Manuscript_ WSSV.docx

By Mohammed Aljahdali

WORD COUNT 3225
TIME SUBMITTED 21-SEP-2022 05:06AM
PAPER ID 90719021

Identification of phytochemical compounds to inhibit the matrix-like linker protein VP26 to block the white spot syndrome virus (WSSV) and nucleocapsid protein of marine shrimp: In silico approach

Abstract

White spot syndrome virus (WSSV) is an enveloped pathogenic virus of cultured shrimp that has a mortality rate of up to 100% in intensive aquaculture systems. The virulent pathogen causes an economic loss estimated at US$1 billion annually to global shrimp and prawn production. To date, no effective antiviral or vaccine has been progress that can reduce the economic losses caused by the pathogen. The complex genomic structure of the nucleocapsid of a virus, consisting of double-stranded DNA as genetic material surrounded by lipid envelope protein. It has been found that VP26, a major envelope protein interacts with the virus VP51 capsid and makes a bridge between the envelope and nucleocapsid protein resulting in mature virion of the virus. Blocking the interaction by targeting VP26 can hinder the mature virion production and non-infectivity of the virus. Therefore, in order to blocking the function of VP26, an in-silico drug design approach has been utilized to screen th e potential antiviral compounds from the medicinal plant, *Withania somnifera* (WS) of Saudi Arabia. Isolated thirty-nine (39) phytochemicals compounds from the plant have initially been screened by molecular docking method. The best three compounds, namely, withanolide (CID: 118701104), Coagulin Q (CID: 10100411), and withanolide D (CID: 23266161) have been selected based on their docking score -8.2 kcal/mol, -8.1 Kacl/mol and -7.9 Kcal/mol, respectively, which have been further evaluated based on pharmacokinetics including ADME (Absorption, Distribution, Metabolism and Excretion) and toxicity approaches. Additionally, 100 ns molecular dynamics (MD) simulation methods confirmed the structural stability of the compounds to the binding sites of the VP26 protein. Therefore, the selected compounds can be utilized as an antiviral drug to fight against the WSSV virus, which will bring remarkable advancement in the aquaculture sector.

Keywords: In-silico drug design; VP26; virtual screening; *W. somnifera*; molecular dynamics simulation

1. Introduction

Shrimp culture have been increased day by day due to their high demand, which included a rich source of protein and omega-3 fatty acid as well as antioxidant astaxanthin (6, 11, 12, 13) (Karnila et al., 2020). However, the devastating effect of numerous virus disease has brought tremendous negative impacts on the shrimp industry. Among them, white spot syndrome virus (WSSV) is one that have been destroyed the shrimp industry since last three decades. Therefore, it led to great economical losses in every year such as about USD 6 million in 2016 by 100% mortality within a week at many farms (Trang et al., 2019). Only genetic improvement fry namely specific pathogen free (SPF) fry has been established to promote sustainable shrimp farming.

White spot syndrome virus (WSSV) has been spreading white spot syndrome disease which is a highly contagious viral disease for farm based penaeid shrimps as well as wild shrimp (Cavalli et al., 2011). It mainly affects the gills and hepatopancreas of the shrimp, and damaged severely. WSSV is a doublestranded DNA virus (300 kbp) in the Nimaviridae family of the genus Whispovirus. WSSV has 22 capsid proteins that express four most abundant ones such as VP26, VP28, VP24 and VP19 where VP26 and VP28 constituted 60% of the envelope protein (Zhou et al., 2009). There has a little knowledge about the predicted protein

~ 184WSSV because there have no known homology sequences in the repositories. However, In the coding region of WSSV, there have some variations including two genomic deletions between ORFs wsv461/wsv464 (14/15) and ORFs wsv77/wsv502 (23/24), and a variable number of tandem repeats (VNTRs) occurring within wsv129 (ORF75), wsv178 (ORF94) and wsv249 (ORF125) (Liao et al., 2021). The variable regions are not only use for molecular marker but also have some relationships with viral evolution and infection phenotype which is apparent by the externally white sign of shrimp.

In GenBank, V26 has no similarity protein but contains 204 amino acids along with the estimated molecular mass of 22 kDa (Aljahdali et al., 2021a). The N terminal of VP26 is not only hydrophobic but also included a putative transmembrane anchor composed of a helix formed through the amino acid 12 to 34. The two α-amino acid have positive charge region behind the anchor with including C-terminal end of the protein from the cytoplasmic side. The cysteine is placed in the C-terminal domain indicate that the disulfide bond are not formed. It forms complex protein by interacting with other low abundance envelop protein, which occur naturally from VP26 (Robinson and Bulleid, 2020). It has a tegument feature and interacting with others proteins that may help to transport WSSV nucleocapsid to the host through the cytoskeleton. Therefore, V26 has been interacting with variety of protein molecules that plays a key role to virus invasion in later stage.

In the field of drug discovery, phytochemical bioactive compounds have a potential involvement in many cases to inhibit the disease related protein (Ling et al., 2014). These have been used because of easy availability, reasonable cost, and not harmful at all as well as the feasibility of oral administration. *Withania somnifera* (WS) has the anti-tumorigenic properties that act against the cancerous cell. It has also potential inhibitor for the treatment of neurodegenerative diseases. Moreover, In-silico drug design techniques is the best key for saving both the time and cost by predicting approximate protein and selecting small bioactive compounds (Aljahdali et al., 2021b). However, the viable phytochemical compounds are the critical objective for the drug design by CAAD. Therefore, molecular docking-based scoring function selected by a specific target, and the dynamic simulation confirm the stability of a drug candidate to the target receptor (Hasan et al., 2022). As a result, In-silico approaches of drug design will predict the new drug candidate against the specific virus by targeting suitable protein.

2. Materials and Methods

2.1. Protein preparation

The structure of V26 protein has been taken from the large RCSB data bank (PDB). The WSSV protein V26 having a PDB identity 2EDM with molecular weight of 22 kDa was used. The protein was recovered and processed using AutoDockTools, which computed gasteiger charges, merged nonpolar hydrogen, added hydrogen atoms, and eliminated metal ions and cofactors. (Aljahdali et al., 2021b).

2.2. Compounds retrieval and preparation

The bioactive compounds have been derived from the medicinal plant for drug design and discovery. Therefore, the popular database (https://cb.imsc.res.in/imppat/home) of Indian Medicinal Plant, Phytochemistry and Therapeutics (IMPPAT) has been surfing for the plant *Withania somnifera*. The database consisting of >1742 Indian medical plants and about >9500 phytochemical compounds which were utilized for natural based drug discovery.

2.3. Grid generation and active site identification

Active site (AS) of VP26 protein has been documented from the specific shape of protein through the CASTp 3.0 web server (http://sts.bioe.uic.edu/). The chemical reaction of protein was binding with a specific molecular substrate. Auto Dock vina, a virtual screening tool developed by PyRx, has been used to determine the receptor grid generation process and binding site from the complex AS. (Samad et al., 2020).

2.4. Molecular Docking

The PyRx virtual screening tool AutoDock Vina has been used for the molecular docking to detect the binding mode from the chosen receptor and designated phytochemical compounds. Entirely, the BIOVIS Discovery Studio Visualizer Tools 16.1.0.41 has been used for the observation of binding interaction of protein-ligands complex (Aljahdali et al., 2021b).

2.5. PK properties prediction

The drug absorption, distribution, metabolism, and excretion (ADME) mainly involved in drug to enter and out of the body are which related to the intensity and time course. The popular server SwissADME has been used in the study to evaluate the PK properties that can predict PK and drug-likeness properties of small molecules (Daina et al., 2017).

2.6. Toxicity prediction

Toxicity is the fundamental obstacle to drug design. The AdmetSAR 2.0 (http://lmmd.ecust.edu.cn/admetsar2/) web-based server was utilized to document toxicity from the selected compound.

2.7. Molecular Dynamic Simulations (MD)

The snapshots were taken from molecular dynamic simulation using Schrodinger's Maestro interface version v9.5. Desmond module in the Schrodinger package from Simulation Interaction Diagram (SID) were used for the analysis of simulation even as well as quality of molecular dynamic simulation (Aljahdali et al., 2022).

2.8. RMSD and RMSF analysis

RMSD was used for the estimate of average distance produced through the movement of a particular atom at the specific time compared to the reference time in MD simulation. In the study, time frames were aligned and estimated based on reference time (100 ns) from the RMSD of the protein fit ligand atom. On the other hand, the large number of residues generated by MD simulation, RMSF was employed as the preferred method for observing the conformational changes that occurred within a protein structure.

3. Results

3.1. Phytochemical retrieval and preparation

Initially, Indian natural and medicinal phytochemical compound library (IMPPAT) has been searched for the compounds from the desirable mangrove plant. Thirty-nine compounds has been identified from the mangrove plant Table (S1). The bioactive compounds have not only been retrieved but also saved in a 2D (SDF) file format. Optimized and prepared compounds were converted into pdbqt file format during the ligand preparation steps for further evaluation.

3.2. Active Site Identification and Receptor Grid Generation

The popular web serve CASTp 3.0 was surfing to detect the binding site of protein from the separate AA restudies in a specific region. The complex AS pocket was identified from the retrieved binding site position of VP26 protein. The AS pocket were represented in the different colours by ball shape with 9 AA restudies (PRO102, THR104, PRO106, ASN115, MET116, THR117, TYR153, ASP170, CYS172) (Fig. 1)

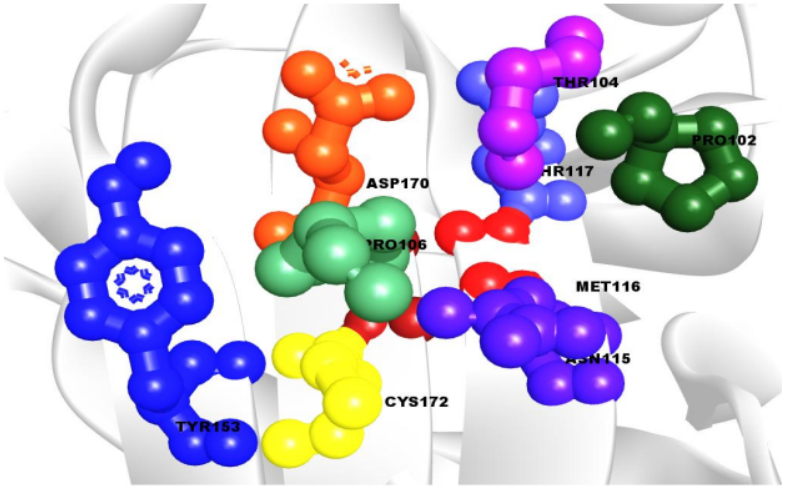


Fig. 1. The active site and corresponding binding site of the VP26 protein are shown. The active sites (AS) of the VP26 protein are represented by a ball with red, green, yellow, and blue colors, which correspond to their binding site positions.

3.3. Molecular docking analysis

The thirty nine phytochemical compounds have been selected from the marine plant of *W. somnifera* by using PyRx tools AutoDock vina wizard. The highest binding affinities range were -8.2 kcal/mol that found in alpha Amyrin. The distribution range of binding affinities between -8.2 and -2.5 kcal/mol were observed from the plant. The best three compounds have been selected by the binding affinities for in silico drug design. The additional table included information on the compound identities, chemical names, and 2D structures of the top three ligands as well as the binding affinities of the controls (S1).

3.4. Interpretations of protein–ligands interaction

The BIOVIA discovery studio visualizer was performed for observing the protein ligand interactions from the selected compounds and protein VP26 based on best binding affinities. The initially compound (CID118701104) was observed and were there had strong hydrogen bond, hydrophobic and Pi-Alkyl bonds. The hydrogen bond was formed at the position of ARG51, TYR46, ARG67 and MET49, and hydrophobic bond at the position of TYR69, MET50, THR71, GLN48 and PRO72 (Fig. 2).

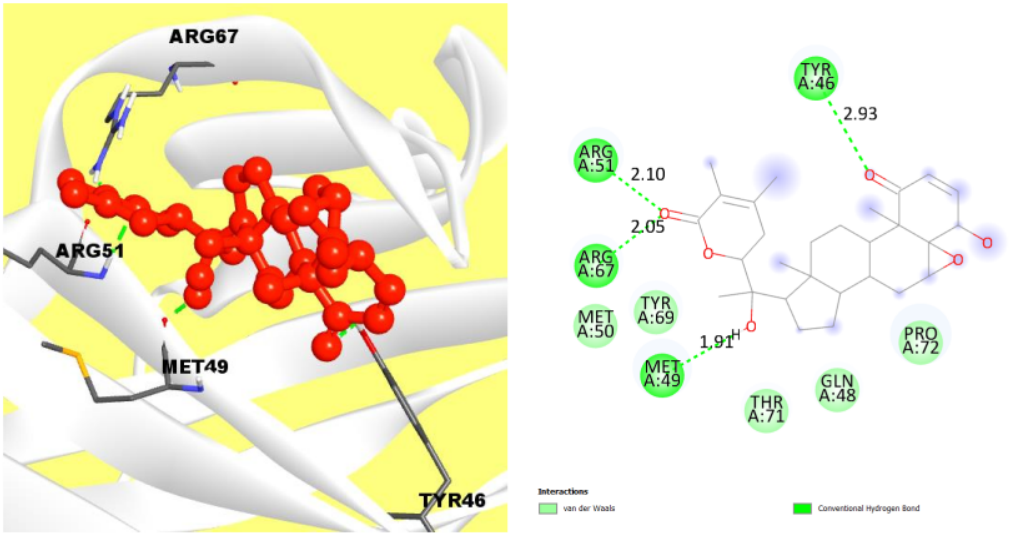


Fig. 2. The interaction of protein and ligand between compound 118701104 (PubChem CID) and VP26 protein and left side representing 2D complex protein-ligand interaction.

Again, for the compound CID 11100411, it was clearly observed that it has one conventional hydrogen bond at the position of ARG136, ASN186, ASP132, PRO72 and TYR46 where there were hydrophobic bonds from the position of THR134, GLY188 and MET49 (Fig.3).

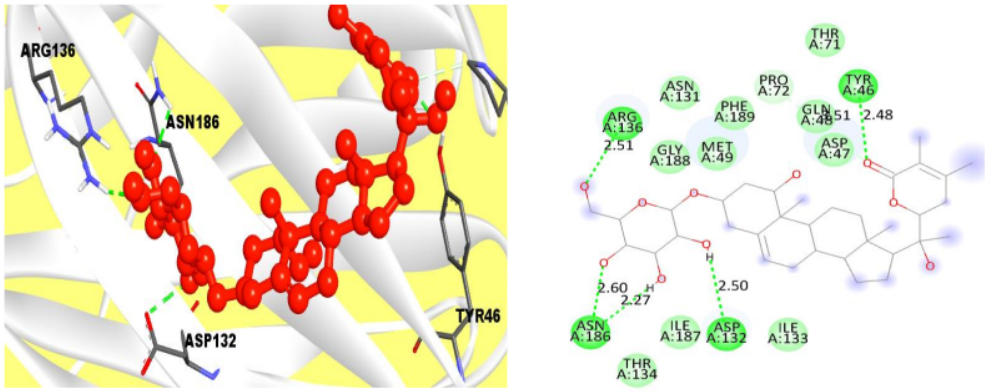
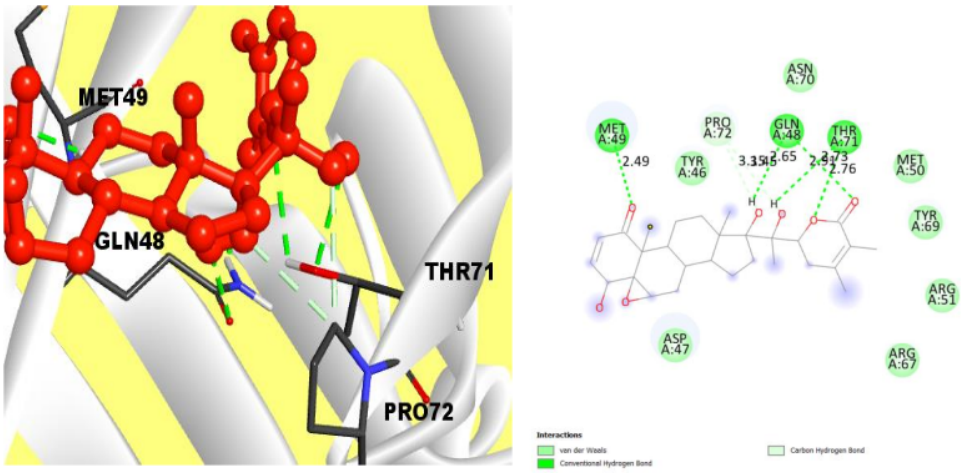


Fig. 3. The interaction of protein and ligand between compound 11100411 (PubChem CID) and VP26 protein, and left side is representing 2D complex protein-ligand interaction.

In case of the compound CID23266161, a conventional strong hydrogen bond and carbon hydrogen bond were found between complex protein-ligand interaction. The conventional hydrogen bond at the position of MET49, GLN48, THR71, and PRO72 were found where the hydrophobic bond was located at the TYR46, ASP47, PRO72 and ASN70 position (Fig. 4).



166 Fig. 4. Interaction between the compounds CID: 23266161 and V26 protein. Left side indicates
167 3D interaction

168 **3.5. AMDE properties**

169 The ADME analysed not only assist the physiochemical properties but also delivered the
170 hypothesis for choosing top treatment candidates. Therefore, the pharmacological efficacy of
171 the three selected compounds of W. *somnifera* has been demonstrated through the use of
172 pharmacophore features such as physicochemical qualities, lipophilicity, water solubility,
173 drug-likeness, and medicinal chemistry W. *somnifera* (Table 1).

174 Table 1. The three chosen compound pharmacokinetic parameters are displayed.

Name	Properties	CID: 118701104	CID: 10100411	CID: 23266161
physicochemical properties	MW (g/mol)	470.60	620.77	486.60
	Heavy atoms	34	44	35
	Rotable bond	2	5	2
	H-bond acceptors	6	10	7
	H-bond donors	2	6	3
Lipophilicity	Log Po/w	3.74	3.76	3.42
Water solubility	Log S (ESOL)	-4.59	-4.88	-3.79
Pharmacokinetics	GI absorption	High	Medium	High
Drug-likeness	Lipinski	Yes	Yes	Yes
Medi. Chemistry	Synth. accessibility	Easy	Easy	Easy

175

176 **3.6. Toxicity prediction**

In silico drug design methods can reduce the large number of biological experimental test and eliminate toxic chemicals compounds in a short period. Therefore, admetSAR 2.0 web server has been used for identifying the toxicity of compounds. It has identified the hepatotoxicity, carcinogenicity, immunotoxicity, mutagenicity, and cytotoxicity through the admetSAR 2.0 web server (Table 2).

Table 2. List of the drug-induced toxicity profile includes hepatotoxicity, carcinogenicity, immunotoxicity, mutagenicity, cytotoxicity of the three selected compounds.

PubChem ID	Hepatotoxicity	Carcinogenicity	Immunotoxicity	Mutagenicity	Cytotoxicity
CID: 118701104	Inactive	No	Inactive	Inactive	Inactive
CID: 10100411	Inactive	No	Lightly active	Inactive	Inactive
CID: 23266161	Inactive	No	Inactive	Inactive	Inactive

3.7. MD Simulation Analysis

In the artificial environment, MD simulation has confirmed the stability of protein-ligand complexes. However, it has analysed the steady nature and conformation by performing MD simulation with 100 ns that observed the stability of the protein - ligand complexes. As a result, simulation trajectories have examined the RMSD, RMSF, intermolecular hydrogen bonds and protein - ligand contact analysis for the possible drug candidate.

3.8. RMSD analysis

The observed value and estimated value have changed with frame to a reference time to 1–3 Å or 0.1–0.3 nm that were entirely admissible. The protein structure (Cα) residues and ligand fit protein were analysed within 100 ns frame in the study. Therefore, VP26 protein from the WSSV and ligand from the compounds were estimated for the considerable fluctuation. All of compounds showed the optimum fluctuation >3.0 Å. However, the fluctuations abnormality happened when the system was not properly equilibrated due to their requirement (Fig. 5).

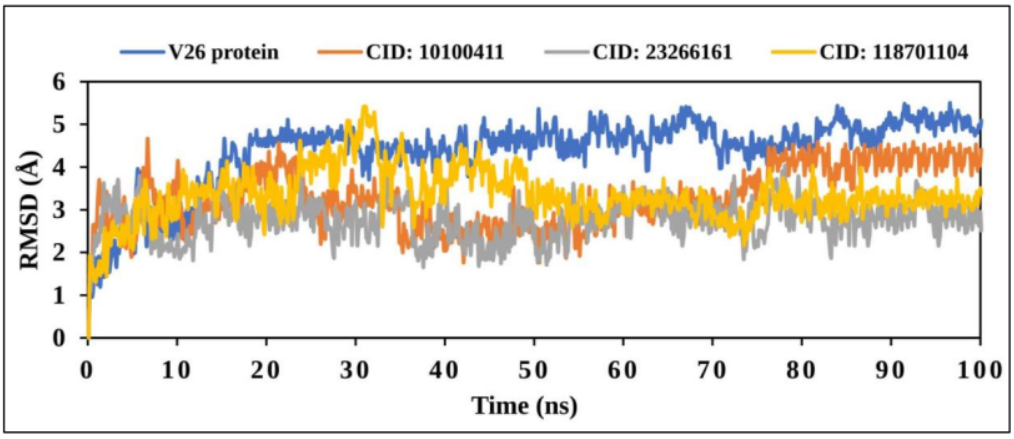


Fig. 5. The graph represents the RMSD values from Cα atoms (blue curves) of VP26 protein and natural compounds, where the compounds have been shown as CID: 10100411 (red), CID: 23266461 (gray) and CID: 118701104 (yellow) with regards of 100 ns simulation time.

3.9. RMSF analysis

RMSF (root mean square fluctuation) was determined the positional difference between the protein and ligands in the particular time frame. It provided the RMSD value, evidence about protein heterogeneity and the movement of macromolecules through the performing data about RMSF. However, RMSF value provided the change of receptor along with amino acid chain for characterizing the protein (Fig. 6).

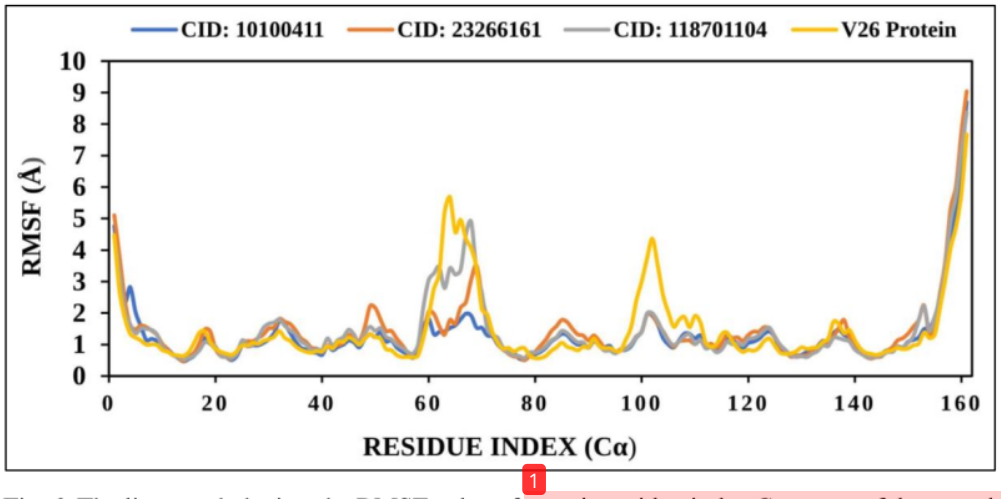


Fig. 6. The line graph depicts the RMSF value of protein residue index Cα atoms of the complex structure. CID 10100411 (blue), CID 23266161 (red), CID118701104 (gray) and VP26 protein (blue) shows in separate colour.

3.10. Protein - Ligand Contact Mapping

The interactions of protein with ligand is the fundamental factor to monitor which happened throughout the simulation during the drug design. It has been categorized by hydrogen bonds, hydrophobic, ionic and water bridge during the simulation (Fig. 7). In the study, Hydrogen bond plays the significant role in the periods of ligand binding, which stimulated the drug specificity, metabolization and adsorption. The backbone acceptor was occurred in 21 times whereas backbone donor was happened in eleven times during the protein and ligand interaction.

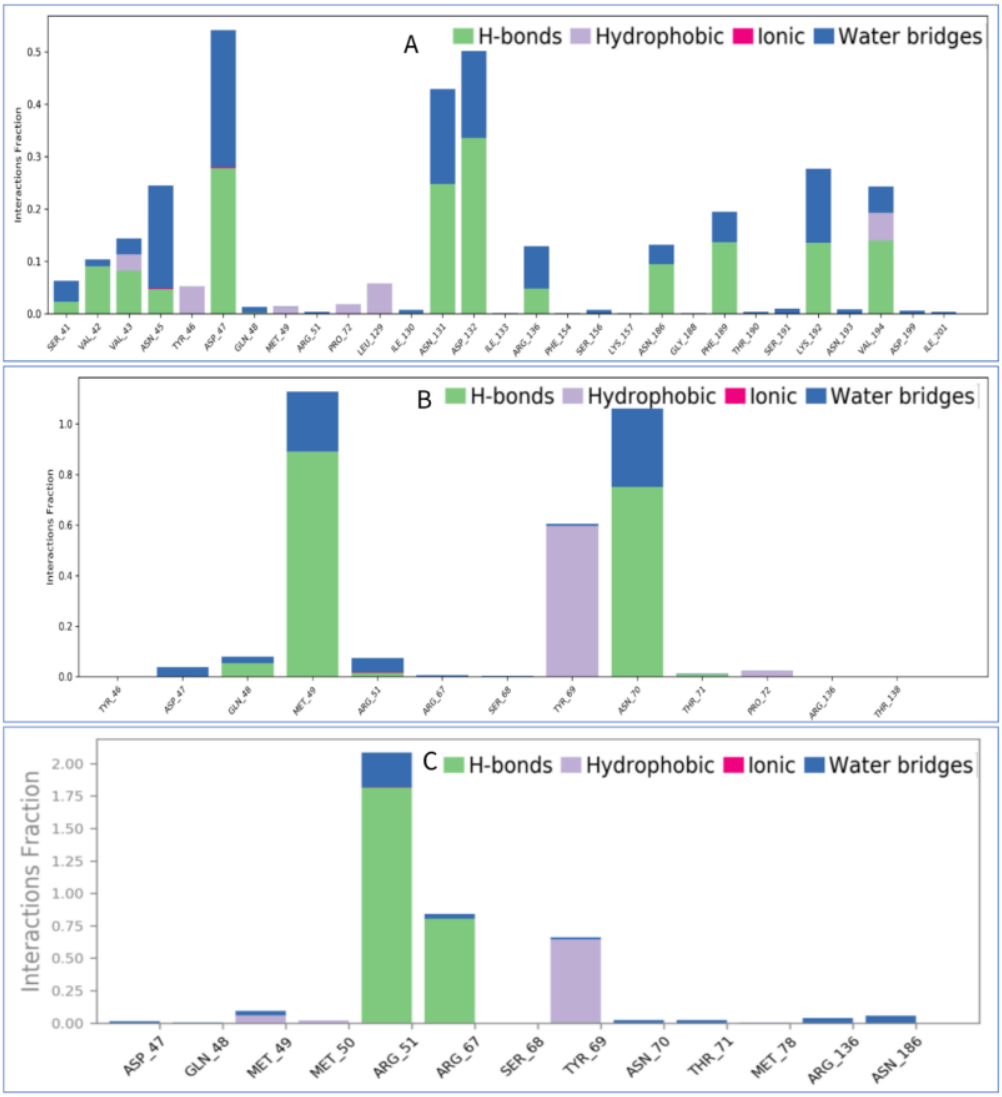


Fig. 7.

The bar charts indicate contact map of VP26 protein with the potential natural compounds from *Withania somnifera*, i.e., (A) CID10100411, (B) CID23266461, and (C) CID118701104 take out from 100 ns simulations.

4. Discussion

In the modern drug design, CADD has been performing as an indispensable tools for reducing cost, time duration and extra employment contribution during the modern drug discovery (Aljahdali and Molla, 2022). Therefore, it speeds up the research fields in broad spectrum through the efforts of biological and synthetic aspect. As a result, critical problems are solved within a short period in the globe such as Covid-19 vaccine invention thereby reducing the infection (Mostaghimi et al., 2021). Moreover, the CADD approach such as molecular docking, ADMET and dynamic simulation help to find out the best biological efficacy during the

selection of drug like small molecules (Islam et al., 2022). CADD approach can not only document the specific diseases by the binding of ligand, and interacting and inhibiting the specific protein but also understand the behaviour mode between the ligand and specific target molecule. Dynamic simulation revealed the mechanism of complicated protein-ligand interaction, and molecular docking detected or predicted the binding mode of ligand and protein during the drug candidate selection process for a certain disease.(JL and MK, 2018).

However, the 39 phytochemical compounds were selected and screened from the mangrove plant *W. somnifera* for targeting VP26 protein to inhibit WSSV infectious disease. Therefore, the highest binding affinity of compound has been chosen by the molecular docking approach after screening. In the present study, three compounds such as CID118701104 (A), CID10100411 (B), and CID23266461 (C) were selected from the 39 compounds on the basis of the higher binding score to the lower binding score -8.2, -8.1 and -7.9 kcal/mol, respectively. The protein and ligand interaction was determined by the binding of strong hydrogen and hydrophobic bond (Bulusu and Desiraju, 2020).

The metabolic kinetics of drug candidate as a small molecule inside of the body were focused by PK. ADME properties included the molecular weight and TPSA that can affected small molecules through the permeability of biological barrier (Daina et al., 2017). The molecular weight has gone down because of permeability, and TPSA has made the lower molecular weight more permeable. Moreover, lipophilicity can focus on the inorganic and aqueous phase of the selected molecules which influenced the absorption, transport, permeability, binding, and distribution in various organs and tissues along with endocrine system (Blokhina et al., 2016). The higher LogP values, resulting from the lower absorption were correlated with each other whereas the lower LogS value may affected the solubility of the selected molecules. The bilayer membranes have been closed when the hydrogen bond donors and acceptors were increased and decreased. The rotable bond should be within 10 that focused on the oral bioavailability of drug candidates. PK properties were documented from all selected compounds and further evaluation was performed during the study (Chung et al., 2015).

5. Conclusions

The infectious viral diseases, caused by white spot syndrome disease (WSSV) has been seriously devastating the shrimp farming industry since the last two decades through the mass killing of shrimps in the culture systems. It not only has wiped out the entire population within few days but also lead to huge economic losses in the world. However, there have not been developed any effective therapeutics against the viral infection till now. Therefore, this in silico technique has been advanced to find the natural and effective antiviral candidate against the hinder mature virion production and non-infectivity in the shrimp. Finally, it will not only provide new information about the matrix-like linker protein VP26 insight in the penaeid shrimp but also will open new avenue for the drug design of shrimp diseases in a significant and worthwhile approach.

Manuscript_ WSSV.docx

ORIGINALITY REPORT

17%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.mdpi.com Internet	224 words — 6%
2	academic.oup.com Internet	122 words — 3%
3	www.ncbi.nlm.nih.gov Internet	77 words — 2%
4	Foysal Ahammad, Rahat Alam, Rasel Mahmud, Shahina Akhter et al. " Pharmacoinformatics and molecular dynamics simulation-based phytochemical screening of neem plant against human cancer by targeting MCM7 protein ", Briefings in Bioinformatics, 2021 Crossref	51 words — 1%
5	Snehal Jamalpure, Gauri Panditrao, Prabir Kumar Kulabhusan, A. S. Sahul Hameed, K. M. Paknikar, Manali Joshi, J. M. Rajwade. "In silico studies on the interaction of phage displayed biorecognition element (TFQAFDLSPFPS) with the structural protein VP28 of white spot syndrome virus", Journal of Molecular Modeling, 2020 Crossref	20 words — 1%
6	Elliasu Y Salifu, Issahaku A. Rashid, James Abugri, Festus Eric Osei, Joseph A Ayariga. "Molecular docking studies of potential inhibitors of acyl carrier protein	19 words — 1%

and acetyl CoA Carboxylase in Plasmodium falciparum.", Cold Spring Harbor Laboratory, 2022
Crossref Posted Content

7 Iqrar Ahmad, Dileep Kumar, Harun Patel. " Computational investigation of phytochemicals from (Indian ginseng/ashwagandha) as plausible inhibitors of GluN2B-containing NMDA receptors ", Journal of Biomolecular Structure and Dynamics, 2021 — 14 words — < 1%
Crossref

8 www.ashdin.com — 12 words — < 1%
Internet

9 Alapide-Tendencia, ELEONOR V.. "The Relation Between Farming Practices, Ecosystem, and White Spot in Syndrome Virus (WSSV) Disease Outbreaks in Penaeus Monodon Farms in the Philippines", Wageningen University and Research, 2021 — 11 words — < 1%
ProQuest

10 Shahina Akhter, Md. Walid Hossain, Sharmin Sultana, Jannatul Ferdous Jharna et al. "Ruellia prostrata Poir. activity evaluated by phytoconstituents, antioxidant, anti-inflammatory, antibacterial activity, and in silico molecular functions", Journal of Saudi Chemical Society, 2022 — 11 words — < 1%
Crossref

11 Desrina, Haryadi. "On the Role of the Polychaete Dendronereis Spp. in the Transmission of White Spot Syndrome Virus in Shrimp Ponds", Wageningen University and Research, 2021 — 9 words — < 1%
ProQuest

12 www.frontiersin.org — 9 words — < 1%
Internet

13 Md.Shahadat Hossain, Anirban Chakraborty, Biju Joseph, S.K. Otta, Indrani Karunasagar, Iddya Karunasagar. "Detection of new hosts for white spot syndrome virus of shrimp using nested polymerase chain reaction", Aquaculture, 2001

8 words — < 1%

Crossref

14 Vikas Jha, Anjali Bhosale, Prakruti Kapadia, Agraj Bhargava et al. "Multitargeted molecular docking study of phytochemicals on hepatocellular carcinoma", Journal of Applied Biology & Biotechnology, 2022

8 words — < 1%

Crossref

15 Reena Kumari, Ravi Rathi, Seema R Pathak, Vikram Dalal. "Structural-based virtual screening and identification of novel potent antimicrobial compounds against YsxC of Staphylococcus aureus", Journal of Molecular Structure, 2022

7 words — < 1%

Crossref

EXCLUDE QUOTES	OFF	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	OFF	EXCLUDE MATCHES	OFF