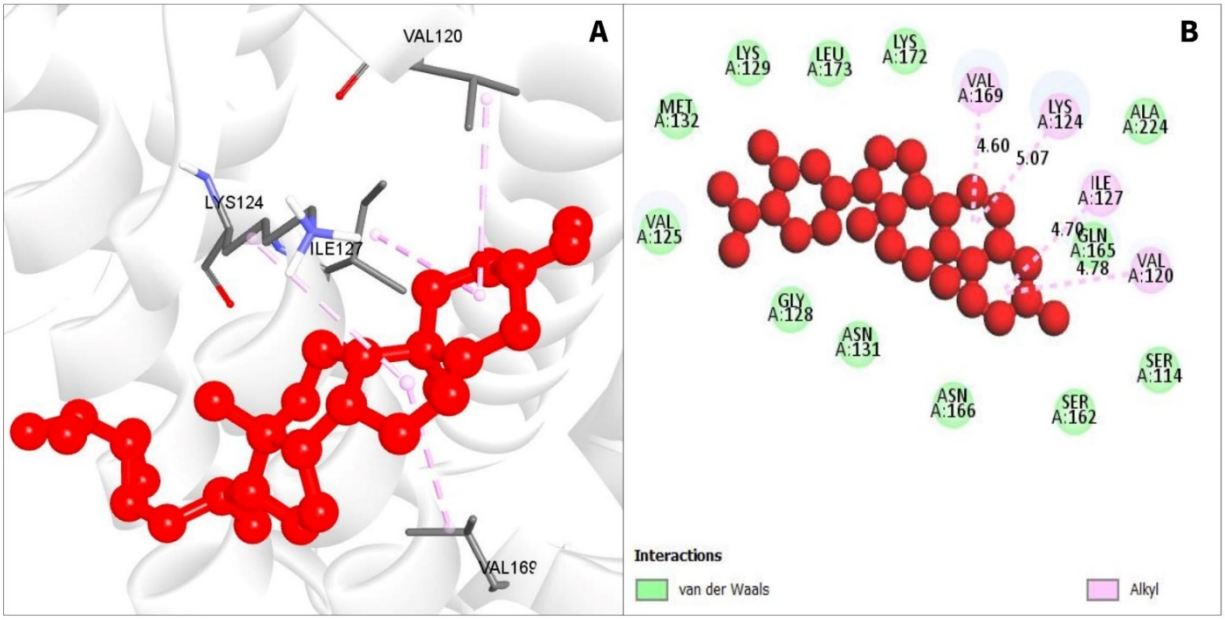


Fig. 4. Interaction among chemical CID12303662 and CRM1 (A). The 3D interaction is shown on the left of the protein ligands, while 2D interaction is shown on the right (B).

3.7. ADME Analysis

The server SwissADME was used to evaluate the pharmacophore characteristics of the three chosen druglike compounds. The druglike compounds possess lipophilic characteristics which enable them to be dissolved in fats, oils, and other nonpolar solvents. Pharmacophore features demonstrated that the compound is effective and druggable, thus it may be employed in the research. The pharmacokinetic parameters discovered for the three chosen drugs are presented in Table 3. The ADME properties of the chosen three phytochemicals were found to be of high quality.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Table 3. The pharmacokinetics list comprises the ADME features of the three chosen drugs. The lists also include the various physicochemical features of the three substances.

Properties	Parameters	CID 107876	CID 12795736	CID 12303662
	MW (g/mol)	594.52 g/mol	412.69 g/mol	414.71 g/mol
	Heavy atoms	43	30	30
	Arom. Heavy atoms	24	0	0
	Rotatable bonds	4	5	6
	H-bond acceptors	13	1	1
	H-bond donor	10	1	1
	Molar Refractivity	147.52	135.75	133.23
Lipophilicity	Log Po/w	1.98	7.94	8.02
Water solubility	Log S (ESOL)	-4.9	-7.48	-7.9
Pharmakokinetics	GI absorption	Low	Low	Low
Drug likeness	Lipinski, Violation	No	Yes	Yes
Medi. Chemistry	Synth. accessibility	5.64	6.03	6.3

Medi. Chemistry = Medicinal chemistry; Synth. Accessibility = Synthetic accessibility.

3.8. Toxicity Prediction

The three compounds PubChem CID107876, CID12795736, and CID12303662 were selected previously through various screening processes. The compounds were subjected to analysis in ProTox-II server, which assessed their cytotoxicity, hepatotoxicity, oral toxicity, mutagenicity and carcinogenicity. However, Table 4 showed the phytochemicals that lack oral or organ toxicity effects. After toxicity testing, the results confirmed that the three selected compounds were either non-toxic or exhibited low toxicity levels.

Table 4. The drug-induced toxicity profile of selected phytochemicals.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

PubChem ID	Hepatotoxicity	Carcinogenicity	Immunotoxicity	Mutagenicity	Cytotoxicity
CID 107876	Inactive	No	Inactive	Inactive	Inactive
CID 12795736	Inactive	No	Light active	Inactive	Inactive
CID12303662	Inactive	No	active	Inactive	Inactive

3.9. MD Simulation Analysis

To investigate the constancy of the protein-ligand docking complexes, molecular dynamics simulations were carried out. This simulation method offers valuable insights into the intermolecular interactions over time. A 100 ns simulation of molecular dynamics was performed for intermolecular interactions among protein and ligand complex structure confirmation. The simulated trajectories were then utilized to assess parameters such as RMSD, Root Mean Square Fluctuation (RMSF), mapping of protein-ligand interactions, and analysis of ligand torsion. These analyses provided more depth in our understanding of the dynamic behavior and stability of the protein-ligand complexes.

3.9.1. RMSD

The RMSD values of three chemicals and the protein were analyzed to assess system equilibration. Specifically, the RMSD values of compounds CID107876, CID12795736, and CID12303662 were compared with the structure of CRM1, as shown in Fig. 5, to identify any variations. The RMSD values for all compounds fell within the range of 1 Å to 2.5 Å, which was considered acceptable when contrasted to the CRM1 structure. Notably, the compound CID12795736 exhibited the highest stability between 65 ns to 100 ns simulation time.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Additionally, compounds CID107876 and CID12303662 displayed initial fluctuations but gradually stabilized over time. The simulation was observed to converge between 40 ns and 100 ns for all compounds, with the RMSD values reaching a stable value. Consequently, the selected phytochemicals are deemed stable for the targeted protein.

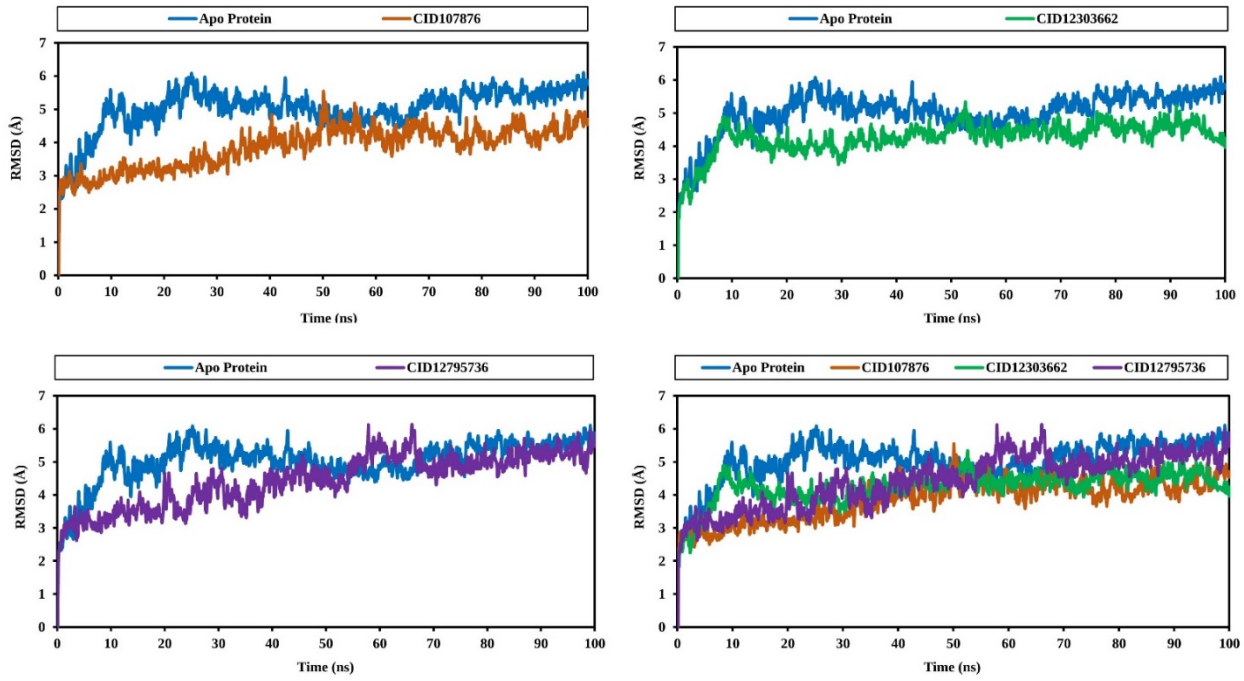


Fig. 5. The RMSD values obtained from the Cα atoms of CRM1 (represented by blue curves) and the naturally occurring compounds were analyzed over a 100 ns simulation time. Specifically, the RMSD values for CID107876 are depicted as brown curves, CID12795736 as violet curves, and CID12303662 as green curves.

3.9.2. RMSF

To assess the displacement of specific atoms during the simulation, analysis of three RMSF complex structures has been done. These values were further compared to the RMSF of the CRM1 structure, as shown in Figure 6, to observe atomic alterations. The highest peak of

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

fluctuation was observed between residue points 184 and 396. Additionally, the entire complex structure exhibited fluctuation between residues 500 and 550, which was attributed to the location of N-terminal domain. Generally, the fluctuation of all the phytochemicals was found to be more optimal than the RMSF of CRM1. This indicates that the phytochemicals are capable of maintaining a steady interaction with no significant alteration of protein's structure.

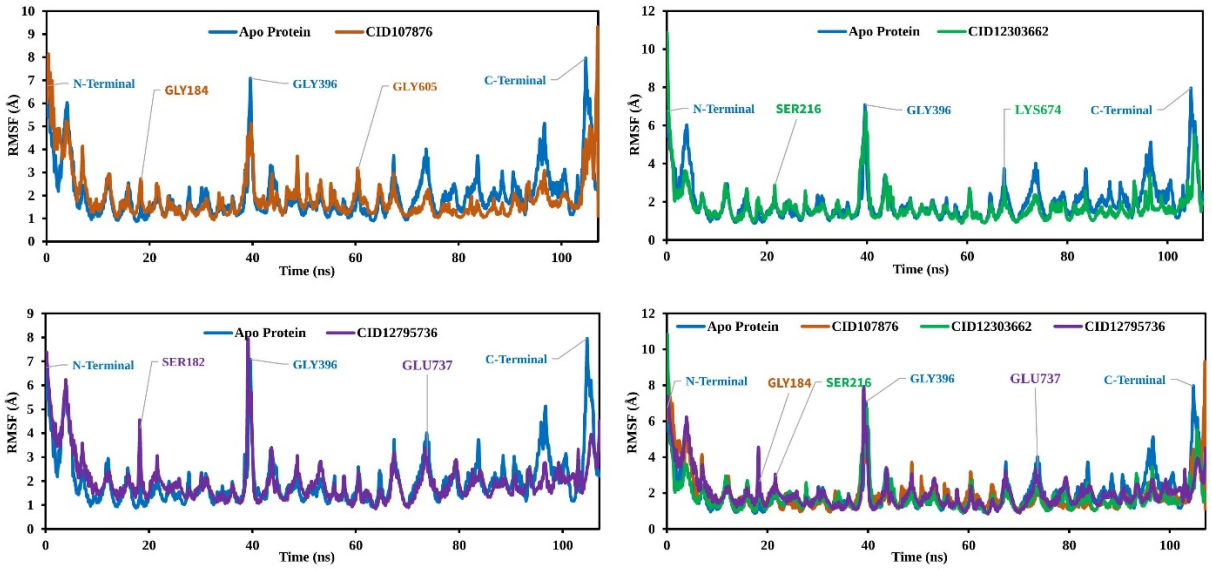


Fig. 6. RMSF values extracted from the protein residue index Cα atoms of the complex structure *viz.*, CID107876 (brown), CID12795736 (violet), and CID12303662 (green) and CRM1 (blue) regarding 100 ns of simulation time.

3.9.3. Protein-Ligand Contact Mapping

Throughout the analysis of the 100ns MD simulation trajectory. The interaction between CRM1 and the three selected phytochemicals (CID107876, CID12795736, and CID12303662) were monitored. These interactions were categorized as hydrogen bonds, hydrophobic bonds, ionic bonds and water bridge bonds. The stacked bar graphs in Fig. 7 shown the summary of those interactions. All three of the selected compounds had a strong interaction with catalytic residues.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

The molecular dynamic simulation interval of the screened compounds was found to contain a significant amount of intermolecular interaction. Active residues were discovered to interact with intermolecular and chosen ligands, and their density was determined for each of the three compounds studied. The active pocket of CRM1 was identified by the screening of natural chemicals and the study of molecular dynamic trajectories. In real life experiments, therefore, the screened natural compounds CID107876, CID12795736 and CID12303662 can be predicted to have a high degree of stability with the CRM1.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

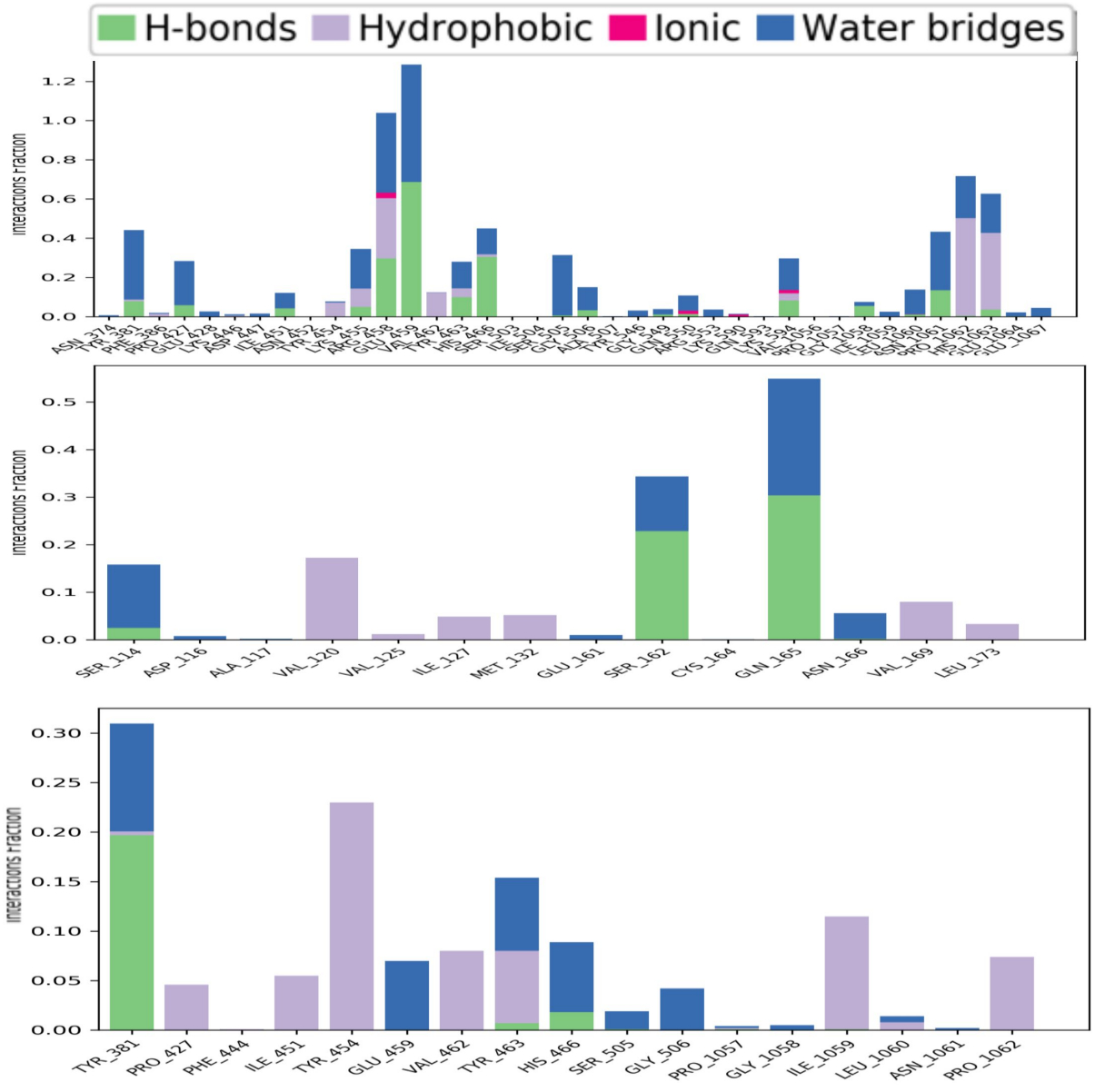


Fig.7. Stacked bars of the contact mapping of CRM1 with potential natural phytochemicals, i.e., (A) CID107876; (B) CID 12795736; and (C) CID 12303662 determined from simulations trajectory of 100 ns**.**

3.9.4. Ligands torsion analysis

The ligand torsions plot provides a summary of the conformational changes observed in each rotatable bond (RB) of the ligand during the entire simulation trajectory (0-100ns), as depicted in

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Figure 8. In the upper section, a two-dimensional representation of the ligand is shown, highlighting the RB (torsion) regions using different colors. Each RB torsion is accompanied by a dial plot and corresponding bar plots in the same color. The dial plots depict the changes in conformation of the torsions during the simulation. The center of the radial plot represents the beginning of the simulation, while the radial direction indicates the passage of time. The bar plots provide a summary of the dial plots, showing the probability density of the torsion at different time points. Furthermore, if information about the torsional potential is available, the plot also includes the potential energy of the RB by adding up the potential energy of the relevant torsions. The values of the potential energy are displayed on the left Y-axis of the chart and are expressed in kcal/mol. Analyzing the histogram and the relationship between torsional potential and torsion values can offer insights into the strain experienced by the ligand in maintaining its conformational bond with the protein.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

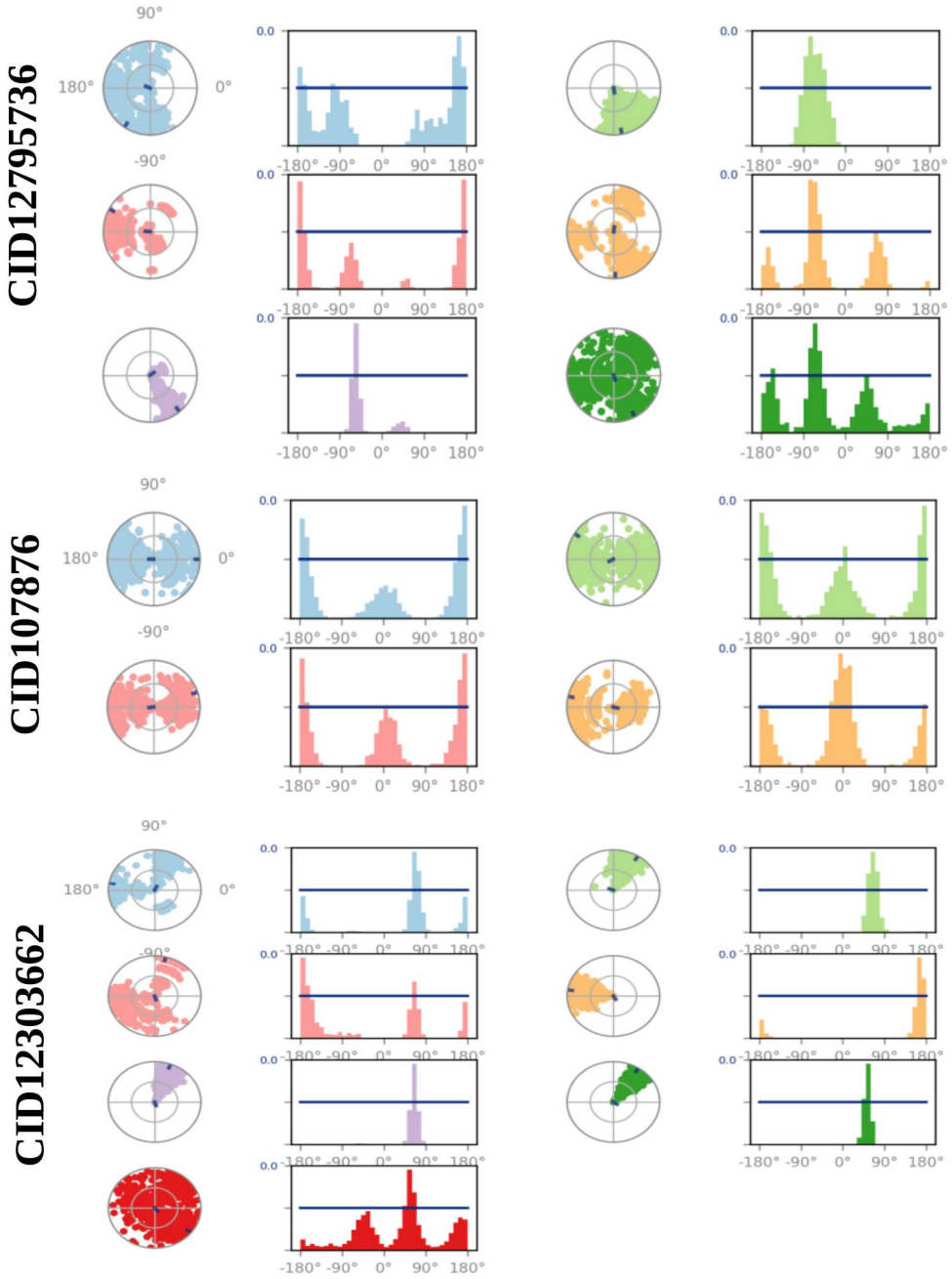


Fig.8. The ligand torsions plots provided a summary of how each RB in the ligand underwent conformational changes throughout the entire simulation trajectory (from 0 to 100 ns).

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

4. Discussion

Since the emergence of TiLV in 2014, eradication of the pathogen infectivity is the major concern of the world scientific community and virus researchers. So far, no anti-viral drug candidate has emerged that can effectively counteract these zoonotic pathogens. It has been found that TiLV's protein 4 functions when CRM1 is inhibited, which plays a key role in the export of TiLV RNA [27]. Hence, this study was targeted to inhibit the CRM1 in TiLV through the identification and characterization of a novel and potentially effective bioactive antiviral drug compound. Computer-Aided Drug Design (CADD) is one of the most promising techniques for selecting novel compounds against a specific protein because it comprises a variety of highly advanced and nuanced characteristics and approaches [28]. CADD provides more affordable methods of designing, decreases recruitment costs, and reduces time in a new era for modern drug development. It has gained popularity and acceptance both in academic circles as well as among pharmaceutical industries. A process involving molecular docking, ADME-Tox prediction, and molecular dynamics simulations is used to identify drug candidates that are predicted to possess the highest biological efficacy [29]. To understand disease mechanisms and facilitate drug design, it is crucial to identify disease-associated proteins and explore their interactions with potential ligands. By identifying compounds that can interact with and inhibit specific proteins essential for virus replication, it becomes possible to block viral infections and their progression. Through CADD, the specific target molecule can be identified based on its behavior and ligand's binding mode. Additionally, molecular docking identifies the dominant modes of binding within a ligand, as well as protein and MD simulations revealing the mechanisms of protein-ligand interaction [30]. As a result, small-molecule candidates can be recognized as suitable drugs against a particular disease including TiLV infection.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

In this study, we first predicted protein structure through Homology Modeling, then refined and validated by utilizing the popular online web portal (including, I-TESSER, GalaxyRefine). A homology-based structural model is widely used in identifying the effects of single nucleotide polymorphisms, as well as in drug design [31]. Homology models contain substantial information about the spatial arrangement of key residues in the protein to investigate the binding site and design/dock drugs that are suited for binding to the molecule. The most highly-ranked protein 3D structure (model 2) with lowest c-score (-0.01) generated from I-TESSER was used. To validate the refine model, the Ramachandran plot and z-score value scoring function was utilized. The model quality was increased after refining from the Galaxy Refine server, and the final refined model showed 97.984 % in the most favored region in the Ramachandran plot and 0.0 % in the disallowed region, indicating good model quality.

Seventeen natural phytochemical compounds were screened, and molecular docking simulations were performed to target the XPO1 to fight and block the replication of TiLV, and among those three most promising compounds were chosen according to highest binding affinity scores for further validation. The compounds with higher docking scores were CID107876, CID12795736 and CID12303662 with binding scores of −8.3, 8.2 and −7.9 kcal/mol, respectively. The value obtained for the selected compounds was highly favorable and exhibited characteristic features indicating their potential as drug candidates. These compounds demonstrated drug-like qualities, as confirmed by their compliance with Lipinski's rule of five (RO5). These selected three compounds have been evaluated and all compounds have shown a good value of the ADMET properties. Generally, the ADMET acts on the drug-related properties of pharmacokinetics (PK). Before developing a promising drug candidate, the PK parameters should be optimized since they need to pass standard clinical trials [32]. ADMET properties were traditionally predicted at

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

the end of the drug discovery process, however, *in-silico* tools can predict them at an early stage. ADMET properties are responsible for 60% of drug molecules failing during the final drug development process. An early prediction of these properties would dramatically reduce the costs of drug research. Further toxicity evaluation has been conducted on the compound that has good ADME properties to evaluate the harmful effects on humans and animals. Toxicology testing is an important element of drug development, but it is costly and time-consuming to conduct on animals [33]. The alternative to *in silico* approaches have been adopted before drug development as they do not need animal trials and are also time and cost effective. Following toxicity testing, we confirmed that the three compounds chosen were non-toxic or low-toxic.

MD simulation has emerged as a valuable tool in computer-aided drug design (CADD) as it allows for the examination of the dynamic behavior of compounds within a specific macromolecular environment. In this study, MD simulation was employed to assess the stability of the selected drug candidates when interacting with the target macromolecule. Key parameters such as RMSD, RMSF, and ligand-protein interactions were analyzed to evaluate the complex system. The results indicated optimal RMSD and RMSF values, along with favorable protein-ligand contacts, for all three compounds [34]. Based on the findings, the three selected compounds have demonstrated favorable characteristics and stability when interacting with the target protein. Consequently, these phytochemicals show promise as potential candidates for further development as antiviral drugs targeting the TiLV.

5. Conclusions

The TiLV poses a substantial threat to the tilapia farming industry and to human food security, as it causes high mortality rates and there is currently no effective antiviral medication available to

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

436 combat the virus. This study aimed to identify potential natural antiviral drugs against the TiLV,
437 with the goal of minimizing the frequency of virus infections. Three potential antiviral
438 candidates, namely Procyanidins, delta7-Avenasterol, and Phytosterols, were selected through
439 computer-aided drug design approaches such as homology modeling, molecular docking,
440 ADMET, and MD simulation methods. These compounds hold promise in mitigating the
441 economic losses in tilapia fisheries caused by the TiLV. However, further *in vivo* and *in vitro*
442 studies are recommended for the assessment of the effectiveness of these phytochemicals against
443 TiLV in a practical manner.

444 **Supplementary Materials:** TablS1: CRM1 active sites.

445 **Author Contributions:** M.A.A.S., A.H.A., M.H.R.M., M.O.A., M.R.H., M.A.R., M.T.H.,
446 T.A.S., M.H.G., M.S.I., B.F., M.M., E.L., and C.L.B.; Conceptualization, M.A.A.S., A.H.A.,
447 M.H.R.M., and M.O.A.; methodology, formal analysis, and writing—original draft preparation,
448 M.A.A.S., A.H.A., M.H.R.M., M.O.A, M.T.H., T.A.S. and C.L.B.; supervision, writing—review
449 and editing and funding acquisition. All authors approved the submitted version.

450 **Acknowledgments:** This research was funded by Institutional Fund Projects under grant no.
451 (IFPIP: 1349-130-1443). The authors gratefully acknowledge technical and financial support
452 provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi
453 Arabia.

454 **Statement of Data Availability:** Data are contained within the article and supplementary
455 material.

456 **References**

457 1. Abd El-Hack, ME, El-Sadonny, MT, Nader, MM, Salem, HM, Tahan, AM, Soliman, SM

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

et al. Effect of environmental factors on growth performance of Nile tilapia (*Oreochromis niloticus*). Int. J. Biometeorol. 2022; 66: 2183-2194.

2. Nicholson, P, Fathi, MA, Fischer, A, Mohan, C, Schieck, E, Mishra, N et al. Detection of Tilapia Lake Virus in Egyptian fish farms experiencing high mortalities in 2015. J. Fish Dis. 2017; 1925-1928. doi:10.1111/jfd.12650.

3. Saranya, SR, Sudhakaran, R. Report on prevalence of tilapia lake virus infection in tilapia fishes (*Oreochromis niloticus*). Biocatal. Agric. Biotechnol. 2020; **27**, 101665. https://doi.org/10.1016/j.bcab.2020.101665

4. Mackey, TK, Liang, BA, Cuomo, R, Hafen, R, Brouwer, KC and Lee, DE. Emerging and reemerging neglected tropical diseases: a review of key characteristics, risk factors, and the policy and innovation environment. Clin. Microbiol. Rev. 2014; 27: 949-979.

5. Eyngor, M, Zamostiano, R, Tsofak, JEK, Berkowitz, A, Berkovier, H, Tinman, S, et al. Identification of a novel RNA virus lethal to tilapia. J. Clin. Microbiol. 2014; 52, 4137. https://doi.org/10.1128/jcm.00827-14

6. Mathew, C, Ghildyal, R. CRM1 inhibitors for antiviral therapy. Front. Microbiol. 2017; **8**, 1171.

7. Sendino, M, Omaetxebarria, MJ, Rodríguez, JA. Hitting a moving target: inhibition of the nuclear export receptor XPO1/CRM1 as a therapeutic approach in cancer. Cancer Drug Resist. 2018;1, 139–163.

8. Pickens, JA, Tripp, RA. Verdinexor targeting of CRM1 is a promising therapeutic approach against RSV and influenza viruses. *Viruses 2*018*;* 10(1)*:*48.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

https://doi.org/10.3390/v10010048

9. Thakar, K, Karaca, S, Port, SA, Urlaub, H, Kehlenbach, RH. Identification of CRM1-dependent nuclear export cargos using quantitative mass spectrometry. Mol. Cell. Proteomics 2013;12, 664. DOI:10.1074/mcp.M112.024877

10. Molla, MHR, Aljahdali, MO. Identification of phytochemical compounds to inhibit the matrix-like linker protein VP26 to block the assembles of white spot syndrome virus (WSSV) envelope and nucleocapsid protein of marine shrimp: *In silico* approach. J. King Saud Univ. Sci. 2022; 34 (8), 102346.

11. Opo, FADM, Alkarim, S, Alrefaei, GI, Molla, MHR, Alsubhi, NH, Alzahrani, F et al. Pharmacophore-model-based virtual-screening approaches identified novel natural molecular candidates for treating human neuroblastoma. Curr. Issues Mol. Biol. 2022; 44: 4838-4858.

12. Hasan, MR, Alsaiari, AA, Fakhurgi, BZ, Molla, MHR, Asseri, AH, Sumon, MAA et al. Application of mathematical modeling and computational tools in the modern drug design and development process. Molecules 2022; 27(13): 4169.

13. Aljahdali, MO, Molla, MHR, Ahammad, F. Immunoinformatics and computer-aided drug design as new approaches against emerging and re-emerging infectious diseases. Antivir. Drugs 2022; doi:10.5772/INTECHOPEN.101367.

14. Islam, MR, Awal, MA, Khames, A, Abourehab, MAS, Hassan, WMI, Alam, R et al. Computational identification of druggable bioactive compounds from *Catharanthus roseus* and *Avicennia marina* against colorectal cancer by targeting thymidylate synthase. Mol. 2022; 27: 2089.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

15. Aljahdali, MO, Molla, MHR, Ahammad, F. Compounds identified from marine mangrove plant (*Avicennia alba*) as potential antiviral drug candidates against WDSV, an *In-Silico* approach. Mar. Drugs 2021; *19: 253*.

16. Islam, SI, Mou, MJ, Sanjida, S, Mahfuj, S. An *In-silico* approach for identifying phytochemical inhibitors against nervous necrosis virus (NNV) in Asian Sea Bass by targeting capsid protein. Genetics of aquatic organisms 2022; 6(2), GA4872022 doi:10.4194/GA487.

17. Samad, A, Haque, F, Nain, Z, Alam, R, noman, MAA, Molla, MHR et al. Computational assessment of MCM2 transcriptional expression and identification of the prognostic biomarker for human breast cancer. Heliyon 2020; **6**: e05087.

18. Islam, SI, Mahfuj, S, Alam, MA, Ara, Y, Sanjida, S, Mou, MJ. Immunoinformatic approaches to identify immune epitopes and design an epitope-based subunit vaccine against emerging tilapia lake virus (TiLV). Aquac. J. 2022; 2(2): 186-202.

19. Zhang, YI. TASSER server for protein 3D structure prediction. BMC Bioinformatics 2008; 9:1–8.

20. Al-Refaei, MA, Makki, RM, Ali, HM. Structure prediction of transferrin receptor protein 1 (TfR1) by homology modelling, docking, and molecular dynamics simulation studies. Heliyon 2020; 6. https://doi.org/10.1016/j.heliyon.2020.e03221

21. Forli, S, Huey, R, Pique, ME, Sanner, M, Goodsell, DS, Olson, AJ. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. Nat. Protoc. 2016;11, 905-919.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

22. Qasaymeh, RM, Rotondo, D, Oosthuizen, CB, Lall, N, Seidel, V. Predictive binding affinity of plant-derived natural products towards the Protein Kinase G enzyme of *Mycobacterium tuberculosis* (MtPknG). Plants 2019; 8(11). doi: 10.3390/plants8110477.

23. Cosconati, S, Forli, S, Perryman, AL, Harris, R, Goodsell, DS, Olson, AJ. Virtual screening with AutoDock: theory and practice. Expert Opin. Drug Discov. 2010; 5, 597-607.

24. 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. 2017; 7:72417.

25. Hemmerich, J, Ecker, GF. *In silico* toxicology: from structure–activity relationships towards deep learning and adverse outcome pathways. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2020;10. doi: 10.1002/wcms.1475

26. Salo-Ahen, OMH, Alanko, I, Bhadane, R, Bonvin, AMJJ, Honorato RV, Hossain, S et al. Molecular dynamics simulations in drug discovery and pharmaceutical development. Process. 2021*; 9(1):71*.

27. De Clercq, E. Antiviral agents active against influenza A viruses. Nat. Rev. Drug Discov. 2006; 5:1015-1025.

28. Tabassum, H, Ahmad, IZ. Molecular docking and dynamics simulation analysis of thymoquinone and thymol compounds from *Nigella sativa* L. that inhibits P38 protein: probable remedies for hepatocellular carcinoma. Med. Chem. 2020;16: 350–357.

29. Elekofehinti, OO, Iwaloye, O, Josiah, SJ, Lawal, AO, Akinjiyan, MO, Ariyo, EO.

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.19.545628; this version posted June 23, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

543 Molecular docking studies, molecular dynamics and ADME/tox reveal therapeutic
544 potentials of STOCK1N-69160 against papain-like protease of SARS-CoV-2. Mol.
545 Divers. 2021; 25: 1761-1773.

546 30. Mathew, SM, Benslimane, F, Althani, AA, Yassine, HM. Identification of potential
547 natural inhibitors of the receptor-binding domain of the SARS-CoV-2 spike protein using
548 a computational docking approach. Qatar Med. J. 2021;1(12).

549 31. Yang J, Zhang, Y. Protein structure and function prediction using I-TASSER. Curr.
550 Protoc. Bioinformatics 2015; **52**, 5.8.1-5.8.15.

551 32. Lipinski, CA. Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov.*
552 Today Technol. 2004;1: 337–341.

553 33. Dong, J, Wang, N-N, Yao, Z-J, Zhang, L, Cheng, Y, Ouyang, D et al. ADMETlab: a
554 platform for systematic ADMET evaluation based on a comprehensively collected
555 ADMET database. J. Cheminform. 2018;10(1): 29.

556 34. Samad, A, Huq, MA, Rahman, MS. Bioinformatics approaches identified dasatinib and
557 bortezomib inhibit the activity of MCM7 protein as a potential treatment against human
558 cancer. Sci. Rep. 2002; 12.

559

560

561