bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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 aCC-BY-ND 4.0 International license.

Screening of anti-*Acinetobacter baumannii* phytochemicals, based on the potential inhibitory effect on OmpA and OmpW functions

Shahab Shahryari[1]*, Parvin Mohammadnejad[2], Kambiz Akbari Noghabi[1]

Running title: Natural products against *A. baumannii*

[1]*Department of Energy & Environmental Biotechnology, National Institute of Genetic Engineering and Biotechnology (NIGEB), P. O. Box 14155-6343, Tehran, Iran,*[2]*Division of Agricultural Biotechnology, National Institute of Genetic Engineering and Biotechnology (NIGEB), P.O. Box: 14965 /161, Tehran, Iran*

*Author for correspondence:

E-mail: sha.shahriari@yahoo.de

Tel: +98- 214478 7301

Fax: +982144787399

For submission to bioRxiv

Resubmission date of revised version: 10.07.2021

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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 aCC-BY-ND 4.0 International license.

Summary

Therapeutic options, including last-line or combined antibiotic therapies for multi-drug resistant (MDR) strains of *Acinetobacter baumannii* are ineffective. The outer membrane protein A (OmpA) and outer membrane protein W (OmpW) are two porins known for their different cellular functions. Identification of natural compounds with the potentials to block these putative porins can attenuate the growth of the bacteria and control the relating diseases. The current work aimed to screen a library of 384 phytochemicals according to their potentials to be used as a drug, and potentials to inhibit the function of OmpA and OmpW in *A. baumannii*. The phytocompounds were initially screened based on their physicochemical, absorption, distribution, metabolism, excretion, and toxicity (ADMET) drug-like properties. Afterward, the selected ligands were subjected to standard docking calculations against the predicted three-dimensional structure of OmpA and OmpW in *A. baumannii*. We identified three phytochemicals (isosakuranetin, aloe-emodin and pinocembrin) possessing appreciable binding affinity towards the selected binding pocket of OmpA and OmpW. Molecular dynamics (MD) simulation analysis confirmed the stability of the complexes. Amongst them, isosakuranetin was suggested as the best phytocompound for further *in vitro* and *in vivo* study.

Key words: Antimicrobial resistance, *Acinetobacter baumannii*, phytochemicals, molecular docking.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

1. Introduction

In recent years, antimicrobial resistance (AMR) by Gram-negative bacteria has become a serious problem with a possible devastating impact on the economy and human life [1]. With the advent of the post-antibiotic era, the need to use more effective and safer antimicrobial compounds is fully felt. [2]. As one of the six superbug ESKAPE pathogens, *Acinetobacter baumannii* is a naturally transformable Gram-negative pathogen with increasing clinical importance [3]. Multi-drug resistant (MDR) strains of *Acinetobacter* are responsible for chronic infections, and can rapidly confer resistance to the most of currently used antibiotics. Thus, *A. baumannii* is in the global priority pathogen list (PPL) to develop effective antimicrobial therapies against it [4]. Due to the rapid changes in the genetic constitution of *A. baumannii* and lack of enough knowledge about resistance mechanisms, no efficient antibiotic against this microorganism has yet been developed by the pharmaceutical industry [5]. This paucity of effective antibiotics has revived scientific interest in finding efficient antimicrobial agents, with the ability to kill, inhibit the growth, or inhibit the activity of some essential virulence factors of *A. baumannii* [6-9].

The outer membrane proteins (OMPs) of Gram-negative bacteria, have attracted some attention as drug targets due to their extensive functions [10, 11]. Outer membrane protein A (OmpA) as the most abundant slow porin in members of genus *Acinetobacter* plays different rules in cells [12, 13]. This protein is closely associated with the virulence and survival of the bacteria under harsh conditions [14-16]. OmpA structure is composed of a β-barrel and a periplasmic (OmpA-like) domain. In addition to a structural role for OmpA of *A. baumannii*, it can play an important role in biofilm formation and adhesion to host cells [12, 17-21]. The absence of OmpA in *A. baumannii* and *E. coli* is associated with significant reduction of virulence [22, 23]. Also, it has been suggested that 4 external loops, connecting N-terminal β-strands,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

participate in some interactions with host-cells (or other surfaces), and can be used as drug-targets [24].

Outer membrane protein W (OmpW) is another porin with a pivotal role in the uptake of nutritional substances such as iron [25]. The role of iron acquisition in the pathogenicity of *A. baumannii* has been reported previously [26], and disruption of this system can be effective in treating infections caused by this microorganism [27-29]. Decreased minimum inhibitory concentrations (MICs) of several antibiotics due to deletion of OmpW from *E. coli* [30], as well as reduced intestinal colonization of *ΔompW* mutant strains of *V. cholerae* [31], are among the indications that OmpW can be served as a potential drug target in Gram-negative bacteria. *In silico* and *in vitro* studies on OmpW resulted in the identification of three inhibitors of this porin in *A. baumannii* [32]. Therefore, any molecule with the ability to disrupt the functions of OmpA and OmpW can help treat the disease caused by this bacterium.

Researchers have reported the benefits of infectious disease treatment using poly-pharmacological drugs [33]. In light of this, the aim of the current study was *in silico* screening of the plant based compounds with potential anti-*A. baumannii* activities. The adhesion property of OmpA and the function of OmpW in the uptake of nutrients were targeted for the screening of compounds. For this purpose, a library consisting of 384 plant-based compounds, known as an anti-bacterial or anti-viral compound, was selected for this study [9, 34-36]. We applied *in silico* approaches to prioritize these biomolecules for toxicological evaluation and safety concerns [37, 38]. The phytocompounds with the highest affinity for both OmpA and OmpW were selected as the multi-target drug candidates. The results for molecular docking were further evaluated using molecular dynamics (MD) simulation. Our results led to the identification of three potential mutual OmpA and OmpW functional inhibitors, isosakuranetin, aloe-emodin and pinocembrin. Isosakuranetin was suggested for further *in vitro* and *in vivo* studies.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

2. Data and methodology

2.1. Tools and servers

A list of all tools and servers used in this work with a brief description is illustrated in Table S1.

2.2. Targets sequences

The complete amino acid sequences of OmpA (BAN86529 accession number) and OmpW (BAN89067 accession number) were retrieved from the GenBank database at National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/protein/) in FASTA format.

2.3. Signal peptide prediction

The secretary nature of selected proteins was predicted using SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/), which integrates a prediction of cleavage sites and signal or non-signal peptide prediction based on a combination of several artificial neural networks. LipoP 1.0 server (http://www.cbs.dtu.dk/services/LipoP) used to predict probable signal peptide within the sequence and signal peptidase I or II cleavage sites within the protein.

2.4. Homology Modeling, model refinement, and energy minimization

As the full three-dimensional (3D) structure of none of the proteins was available in the Protein Data Bank (PDB), a homology modeling was performed to predict the 3D structure of each protein using RaptorX web server (http://raptorx.uchicago.edu/StructurePrediction/predict/). The server predicts the absolute global quality and comparable global quality for each of the residues of the query sequence. In order to check the 3D structure, we used PyMOL molecular graphics system version 1.7.4.4 (Schrödinger, LLC, and Portland, OR, USA). To perform structure refinement and energy minimization of the 3D modeled protein structures, the online

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

server GalaxyRefine (http://galaxy.seoklab.org/) and YASARA software were used, respectively. GalaxyRefine employs the CASP10 assessment to refine the query structure, improving the structural and global quality of the 3D model. This method initially rebuilds side chains and performs side chain repacking and subsequently uses molecular dynamics simulation to achieve overall structure relaxation.

2.5. Validation of the 3D structures

To validate the refined and optimized 3D structures, three freely available web tools of RAMPAGE (http://mordred.bioc.cam.ac.uk/rapper/~rampage.php), ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php), and ERRAT (https://servicesn.mbi.ucla.edu/ERRAT/) were used. The server RAMPAGE evaluates the Ramachandran plot by applying PROCHECK principles. The ProSA validation method evaluates model accuracy and statistical significance with a knowledge-based potential. It plots overall excellence scores of the errors calculated in the query 3D structure. Further, the ERRAT server was used to check the reliability of modeled proteins and their overall quality factors.

2.6. Prediction of potential ligand binding pocket

To avoid blind docking and increase the accuracy of molecular docking, the top three possible binding sites of the proteins were identified using METAPOCKET version 2.0 (https://projects.biotec.tu-dresden.de/metapocket/index.php). As a META server, METAPOCKET is a popular consensus method, which combines eight methods of asLIGSITEcs, PASS, QSiteFinder, SURFNET, Fpocket, GHECOM, Concavity and POCASA to predict binding sites based on 3D protein structures.

2.7. Natural-compound ligand library

A library of 384 phytochemicals (Table S2) with potential anti-infection properties was used in this study [9, 34-36, 39]. The 3D structures of the molecules were retrieved from UniProt

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

darabase. In the case of lacking X-ray crystallography or NMR structure of the compound, a 3D conformer was obtained from PubChem database. The ligands 3D structures were energy minimized using Chem3D software before docking. The ligand files were finally saved in PDB format for molecular docking studies.

2.8. Pre-filtration and pharmacokinetic analysis of phytochemicals

A compound must pass through multiple filters to be considered as a novel drug. All of the bioactive candidates used in this study were evaluated based on their important physicochemical properties using SwissADME (http://www.swissadme.ch/) online server. Initially, those compounds satisfying Lipinski's rule of five (RO5) were selected and those that violated the rule were eliminated from downstream analysis. Pharmacokinetic properties of absorption, distribution, metabolism, excretion, and toxicity (ADMET) with a crucial role in the development of drug design were predicted using the pkCSM webserver at (http://biosig.unimelb.edu.au/pkcsm/) to decrease the failure rate of the compound for further analysis. For further screening PreADME (https://preadmet.bmdrc.kr/) server was used to filter the compounds based on their drug-likeliness, ADME properties and toxicity.

2.9. Molecular docking analysis

Molecular docking analysis was performed using AutoDock Vina software [40]. The .pdbqt file of ligands was prepared after specifying of rotatable, non-rotatable bonds, and torsion angles in the compounds using AutoDock tools. Grid boxes were defined according to the selected binding pockets and the aim of this study. The protein (receptor) files in .pdbqt format were prepared for a flexible body docking using AutoDock Tools software. As the target for OmpA was external loops because of their potential for adhesion, all the residues of external loops in the predicted binding pocket, were chosen for torsions. Concerning OmpW, only the residues of the binding pocket located toward the β-barrel were selected for torsion, because of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

the aim of this study for screening of compounds with the ability to block the OmpW channel. The docking process was performed after preparing a configuration file. The obtained interactions and generated poses were visualized using PyMOL. The docking results were evaluated for their orientation, number of hydrogen bonds, number of interacting residues, binding affinity (kcal/mol), and RSMD.

2.10. Drug-Target Interactions and Mechanism of Binding

The mechanism of binding in drug-target complexes was profiled using the LigPlot^{+} program (https://www.ebi.ac.uk/thornton-srv/software/LigPlus/). Hydrogen bonds, as well as hydrophobic interactions between potential drug molecules and the protein target were studied via 2D interaction diagrams of the docked drug-target complexes.

2.11. BOILED-Egg Model for Prediction of gastrointestinal absorption (GI) and blood-brain barrier penetration (BBB)

To reveal the capability of GI absorption and permeability of the Blood-Brain Barrier (BBB), the BOILED-Egg model of the potential drugs was predicted using SwissADME (http://www.swissadme.ch/).

2.12. Molecular dynamics (MD) simulation:

MD simulation was used to refine the 4 complexes obtained from molecular docking simulation. MD analysis was performed using a GROMACS 4.6.5 with GROMOS96 force field and simple point charge (SPC) as a water model [41].

To mimic the experimental conditions the simulation was carried out at 343 K. Some molecules of water were randomly replaced by ions for neutralization. All systems were simulated under the NpT ensemble within periodic boundary conditions. The neighbor list was updated with 10 step frequency using the grid search method. The leap-frog algorithm was utilized with a 2 fs

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

time step. Protein and solvent were coupled separately to a heat bath at the desired temperature with time constant τT = 0.1 ps applying V-rescale thermostat [42]. After being neutralized, the system was submitted to energy minimization applying the steepest descent algorithm. After that, the equilibrations of systems were done under NVT up to 100 ps at 300 K with restraint forces of 1000 kJ/mol, followed by 100 ps under NPT at the pressure of 1 bar and with restraint forces of 1000 kJ/mol using modified Berendsen thermostat and Parinello-Rahman barostat algorithms respectively. The electrostatic interactions were analyzed using the Particle Mesh Ewald (PME) method .Ultimately, the MD run was performed with no restraint for 50 ns.

3. Results and discussion

3.1. Homology modeling, model refinement, and energy minimization

Multiple templates were used to predict the 3D structure of OmpA and OmpW (Figure 1 and 2). Crystal structure of OmpA-like domain from *A. baumannii* (3TD3) with p-value 6.35e-06 was used as a template for prediction of OmpA periplasmic domain (Figure 1). Five β-barrel crystal structures (1P4T, 2X4M, 3NB3, 3QRA, and 4RLC) were suggested for3D structure prediction of this domain in *A. baumannii* OmpA (Figure 1). The crystal structure 4RLC with p-value 2.05e-14 was used as the best template. The tertiary structure of OmpW predicted to contain only one domain (β-barrel domain) and crystal structure 2F1T with p-value 3.48e-07 was employed as the best template for this prediction (Figure 2). Refinement of the obtained "crude" models from the GalaxyRefine server resulted in five models for each protein (Table S3). Based on model quality scores for all refined models, model two with GDT-HA value of 0.9633, RMSD of 0.384, and MolProbity of 1.922, clash score of 17.7, and poor rotamers score of 0.7 was found to be the best model of OmpA. Its Ramachandran plot score was 97.0%. For OmpW, model three was found to be the best model based on GDT-HA of 0.9684, RMSD of 0.392, MolProbity of 1.746, clash score of 11.8, and poor rotamers of 0.0. Ramachandran plot

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

score was 97.1% for this model. YASARA energy minimization tool was used to further improve the quality of the models.

3.2. Tertiary structure validation

The validity of the refined 3D modeled structures of OmpA and OmpW was checked by using RAMPAGE, ProSA-web, and ERRAT servers (Figure 3). The Ramachandran plot checked the stereochemical quality of the refined proteins. Analysis of the plot for the modeled OmpA revealed that 321 residues (96.7%) of the protein are in favored regions, which confirms that the model is characterized by stereochemical parameters of a stable structure. This is consistent with the 97.0% score predicted by the GalaxyRefine analysis. Additionally, five residues (1.5%) were predicted to be in allowed regions and six residues (1.8%) in disallowed region. The same analysis for OmpW verified that 167 residues (97.1%) are in the favored region, four residues (2.3%) are in allowed and only one residue (0.6%) is in the outlier area. The obtained result for residues of the favored region is the same as the score obtained by the GalaxyRefine analysis. These results indicate the correct geometry and three-dimensional arrangement of the models. The server ProSA-web further evaluated the overall quality of the models by providing a Z-score, as an indicative of overall model quality [43]. The reference ranges for this score change depending on the size of the protein, where the more negative Z-score, the better is the protein model. The values for OmpA and OmpW were -5.03 and -4.21, respectively. The Z-score of both input structures was within the range of the scores typically found for native proteins of similar sizes, which shows the modeled proteins fall within the range of X-ray solved protein structures. Further verification with PROCHECK showed a good resolution (normality) of 1.5 Å for the structures of both proteins, since most high-resolution X-ray structures have a resolution within 1.5 and 2.0 Å. The ERRAT server was used to check the statistics of non-bonded interactions between different atom types, the reliability of model proteins, and the overall quality factor; where higher scores indicate higher quality of protein

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

structure (acceptable range= score >50). The overall quality factors were 75.44% for OmpA and 80.62% OmpW.

3.3. Pre-filtration and pharmacokinetic analysis of phytochemicals

Due to the large attrition rate of the molecules under clinical trials, and to ensure that selected compounds can be used as a drug, the putative antimicrobial phytocompounds were screened using PreADMET, SwissADME and pkCSM servers. Different filters, including Lipinski's rule (Molecular weight ≤500 Da, logP (lipophilicity log) ≤5, number of Hydrogen Bond Donor (HBD) ≤5 and Hydrogen Bond Acceptor (HBA) ≤10, number of rotatable bonds (nRot) ≤ 10, topological polar surface area (TPSA) ≤ 150 Å2) [44], CMC like rule, MDDR like rule, lead like rule, and WDI like rule, were used to select compounds. Moreover, one important parameter for oral bioavailability is Molar Refractivity (MR) described by the Ghose filter [45]. The Ghose filter quantitatively characterizes small molecules based on computed physicochemical property profiles that include log P, molar refractivity (MR), molecular weight (MW), and the number of atoms. According to this rule, the value of MR should lie between 40 and 130 for drug-likeness.

Effective and safe drugs exhibit a good combination of pharmacodynamics (PD) and pharmacokinetics (PK) including adequate absorption, distribution, metabolism, excretion and tolerable toxicity (ADMET), which reduces the number of synthesis-evaluation cycles and more expensive late-stage failures through by excluding inappropriate drug candidates [46, 47]. Regarding this, a compound should be absorbed easily through the gastrointestinal tract so that it can be available in the systemic circulation [48]. To predict the absorption of orally administered drugs, permeability coefficient across monolayers of the human colon carcinoma cell line Caco-2 is commonly used. This permeability as log Papp (10^{-6} cm/s) rate is considered high if log Papp > 0.9 and low if < 0.9. For the achievement of an optimal clinical drug, a

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

molecule should have high log Papp. The blood-brain barrier (BBB) and central nervous system (CNS) permeability values were also predicted for the candidates. It has been reported that those with log BB value greater than 0.3 have the potential to pass the BBB, while those with values less than -1 are poorly distributed to the brain. In addition, candidates that have log PS value > −2 are considered to penetrate the CNS, while those with values < −3 have difficulty in penetrating the CNS.

Considering the binding status of a drug with the blood proteins, the efficiency of a drug is affected by the binding efficacy of the molecule with whole blood proteins [48]. In this regard, we evaluated the fraction unbound (Fu) and steady-state volume of distribution (VDss) for the candidates. In distribution properties, VDss represents the total volume required by the drug concentration to be uniformly distributed, to give the same concentration at the target site as in the plasma. Higher values of VDss indicate that most of the drug is contained in tissues rather than in plasma (VDss (human) ($\log_{L/kg}$); low if $\log_{L/kg} < -0.15$ and high if $\log_{L/kg} > 0.45$).

Early *in silico* prediction of toxicity endpoints of the candidates was performed. These end points consist of the inhibition of cytochrome P450 (CYPs) monooxygenase enzymes [49]. Metabolism enzymes should metabolize a drug candidate to prevent undesirable adverse effects. In this regard, inhibition of the major human cytochrome P450 isoform of CYP2D6, which is involved in drug metabolism, as well as the renal OCT2 substrate was assessed categorically.

The factors of AMES toxicity and LD_{50} (lethal rat acute toxicity) value of the phytocompounds were investigated using the pkCSM *in silico* screening approach. Among major toxicity end-points, LD_{50} is an index determination of medicine and poison's virulence Hepatotoxicity, skin sensitization, and inhibition of hERG potassium ion channel effects were also determined for the evaluation of drug-drug interactions (DDIs). Moreover, the solubility of the compounds

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

was evaluated by topological methods of ESOL, Ali, and fragmental method of SILICOST-IT [48, 50].

Results showed that out of 384 investigated compounds, only 27 compounds satisfy the key parameters of these physicochemical properties (Tables S4 and S5). They fulfilled Lipinski's rule of five, CMC like rule, lead like rule, and WDI like rule, with no violation, which are beneficial to assess *in vivo* abilities of any compounds. The ADMET profile of these candidates showed that all of the screened compounds possess high intestinal absorption. Near 63% of them showed high log Papp values. Moreover, no hepatotoxicity characterized by disrupted normal liver functions, skin sensitization, hERGI inhibition was observed for them. Our results demonstrated an acceptable prediction of physicochemical properties in comparison to Lipinski's RO5. It was observed that selected molecules fall within most of the accepted range of the parameters of ADMET. Therefore, they were further screened against target proteins (OmpA and OmpW).

3.4. Protein-Ligand Interactions and Mechanisms of Binding

Previous studies have suggested that a cyclic hexapeptide (AOA-2) and dichlorophen (D6) can fight against *A. baumannii* by affecting the activity of OmpA and OmpW, respectively [24, 32, 51]. Therefore, we used these compounds as controls to evaluate the results of molecular docking. AOA-2 prevents invasion of host cells by potential binding to external loops of OmpA β-barrel structure. D6 can locate in the inner side of the OmpW β-barrel structure and interact with its residues. This can block the channel provided by this porin which consequently prevents nutrients from entering the cell.

Top-three binding cavities in the energy minimized homology models of OmpA and OmpW were predicted by the METAPOCKET target prediction tool (Figure 4). Results showed that from eight methods used by the server, only Q-SITEFINDER and POCASA algorithms failed

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

to identify binding sites within the top predictions for both proteins. The cavity ID 03, predicted as one of OmpA binding pockets, was selected as a potential target, based on AOA-2 interaction with external loops of β-barrel structure [24]. In the case of OmpW, the cavity ID 02, containing the residues of the inner side of the β-barrel structure, was selected based on its potential interaction with D6 [32].

The molecular docking results showed that isosakuranetin, aloe-emodin, and pinocembrin, with the lowest biding energy (-7.8, -7.7 and 7.5, respectively), are top three ligands to interact with the selected binding pocket of OmpA (Table 1). Although the affinity of these compounds for OmpA is lower than AOA-2, but in this regard, they can be served as stronger interacting ligands than episterol (an *A. baumannii* growth inhibitor with the ability of binding to OmpA [9]). Both isosakuranetin and pinocembrin interact with residues Gln14, Asp15, Leu27, Thr28, Ser30, Glu158, and Trp160 (Figure 5). Both compounds can interact with Asp15 by forming hydrogen bonds. In this regard, aloe-emodin is slightly different from isosakuranetin and pinocembrin. Aloe-emodin has no interaction with Ser30 and forms hydrogen bonds with Asp15 and Gln14.

Compared to D6, all the above mentioned compounds possess a higher affinity for OmpW (Table 2). Isosakuranetin interacts with the highest number of residues in the selected binding pocket of OmpW. Four out of 14 residues in OmpW (Asn2, Gln42, Thr44, and Lys 84) form hydrogen bonds with isosakuranetin. Aloe-emodin and pinocembrin interact with 10 and 11 residues of OmpW, respectively. Aloe-emodin forms hydrogen bonds with 3 residues (Asn2, Gln4, and Trp134), while pinocembrin only possesses one H-bond connected to Gln4 (Figure 6).

3.5. Pharmacokinetic properties of selected hit phytochemicals

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Physicochemical, ADMET, and safety endpoints of the selected hit phytochemicals were predicted (Table 2-4). Results showed that all of these compounds satisfy the key parameters of a drug molecule via passing in Lipinski's RO5 (Table 3). Analyzing their ADMET profile, we observed that all of the hit inhibitors possess good gastrointestinal absorption (74.179-92.417%). Among them, the compounds isosakuranetin and pinocembrin with log Papp values of 1.1 and 1.52, respectively, showed high Caco-2 permeability. LD_{50} values of the compounds were in the range of 1.586-2.329 mol kg^{-1}. None of the compounds was observed as OCT2 substrate. None of them acts as CYP2D6 substrates/inhibitors (Table 4). Additionally, the selected hits were predicted as compounds with no inhibitory effect on hERGI. They also showed no hepatotoxicity and skin sensitization activity. The high value of VDss predicted for aloe-emodin (log L/kg = 0.671 > 0.45), indicates that most amount of the drug can be contained in tissues (rather than in plasma). Since only the unbound (free) drug can diffuse between plasma and tissues, the ability of the selected compounds to interact with pharmacological target proteins such as channels was also examined. The values (Fu) were in the range of 0.022-0.226 (Table 5).

Apart from efficacy and toxicity, poor pharmacokinetics and bioavailability are among important factors of drug development failures. Among these, gastrointestinal absorption and brain access are crucial pharmacokinetic features, which need to be estimated in drug discovery processes. Permeation from the blood–brain barrier (BBB), as a physical and biochemical shield protecting the brain, is the basic of the distribution of central-acting molecules. Moreover, regarding the preference of oral bioavailability of a drug in drug administration, Human Intestinal Absorption (HIA) is an important roadblock in drug research. The accurate predictive model of Brain Or IntestinaL EstimateD permeation method (BOILED-Egg) was used to predict these two important behaviors for selected phytocompounds (Figure 7). The model calculates the lipophilicity and polarity of small molecules. It allows for the evaluation

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

of HIA as a function of the position of the small molecules in the WLOGP-versus-TPSA referential. The white region of the "BOILED-Egg" is for a high probability of passive absorption of a compound by the gastrointestinal tract, and the yellow region (yolk) is for a high probability of brain penetration. Yolk and white areas are not mutually exclusive. In addition, the points are colored in blue if a compound predicted as actively effluxed by P-glycoprotein (PGP+) and in red if a compound predicted as non-substrate of P-gp (PGP−). Our results showed that among the selected compounds, isosakuranetin and pinocembrin were predicted to permeate through BBB by passive diffusion. According to the prediction, aloe-emodin can be passively absorbed by the gastrointestinal tract. The assessment of P-gp efflux for the selected compounds showed all of the selected compounds are not substrates of P-gp. Indeed, P-glycoprotein is a multidrug transporter, which is apically expressed in the gastrointestinal tract, liver, kidney and brain endothelium [52]. Consequently, P-glycoprotein plays an important role in the oral bioavailability, CNS distribution biliary and renal elimination of drugs which are substrates of this transporter.

HIA and BBB are dependent on the water solubility and lipophilicity of the drug. Two topological methods of the ESOL [53] and Ali [54] are included in SwissADME to predict water solubility. SwissADME third predictor for solubility was developed by SILICOS-IT. All predicted values are the decimal logarithm of the molar solubility in water (log S) (Table 5). Three predictors classified aloe-emodin and pinocembrin as soluble molecules. Isosakuranetin is predicted to be soluble adopted by Esol predictor, but moderately soluble by Ali and SILICOS-IT predictors. According to Log P, isosakuranetin (Log p= 2.81) and pinocembrin (Log p= 2.8) are more lipophilic than aloe-emodin (Log p= 1.36) (Table 6).

3.6 Molecular dynamics simulation analysis:

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Given the results, it seems isosakuranetin and aloe-emodin can be served as suitable ligands for both OmpA and OmpW to inhibit their targeted functions. Therefore, both of the compounds were selected to evaluate their molecular docking results using MD simulation (Figure 8). Backbone deviation of complexes were plotted as a function of time. The results showed stability of complexes. The aloe emodin-OmpA complex showed a RSMD value ranging from 0.1 to 0.65 nm during simulation. RMSD value for the isosakuranetin-OmpA complex were in the range of 0.1 to 0.65 nm. Both the isosakuranetin- and aloe emodin-OmpW complex demonstrated a similar value, ranging from 0.1 to 0.4 nm. The flexibility of backbone was evaluated using RMSF plot. The higher RMSF value means the higher flexibility, whereas the lower value represents the lower mobility during simulation. In general, β-strand residues showed less fluctuations than loops. Compared to the aloe emodin-OmpA complex, the isosakuranetin-OmpA complex demonstrated lower fluctuations. In regard to the OmpW-ligand complexes, the isosakuranetin-OmpW complex showed slightly higher fluctuations than the aloe emodin-OmpW complex.

We performed MM/PBSA analysis to evaluate free-binding energy of the ligand-protein complexes. Compared to the aloe-emodin, isosakuranetin interacts with OmpA with larger Van der Waals and electrostatic energy value. The binding energies for isosakuranetin and aloe-emodin, in the form of OmpA-ligand complex, were -147.097 and -106.871 (kcal/mol), respectively. According to OmpW-ligand complex, both isosakuranetin and aloe-emodin form complexes with the same binding energy (-51.49 kcal/mol). Overall, although the results show that both compounds are suitable for the purpose of our study, isosakuranetin appears to establish more stable interactions with both OmpA and OmpW.

4. Conclusion

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

The global antimicrobial resistance problem raises the necessity of novel approaches for the development of new drugs against emerging superbugs. A major contributor to this resistance is the outer membrane of Gram-negative bacteria with outer membrane proteins (OMPs) as promising drug targets due to their extensive functions. Therefore, any molecule which has the potential to disrupt the normal activity of these porins can help the host defense system overcoming the pathogens. Our results suggest three out of 384 phytocompounds (isosakuranetin, aloe-emodin, and pinocembrin), can control *A. baumannii* infection by targeting the function of two proteins, OmpA, based on its important roles in biofilm formation and bacterial adhesion to host cells, and OmpW, based on its important role in the transport of essential nutrients. Also, these compounds showed promising pharmacokinetic features as potential drugs with possible anti-*A.baumannii* activities. Amongst them, isosakuranetin was introduced as a top inhibitor for the targeted functions of OmpA and OmpW. Therefore, this biomolecule can be further characterized experimentally to corroborate its inhibitory activity. In addition, the skeletons of these active compounds could be adopted as starting potential scaffolds for the design of future anti-*A.baumannii* drugs.

Data accessibility. The data supporting the results in this article can be accessed at the Dryad Digital Repository: https://doi.org/10.5061/dryad.k6djh9w6f

https://datadryad.org/stash/share/joYqZpWMp7ssFCmDdVSVy3lMdnYlBQ6qyK2r_bJaubg

Competing interests. The authors declare no competing interests.

Funding. The authors received no financial support for this study.

References

1 Nathan, C. 2020 Resisting antimicrobial resistance. *Nature Reviews Microbiology*. **18**, 259-260.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

2 Alanis, A. J. 2005 Resistance to antibiotics: are we in the post-antibiotic era? *Archives of medical research*. **36**, 697-705.

3 Peleg, A. Y., Seifert, H., Paterson, D. L. 2008 Acinetobacter baumannii: emergence of a successful pathogen. *Clinical microbiology reviews*. **21**, 538-582.

4 Dijkshoorn, L., Nemec, A., Seifert, H. 2007 An increasing threat in hospitals: multidrug-resistant Acinetobacter baumannii. *Nature reviews microbiology*. **5**, 939-951.

5 Karageorgopoulos, D. E., Falagas, M. E. 2008 Current control and treatment of multidrug-resistant Acinetobacter baumannii infections. *The Lancet infectious diseases*. **8**, 751-762.

6 Moriel, D. G., Beatson, S. A., Wurpel, D. J., Lipman, J., Nimmo, G. R., Paterson, D. L., Schembri, M. A. 2013 Identification of novel vaccine candidates against multidrug-resistant Acinetobacter baumannii. *PloS one*. **8**, e77631.

7 Piddock, L. J. 2012 The crisis of no new antibiotics—what is the way forward? *The Lancet infectious diseases*. **12**, 249-253.

8 Skariyachan, S., Manjunath, M., Bachappanavar, N. 2019 Screening of potential lead molecules against prioritised targets of multi-drug-resistant-Acinetobacter baumannii–insights from molecular docking, molecular dynamic simulations and in vitro assays. *Journal of Biomolecular Structure and Dynamics*. **37**, 1146-1169.

9 Skariyachan, S., Muddebihalkar, A. G., Badrinath, V., Umashankar, B., Eram, D., Uttarkar, A., Niranjan, V. 2020 Natural epiestriol-16 act as potential lead molecule against prospective molecular targets of multidrug resistant Acinetobacter baumannii-Insight from in silico modelling and in vitro investigations. *Infection, Genetics and Evolution*. **82**, 104314.

10 Lee, C.-R., Lee, J. H., Park, M., Park, K. S., Bae, I. K., Kim, Y. B., Cha, C.-J., Jeong, B. C., Lee, S. H. 2017 Biology of Acinetobacter baumannii: pathogenesis, antibiotic resistance mechanisms, and prospective treatment options. *Frontiers in cellular and infection microbiology*. **7**, 55.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

11 Masi, M., Winterhalter, M., Pagès, J.-M. 2019 Outer membrane porins. In *Bacterial cell walls and membranes*. (ed.^eds. pp. 79-123: Springer.

12 Nie, D., Hu, Y., Chen, Z., Li, M., Hou, Z., Luo, X., Mao, X., Xue, X. 2020 Outer membrane protein A (OmpA) as a potential therapeutic target for Acinetobacter baumannii infection. *Journal of Biomedical Science*. **27**, 26.

13 Shahryari, S., Talaee, M., Haghbeen, K., Adrian, L., Vali, H., Zahiri, H. S., Noghabi, K. A. 2021 New Provisional Function of OmpA from Acinetobacter sp. Strain SA01 Based on Environmental Challenges. *Msystems*. **6**,

14 McConnell, M. J., Actis, L., Pachón, J. 2013 Acinetobacter baumannii: human infections, factors contributing to pathogenesis and animal models. *FEMS microbiology reviews*. **37**, 130-155.

15 Confer, A. W., Ayalew, S. 2013 The OmpA family of proteins: roles in bacterial pathogenesis and immunity. *Veterinary microbiology*. **163**, 207-222.

16 Smith, S. G., Mahon, V., Lambert, M. A., Fagan, R. P. 2007 A molecular Swiss army knife: OmpA structure, function and expression. *FEMS microbiology letters*. **273**, 1-11.

17 Sugawara, E., Nikaido, H. 1992 Pore-forming activity of OmpA protein of Escherichia coli. *Journal of Biological Chemistry*. **267**, 2507-2511.

18 Saint, N., De, E., Julien, S., Orange, N., Molle, G. 1993 Ionophore properties of OmpA of Escherichia coli. *Biochimica et Biophysica Acta (BBA)-Biomembranes*. **1145**, 119-123.

19 Zakharian, E., Reusch, R. 2003 Outer membrane protein A of Escherichia coli forms temperature-sensitive channels in planar lipid bilayers. *FEBS letters*. **555**, 229-235.

20 Arora, A., Rinehart, D., Szabo, G., Tamm, L. K. 2000 Refolded Outer Membrane Protein A of Escherichia coliForms Ion Channels with Two Conductance States in Planar Lipid Bilayers. *Journal of Biological Chemistry*. **275**, 1594-1600.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

21 Skariyachan, S., Taskeen, N., Ganta, M., Venkata Krishna, B. 2019 Recent perspectives on the virulent factors and treatment options for multidrug-resistant Acinetobacter baumannii. *Critical reviews in microbiology*. **45**, 315-333.

22 Sánchez-Encinales, V., Álvarez-Marín, R., Pachón-Ibáñez, M. E., Fernández-Cuenca, F., Pascual, A., Garnacho-Montero, J., Martínez-Martínez, L., Vila, J., Tomás, M. M., Cisneros, J. M. 2017 Overproduction of outer membrane protein A by Acinetobacter baumannii as a risk factor for nosocomial pneumonia, bacteremia, and mortality rate increase. *The Journal of infectious diseases*. **215**, 966-974.

23 Weiser, J. N., Gotschlich, E. C. 1991 Outer membrane protein A (OmpA) contributes to serum resistance and pathogenicity of Escherichia coli K-1. *Infection and Immunity*. **59**, 2252-2258.

24 Vila-Farrés, X., Parra-Millán, R., Sánchez-Encinales, V., Varese, M., Ayerbe-Algaba, R., Bayó, N., Guardiola, S., Pachón-Ibáñez, M. E., Kotev, M., García, J. 2017 Combating virulence of Gram-negative bacilli by OmpA inhibition. *Scientific reports*. **7**, 1-11.

25 Catel-Ferreira, M., Marti, S., Guillon, L., Jara, L., Coadou, G., Molle, V., Bouffartigues, E., Bou, G., Shalk, I., Jouenne, T. 2016 The outer membrane porin OmpW of Acinetobacter baumannii is involved in iron uptake and colistin binding. *FEBS letters*. **590**, 224-231.

26 López-Rojas, R., Smani, Y., Pachón, J. 2013 Treating multidrug-resistant Acinetobacter baumannii infection by blocking its virulence factors. *Expert review of anti-infective therapy*. **11**, 231-233.

27 Antunes, L. C., Imperi, F., Minandri, F., Visca, P. 2012 In vitro and in vivo antimicrobial activities of gallium nitrate against multidrug-resistant Acinetobacter baumannii. *Antimicrobial agents and chemotherapy*. **56**, 5961-5970.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

28 Wu, X.-B., Tian, L.-H., Zou, H.-J., Wang, C.-Y., Yu, Z.-Q., Tang, C.-H., Zhao, F.-K., Pan, J.-Y. 2013 Outer membrane protein OmpW of Escherichia coli is required for resistance to phagocytosis. *Research in microbiology*. **164**, 848-855.

29 Vila, J., Martí, S., Sanchez-Céspedes, J. 2007 Porins, efflux pumps and multidrug resistance in Acinetobacter baumannii. *Journal of antimicrobial chemotherapy*. **59**, 1210-1215.

30 Lin, X.-M., Yang, J.-N., Peng, X.-X., Li, H. 2010 A novel negative regulation mechanism of bacterial outer membrane proteins in response to antibiotic resistance. *Journal of proteome research*. **9**, 5952-5959.

31 Nandi, B., Nandy, R. K., Sarkar, A., Ghose, A. C. 2005 Structural features, properties and regulation of the outer-membrane protein W (OmpW) of Vibrio cholerae. *Microbiology*. **151**, 2975-2986.

32 Soojhawon, I., Pattabiraman, N., Tsang, A., Roth, A. L., Kang, E., Noble, S. M. 2017 Discovery of novel inhibitors of multidrug-resistant Acinetobacter baumannii. *Bioorganic & medicinal chemistry*. **25**, 5477-5482.

33 Lipinski, C. A. 2004 Lead-and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies*. **1**, 337-341.

34 Silva, L. N., Zimmer, K. R., Macedo, A. J., Trentin, D. S. 2016 Plant natural products targeting bacterial virulence factors. *Chemical reviews*. **116**, 9162-9236.

35 Sun, Y., Yang, A. W. H., Hung, A., Lenon, G. B. 2020 Screening for a Potential Therapeutic Agent from the Herbal Formula in the 4th Edition of the Chinese National Guidelines for the Initial-Stage Management of COVID-19 via Molecular Docking. *Evidence-Based Complementary and Alternative Medicine*. **2020**,

36 Rolta, R., Yadav, R., Salaria, D., Trivedi, S., Imran, M., Sourirajan, A., Baumler, D. J., Dev, K. 2020 In silico screening of hundred phytocompounds of ten medicinal plants as potential

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

inhibitors of nucleocapsid phosphoprotein of COVID-19: an approach to prevent virus assembly. *Journal of Biomolecular Structure and Dynamics*. 1-18.

37 Chan, H. S., Shan, H., Dahoun, T., Vogel, H., Yuan, S. 2019 Advancing drug discovery via artificial intelligence. *Trends in pharmacological sciences*. **40**, 592-604.

38 Moroy, G., Martiny, V. Y., Vayer, P., Villoutreix, B. O., Miteva, M. A. 2012 Toward in silico structure-based ADMET prediction in drug discovery. *Drug discovery today*. **17**, 44-55.

39 Petit, L., Vernès, L., Cadoret, J.-P. 2021 Docking and in silico toxicity assessment of Arthrospira compounds as potential antiviral agents against SARS-CoV-2. *Journal of applied phycology*. 1-24.

40 Trott, O., Olson, A. J. 2010 AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*. **31**, 455-461.

41 Fan, H., Mark, A. E. 2004 Refinement of homology-based protein structures by molecular dynamics simulation techniques. *Protein Science*. **13**, 211-220.

42 Zamani Amirzakaria, J., Malboobi, M. A., Marashi, S.-A., Lohrasebi, T. 2020 In silico prediction of enzymatic reactions catalyzed by acid phosphatases. *Journal of Biomolecular Structure and Dynamics*. 1-12.

43 Wiederstein, M., Sippl, M. J. 2007 ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research*. **35**, W407-W410.

44 Pollastri, M. P. 2010 Overview on the Rule of Five. *Current protocols in pharmacology*. **49**, 9.12. 11-19.12. 18.

45 Ghose, A. K., Viswanadhan, V. N., Wendoloski, J. J. 1999 A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *Journal of combinatorial chemistry*. **1**, 55-68.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

46 Hodgson, J. 2001 ADMET—turning chemicals into drugs. *Nature Biotechnology*. **19**, 722-726.

47 Pellicciari, R. 2017 Attrition in the Pharmaceutical Industry—Reasons, Implications, and Pathways Forward. Edited by Alexander Alex, C. John Harris, Dennis A. Smith. *ChemMedChem*. **12**, 1097-1098.

48 Zafar, F., Gupta, A., Thangavel, K., Khatana, K., Sani, A. A., Ghosal, A., Tandon, P., Nishat, N. 2020 Physicochemical and Pharmacokinetic Analysis of Anacardic Acid Derivatives. *ACS omega*. **5**, 6021-6030.

49 Dearden, J. C. 2003 In silico prediction of drug toxicity. *Journal of computer-aided molecular design*. **17**, 119-127.

50 Ndombera, F. T., Maiyoh, G. K., Tuei, V. C. 2019 Pharmacokinetic, Physicochemical and Medicinal Properties of N-glycoside Anti-cancer Agent More Potent than 2-Deoxy-D-Glucose in Lung Cancer Cells. *Journal of Pharmacy and Pharmacology*. **7**, 165-176.

51 Ayerbe-Algaba, R., Bayó, N., Verdú, E., Parra-Millán, R., Seco, J., Teixidó, M., Pachón, J., Giralt, E., Smani, Y. 2021 AOA-2 Derivatives as Outer Membrane Protein A Inhibitors for Treatment of Gram-Negative Bacilli Infections. *Frontiers in microbiology*. **12**, 203.

52 Finch, A., Pillans, P. 2014 P-glycoprotein and its role in drug-drug interactions. *Aust Prescr*. **37**, 137-139.

53 Delaney, J. S. 2004 ESOL: estimating aqueous solubility directly from molecular structure. *Journal of chemical information and computer sciences*. **44**, 1000-1005.

54 Fagerberg, J. H., Karlsson, E., Ulander, J., Hanisch, G., Bergström, C. A. 2015 Computational prediction of drug solubility in fasted simulated and aspirated human intestinal fluid. *Pharmaceutical research*. **32**, 578-589.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Table 1. Molecular docking analysis of the top three compounds which have highest affinity for OmpA. The asterisks represent the compounds with the potential to inhibit OmpA according to previous studies.

Compounds	Binding energy (kcal/mol)	RMSD	Number of H-bounds	Number of interacting residues	Interacting residues
***AOA-2**	-8.5	0	4	12	Asp15, Lys23, Gly25, Asn26, Leu27, Thr28, Asn29, Ser30, Pro31, Glu157, Glu158, Phe159
***Epiestriol**	-6.6	0	1	6	Gly40, Tyr57, Asn58, Gln59, Gln75, Lys76
Isosakuranetin	-7.8	0	1	7	Gln14, Asp15, Leu27, Thr28, Ser30, Glu158, Trp160
Aloe-emodin	-7.7	0	2	6	Gln14, Asp15, Leu27, Thr28, Glu158, Trp160
Pinocembrin	-7.5	0	1	7	Gln14, Asp15, Leu27, Thr28, Ser30, Glu158, Trp160

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Table 2. The ligand-OmpW molecular docking analysis of selected compounds. . The asterisks represent the compounds with the potential to inhibit OmpW according to previous studies.

Compounds	Binding energy (kcal/mol)	RMSD	Number of H-bounds	Number of interacting residues	Interacting residues
***D6**	-8.4	0	1	12	Asn2, Gln4, Phe40, Gly42, Gln43, Thr44, His82, Lys84, Asn85, Thr90, Asn125, Gln127
***Epiestriol**	-6.7	0	0	6	Gly1, Trp3, Tyr39, Phe40, Phe41, Ala48
Isosakuranetin	-9.8	0	4	14	Asn2, Gln4, Phe40, Gly42, Gln43, Thr44, His82, Lys84, Asn85, Thr90, Gln127, Asn133, Trp134, Ser171
Aloe-emodin	-9.3	0	3	10	Asn2, Gln4, Phe40, Gln43, Thr44, Lys84, Asn85, Thr90, Gln127, Trp134
Pinocembrin	-10.2	0	1	11	Asn2, Gln4, Phe40, Gly42, Gln43, Asn85, Asn125, Gln127, Asn133, Trp134, Ser171

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Table 3. The physicochemical properties of the best selected compounds with mutual inhibitory effect on both OmpA and OmpW.

Compound name	n_{Rot}[a]	HBD[b]	HBA[c]	TPSA[d] (Å2)	logP[e]	MR[f]	Lipinski violation	CMC-like rule	Lead-like rule
Isosakuranetin	2	2	5	120.92	2.81	76.04	0	Qualified	Suitable
Aloe-emodin	1	3	5	113.28	1.36	69.92	0	Qualified	Suitable
Pinocembrin	1	2	4	109.44	2.8	69.55	0	Qualified	Suitable

[a] Number of rotatble bonds, [b] Hydrogen Bond Donor, [c] Hydrogen Bond Acceptor, [d] Topological Polar Surface Area, [e] Lipophilicity, [f] Molar Refractivity.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Table 4. The absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties of the best selected compounds with mutual inhibitory effect on both OmpA and OmpW.

Compound name	ID	Absorption			Distribution		Metabolism	Excretion	Toxicity	
		GI (%)	Water solubility (log mol/L)	Log P_{app}	Log BB	Log PS	CYP2D6 Subs./Inh.	OCT2 Subs	AMES Toxicity	LD50 (mol/kg)
Isosakuranetin	160481	92.37	-3.092	1.1	0.194	-2.17	No/No	No	No	1.843
Aloe-emodin	10207	74.179	-3.104	-0.233	-0.729	-2.466	No/No	No	Yes	2.329
Pinocembrin	68071	92.417	-3.538	1.152	0.42	-2.047	No/No	No	No	1.586

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Table 5. The computed safety endpoints of the best selected compounds with mutual inhibitory effect on both OmpA and OmpW.

Compound name	Total clearance	hERGI inhibitor	Oral rat chronic toxicity (LOAEL)	Hepatotoxicity	Skin sensitization	**VDss** [a]	**F_u** [b]
Isosakuranetin	0.111	No	2.148	No	No	0.217	0.077
Aloe-emodin	0.008	No	1.878	No	No	0.671	0.226
Pinocembrin	0.122	No	2.059	No	No	-0.386	0.022

[a] Steady state volume of distribution (VDss), [b] fraction unbound (F_u)

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Table 6. The water solubility evaluation of the best selected compounds with mutual inhibitory effect on both OmpA and OmpW.

Molecule	ESOL LogS	Ali LogS	SILICOS-IT LogS	ESOL Class	Ali Class	SILICOS-IT Class
Isosakuranethin	-3.7	-4.1	-4.12	Soluble	Moderately soluble	Moderately soluble
Aloe-emodin	-3.04	-3.43	-3.92	Soluble	Soluble	Soluble
Pinocembrin	-3.64	-3.94	-4.00	Soluble	Soluble	Soluble

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

Figure captions:

Figure 1. Cartoon representation of the best structures of OmpA modeled by RaptorX and refined by GalaxyRefine (**A**). Amino acid sequence alignment of OmpA and its 3 dimentional (3D) structure template for the periplasmic domain (**B**). Multiple sequence alignment of OmpA and its 3D structure templates for the β-barrel structure (**C**).

Figure 2. Cartoon representation of the best structures of OmpW modeled by RaptorX and refined by GalaxyRefine (**A**). Multiple sequence alignment of OmpA and its 3D structure templates (**B**).

Figure 3. Ramachandran plot, the PROSA-web, and the ERRAT server validation of the modeled 3D structures of OmpA and OmpW. Distribution of each amino acid in the favored, allowed, and disallowed regions are shown in corresponding Ramachandran plots of OmpA (**A**), and OmpW (**B**). Ramachandran plot summary shows 96.7% residues of OmpA are in most favorable regions, 1.5% in allowed regions and 1.8% in disallowed regions. For OmpW, 97.1% residues are in most favorable regions, 2.3% in allowed regions and 0.6% in disallowed regions. Validation of OmpA (**C**) and OmpW (**D**) modeled structures using PROSA-web shows Z-score values as -5.03 and -4.21 for modeled proteins. The Overall quality factors obtained by ERRAT for OmpA (**E**) and OmpW (**F**) were 75.44% and 80.62%, respectively.

Figure 4. Predicted top-three potential binding pockets with corresponding active site residues for OmpA (**A**) and OmpW (**B**) by METAPOCKET 2.0. The selected potential binding pockets are emphasized by circle and their residues are shown in table.

Figure 5. 3D representation of the OmpA-ligand complexes. (**A**) represents isosakuranetin and OmpA complex, (**B**) aloe-emodin and OmpA complex, and (**C**) pinocembrin and OmpA complex. (**D**), (**E**) and (**F**) represent 2D interaction diagram of isosakuranetin, aloe-emodin and pinocembrin in the binding pockets of the target, respectively. Ligands, H-bond interacting

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

residues and hydrophobic interacting residues are shown in purple, red, and golden, respectively. Green bonds represent hydrogen bonds and golden bonds represent hydrophobic interactions.

Figure 6. 3D representation of the compound-OmpW complexes. (**A**) represents isosakuranetin and OmpW complex, (**B**) aloe-emodin and OmpW complex, and (**C**) pinocembrin and OmpW complex. (**D**), (**E**) and (**F**) represent 2D interaction diagram of isosakuranetin, aloe-emodin and pinocembrin in the binding pockets of the target, respectively. Ligands, H-bond interacting residues and hydrophobic interacting residues are shown in purple, red, and golden, respectively. Green bonds represent hydrogen bonds and golden bonds represent hydrophobic interactions.

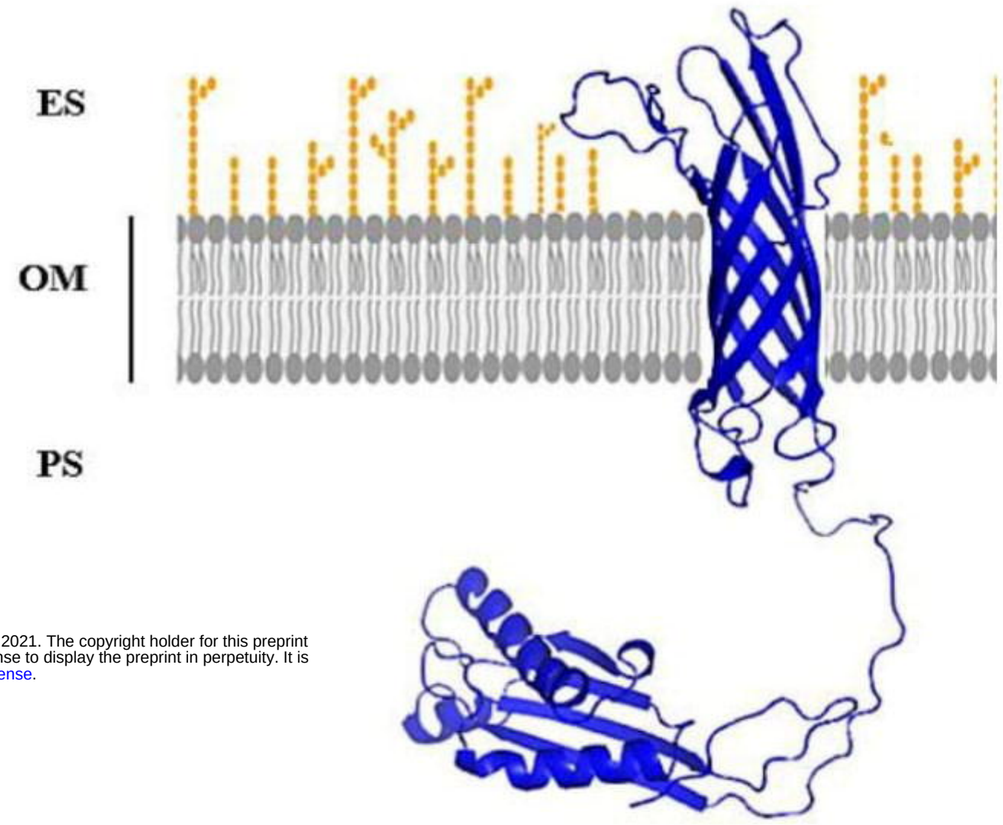
Figure 7. The BOILED-Egg model of the selected best potential inhibitors of OmpA and OmpW. The compounds located in the yellow region were predicted to be passively penetrate from BBB. The compound in the white region is predicted to be passively absorbed by gastrointestinal tract. The molecules predicted to be effluated from the central nervous system by p-glycoproteins are shown by blue dots. Those which do not effluated from the central nervous system by P-glycoproteins are shown by red dots.

Figure 8. Detection of the stability of the protein-ligand complexes. (**A**) RMSD of the protein-ligand complexes. (**B**) Root mean square fluctuation (RMSF) plots of the protein-ligand complexes. Isosakuranetin bounded to OmpA (black) and bounded to OmpW (red). Aloe-emodin bounded to OmpA (green) and bounded to OmpW (blue).

Figure 9. (**A**) Binding energies (including its components) of isosakuranetin-OmpA and aloe emodin-OmpA (black and green bars, respectively). (**B**) Binding energies (including its components) of isosakuranetin-OmpW and aloe emodin-OmpW (red and blue bars, respectively).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.

A)



B)

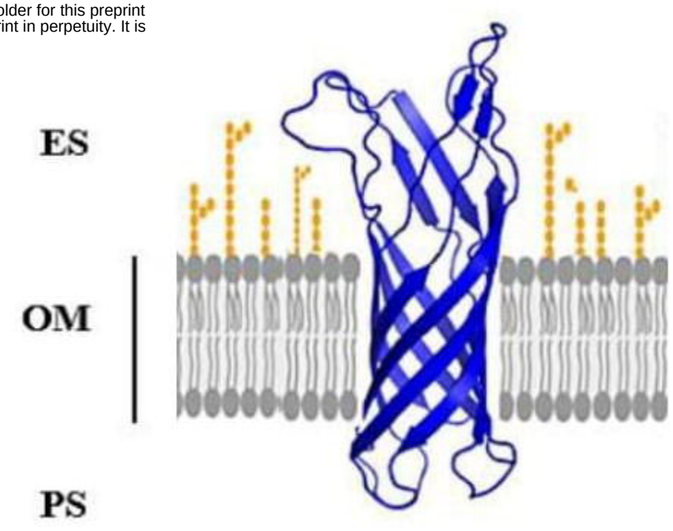
```
                    190       200       210       220       230       240       250       260       270
OmpA        PVVEVAPVEPTPVAPQPQELTEDLNMELRVFFDTNKSNIKDQYKPEIAKVAEKLSEYPNATARIEGHTDNTGPRKLNERLSLARANSVKS
3TD3        ................HMELTEDLNMELRVFFDTNKSNIKDQYKPEIAKVAEKLSEYPNATARIEGHTDNTGPRKLNERLSLARANSVKS
consensus>70 ..................ELTEDLNMELRVFFDTNKSNIKDQYKPEIAKVAEKLSEYPNATARIEGHTDNTGPRKLNERLSLARANSVKS

                    280       290       300       310       320       330
OmpA        ALVNEYNVDASRLSTQGFAWDQPIADNKTKEGRAMNRRVFATITGSRTVVVQPGQEAAAPAAAQ
3TD3        ALVNEYNVDASRLSTQGFAWDQPIADNKTKEGRAMNRRVFATITGSR................
```

C)

```
                    30                  40                  50                  60          70
OmpA        .....NLTNSPELQ...............DDLFVGAA.LG........IELTP...WLGF...EAEYNQ.....VKGDVDGASAGAEYKQ
1P4T        ............AK...............ASSSLGSA.KGFSPRISAGYRI.N...DLRF...AVDYTR.....YK............NY
2X4M        SGNMDDYAWMNENQSEWTDHSSHPATNVNHANEYDLNVKG........WLLQDENYKAGI...TAGYQE.....TRFSWTATGGSYSYNN
3NB3        ............HE...............NKLGAGAF.GG........YQVNP...YVGF...EMGYDWLGRMPYKGSVENGAYKAQGVQ
3QRA        ............ED...............GYKLDKNP.RGFNL..KYRYEFNN...DWGV...IGSFAQ.....TRRGFEESV....DGF
4RLC        ..........RNMK...............NADLYGGS.IG........YFLTD...DVELALSYGEYHD.....VRGT.......YETGN

                    80          90          100                         110         120         130
OmpA        KQIN......GNF..YVTSDLITKNYD...SKIKPYVLLGAG....................HYKYDFDGVNRGTRGTSEEGTLGNAGVGA
1P4T        KAPS......TDFKLYSIGASAIYDFDT.QSPVKPYLG..AR....................LSLNRASVDLGGSDS.FSQTSIG.LG..V
2X4M        GAYT......GNF..PK.G.VRVIGYN..QRFSMPYIGL.AGQYRINDFELNALFKFSDWVRAHDNDEHYMRDLTF.REKTSG.SRYYGT
3NB3        LTAKLGYPITDDLDIYTRLGGMVWRADTY.SNVYGK.....................................NH.DTGVSPV.FAG..
3QRA        KLID......GDFKYYSVTAGPVFRINE.YVSLYGLLG..AG....................HGKAKFSSIFGQSES.RSKTSLA.YG...
4RLC        KKVH......GNL....TSLDAIYHFGTPGVGLRPYVS..AG....................LAHQNITNINSDSQG.RQQMTMANIGA..

                    140         150         160                                 170         180         190
OmpA        FWRLNDALSLRTE....ARATYN..A.DEEFWN.........................YTALAGLNVVLGGHLKPAAPVVEVAPVEPTP
1P4T        LTGVSYAVTPNVD....LDAGYRYNY.IGKV...............NTVK......NVRSGELSAGVRVKF.....................
2X4M        VINAGYYVTPNAK....VFAEFTYSK.YDEGKGGTQTIDKNSGDSVSIGGDAAGISNKNYTVTAGLQYRFGG........GGGG......
3NB3        ..GVEYAITPEIA....TRLEYQWTNNIGDA...............HTIG.....TRPDNGMLSLGVSYRFGQ..GEAAPVVAPA......
3QRA        .AGLQFNPHPNFV....IDASYEYSK.LD.........................DVKVGTWMLGAGYRF.....................
4RLC        ..GLKYYFTENFFAKASLDGQYGLEK.RD.........................NGHQGEWMAGLGVGF.....................
```

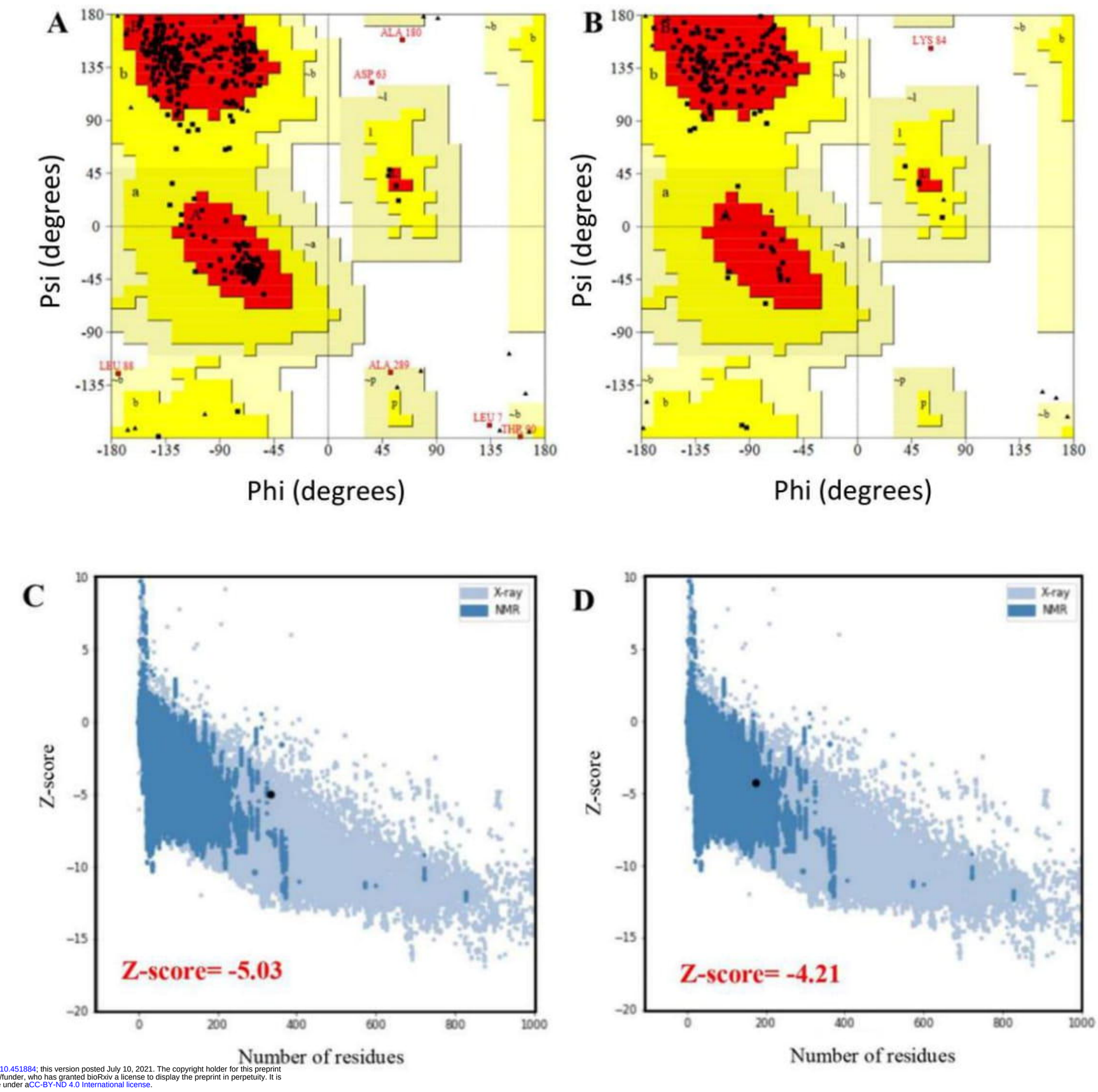
bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

A)

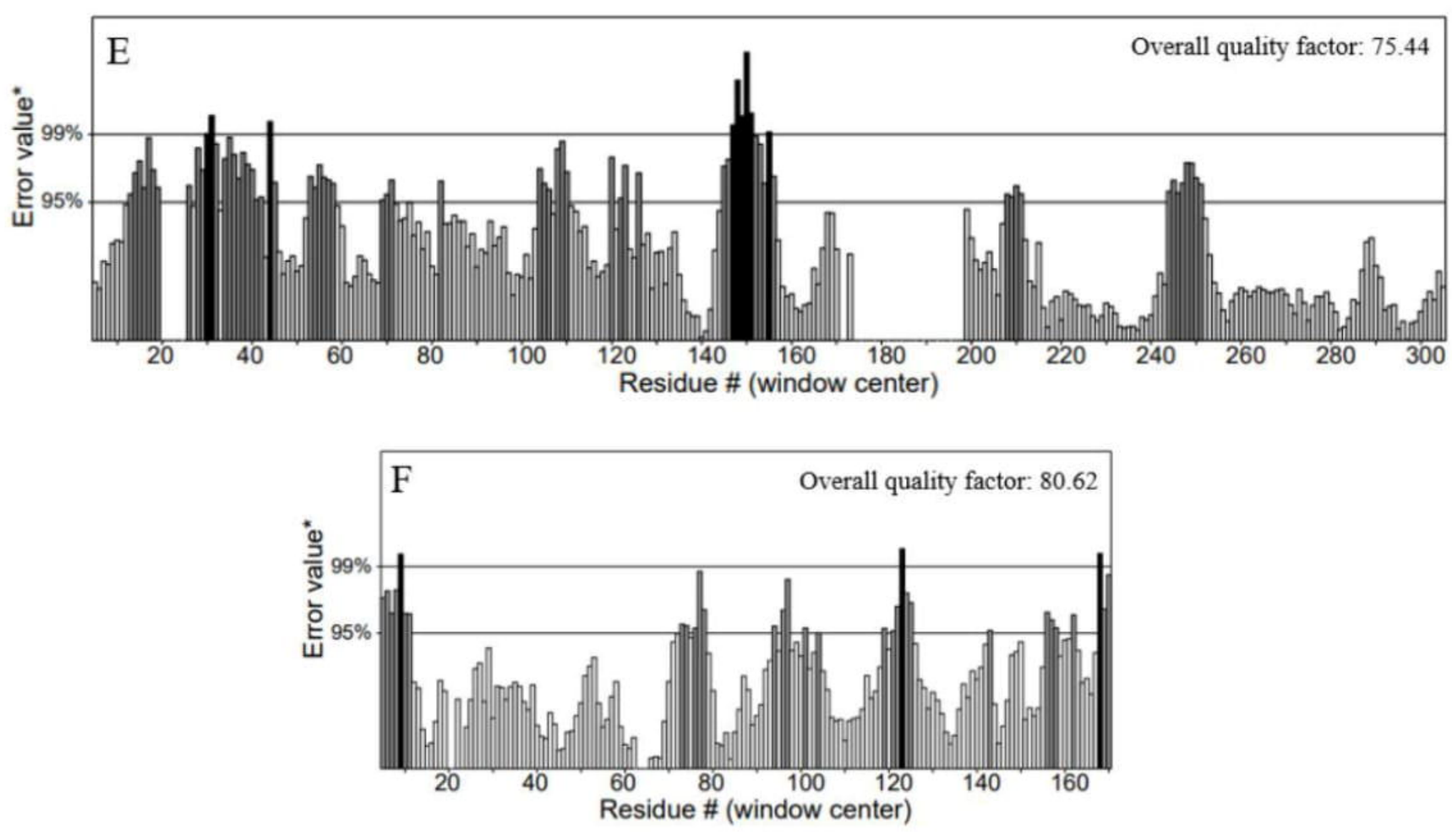


B)

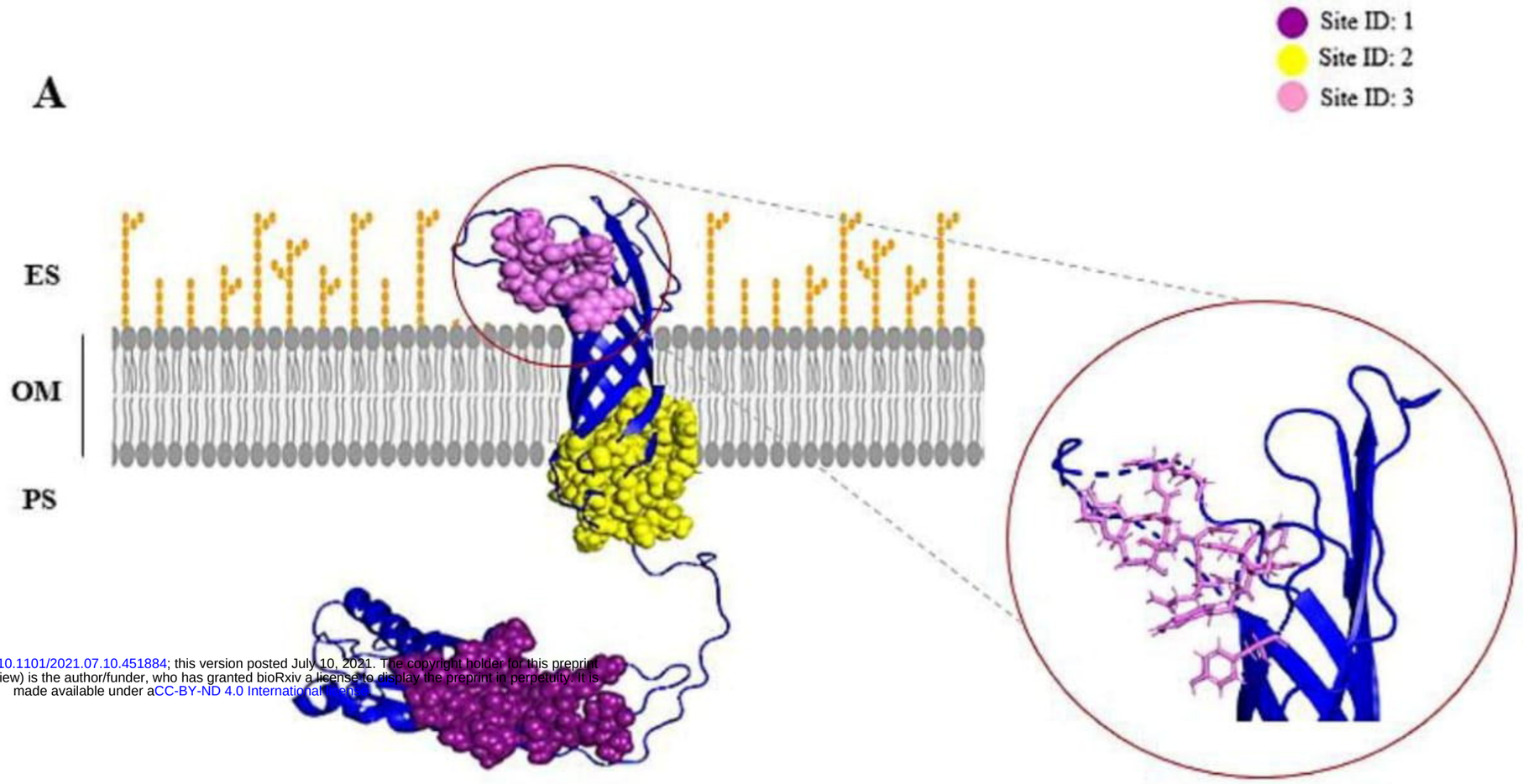
```
                         1         10                  20                  30                  40                  50                  60
OmpW     ................GNWQVKFGG.SVIAPSEDT......TTA...LGVVKADHEYAFTP..SVEYFFGQTPFSAELLLATPVNHDVL.
1P4T     EGA.............SGFYVQADA.AHAK...A......SS........SLGSAKGFSPRISAGYRI.N.DLRF..AVDYTRYK....
2F1T     HEA.............GEFFMRAGS.ATVRPTEGA......GGTLGSLGGFSVTNNTQLGL..TFTYMATD.NIGVELLAATPFRHKIG.
3QRA     ME..............GESSISIGY.AQSRVKEDG......YK........LDKNPRGFNL..KYRYEFNN.DWGV..IGSFAQTR.RGF
4RLC     ANVRLQHHHHHHHLEQGQNSVEIEAFGKRYFTD...............SVRNMKNADLYGG..SIGYFLTD.DVEL..ALSYGEYHDVRG
6QAM     MNENYPAKS..AGYNQGDWVASFNF.SKVYVGEELGDLNVGGGAL.PNADVSIGNDTTLTF..DIAYFVSS.NIAVDFFVGVPARAKFQG

                    70                  80                  90                 100                 110                 120                 130
OmpW     ...LDG.QKVARIKQ..LPPTITAKYHF.KN.STRFTPYIGIGATAFIPWD........EQGVAD.KVKEDF..GV.AGQVGFNFQPADA
1P4T     .....N.YKAPSTDFKLYSIGASAIYDF.DT.QSPVKPYLGARLSLNRASV........DLGGS..DSFSQTSIGL.GVLTGVSYAVT..
2F1T     ...TRATGDIATVHH..LPPTLMAQWYFGDA.SSKFRPYVGAGINYTTFFDNGFNDHGKEAGLSDLSLKDSW..GA.AGQVGVDYLIN..
3QRA     EESVDG.FKLIDGDFKYYSVTAGPVFRI.NE.YVSLYGLLGAGHGKAKFSS........IFGQS..ESRSKT..SL.AYGAGLQFNPH..
4RLC     ...TYETGN.KKVHG..NLTSLDAIYHF.GTPGVGLRPYVSAGLAHQNITN........INSDS..QGRQQM..TMANIGAGLKYYFT..
6QAM     EKSISSLGRVSEVDY..GPAILSLQYHY.DS.FERLYPYVGVGVGRVLFFDKT......DGALSSFDIKDKW..AP.AFQVGLRYDLG..

                   140                 150                 160                 170
OmpW     KNWGVFVDVRYADISPEVTLTNGA...KFDLDINPFVYTLGYSYKF................
1P4T     PNVDLDAGYRYNYI........GKVNTVKNVRSG..ELSAGVRVKF................
2F1T     RDWLVNMSVWYMDIDTTANYKLGGAQQHDSVRLDPWVFMFSAGYRFH.........HHHHH
3QRA     PNFVIDASYEYS...........KLDDVKVG..TWMLGAGYRF................
4RLC     ENFFAKASLDG.......QY..G.LEKRDNGHQGEWMAGLGVGFNF................
6QAM     NSWMLNSDVRYIPFKTDVTGTLGPVPVSTKIEVDPFILSLGASYVFKLAAALEHHHHHHH
```



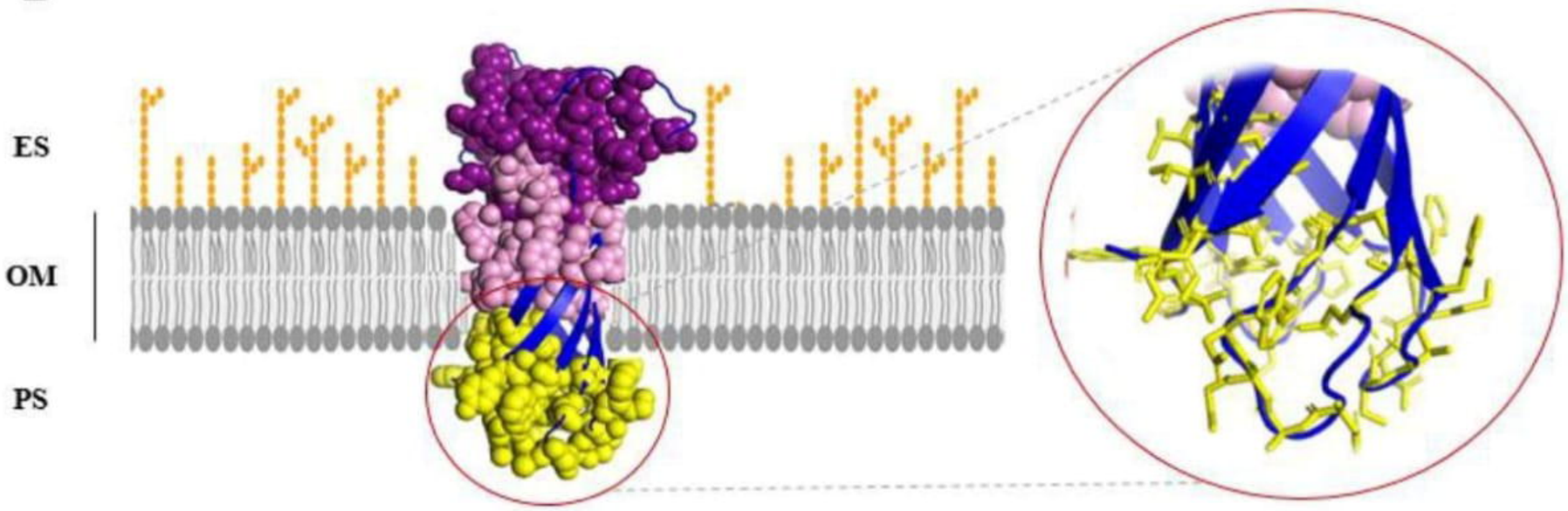
bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



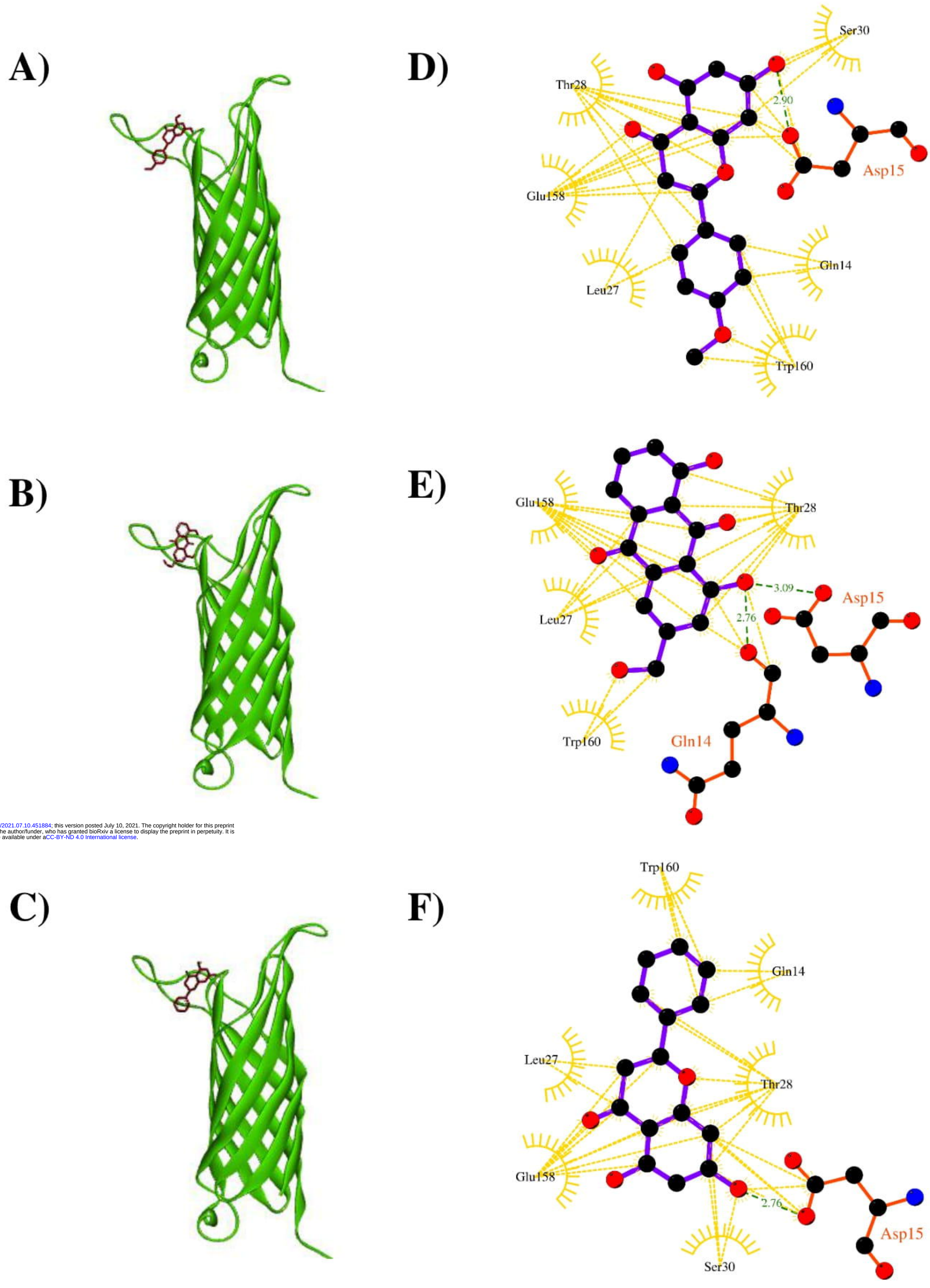
bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



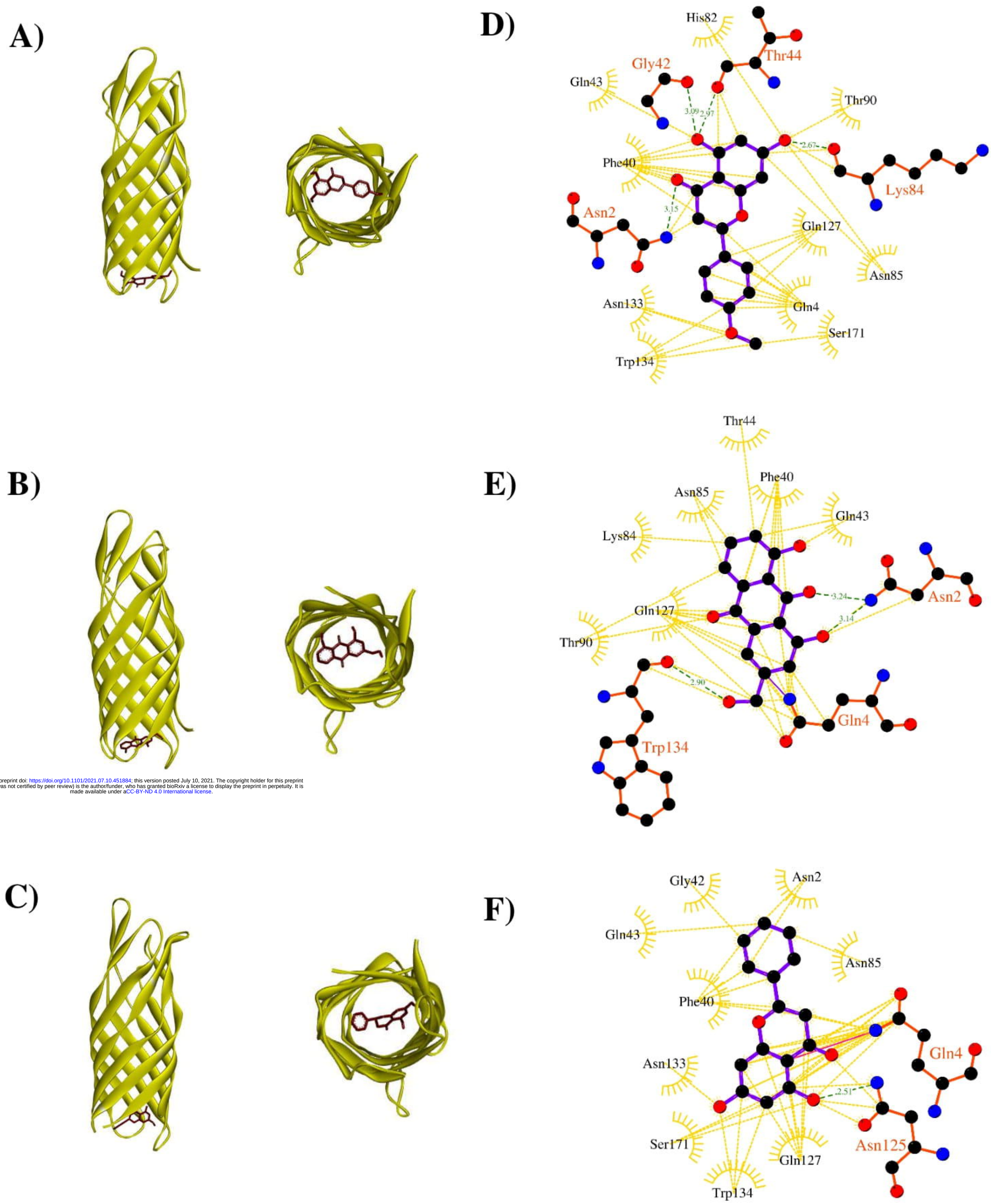
Active site residues						
PHE 13	GLN 14	ASP 15	SER 16	GLY 25	ASN 26	LEU 27
THR 28	ASN 29	SER 30	TYR 153	GLU 158	PHE 159	TRP 160



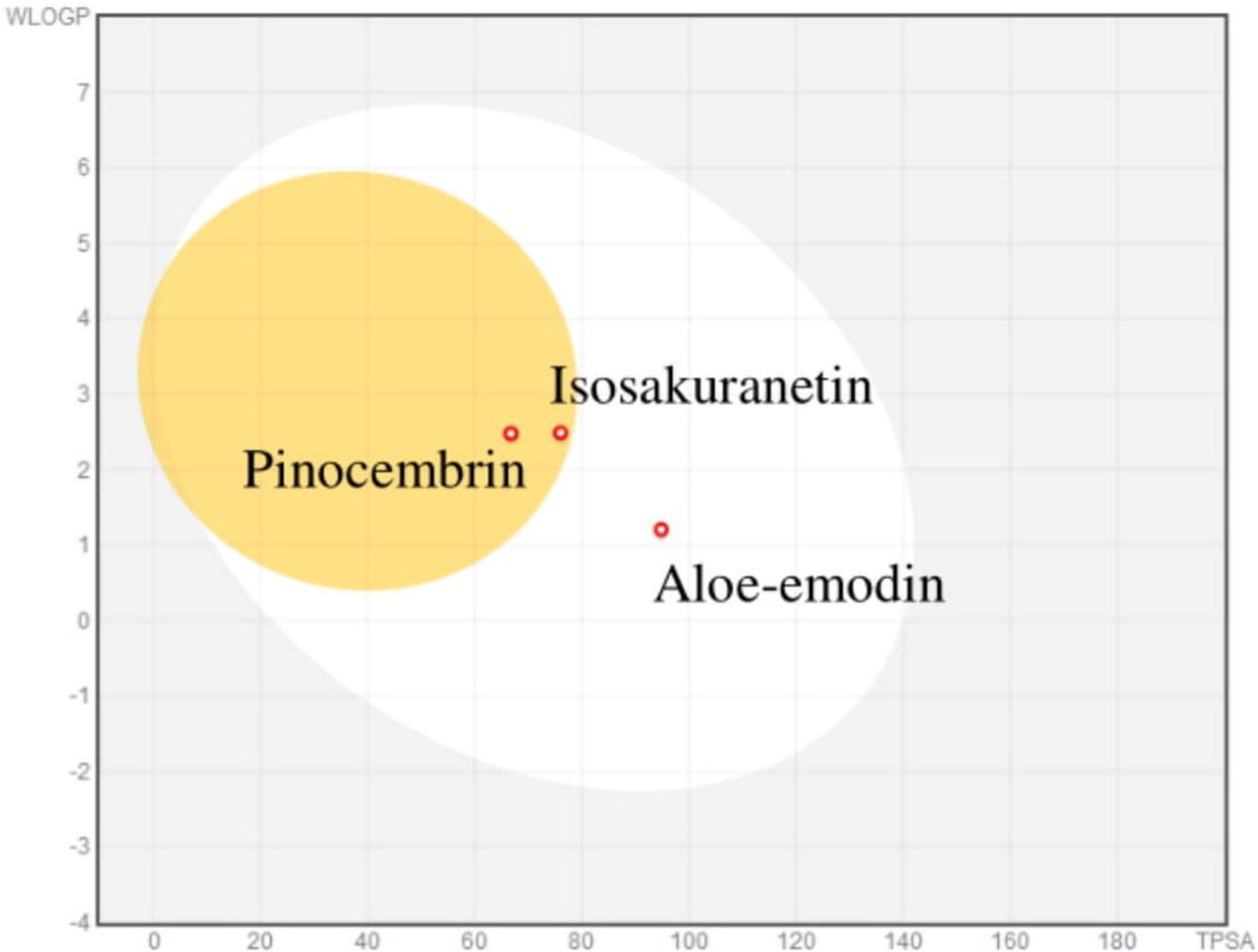
Active site residues						
GLY 1	ASN 2	TRP 3	GLN 4	VAL 5	LYS 6	SER 36
VAL 37	GLU 38	PHE 40	PHE 41	GLY 42	GLN 43	THR 44
PRO 45	PHE 46	GLU 49	HIS 82	PHE 83	LYS 84	ASN 85
SER 86	THR 87	PHE 89	THR 90	ASN 125	GLN 127	ASP 130
ALA 131	LYS 132	ASN 133	TRP 134	GLY 135	SER 171	TYR 172
LYS 173						



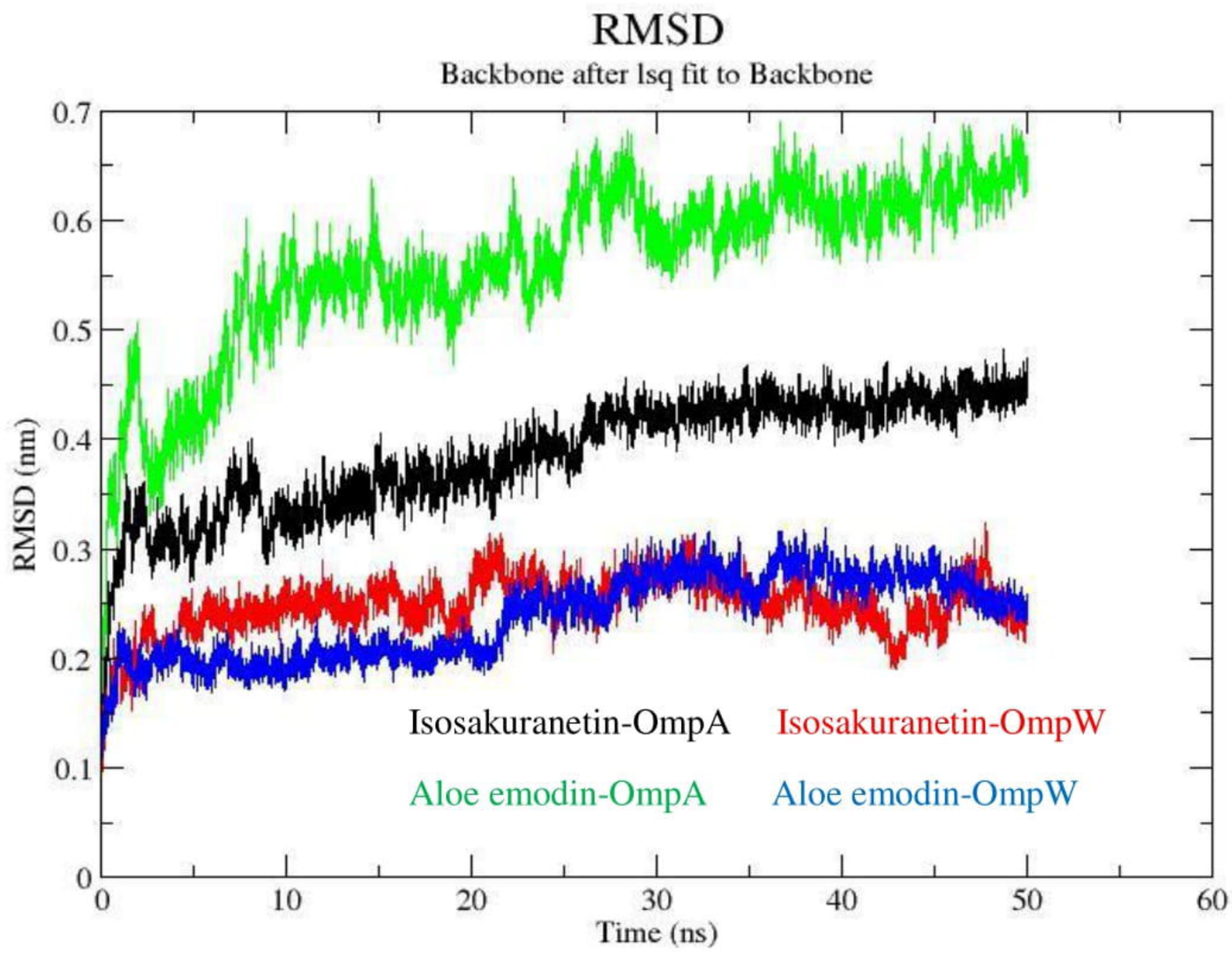
bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



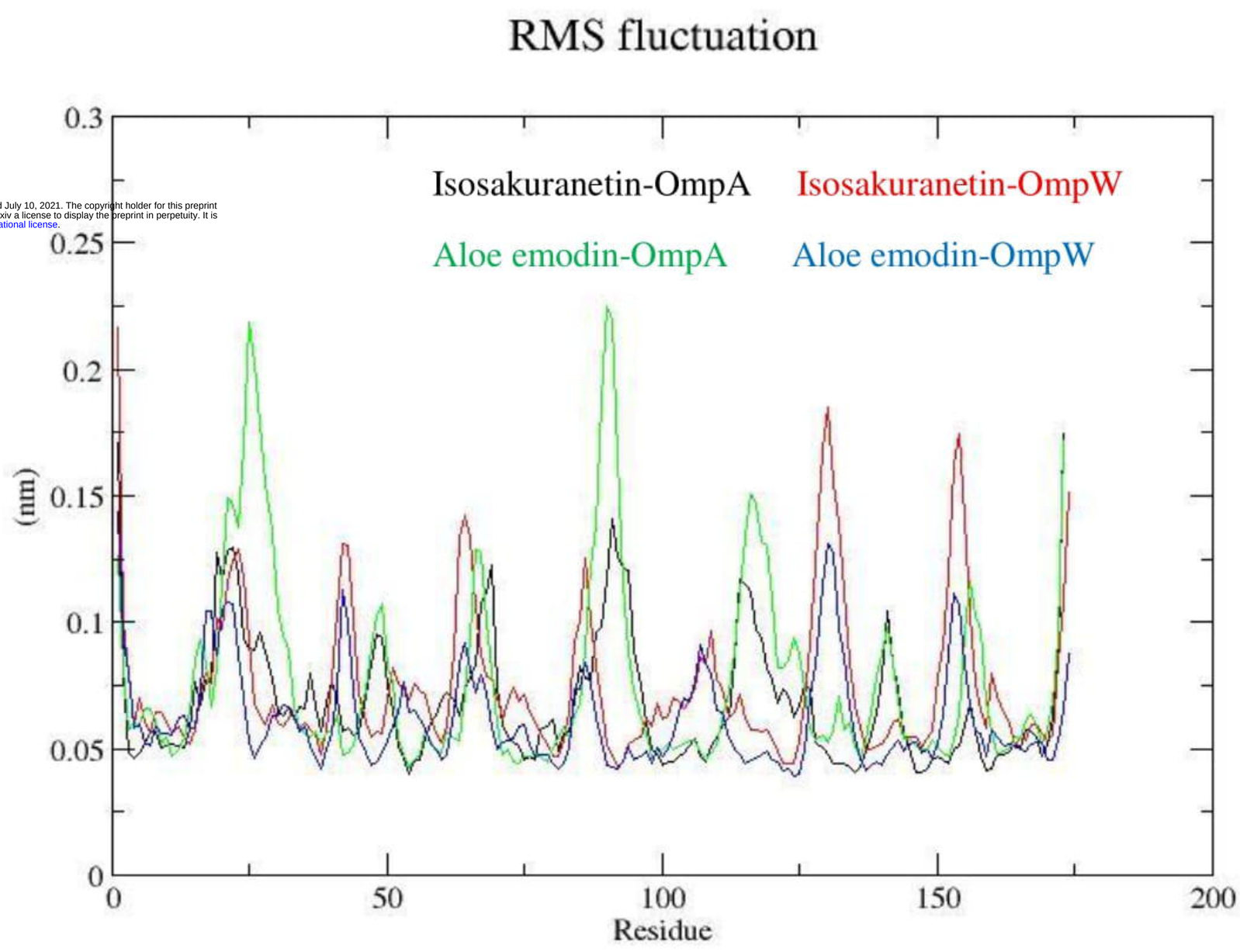
bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.


WLOGP
7
6
5
4
3
2
1
0
-1
-2
-3
-4
0
20
40
60
80
100
120
140
160
180
TPSA
Isosakuranetin
Pinocembrin
Aloe-emodin

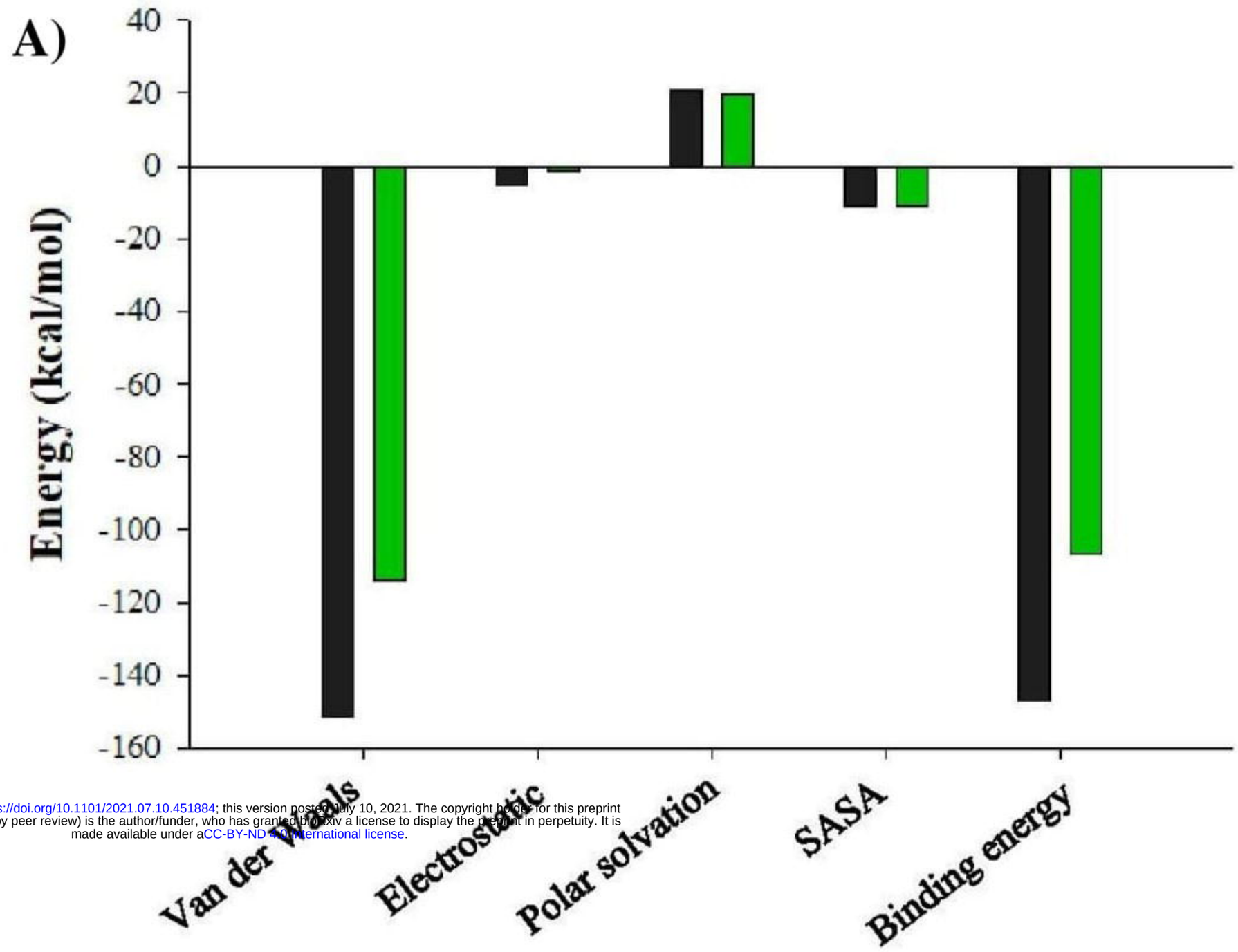
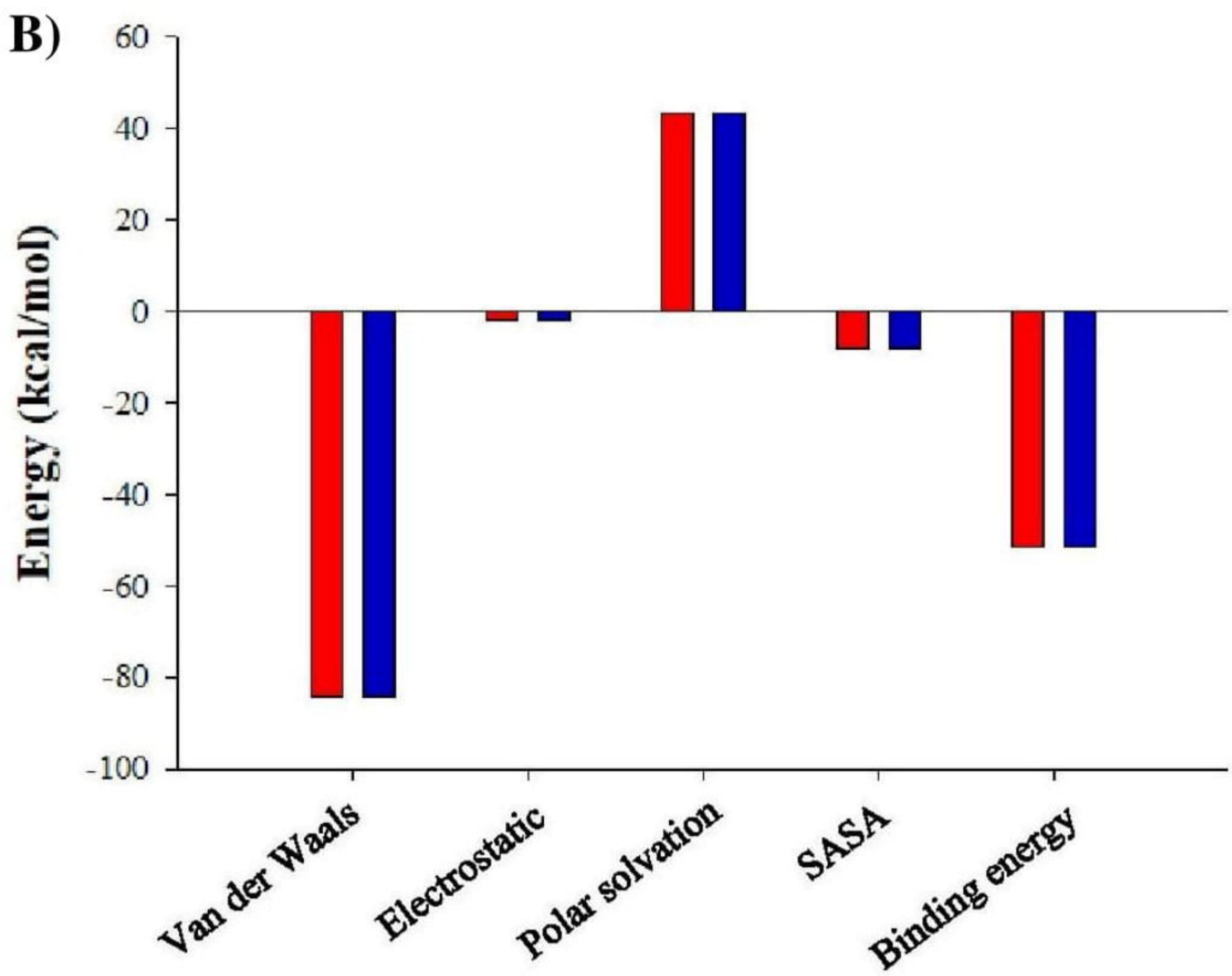
A)



B)



bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.07.10.451884; this version posted July 10, 2021. 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 aCC-BY-ND 4.0 International license.