Research Article

Identification of Putative Cell-entry-inhibitory Peptides against SARS-CoV-2 from Edible Insects: An in silico Study

Fai-Chu Wong1,2, Joe-Hui Ong1, Tsun-Thai Chai1,2*

1Department of Chemical Science, Faculty of Science, Universiti Tunku Abdul Rahman, Kampar 31900, Malaysia
2Center for Biodiversity Research, Universiti Tunku Abdul Rahman, Kampar 31900, Malaysia

ARTICLE INFO

Article History
Received 09 August 2020
Accepted 15 September 2020

Keywords
COVID-19
edible insect
gastrointestinal digestion
peptide
spike protein

ABSTRACT

Coronavirus Disease 2019 (COVID-19), an infectious disease caused by SARS-CoV-2, has become the most contagious pandemic in the world. The binding of SARS-CoV-2 spike glycoprotein Receptor-binding Domain (RBD) to human angiotensin-converting enzyme 2 receptor precedes cell entry. Thus, RBD of SARS-CoV-2 spike glycoprotein is a key target for the development of drugs and vaccines to curb COVID-19. The aim of this in silico study was to identify potential SARS-CoV-2 cell entry inhibitors from peptides derived from edible insects. Twenty-four major proteins from mealworms, silkworm cocoons and housefly larvae were subjected to in silico gastrointestinal (GI) digestion, yielding 3560 fragments. Further screening led to 82 high-GI-absorption peptides with unique sequences. Molecular docking revealed 10 promising peptides (VPW, PPY, PIF, VW, PSF, PGF, PAY, VGF, PF and TW), predicted to interact with at least one key binding residue on RBD. Notably, VPW had the lowest docking energy score (−144.359) and binding affinity (−7.0 kcal/mol), highlighting its potency among the ten peptides. Binding affinities of the insect peptides were superior to some reported natural products. Our study suggests that when consumed, edible insects may be a source of putative SARS-CoV-2 cell entry inhibitors in the form of RBD-binding peptides.

GRAPHICAL ABSTRACT

© 2020 The Authors. Publishing services by Atlantis Press International B.V.
This is an open access article distributed under the CC BY-NC 4.0 license (http://creativecommons.org/licenses/by-nc/4.0/).

1. INTRODUCTION

The current outbreak of Coronavirus Disease 2019 (COVID-19), caused by the newly discovered Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), has driven the world to a public health emergency. Globally, as of 5 August 2020, 18,354,342 confirmed cases of COVID-19, including 696,147 deaths, have been reported to the World Health Organization (https://covid19.who.int). People infected by SARS-CoV-2 might suffer from severe respiratory illness, along with headache, runny nose and fever; severe cases of COVID-19 may even lead to multi-organ failure and deaths [1]. SARS-CoV-2 is a single-stranded, positive sense RNA virus with four major structural proteins, namely the spike, membrane, envelope and nucleocapsid proteins [2]. The binding between the Receptor-binding Domain (RBD) of the spike glycoproteins and the Human Angiotensin Converting Enzyme 2 (hACE2) facilitate viral attachment onto human cells, which precedes membrane fusion and the subsequent SARS-CoV-2 infection [1,3,4]. Receptor recognition is a critical determinant of the entry of SARS-CoV-2 into the human host cells [5]. Thus, blocking the binding of RBD of SARS-CoV-2 spike glycoprotein to hACE2 could be a promising strategy toward the prevention and/or treatment of COVID-19 [4,6].

In recent years, molecular docking has become an increasingly popular drug discovery tool on account of its quick, inexpensive and accurate prediction of biological activity and interactions of ligand compounds with target molecules [7,8]. With this approach, information on the bond types, binding affinity and interactions between two molecules can be predicted from the best-fit conformation [9]. With the availability of the crystal structure information...
of SARS-CoV-2-hACE2 complex, there has been an increasing number of molecular docking studies that screen and identify natural and synthetic inhibitors as well as repurposed drugs that could target the interactions between the SARS-CoV-2 and hACE2 [1,2,8,10]. To the best of our knowledge, in contrast to the studies on phytochemicals, few molecular docking studies have investigated the potential of food-derived peptides in blocking the SARS-CoV-2–hACE2 interactions.

Food-derived peptides are known to exhibit diverse bioactivities, including antiviral properties [11]. Furthermore, peptides are often the preferred choice of therapeutic agents over conventional drugs because of their higher specificity, efficiency, and biodegradability, as well as lower molecular weight and toxicity [12]. In this regard, food-derived peptides which could serve as host cell entry inhibitors against SARS-CoV-2 are of specific interest in this study. Specifically, we are interested in peptides that are released from edible insects upon gastrointestinal (GI) digestion in the human body. Edible insects are good sources of bioactive compounds with great potential for pharmaceutical uses [13]. In this study, the mealworm, silkworm cocoon and housefly larvae were investigated. The three insect samples were chosen because previous studies [14–16] have highlighted their nutritional and medicinal values. The aim of this study was to identify potential SARS-CoV-2 cell entry inhibitors from edible insects through in silico analysis, focusing on high GI-absorption peptides that could bind to the RBD of the SARS-CoV-2 spike proteins. The outcome of this research could contribute towards the future development of food-derived peptides, specifically edible insect peptides, into therapeutic agents that target SARS-CoV-2.

2. MATERIALS AND METHODS

2.1. Materials

Figure 1 shows a flowchart of the study. A total of 24 major proteins found in three edible insects, namely mealworms, housefly larvae and silkworm cocoons [14–16], were studied. Sequences of the proteins were downloaded from the UniProt web server (http://www.uniprot.org) [17] on 14 July 2020.

2.2. In silico GI Digestion and Screening for GI Absorption

In silico GI digestion was performed by using the BIOPEP-UWM web server [18] as previously described [19,20]. The aforementioned 24 protein sequences were subjected to hydrolysis by pepsin (EC 3.4.23.1), trypsin (EC 3.4.21.4) and chymotrypsin A (EC 3.4.21.1) to simulate in silico GI digestion. The number of residues and the distribution of peptide length among the resultant peptide fragments were recorded. Next, unique peptide sequences comprising at least...
two amino acid residues were converted into Simplified Molecular-Input Line-Entry System (SMILES) strings by using the online tool PepSMI (https://www.novoprolabs.com/tools/convert-peptide-to-smiles-string). Using the SMILES strings generated as input, the potential GI absorption of the GI digestion-released peptide fragments was analyzed by using SwissADME (http://www.swissadme.ch). SwissADME is a free bioinformatics web tool that can be used to evaluate the penetration of molecules through the GI wall [21]. Following the screening step, peptide fragments predicted to have high GI absorption were taken to the following step of molecular docking analysis.

2.3. Molecular Docking Analyses with HPEPDOCK

The crystal structure of SARS-CoV-2 Spike RBD in complex with the hACE2 (PDB ID: 6LZG) [3] was obtained from the Protein Data Bank (http://www.rcsb.org/pdb) [22,23]. Next, the preparation of A chain (hACE2) and B chain (RBD) were performed by using BIOVIA Discovery Studio Visualizer (BIOVIA, Dassault Systèmes, BIOVIA Discovery Studio Visualizer, Version 20.1.0.192, San Diego: Dassault Systèmes, 2020), adding in polar hydrogens, and removing water, ligands and heteroatoms. Both structures were then saved in the PDB format and used for analysis on the HPEPDOCK server [24]. The prepared B chain was uploaded as receptor input, whereas the prepared A chain was uploaded as binding site reference. High-GI-absorption peptides released from our in silico GI digestion and four reference peptides that were previously reported to bind to the SARS-CoV-2 Spike RBD were entered in the FASTA format as peptide input. The resultant docking energy scores of peptide samples were tabulated. Interactions of the peptides with the SARS-CoV-2 Spike RBD in the docked models were visualized by using the LigPlot+ v.2.2 [25,26]. The docked models were checked for interactions, especially hydrogen bond and hydrophobic interactions, between the peptides and RBD residues. Based on the best-score models computed by HPEPDOCK, the binding poses of selected insect peptides on the surface of the RBD of SARS-CoV-2 spike proteins were visualized by using the ProteinsPlus web service (https://proteins.plus) [27].

2.4. Molecular Docking Analyses with AutoDock Vina Tool in PyRx

Following HPEPDOCK analysis, high-GI-absorption peptides predicted to bind to at least one of the nine known key binding residues (Leu455, Phe456, Ser459, Gln474, Ala475, Phe486, Phe490, Gln493 and Pro499) on RBD [4] were selected for further comparison with 26 phytochemicals that were previously reported to bind to the SARS-CoV-2 Spike RBD. Molecular docking analysis of the selected insect peptides and reference phytochemicals was performed by using the AutoDock Vina option in PyRx 0.8 [28,29]. For macromolecule preparation, the RBD protein was pre-processed by using the AutoDockTools 1.5.6 [30,31], which involved the removal of water as well as addition of Kollman charges and polar hydrogens, and then saved in the PDBQT format file. For the ligand preparation, the 3D structures of reference phytochemicals, where available, were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov) [32] in the SDF format and converted into PDB format with the Open Babel tool in PyRx 0.8. Where only 2D information is found on PubChem for a reference phytochemical, the 2D-to-3D structure conversion was accomplished by using the MarvinSketch software (version 20.13.0) (ChemAxon, https://www.chemaxon.com), saved in the PDB format. The 3D structures of the selected insect peptides were obtained from the collection of 3D structures of dipeptides and tripeptides made available by Prasasty and Istyastono [33]. The docking process was performed in which the binding site of protein was set with the grid box center x-, y-, z-axis values of −25.0426, 15.8143, −25.4535 with dimension of 100 × 100 × 90 Å, encompassing 21 binding site residues (Lys417, Gly446, Tyr449, Tyr453, Leu455, Phe456, Tyr473, Ala475, Gly476, Glu484, Phe486, Asn487, Tyr489, Phe490, Gln493, Gly496, Gln498, Thr500, Asn501, Gly502 and Tyr505) [3] and three other RBD residues (Ser459, Gln474 and Pro499) that were reported to be important for RBD-hACE2 binding [4]. The resultant binding affinities of peptide samples and reference phytochemicals were tabulated.

3. RESULTS AND DISCUSSION

In order to discover insect peptides that are potentially bioavailable upon oral consumption, simulated GI digestion was performed on insect proteins in this in silico study. A total of 3560 peptide fragments were released from the 24 insect proteins analyzed following in silico GI digestion; 1771 peptide fragments were released from the silkworm cocoon, 1006 fragments from the mealworm, and 783 fragments from housefly larvae. Details of each insect protein and the outcome of in silico GI digestion are presented in Table 1. The distribution of peptide length in the GI digestion-released fragments varied among the three insect samples. When 2- and 3-residue peptide fragments generated by in silico GI digestion are considered collectively, the percentages of such peptides released from the proteins of the silkworm cocoon, housefly larvae, and mealworm are in the range of 25.53–57.14%, 28.89–45.45% and 27.78–43.08%, respectively. Low-molecular-weight peptides are absorbed more readily through the GI tract than the high-molecular-weight peptides. Most often, di- and tri-peptides can be absorbed intact through the epithelium of the GI tract [34]. In accordance with this, the greatest number of high-GI-absorption peptides was predicted for the silkworm cocoon (172 fragments) by the SwissADME web server [21], in contrast with 140 high-GI-absorption peptides for mealworm and 105 for housefly larvae.

More specifically, in silico GI digestion of the proteins of each insect released the largest number of peptide fragments from sericin 2 (silkworm cocoon; 652 fragments), alpha-actinin-4 (mealworm; 390 fragments), and catalase (housefly larvae; 197 fragments). Unsurprisingly, the three insect proteins were also the key contributors of high-GI-absorption peptides, as predicted by the SwissADME web server [21], in the descending order of sericin 2 (74 fragments) > alpha-actinin-4 (58 fragments) > catalase (27 fragments). Notably, the combination of sericin 2 and fibroin heavy chain contributed to 76% of the 172 high-GI-absorption peptides generated from silkworm cocoon major proteins upon in silico GI digestion. Taken together, our in silico analysis suggests that silkworm cocoon proteins, when consumed, are more likely to produce GI-absorbable peptides which could be bioavailable in vivo. Specifically, the sericin 2 and fibroin heavy chain of the silkworm cocoon deserve future attention in the search of natural, edible sources of bioavailable peptides.
In this study, we also compared the docking energy scores of the high-GI-absorption insect peptides, implying its relative potency among the whole set of 82 high-GI-absorption insect peptides, suggesting its potential as an entry inhibitor from edible insect peptides that are potentially bioavailable. Since the aim of this study was to identify putative SARS-CoV-2 cell entry inhibitors from edible insect peptides that are potentially bioavailable, we analyzed only the 82 high-GI-absorption insect peptides in our molecular docking studies. In HPEP Dock analysis, all the high-GI-absorption peptides were docked to the spike glycoprotein RBD via the local peptide docking mode, taking hACE2 as the reference for binding sites. The docking energy scores of the peptides ranged from −53.873 to −144.359, as documented in Table 1.

Recently, Yang et al. [35] reported a strong linear correlation between the in vitro activity of monoamine oxidase A inhibitory peptides and the HPEP Dock docking energy scores predicted for the peptides when docked to the enzyme. Their finding highlights the possible application of HPEP Dock scores for ranking or comparing predictively the relative bioactivity of peptides. In this study, based on docking energy scores, the five top-score peptides of mealworms in descending order are VPW, PPY, PW, IW, and PAY. For silkworm cocoon, the five top-score peptides in descending order are PIF, VW, IPY, PGF and TW. In the case of housefly larvae, the five top-score peptides in descending order are VW, PSF, PGF, VGF, and PF. The same high-GI-absorption peptide sequences might be released from different proteins of same insect, or from proteins of different insects. Thus it is unsurprising that VW and PGF were both listed in the five top-score peptides of silkworm cocoon and housefly larvae. Importantly, VPW was the peptide with the lowest (most negative) docking energy score among the whole set of 82 high-GI-absorption insect peptides, implying its relative potency in binding with SARS-CoV-2 spike protein RBD.

In this study, we also compared the docking energy scores of the high-GI-absorption insect peptides with reference peptides previously reported or predicted to bind to SARS-CoV-2 RBD, namely IEEQAKTFLDKFKNHEAEDLFYQYS, FLDKFKNHEAEDLFYQSSL, LP1Y and CSNAIPEL. IEEQAKTFLDKFKNHEAEDLFYQYS, also named as SBP1, is a 23-mer peptide designed from the peptidase domain of hACE2 [36]. FLDKFKNHEAEDLFYQSSL is an artificially-designed 18-residue peptide [37]. LP1Y and CSNAIPEL are peptides identified from the myosin of scallop Mizuhopecten yessoensis after in silico enzymatic hydrolysis [38]. The docking energy scores of IEEQAKTFLDKFKNHEAEDLFYQYS and FLDKFKNHEAEDLFYQSSL are both lower than that computed for all high-GI-absorption insect peptides (Table 2). On the other hand, the docking energy scores of LP1Y and CSNAIPEL are both slightly higher than VPW, although still lower than the rest of the insect peptides. Notably, although the four reference peptides are either comparable or superior to the insect peptides in docking energy scores, all four were predicted to be susceptible to enzymatic hydrolysis after in silico GI digestion by BIOPEP server and to have low GI absorption by SwissADME server (data not shown). In short, where oral bioavailability is concerned, the insect peptides presented in Table 2 would still have an advantage over the four reference peptides. In particular, the mealworm-derived VPW is the most promising peptide in our high-GI-absorption insect peptide dataset to play the role of SARS-CoV-2 cell entry inhibitor.

Since the aim of this study was to identify putative SARS-CoV-2 cell entry inhibitors from edible insect peptides that are potentially bioavailable, we analyzed only the 82 high-GI-absorption insect peptides in our molecular docking studies. In HPEP Dock analysis, all the high-GI-absorption peptides were docked to the spike glycoprotein RBD via the local peptide docking mode, taking hACE2 as the reference for binding sites. The docking energy scores of the peptides ranged from −53.873 to −144.359, as documented in Table 1. Recently, Yang et al. [35] reported a strong linear correlation between the in vitro activity of monoamine oxidase A inhibitory peptides and the HPEP Dock docking energy scores predicted for the peptides when docked to the enzyme. Their finding highlights the possible application of HPEP Dock scores for ranking or comparing predictively the relative bioactivity of peptides. In this study, based on docking energy scores, the five top-score peptides of mealworms in descending order are VPW, PPY, PW, IW, and PAY. For silkworm cocoon, the five top-score peptides in descending order are PIF, VW, IPY, PGF and TW. In the case of housefly larvae, the five top-score peptides in descending order are VW, PSF, PGF, VGF, and PF. The same high-GI-absorption peptide sequences might be released from different proteins of same insect, or from proteins of different insects. Thus it is unsurprising that VW and PGF were both listed in the five top-score peptides of silkworm cocoon and housefly larvae. Importantly, VPW was the peptide with the lowest (most negative) docking energy score among the whole set of 82 high-GI-absorption insect peptides, implying its relative potency in binding with SARS-CoV-2 spike protein RBD.

In this study, we also compared the docking energy scores of the high-GI-absorption insect peptides with reference peptides previously reported or predicted to bind to SARS-CoV-2 RBD, namely IEEQAKTFLDKFKNHEAEDLFYQYS, FLDKFKNHEAEDLFYQSSL, LP1Y and CSNAIPEL. IEEQAKTFLDKFKNHEAEDLFYQYS, also named as SBP1, is a 23-mer peptide designed from the peptidase domain of hACE2 [36]. FLDKFKNHEAEDLFYQSSL is an artificially-designed 18-residue peptide [37]. LP1Y and CSNAIPEL are peptides identified from the myosin of scallop Mizuhopecten yessoensis after in silico enzymatic hydrolysis [38]. The docking energy scores of IEEQAKTFLDKFKNHEAEDLFYQYS and FLDKFKNHEAEDLFYQSSL are both lower than that computed for all high-GI-absorption insect peptides (Table 2). On the other hand, the docking energy scores of LP1Y and CSNAIPEL are both slightly higher than VPW, although still lower than the rest of the insect peptides. Notably, although the four reference peptides are either comparable or superior to the insect peptides in docking energy scores, all four were predicted to be susceptible to enzymatic hydrolysis after in silico GI digestion by BIOPEP server and to have low GI absorption by SwissADME server (data not shown). In short, where oral bioavailability is concerned, the insect peptides presented in Table 2 would still have an advantage over the four reference peptides. In particular, the mealworm-derived VPW is the most promising peptide in our high-GI-absorption insect peptide dataset to play the role of SARS-CoV-2 cell entry inhibitor.

Five insect peptides having the lowest docking energy scores were chosen from each insect sample and further analyzed by using LigPlot+. Figure 2 shows the binding poses of the insect peptides on the RBD surface. Table 3 summarizes the RBD residues that interact with each peptide, in addition to the types and the number of interactions, as computed by LigPlot+. Figures 3 and 4 present the graphic representations that reveal the molecular interactions at the reference peptide- and insect peptide-RBD interfaces, respectively. According to Wang et al. [3] and Yi et al. [4], 24 residues on RBD are known
Table 2  Docking energy scores of the top peptide binding models generated from the high-GI-absorption insect peptides and the SARS-CoV-2 spike protein RBD

| Insect peptides/Reference peptides | Docking energy scores | Sources |
|-----------------------------------|-----------------------|---------|
| IEEQAKTFDLKFNEAHEDLFYQSY         | −223.534              | Mealworm Silkworm cocoon Housefly larve |
| FLDKFNFHEADELFYQSSL               | −216.761              | - - - |
| VPW                               | −144.359              | √ - - |
| LPJY                              | −139.134              | - - - |
| CSNAIPEL                          | −135.411              | - - - |
| PIP                               | −131.773              | √ - - |
| PLF                               | −131.740              | - √ - |
| PV                                   | −128.975              | √ - - |
| VV                                   | −127.336              | √ √ - |
| PSF                               | −126.842              | - - √ |
| JW                               | −125.121              | √ - - |
| IPY                               | −124.618              | - √ - |
| PGF                               | −124.263              | - √ √ |
| PAY                               | −124.041              | √ - - |
| VGW                               | −122.073              | √ - - |
| VGF                               | −117.686              | - - √ |
| PF                                | −117.215              | - √ - |
| TW                                | −116.374              | - - √ |
| GPF                               | −116.337              | √ √ - |
| AVF                               | −114.853              | √ √ - |
| PY                                | −113.683              | √ - - |
| GW                                | −112.521              | - √ √ |
| GPF                               | −111.835              | - - - |
| AF                                | −109.368              | √ - - |
| AGF                               | −108.132              | - √ √ |
| VY                                | −106.424              | - - - |
| PH                                | −106.229              | √ - - |
| IF                                | −105.689              | - √ - |
| TF                                | −103.652              | √ √ √ |
| PIL                               | −102.956              | - - √ |
| IF                                | −102.463              | √ √ - |
| TY                                | −102.385              | √ - - |
| PPK                               | −101.851              | - - - |
| CY                                | −100.116              | √ √ √ |
| QF                                | −99.749               | - √ - |
| AF                                | −99.248               | √ √ - |
| CF                                | −99.034               | - √ √ |
| EF                                | −98.684               | √ - - |
| SF                                | −98.372               | √ √ - |
| AF                                | −97.583               | √ √ √ |
| VF                                | −97.546               | - √ - |
| GF                                | −97.025               | - - - |
| PGL                               | −95.628               | - - √ |
| PV                                   | −95.210               | √ - - |
| IL                                 | −93.084               | √ - - |
| VIL                               | −92.851               | - - - |
| IH                                 | −92.003               | √ - - |
| DF                                 | −91.415               | √ √ - |
| AH                                 | −86.858               | - - √ |
| PM                                 | −88.607               | - - - |
| AIL                               | −85.500               | √ - - |
| IGV                               | −84.507               | - - √ |
| VH                                 | −83.877               | - - √ |
| IM                                 | −83.596               | √ - - |
| PL                                 | −81.722               | √ √ - |
| VVL                               | −81.706               | √ - - |
| GVK                               | −81.566               | √ - - |

Table 2  Docking energy scores of the top peptide binding models generated from the high-GI-absorption insect peptides and the SARS-CoV-2 spike protein RBD—Continued

| Insect peptides/Reference peptides | Docking energy scores | Sources |
|-----------------------------------|-----------------------|---------|
| IGL                               | −81.030               | - - √ |
| VPL                               | −80.901               | - - √ |
| GIL                               | −80.882               | √ √ - |
| GH                                | −80.350               | - - - |
| IAL                               | −80.321               | - - - |
| VM                                 | −79.125               | √ - √ |
| PK                                 | −78.120               | √ √ - |
| VAL                               | −75.726               | √ - - |
| AAL                               | −75.648               | - √ - |
| IK                                 | −74.883               | √ √ √ |
| GAL                               | −73.325               | - - - |
| VGL                               | −73.179               | - - - |
| EL                                 | −71.453               | √ - √ |
| AM                                 | −71.421               | - - √ |
| IL                                 | −70.938               | √ - √ |
| GGL                               | −70.675               | - √ - |
| AGL                               | −69.661               | - - - |
| GM                                 | −69.515               | √ - - |
| TL                                 | −68.382               | √ √ - |
| VL                                 | −63.197               | √ √ - |
| VDF                               | −62.584               | √ - √ |
| VB                                 | −61.557               | √ √ - |
| AL                                 | −61.453               | √ - - |
| CL                                 | −59.264               | √ √ - |
| SL                                 | −58.303               | √ √ - |
| GL                                 | −57.232               | √ - - |
| EV                                 | −56.956               | √ - - |
| DL                                 | −56.133               | √ √ - |
| AK                                 | −53.873               | √ √ √ |

*Indicates reference peptides.

to bind to hACE2. In addition, Yi et al. [4] reported that the nine residues Leu455, Phe456, Ser459, Gln474, Ala475, Phe486, Phe490, Gln493 and Pro499 are the key binding sites on RBD; binding to these residues can significantly affect the affinity of RBD for hACE2. Peptides that are able to interact with any of the nine key binding residues would be potential inhibitors that block the binding of the RBD on spike glycoprotein to hACE2, hence preventing virus entry into the host cell and the initiation of infection [4]. In our LigPlot+ analysis, the two reference peptides IEEQAKTFDLKFNEAHEDLFYQSY and FLDKFNFHEADELFYQSSL were found to interact with more key binding residues (Leu455, Phe456, Ala475 and Gln493) than were the insect peptides or the other two reference peptides LPJY and CSNAIPEL. Moreover, in comparison with the other peptides, IEEQAKTFDLKFNEAHEDLFYQSY and FLDKFNFHEADELFYQSSL also interact with a greater number of other RBD binding site residues, mainly via hydrophobic interactions. The two reference peptides each establish at least 15 hydrophobic interactions with the RBD structure, surpassing all other peptides we analysed (Table 3 and Figure 3). This observation substantiates the much lower docking energy scores calculated by HPEPDOCK for the two peptides, relative to the rest of the peptides analysed (Table 2). Further supporting the importance of hydrophobic interactions at the RBD-peptide interface is our observation that despite the lack of hydrogen
Figure 2. Binding poses of 10 selected insect peptides on the surface of the RBD of SARS-CoV-2 spike protein. (A) VPW, (B) PPY, (C) PIF, (D) VW, (E) PSF, (F) PGE, (G) PAY, (H) VGE, (I) PF, and (J) TW.
Table 3  Interactions between the SARS-CoV-2 RBD residues and five peptides having the lowest docking energy scores from each insect sample

| Insect peptides/Reference peptides | Sources | Hydrogen bond | Hydrophobic interaction | Hydrogen bond | Hydrophobic interaction |
|-----------------------------------|---------|---------------|-------------------------|---------------|-------------------------|
|                                   | Mealworm| Silkworm cocoon| Housefly larvae         |               |                         |
| IEEQAKTFLDKFN-                   |         |               |                         |               |                         |
| HEADLYQYS                        |         |               |                         |               |                         |
| FLDFKNEADLYQSSL                  |         |               |                         |               |                         |
| VPW                              | √       |               |                         |               |                         |
| LPIY                             |         |               |                         |               |                         |
| CSNAIPEL’                        |         |               |                         |               |                         |
| PPy                              | √       |               |                         |               |                         |
| PIF                              | -       | √             |                         |               |                         |
| PW                               | √       | -             |                         |               |                         |
| VW                               | -       | √             | √                       |               |                         |
| PSF                              | -       | -             | √                       |               |                         |
| TW                               | √       | -             | √                       |               |                         |
| IPY                              | -       | √             | √                       |               |                         |
| PFG                              | -       | √             | √                       |               |                         |
| PAY                              | √       | -             | √                       |               |                         |
| VGF                              | -       | -             | √                       |               |                         |
| PF                               | -       | -             | √                       |               |                         |
| TW                               | -       | √             |                         |               |                         |

*Residues in **bold** are among the nine known key binding residues on RBD that hACE2 binds to, whereas the underlined residues are other binding sites for hACE2. *Indicates reference peptides.
Figure 3  Two dimensional visualization of the RBD-peptide binding interfaces as computed by LigPlot+ for the reference peptides (A) IEEQAKTFLDKFNHEAEDLFYQS, (B) FLDKFNHEAEDLFYQSSL, (C) LPIY and (D) CSNAIPEL. Ligands (peptides) are in purple, whereas the bonds of protein (RBD of SARS-CoV-2 spike protein) are represented in orange color. Hydrogen bonds are displayed in green dashed lines, whereas hydrophobic interactions are represented by red spoked arcs.
Figure 4  Two dimensional visualization of the RBD-peptide binding interfaces as computed by LigPlot+ for the insect peptides: (A) VPW, (B) PPy, (C) PIF, (D) VW, (E) PSF, (F) PGF, (G) PAY, (H) VGF, (I) PF, and (J) TW. Descriptions for ligands (peptides), protein (RBD of SARS-CoV-2 spike protein), and bonds are the same as in Figure 3.
bond formation, the lowest docking energy score was predicted for the RBD-VPW docked model relative to the rest of the insect peptides. All other insect peptides form at least one hydrogen bond with a key binding site residue and/or other binding site residues as revealed by our LigPlot+ analysis (Table 3).

In this study, VPW was predicted to interact with the SARS-CoV-2 spike protein through 12 hydrophobic interactions, which involve one key binding residue Leu455 and five other residues previously reported to bind to hACE2 (Lys417, Tyr453, Gly496, Asn501, Tyr505) among others (Table 3 and Figure 4A). This clearly indicates the potential of VPW to bind to the RBD of SARS-CoV-2 spike proteins and to block its interaction with the human cell surface receptor hACE2. PPY and PIF are the peptides computed by HPEPDOCK to have the second and the third lowest docking energy scores. The two peptides each form nine hydrophobic interactions at the RBD-peptide interface, similarly interacting with one key binding residue Leu455 and three other binding site residues (Lys417, Tyr453 and Tyr505). The slightly lower docking energy score of PPY, relative to PIF, is apparently attributable to PPY forming four hydrogen bonds with the RBD of the spike protein, whereas PIF forms only one (Figures 4B and 4C). LPIY can form both hydrophobic interaction and hydrogen bond with the key binding residue Gln493. We also observed the same for insect peptides VW, PSF and PAY (Figures 4D, 4E and 4G). In addition, like CSNAIPEL, insect peptides VPW, PPY, PIF, and PAY are each able to form hydrophobic interaction with the key binding residue Leu455 (Figures 4A–4C and 4G). Thus Leu455 and Gln493 are two of the nine key binding residues targeted by the insect peptides in our dataset.

At present, the number of studies concerning natural or food-derived peptides which potentially bind to SARS-CoV-2 spike proteins is limited. Hence we decided to further characterise the insect peptides by comparing their RBD-binding affinities with those of phytochemicals. Comparison was made with 26 natural products derived from edible and medicinal plants that were reported in 10 studies to have the ability to bind to SARS-CoV-2 spike proteins; such natural products represent potential cell entry inhibitors against SARS-CoV-2 [2,10,39–46]. To this end, we focussed on 10 insect peptides (VPW, PPY, PIF, VW, PSF, PGF, PAY, VGF, PF and TW) which were predicted by LigPlot+ to bind to at least one key binding residue on the RBD of SARS-CoV-2 spike proteins. Molecular docking analysis by using the PyRx software (AutoDock Vina option) found that among the insect peptides, VPW displayed the lowest binding affinity value (−7.0 kcal/mol), suggesting that it can bind more strongly to the SARS-CoV-2 spike proteins than the other nine insect peptides (Figure 5). In contrast with the docking energy scores computed by HPEPDOCK, the difference of binding affinities between peptides as analyzed by PyRx was less distinct, for example, between PIF, VW and PSF. This may be attributed to the greater accuracy of HPEPDOCK in analyzing small peptides when compared with traditional small-molecule docking programs, including AutoDock Vina [47]. All in all, the trends of docking energy scores and binding affinities generally appear similar for the 10 insect peptides, both suggesting the strongest effectiveness of VPW in binding to the SARS-CoV-2 spike protein. When compared with the phytochemicals, the binding affinity of VPW was found to be 4.11–17.65% weaker than digitoxin, hesperidin, glycyrrhizic acid, taraxerol and asparoside-C. The binding affinity of VPW is comparable to those of dihydrotanshinone I, fisetin, frie-delin and asparoside-D. By contrast, the binding affinity of VPW was 2.86–42.86% stronger than 17 of the phytochemicals analyzed. The binding affinities of nine natural products, namely curcumin, carvacrol, thymol, cinnamyl acetate, l-4-terpineol, pulegone, cin-namaldehyde, anethole and geraniol, were weaker than the 10 insect peptides. Taken together, the 10 insect peptides in Figure 5 are comparable or superior to some phytochemicals previously reported to have binding potential toward the SARS-CoV-2 spike proteins.
Ten peptides were identified from the *in silico* GI digestion of major proteins from edible insects, which were predicted *in silico* to be readily absorbed by the GI tract. Importantly, the peptides could potentially serve as cell entry inhibitors against SARS-CoV-2. Molecular docking suggests mealworm-derived peptide VPW is the most promising among the ten peptides. Our study highlights that edible insects, especially mealworms, are promising sources of potential anti-COVID-19 peptides, which should be explored more intensively in future. Importantly, future *in vivo* experiments are necessary to confirm our findings; despite high GI absorption, *in vivo* bioavailability could also be influenced by alteration or inactivation by the liver.

**REFERENCES**

[1] Choudhary S, Malik YS, Tomar S. Identification of SARS-CoV-2 cell entry inhibitors by drug repurposing using *in silico* structure-based virtual screening approach. Front Immunol 2020;11:1664.

[2] Subbaiyan A, Ravichandran K, Singh SV, Sankar M, Thomas P, Dhama K, et al. *In silico* molecular docking analysis targeting SARS-CoV-2 Spike protein and selected herbal constituents. J Pure Appl Microbiol 2020:14:989–98.

[3] Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. Cell 2020;181:894–904.

[4] Yi C, Sun X, Ye J, Ding L, Liu M, Yang Z, et al. Key residues of the receptor binding motif in the spike protein of SARS-CoV-2 that interact with ACE2 and neutralizing antibodies. Cell Mol Immunol 2020;17:621–30.

[5] Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, et al. Cell entry mechanisms of SARS-CoV-2. Proc Natl Acad Sci U S A 2020;117:11727–34.

[6] Huang X, Pearce R, Zhang Y. *De novo* design of protein peptides to block association of the SARS-CoV-2 spike protein with human ACE2. Aging (Albany NY) 2020;12:11263–76.

[7] Ahmad I, Azminah A, Mulia K, Yanuar A, Mun'im A. Angiotensin-converting enzyme inhibitory activity of polyphenolic compounds from *Peperomia pellucida* (L) Kunth: an *in silico* molecular docking study. J Appl Pharm Sci 2019;9:25–31.

[8] Poochi SP, Easwaran M, Balasubramanian B, Anbuselvam M, Meyyazhagan A, Park S, et al. Employing bioactive compounds derived from *Ipomoea obscura* (L) to evaluate potential inhibitor for SARS-CoV-2 main protease and ACE2 protein. Food Front 2020;1:168–79.

[9] Akhter M. Challenges in docking: mini review. JMS Chem 2016;4:1025.

[10] Pandey P, Rane JS, Chatterjee A, Kumar A, Khan R, Prakash A, et al. Targeting SARS-CoV-2 spike protein of COVID-19 with naturally occurring phytochemicals: an *in silico* study for drug development. J Biomol Struct Dyn 2020 (In Press).

[11] Harnedy PA, FitzGerald RJ. Bioactive peptides from marine processing waste and shellfish: a review. J Funct Foods 2012;4:6–24.

[12] Castel G, Chiéoui M, Heyd B, Tordo N. Phage display of combinatorial peptide libraries: application to antiviral research. Molecules 2011;16:3499–518.

[13] Ratcliffe NA, Mello CB, Garcia ES, Butt TM, Azambuja P. Insect natural products and processes: new treatments for human disease. Insect Biochem Mol Biol 2011;41:747–69.

[14] Jiao M, Chen L, He Y, Wu L, Mei H. Identification of proteins in housefly (*Musca domestica*) larva powder by LC-MS/MS and their potential medical relevance. RSC Adv 2019;9:30545–55.

[15] Yi L, Van Boekel MAJS, Boeren S, Lakemond CMM. Protein identification and *in vitro* digestion of fractions from *Tenebrio molitor*. Eur Food Res Technol 2016;242:1285–97.

[16] Zhang Y, Zhao P, Dong Z, Wang D, Guo P, Guo X, et al. Comparative proteome analysis of multi-layer cocoon of the silk-worm, *Bombyx mori*. PLoS One 2015;10:e0123403.

[17] The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 2019;47:D506–D15.

[18] Minkiewicz P, Iwaniaik A, Darewicz M. BIOPEP-UWM database of bioactive peptides: current opportunities. Int J Mol Sci 2019;20:5978.

[19] Barati M, Javanmardi F, Jabbari M, Mokari-Yamchi A, Farahmand F, Eq I, et al. An *in silico* model to predict and estimate digestion-resistant and bioactive peptide content of dairy products: a primarily study of a time-saving and affordable method for practical research purposes. IWT 2020;130:109616.

[20] Chai TT, Soo ZY, Hsu KC, Li JC, Abd MF, Wong FC. Antioxidant activity of semen cassiae protein hydrolysate: thermal and gastrointestinal stability, peptide identification, and *in silico* analysis. Mod Food Sci Technol 2019;35:38–48.

[21] Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 2017;7:42717.

[22] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res 2000;28:235–42.

[23] Burley SK, Berman HM, Bhikadiya C, Bi C, Chen L, Di Costanzo L, et al. RCSB protein data bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. Nucleic Acids Res 2018;47:D464–D74.

[24] Zhou P, Jin B, Li H, Huang SY. HPEPDOCK: a web server for blind peptide–protein docking based on a hierarchical algorithm. Nucleic Acids Res 2018;46:W443–W50.

[25] Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. Protein Eng 1995;8:127–34.
[26] Laskowski RA, Swindells MB. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. J Chem Inf Model 2011;51:2778–86.

[27] Schöning-Stierand K, Diedrich K, Fährrolfes R, Flachsenberg F, Meyder A, Nittinger E, et al. ProteinsPlus: interactive analysis of protein–ligand binding interfaces. Nucleic Acids Res 2020;48:W48–W53.

[28] Trot O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 2010;31:455–61.

[29] Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. In: Hempel JE, Williams CH, Hong CC, editors. Chemical biology. Methods in molecular biology. Vol. 1263. New York, NY: Springer; 2015, pp. 243–50.

[30] Prasasty VD, Istyastono EP. Data of small peptides in SMILES and three-dimensional formats for virtual screening campaigns. Nucleic Acids Res 2016;44:D1202–D13.

[31] Wang B, Xie N, Li B. Influence of peptide characteristics on their stability, intestinal transport, and in vitro bioavailability: a review. J Food Biochem 2019;43:e12571.

[32] Yang X, Wang K, Liu Q, Zhang X. Discovery of monoamine oxidase A inhibitory peptides from hairtail (Trichiurus japonicus) using in vitro simulated gastrointestinal digestion and in silico studies. Bioorg Chem 2020;101:104032.

[33] VanPatten S, He M, Altiti A, Cheng KF, Ghemani MH, Al-Abed Y. Evidence supporting the use of peptides and peptidomimetics as potential SARS-CoV-2 (COVID-19) therapeutics. Future Med Chem 2020 (In Press).

[34] Wu C, Liu Y, Yang Y, Zhong R, Zhong W, Wang Y, et al. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. Acta Pharm Sin B 2020;10:766–88.

[35] Zhang DH, Wu KL, Zhang X, Deng SQ, Peng B. In silico screening of Chinese herbal medicines with the potential to directly inhibit 2019 novel coronavirus. J Integr Med 2020;18:152–8.

[36] Zhou P, Li B, Yan Y, Jin B, Wang L, Huang SY. Hierarchical flexible peptide docking by conformer generation and ensemble docking of peptides. J Chem Inf Model 2018;58:1292–302.