Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
SARS-CoV-2 spike protein-, main protease- and papain-like-protease-targeting peptides from seed proteins following gastrointestinal digestion: An in silico study

Fai-Chu Wong a,b, Joe-Hui Ong a, Tsun-Thai Chai a,b,*

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

A R T I C L E   I N F O

Keywords:
COVID-19
Gastrointestinal digestion
Peptide
SARS-CoV-2
Seed protein

A B S T R A C T

Background: The anti-COVID-19 potential of phytochemicals was investigated in numerous studies, but efficacy of peptides released by seed proteins upon gastrointestinal (GI) digestion is underexplored.

Purpose: This study investigated whether multi-target anti-COVID-19 peptides could be released from edible seeds following GI digestion, by using in silico and molecular docking approaches.

Methods: Nineteen seed storage proteins from Chenopodium quinoa (quinoa), Sesamum indicum (sesame), Brassica napus (rape), Helianthus annuus (sunflower) and Cucurbita maxima (pumpkin) seeds were subjected to in silico GI digestion, in order to detect the released peptides with high GI absorption that concurrently target the spike protein, main protease and papain-like protease of SARS-CoV-2.

Results: Molecular docking study revealed that 36 peptides with high GI absorption, out of the 1593 peptides released from seed proteins, could bind to the binding or catalytic sites of the spike protein, main protease and papain-like protease of SARS-CoV-2, after GI digestion. Among the five seeds, quinoa was predicted to release the largest number (27) of multi-target peptides. When compared with PIY (Pro-Ile-Tyr), a high GI-absorption fragment released from a potential anti-COVID-19 peptide, pumpkin-derived peptide PW (Pro-Trp) could bind more strongly to SARS-CoV-2 spike protein. PW was superior to some previously reported anti-SARS-CoV-2 phytochemicals when binding affinities towards the three viral targets were compared.

Conclusion: Edible seeds are a potential source of anti-COVID-19 peptides upon GI digestion, hence they should be considered as an alternative to assist in the treatment and management of COVID-19.

Introduction

The pandemic of the coronavirus disease 2019 (COVID-19), caused by a novel pathogenic virus named as Severe Acute Respiratory Syndrome Coronavirus type 2 (SARS-CoV-2), has escalated globally. The great pandemic potential of SARS-CoV-2 could be attributed to its high human-to-human transmissible efficiency (Zehra et al., 2020). Patients affected by COVID-19 most commonly experience cough, runny nose, fever and headache, although some suffer from serious complications such as severe pneumonia, bronchitis, multi-organ failure, respiratory illness, and even deaths (Zehra et al., 2020). SARS-CoV-2 is a beta-coronavirus whose genome codes for four major structural proteins, namely spike, envelope, membrane and nucleocapsid proteins, as well as non-structural proteins, for instance the main protease (M pro) and papain-like protease (PL pro) (Zehra et al., 2020). The interaction between the receptor-binding domain (RBD) of SARS-CoV-2 spike glycoproteins and human angiotensin converting enzyme 2 (hACE2) facilitates the entry of viral particles into the host cell, which precedes infection (Naqvi et al., 2020; Wang et al., 2020a; Yi et al., 2020). Nine residues, Leu455, Phe456, Ser459, Gln474, Ala475, Phe486, Phe490, Gln493 and Pro499 form the key binding sites on RBD that are crucial for the affinity of SARS-CoV-2 spike glycoprotein RBD for hACE2 (Yi et al., 2020). Blocking the binding of SARS-CoV-2 spike glycoproteins to hACE2 is a promising approach towards the prevention and/or treatment of COVID-19 (Yi et al., 2020). Besides, viral proteases critical in genomic transcription and replication are attractive targets for drug development against SARS-CoV-2 (Naqvi et al., 2020). Following the release of the coronavirus genome into the host cell, polypeptides are translated and subsequently cleaved by M pro (mediated by catalytic-dyad His41 and Cys145) and PL pro (mediated by catalytic-tryad Cys11, His272 and Asp286) to produce functional proteins for folding and packaging of new virions which are essential for viral replication (Naqvi et al., 2020). Thus, inhibition of M pro and PL pro catalytic activities could potentially suppress SARS-CoV-2 replication and the spread of infection. Notably, there is no cleavage motif characteristic of SARS-CoV-2 proteases that occurs in the human host. Hence M pro and PL pro...
could serve as specific targets in the formulation of inhibition strategies and pharmacological treatments against SARS-CoV-2 (Ortega et al., 2020).

In response to the COVID-19 pandemic, there is a need to develop countermeasures urgently. This has fueled current interest among researchers to adopt molecular docking and other in silico approaches to evaluate the potential of existing antiviral drugs and bioactive natural products in targeting SARS-CoV-2 (Fakhri, 2020). Molecular docking has been widely exploited as structure-based drug discovery tool owing to its quick, low-cost and precise prediction of the biological activity and interactions of ligands with target molecules. With this approach, information on the bond types, binding affinity and interactions between two molecules can be predicted from the best-matching conformation (Tao et al., 2020). To date, molecular docking studies have identified some natural and synthetic inhibitors that could potentially block the binding between SARS-CoV-2 and hACE2 and potentially inhibit the catalytic activities of SARS-CoV-2 MPro and PLpro, as presented in Tables S1–S6. Notwithstanding the foregoing, studies focusing on food-derived and other natural peptides, specifically those that could concurrently target the spike glycoproteins, MPro and PLpro of SARS-CoV-2 are scarce.

Food peptides exhibit diverse bioactivities, including antiviral activity (Agarwal and Gabrani, 2020). The potential of food-derived peptides in the development of nutraceuticals and functional foods is well-recognized (Chai et al., 2017; Wong et al., 2020b). Peptides are also preferred as therapeutic agents due to their higher specificity, efficiency and biodegradability, as well as lower molecular weight and toxicity (Haggag et al., 2018). Considering the aforementioned, in this study, we investigated whether food-derived peptides could serve as multi-target inhibitors against host cell entry and viral replication of SARS-CoV-2. Specifically, we explored the potential release of such multifunctional peptides from the seed storage proteins of Cucurbita maxima L. (Cucurbitaceae) (pumpkin), Chenopodium quinoa Willd. (Amaranthaceae) (quinoa), Brassica napus L. (Brassicaceae) (rape), Helianthus annuus L. (Asteraceae) (sunflower) and Sesamum indicum L. (Pedaliaceae) (sesame) seeds by using in silico gastrointestinal (GI) digestion, followed by molecular docking analysis. Seed storage proteins are the major proteins in the five seeds and are likely to be the main sources of bioactive peptides upon oral consumption and GI digestion. The nutritional values and the bioactivity of peptides derived from the five seeds have been reported (Bao et al., 2020; Guo et al., 2020; Han et al., 2019; Innocent-Ukachi, 2019; Vilacundo et al., 2018; Wang et al., 2020b). Currently, there is still no effective approved treatment or vaccine that can be used to tackle or prevent COVID-19. The outcome of this research could relevant information that could be used by medical professionals for formulating strategy to assist in the treatment and/or management of COVID-19. Importantly, this research could serve as the basis for future studies to elucidate the health-promoting effects of seed peptides upon oral ingestion, especially in relation to coronavirus diseases.

Methods

In silico GI digestion and screening for GI absorption

Nineteen seed storage proteins found in pumpkin, quinoa, rape, sunflower and sesame seeds (Guo et al., 2020; Han et al., 2019), were targeted in this study. Quinoa seed protein sequences were retrieved from NCBI (https://www.ncbi.nlm.nih.gov) (Clark et al., 2015) on 5 August 2020. The protein sequences of the other four seeds were downloaded from the UniProt web server (http://www.uniprot.org) (The UniProt Consortium, 2019) on 5 August 2020. In silico GI digestion of peptides and the screening of GI absorption were performed by using BIOPEP-UWM (http://www.uwm.edu.pl/biochemia/index.php/en/biopep) (Minkiewicz et al., 2019) and SwissADME (http://www.swissadme.ch) (Daina et al., 2017) web servers, respectively, as previously described (Wong et al., 2020a). Besides, phytochemicals used as references in molecular docking studies were also screened for GI absorption, entering their canonical Simplified Molecular-Input Line-Entry System (SMILES) from PubChem (https://pubchem.ncbi.nlm.nih.gov) (Kim et al., 2016) as input to be analyzed by SwissADME. Following screening, peptides and reference phytochemicals predicted to have

Table 1

| Seeds       | Seed storage proteins | Accession               | Number of fragments released | Distribution of peptide length (%) | Number of peptides with high GI absorption |
|-------------|-----------------------|-------------------------|-------------------------------|-----------------------------------|-----------------------------------------|
|             |                       |                         |                               | 2 residues | 3 residues | 4 residues | 5 residues | > 5 residues |                          |                          |
| Pumpkin     | 25 albumin large chain | Q39649                  | 25                            | 20.00      | 20.00      | 8.00       | 16.00      | 4.00        |                          | 2                         |
|             | 11S globulin gamma chain | P13744                  | 95                            | 17.89      | 16.84      | 8.42       | 5.26       | 14.74       |                          | 9                         |
|             | 11S globulin delta chain | P13744                  | 77                            | 20.78      | 16.88      | 12.99      | 7.79       | 2.60        |                          | 10                        |
| Quinoa      | 25 albumin-like        | XP_021758596             | 60                            | 18.33      | 15.00      | 10.00      | 1.67       | 6.67        |                          | 4                         |
|             | 11S seed storage globulin | AA567036                | 177                           | 20.90      | 17.51      | 8.47       | 8.47       | 9.60        |                          | 24                        |
|             | 11S globulin seed storage protein 2-like | XP_021770184 | 185                          | 25.41      | 14.05      | 6.49       | 7.03       | 9.19        |                          | 27                        |
|             | 13S globulin seed storage protein 1-like | XP_021752233 | 175                          | 29.14      | 13.14      | 9.71       | 4.57       | 8.00        |                          | 30                        |
|             | 13S globulin seed storage protein 2-like | XP_021752668 | 204                          | 23.04      | 15.20      | 7.35       | 8.33       | 9.80        |                          | 23                        |
| Rape        | Cruciferin subunit alpha | P11090                  | 87                            | 26.44      | 14.94      | 11.49      | 6.90       | 12.64       |                          | 13                        |
|             | Cruciferin subunit beta | P11090                  | 69                            | 27.54      | 13.04      | 14.49      | 4.35       | 10.14       |                          | 13                        |
|             | Nanpod small chain    | P17333                  | 15                            | 13.33      | 13.33      | 6.67       | 20.00      | 0.00        |                          | 1                         |
|             | Nanpod large chain    | P17333                  | 23                            | 21.74      | 13.04      | 13.04      | 17.39      | 17.39       |                          | 4                         |
| Sunflower   | 25 seed storage protein | P15461                  | 36                            | 25.00      | 11.11      | 8.33       | 11.11      | 22.22       |                          | 6                         |
|             | 11S globulin seed storage protein G3 acidic chain | P19084 | 88                          | 15.91      | 17.05      | 14.77      | 5.68       | 13.64       |                          | 5                         |
|             | 11S globulin seed storage protein G3 basic chain | P19084 | 72                          | 23.61      | 15.28      | 5.56       | 2.78       | 12.50       |                          | 6                         |
| Sesame      | 25 seed storage protein 1 small subunit | Q9XHP1                  | 14                            | 28.57      | 21.43      | 14.29      | 0.00       | 0.00        |                          | 0                         |
|             | 25 seed storage protein 1 large subunit | Q9XHP1                  | 32                            | 34.38      | 12.50      | 15.63      | 3.13       | 0.00        |                          | 3                         |
|             | 11S globulin seed storage protein 2 acidic chain | Q9XHP0 | 87                          | 28.74      | 11.49      | 9.20       | 0.00       | 16.09       |                          | 10                        |
|             | 11S globulin seed storage protein 2 basic chain | Q9XHP0 | 72                          | 22.22      | 11.11      | 11.11      | 13.89      | 2.78        |                          | 9                         |
high GI absorption (Tables S4–S6) were taken for molecular docking analysis.

**Molecular docking analyses with HPEPDock**

The crystal structures of SARS-CoV-2 proteins were obtained from the Protein Data Bank (http://www.rcsb.org/pdb) (Berman et al., 2000; Burley et al., 2018). Three SARS-CoV-2 viral proteins were chosen for this study, namely spike glycoprotein RGB (PDB ID: 6LZG) (Wang et al., 2020a), MPro (PDB ID: 6LU7) (Jin et al., 2020) and PLPro (PDB ID: 6W9C) (Osipiuk et al., 2020). Proteins and ligands were separated and prepared for molecular docking by using BIOVIA Discovery Studio Visualizer (BIOVIA, Dassault Systèmes, BIOVIA Discovery Studio Visualizer, Version 20.1.0.192, San Diego: Dassault Systèmes, 2020), as previously reported (Wong et al., 2020a). Crystallographic water molecules were removed when the target protein structures were prepared for molecular docking analysis. There is no evidence in the literature that such water molecules are part of the viral proteins structure and tightly bound to viral proteins. Hence, maintaining the water molecules in the target protein structures for molecular docking simulation is not justified.

### Table 2

Docking energy scores of high-GI-absorption peptides, types of seeds they were released from, and capability to bind to key binding residues on RBD as well as to key catalytic residues of Mpro and PLpro.

| Peptides | Docking energy scores | Seeds | Binding to key residues<sup>a</sup> |
|----------|------------------------|-------|-----------------------------------|
|          | MPro<sup>b</sup>       | PLPro | Pumpkin | Quinoa | Rape | Sunflower | Sesame |
| PW (Pro-Trp) | −144.652 | −132.325 | −128.975 | ✓ | | | | |
| PSF (Pro-Ser-Phe) | −138.782 | −119.408 | −126.842 | ✓ | ✓ | | | |
| IVF (Ile-Val-Phe) | −130.703 | −109.135 | −111.110 | ✓ | | | | |
| PFG (Pro-Gly-Phe) | −127.857 | −114.371 | −124.263 | ✓ | ✓ | | | |
| PLL (Pro-Ile-Leu) | −127.591 | −103.168 | −102.956 | ✓ | ✓ | ✓ | ✓ |
| PY (Pro-Tyr) | −127.585 | −109.013 | −113.683 | ✓ | | ✓ | ✓ |
| APY (Ala-Pro-Tyr) | −124.257 | −102.990 | −119.964 | ✓ | | | | |
| IV (Ile-Ile) | −123.809 | −91.420 | −105.689 | ✓ | ✓ | ✓ | ✓ |
| VAF (Val-Ala-Phe) | −122.961 | −96.303 | −110.495 | ✓ | | | | |
| PPL (Pro-Pro-Leu) | −116.764 | −102.890 | −108.089 | ✓ | | ✓ | ✓ |
| PH (Pro-His) | −113.926 | −96.694 | −106.229 | ✓ | | | | |
| IAF (Ile-Ala-Phe) | −113.811 | −101.035 | −116.402 | ✓ | | | | |
| PF (Pro-Phe) | −112.988 | −101.396 | −117.215 | ✓ | ✓ | ✓ | ✓ |
| VIL (Val-Ile-Leu) | −112.072 | −79.173 | −92.851 | ✓ | | | | |
| IVL (Ile-Ile-Ala) | −111.314 | −77.986 | −90.589 | ✓ | | | | |
| IF (Ile-Phe) | −109.039 | −91.879 | −102.463 | ✓ | | ✓ | ✓ |
| YY (Val-Tyr) | −107.521 | −93.618 | −106.424 | ✓ | | | | |
| TY (Thr-Tyr) | −106.848 | −97.157 | −102.385 | ✓ | | ✓ | ✓ |
| VVL (Val-Val-Leu) | −102.353 | −75.203 | −81.706 | ✓ | | | | |
| VF (Val-Phe) | −99.947 | −92.526 | −97.546 | ✓ | | | | |
| TF (Thr-Phe) | −98.941 | −86.242 | −101.645 | ✓ | | | | |
| VH (Val-His) | −97.707 | −84.668 | −83.877 | ✓ | | | | |
| PL (Pro-Leu) | −96.892 | −83.195 | −91.722 | ✓ | ✓ | | | |
| VPL (Val-Pro-Leu) | −96.483 | −79.945 | −90.801 | ✓ | | | | |
| VAL (Val-Ala-Leu) | −95.447 | −67.272 | −75.728 | ✓ | | | | |
| AVL (Ala-Val-Leu) | −95.243 | −67.020 | −78.768 | ✓ | | | | |
| PK (Pro-Lys) | −94.781 | −80.196 | −78.120 | ✓ | | ✓ | ✓ |
| AY (Ala-Tyr) | −93.611 | −93.080 | −99.248 | ✓ | | | | |
| PI (Pro-Le) | −92.666 | −78.205 | −73.685 | ✓ | | | | |
| SF (Ser-Phe) | −91.606 | −80.301 | −98.372 | ✓ | | ✓ | ✓ |
| VGL (Val-Gly-Leu) | −90.379 | −71.666 | −73.179 | ✓ | | | | |
| VM (Val-Met) | −89.777 | −70.839 | −79.125 | ✓ | | | | |
| CY (Gly-Tyr) | −88.814 | −95.044 | −100.116 | ✓ | | ✓ | ✓ |
| GVL (Gly-Val-Leu) | −87.755 | −60.485 | −75.297 | ✓ | | ✓ | ✓ |
| AF (Ala-Phe) | −86.025 | −82.597 | −97.583 | ✓ | | | | |
| AIV (Ala-Ile-Val) | −85.988 | −63.712 | −79.137 | ✓ | | | | |
| GH (Gly-His) | −85.738 | −76.448 | −80.350 | ✓ | | | | |
| AAL (Ala-Ala-Leu) | −83.704 | −71.858 | −75.648 | ✓ | | | | |
| EF (Glu-Phe) | −83.207 | −80.150 | −98.684 | ✓ | | | | |
| EL (Ile-Leu) | −82.932 | −65.242 | −70.938 | ✓ | | | | |
| GF (Gly-Phe) | −82.760 | −84.622 | −97.025 | ✓ | | ✓ | ✓ |
| AH (Ala-His) | −81.755 | −79.938 | −86.858 | ✓ | | ✓ | ✓ |
| TL (Thr-Leu) | −78.737 | −64.754 | −68.382 | ✓ | | ✓ | ✓ |
| EL (Glu-Leu) | −78.689 | −52.189 | −71.453 | ✓ | | ✓ | ✓ |
| CL (Cys-Leu) | −74.533 | −61.469 | −59.264 | ✓ | | ✓ | ✓ |
| VL (Val-Leu) | −72.686 | −66.248 | −63.197 | ✓ | | | | |
| AM (Ala-Met) | −69.435 | −64.818 | −71.421 | ✓ | | ✓ | ✓ |
| VK (Val-Lys) | −69.168 | −60.564 | −62.584 | ✓ | | ✓ | ✓ |
| AL (Ala-Leu) | −68.432 | −55.237 | −61.453 | ✓ | | | | |
| SL (Ser-Leu) | −68.152 | −60.409 | −58.303 | ✓ | | | | |
| GM (Gly-Met) | −67.578 | −64.821 | −69.515 | ✓ | | ✓ | ✓ |
| AK (Ala-Lys) | −59.570 | −58.362 | −53.873 | ✓ | | ✓ | ✓ |
| GK (Gly-Lys) | −56.566 | −50.684 | −61.557 | ✓ | | ✓ | ✓ |
| GL (Gly-Leu) | −55.886 | −48.038 | −57.232 | ✓ | | ✓ | ✓ |
| DL (Asp-Leu) | −55.580 | −49.046 | −56.133 | ✓ | | ✓ | ✓ |
| AA (Ala-Ala) | −49.066 | −36.734 | −38.642 | ✓ | | ✓ | ✓ |

<sup>a</sup> Peptides predicted to bind to at least one key binding residue on RBD as well as to at least one key catalytic residue of MPro and PLPro.
To perform molecular docking on the HPEPDOCK server (http://huanglab.phys.hust.edu.cn/hpepdock) (Zhou et al., 2018), the three prepared viral proteins were uploaded as receptor inputs. The prepared hACE2 and N3 were uploaded as binding site references for RBD and M\textsuperscript{Pro} respectively. For PL\textsuperscript{Pro}, three receptor binding residues Cys111, His272 and Asp286 were entered in the docking option for HPEPDOCK analysis (Wlodawer et al., 2020). High-GI-absorption seed peptides released from in silico GI digestion and selected reference peptides that were previously reported to bind to SARS-CoV-2 viral proteins, where 3D structures are not available, were entered in the FASTA format as peptide input. Tables S1–S3 list the reference peptides used and sources of their crystal structures, where available. The peptide model giving the lowest (most negative) docking score as indicated by the HPEPDOCK analysis was chosen to combine with the respective SARS-CoV-2 viral protein to form a docked model. The protein-peptide interactions of the docked model were analyzed by using LigPitPlot + v.2.2 (Laskowski and Swindells, 2011; Wallace et al., 1995). The binding poses of selected peptides on the surface of the viral proteins were visualized by using BIOVIA Discovery Studio Visualizer.

Molecular docking analyses with AutoDock Vina tool in PyRx

Following HPEPDOCK analysis, the high-GI-absorption seed peptide with the lowest docking energy score when docked on three SARS-CoV-2 target proteins was further compared with selected phytochemicals (Tables S4–S6). The reference phytochemicals were chosen from those reported to bind to the three viral proteins and also predicted in this study to have high GI absorption. Molecular docking analysis of the selected seed peptide and reference phytochemicals was performed by using the AutoDock Vina option in PyRx 0.8 (Dallakyan and Olson, 2015; Trott and Olson, 2010) as previously described (Wong et al., 2020a). Docking was carried out for the spike protein RBD, with the grid box set to center x-, y-, z-axis values of −40.34, 30.85, 6.61 with dimension of 35 Å x 50 Å x 30 Å, encompassing 21 binding site residues (Lys417, Gly446, Tyr449, Tyr453, Leu455, Phe456, Tyr473, Ala475, Gly476, Glu484, Phe486, Asn487, Tyr489, Phe490, Gly496, Gly498, Thr500, Asn501, Gly502 and Tyr505) (Wang et al., 2020a) and three other RBD residues (Ser459, Glu474 and Pro499) reported to be important for RBD-hACE2 binding (Yi et al., 2020). For M\textsuperscript{Pro}, grid box was set with center x-, y-, z-axis values of −10.71, 12.41, 68.83 with dimension of 15 Å x 25 Å x 15 Å, surrounding the catalytic-dyad His41 and Cys145 (Ortega et al., 2020). For the docking of PL\textsuperscript{Pro}, the grid box was set with center x-, y-, z-axis values of −37.19, 24.28, 37.99 with each dimension of 20 Å, encircling the catalytic-triad Cys111, His272 and Asp286 (Wlodawer et al., 2020). The resultant binding affinities of the peptide and reference phytochemicals were tabulated.

Data analysis

For binding affinities, data are presented as mean ± standard errors based on analysis in triplicates. Statistical analysis was performed by using the SAS University Edition Software (Version 9.4) (https://www.sas.com/en/us/software/university-edition.html). P-value of less than 0.05 was considered as statistically significant according to Tukey’s test.

Results and discussion

In silico GI digestion generated a total of 1593 peptides fragments from 19 seed storage proteins taken from the five seed samples. The largest number of peptide fragments (801 fragments) were released from quinoa seed storage proteins, whereas the smallest number of peptide fragments (194 fragments) were from rape seed (Table 1). Regardless of seed type, more peptide fragments were released from the globulin-type proteins than from the albumin-type. For example, the 13S and 11S globulins of the quinoa seed collectively released about 93% of all GI-released peptide fragments. Our finding agrees with the observation that most peptide fragments found in the gastrointestinal digest of quinoa proteins were derived from 11S globulin of the seed (Vilcacundo et al., 2018). The distribution of peptide length in the in silico GI digestion-released fragments varied among the five seeds. For each seed storage protein listed in Table 1, the total pool of 2- and 3-residue peptides accounts for about 27–50% of the total number of peptides released. Such short peptides are of great interest as they may be absorbed intact through the epithelium of GI tract, entering the blood stream and producing biological effects at the tissue level (Wang et al., 2019). Thus, our findings suggest the five seeds analyzed in this study are likely to release a significant portion of potentially bioavailable peptides upon GI digestion. Besides resistance to GI digestion, GI absorption is another determinant of the bioavailability of short peptides. As predicted by the SwissADME tool, the greatest number of high-GI-absorption peptides was released from quinoa (108 fragments), in contrast with the smallest number from sunflower seeds. For all five seeds, the globulin-type seed storage proteins are the main contributors of high-GI-absorption peptides (Table 1). Together, our findings suggest that the quinoa seed, in particular its globulin-type storage proteins, is a promising source of potentially bioavailable peptides, which deserves more research attention in future.

The same peptide sequences might be released from different proteins of same seed, or from proteins of different seeds. Hence, the 199

![Fig. 1. Binding poses of five top-score seed peptides on the surface of M\textsuperscript{Pro}: (A) PW, (B) PSF, (C) IVF, (D) PGF, (E) PIIL; on PL\textsuperscript{Pro}: (F) PW, (G) PSF, (H) PGF, (I) IVF, (J) PY; and on spike protein RBD: (K) PW, (L) PSF, (M) PGF, (N) APY, (O) IVF.](image-url)
Table 3  
Interactions of five top-score seed peptides with Mpro, PLpro and spike glycoprotein RBD.

| Target protein | Peptides | Interactions with viral protein residues | Total number of interactions |
|----------------|----------|----------------------------------------|-----------------------------|
|                |          | Hydrogen bond                           | Hydrophobic interaction     |                          |
| M\textsuperscript{pro} | PW      | His41, His164, Gln166                    | 3                           | 13                        |
|                | PSF     | His41, G143, Cys145                     | 3                           | 14                        |
|                | IVF     | Gln166                                  | 2                           | 13                        |
|                | PGF     | Gln166, Gln189                          | 3                           | 13                        |
|                | PL      | His164                                  | 1                           | 14                        |
| PL\textsuperscript{pro} | PW      | Gln269, Gly271                          | 0                           | 8                         |
|                | PSF     | Gln269, Gly271                          | 3                           | 11                        |
|                | IVF     | His272, Asp286                          | 2                           | 9                         |

Spike protein RBD

|                | PW      | Arg403, Tyr453                          | 1                           | 7                         |
|                | PSF     | Arg403, Gln493                          | 2                           | 10                        |
|                | IVF     | Arg403, Gln493                          | 1                           | 10                        |
|                | APY     | Arg403, Gln496, Tyr505                  | 3                           | 9                         |

\* Residues in bold are the catalytic site residues of M\textsuperscript{pro} and PL\textsuperscript{pro}, or key binding sites on the spike protein RBD. The underlined residues are other RBD residues that also bind to hACE2.

High-GI-absorption peptides in Table 1 can be consolidated into a set of 56 unique sequences. Molecular docking analysis by HPEPDock revealed that the docking energy scores of the 56 peptides range from −49.066 to −144.652, −36.734 to −132.325, and −38.642 to −128.975, corresponding to viral proteins M\textsuperscript{pro}, PL\textsuperscript{pro} and spike glycoprotein RBD, respectively (Table 2). In general, lower (more negative) scores suggest more stable binding between a peptide and its target protein (receptor). Based on docking energy scores, the five most promising peptides, in descending order, are PW, PSF, IVF, PGL and PL, when docked with M\textsuperscript{pro}, PW, PSF, PGF and IVF, when docked with M\textsuperscript{pro}; and PW, PSF, PGF, APY and IVF when docked with spike glycoprotein RBD (Table 2). Overall, our analysis implies that all five seeds analyzed in this study could potentially release not only high-GI-absorption peptides, but some such peptides may concurrently target the M\textsuperscript{pro}, PL\textsuperscript{pro} and spike glycoprotein RBD of SARS-CoV-2.

We also analyzed potential anti-SARS-CoV-2 peptides reported in the literature for comparison and found that unlike PW, the bivalve mollusk *Mizuhopecten yessoensis* myosin-derived QRPR (Gln-Arg-Arg-Pro) and LPIY (Leu-Pro-Ile-Tyr) were predicted to be susceptible to GI digestion by BIOPEP-UWM (data not shown). Fragments released from GI-digested QRPR were predicted to have low GI absorption by SwissADME. Only PIY (Pro-Ile-Tyr) released from GI-digested LPIY was predicted to exhibit high GI absorption. The docking energy score computed for PIY was lower than that of PW, suggesting stronger M\textsuperscript{pro}-binding potential of PIY relative to PW (data not shown). Our analysis found that for binding to PL\textsuperscript{pro}, bacterial azurin-derived p18 and p28 peptides had similar or superior docking scores relative to the seed peptides in our study. However, p18 and p28 were predicted to be susceptible to GI degradation and absorbed by the GI tract poorly (data not shown). While *M. yessoensis* myosin-derived CSNAIPEL (Cys-Ser-Asn-Ala-Ile-Pro-Glu-Leu) and LPIY could potentially bind to spike protein RBD, they were predicted to be susceptible to GI digestion. The fragments released from GI-digested CSNAIPEL were still predicted to be poorly absorbed. Although PIY released from GI-digested LPIY potentially has high GI absorption, its docking energy score when bound to RBD was weaker (less negative) than that of PW. Thus, comparison with other natural peptides in the literature highlights the promising potential of seed peptides in this study, particularly when considering their GI resistance and hence potential bioavailability upon oral consumption. In particular, pumpkin seed-derived PW is the most promising peptide in our high-GI-absorption peptide dataset, being a putative multi-target antagonist against SARS-CoV-2 M\textsuperscript{pro}, PL\textsuperscript{pro}, and spike glycoproteins.

Visualization of protein-peptide docking results by using LigPlot+ revealed that molecular interactions with SARS-CoV-2 viral proteins varied among the 56 high-GI-absorption seed peptides (data not shown). Interestingly, 36 of 56 high-GI-absorption peptides analyzed were found to be capable of binding to at least one key binding residue on spike glycoprotein RBD, as well as binding to at least one residue in the catalytic dyad of M\textsuperscript{pro} and in the catalytic triad of PL\textsuperscript{pro}. The 36 peptides include 19 dipeptides and 17 tripeptides (Table 2). The 36 peptides may have the versatility of targeting different phases that are critical to the development of COVID-19 infection. Remarkably, 27 of the 36 multitarget, high-GI-absorption peptides were released from quino seed storage proteins. Thus the quinoa seed could be a very promising source of anti-COVID-19 peptides upon oral consumption and GI digestion.

The binding poses of the five top-score seed peptides on the surfaces of the M\textsuperscript{pro}, PL\textsuperscript{pro} and spike glycoprotein RBD are shown in Fig. 1. Detailed interactions between the three viral proteins and the five top-score seed peptides, as revealed by LigPlot+ analysis, are presented in Table 3 and Fig. S1–S3. Seed peptides PW, PSF, IVF and PGL could interact with both catalytic-dyad residues His41 and Cys145 on M\textsuperscript{pro} either via hydrophobic interactions, hydrogen bond, or both. The same ap-
plies to the PIY fragment released from reference peptide LPIY following GI digestion (data not shown). The seed peptides were predicted to also interact with the M\textsuperscript{pro} active site residues Met\textsubscript{49}, Leu\textsubscript{141}, Asn\textsubscript{142}, His\textsubscript{163}, His\textsubscript{164}, Met\textsubscript{165}, Glu\textsubscript{166}, Asp\textsubscript{187} and Glu\textsubscript{189} (Narkhede et al., 2020); such interactions were also observed between SARS-CoV-2 M\textsuperscript{pro} and its inhibitor N3 (Jin et al., 2020). LigPlot+ analysis on PL\textsuperscript{pro} revealed that PW, PSF, PGF, IVF and PY could interact with catalytic-triad residues Cys\textsubscript{111}, His\textsubscript{272} and Asp\textsubscript{286}. Other than that, the main interactions of peptides with PL\textsuperscript{pro} involved residues Trp\textsubscript{106}, Asn\textsubscript{109}, Leu\textsubscript{162} and Gly\textsubscript{271} (Table 3). The four aforementioned residues are part of the substrate-binding pocket of SARS-CoV-2 PL\textsuperscript{pro} (Ortega et al., 2020). The same interactions with these four residues were also be detected between PL\textsuperscript{pro} and reference peptide p28 (data not shown). Seed peptides PSF, PGF and IVF could each form at least one interaction with Glu\textsubscript{493}, a key binding residue on RBD, besides binding to Tyr\textsubscript{453}, Gly\textsubscript{496} and Tyr\textsubscript{505} that were previously reported to participate in binding to hACE2 (Wang et al., 2020a). PW and APY were also found to interact with Tyr\textsubscript{453}, Gly\textsubscript{496}, Asn\textsubscript{501} and Tyr\textsubscript{505} which participates in binding to hACE2 (Wang et al., 2020a), despite no interactions with any key binding residues on RBD. In short, all the five top-score peptides analyzed in this study were predicted to bind to the catalytic site residues on M\textsuperscript{pro} and PL\textsuperscript{pro}, as well as to the hACE2-binding residues on the spike glycoprotein RBD.

The top-score peptides corresponding to M\textsuperscript{pro}, PL\textsuperscript{pro} and spike glycoprotein RBD, namely PW, PSF, PGF, PIL, PY and APY, comprise mainly hydrophobic residues (Ala, Ile, Leu, Phe, Pro, Trp and Val). At least 50% of the residues in the aforementioned peptides are hydrophobic. Hydrophobic residues also make up of 25–75% of the reference peptides PIY, QPRR, CSNAIPEL, p18 and p28. Notably, Pro is the most common residue among the top-score seed peptides and the reference peptides investigated in this study. Hydrophobic interactions contributed about 80% of the total number of interaction between the seed peptides and the viral proteins (Table 3), suggesting their significance in forming stable protein-peptide complexes. We also observed that hydrophobic interactions account for more than 60% of the total number of protein-peptide interaction between all the reference peptides and the three viral proteins (data not shown).

PW, the multi-target, high-GI-absorption seed peptide with the lowest (most negative) docking energy score as computed by HPEPDock, was also compared with phytochemicals from edible and/or medicinal plants. For fair comparison, we focused on only phytochemicals that were predicted to show high GI absorption. In a preliminary step, we screened 28, 25 and 26 phytochemicals previously reported as M\textsuperscript{pro}, PL\textsuperscript{pro} and spike glycoprotein RBD inhibitors, respectively, on the SwissADME server for high GI absorption. Consequently, 15 high-GI-absorption phytochemicals were chosen for M\textsuperscript{pro} docking analysis, 17

![Fig. 2. Binding affinities of seed peptide PW and selected phytochemicals based on docking analysis on (A) M\textsuperscript{pro}, (B) PL\textsuperscript{pro} and (C) spike protein RBD. Bars represent mean ± standard errors (N = 3). Different lowercase letters on the left of the bars indicate statistically significant difference (P < 0.05) between mean values. Black bars represent PW. The numbers on the vertical axis in (A) represent: 1, Dihomo-γ-linolenic acid; 2, p-Coumaroyletryamine; 3, Lignan; 4, Sugirol; 5, PW; 6, Ursodeoxycholic acid; 7, N-cis-feruloyltyramine; 8, Deserpidine; 9, Quercetin; 10, Neoandrographolid; 11, Rhein; 12, Tanshinone IIA; 13, Berberine; 14, Kaempferol; 15, Cryptotanshinone; 16, Absinthin. The numbers on the vertical axis in (B) represent: 1, N-cis-feruloyltyramine; 2, Magnolol; 3, 14-deoxyandrogapholid; 4, Picatanol; 5, Kaempferol; 6, Chrys; 7, PW; 8, 14-deoxy-11,12-didehydroandrographolide; 9, Quercetin; 10, p-Coumaroyletryamine; 11, Moupinamide 12, Androgapholid; 13, Neoandrographolid; 14, Phaitanthrin D; 15, GRL0617; 16, Cryptotanshinone; 17, Tanshinone IIA; 18, Withanolide A. The numbers on the vertical axis in (C) represent: 1, Geraniol; 2, Anethole; 3, Cinnamaldehyde; 4, Pulegone; 5, 1-4-terpineol; 6, Cinnamyl acetate; 7, Thymol; 8, Carvacrol; 9, Curcumin; 10, Chrysophanol; 11, Neoandrographolid; 12, Quercetin; 13, Apigenin; 14, Kaempferol; 15, PW; 16, Dihydotanshinone I; 17, Fisetin.](image-url)
for PLP^29 and 16 for spike glycoprotein RBD (Tables S4-S6). The binding affinity of PW (−6.9 kcal/mol) when docked on M^P^29 was 28% stronger than dihom-o-linolenic acid, but 26% weaker than absinthin (Fig. 2A). When docked on PLP^29, the binding affinity of PW (−5.8 kcal/mol) was up to 16% stronger than magnolol and N-cis-feruloyltyramine, but 29% weaker than 11 phytochemicals (Fig. 2B). Thus, when compared with other phytochemicals, PW has moderate binding affinities to M^P^29 and PLP^29. Evaluation of binding to the spike protein RBD (Fig. 2C) highlights the superior binding affinity of PW over most of the reference phytochemicals analyzed. Specifically, the binding affinity of PW towards spike protein RBD was comparable to those of kaempferol and apigenin, surpassing 12 other reference phytochemicals. Taken together, seed peptide PW is comparable or superior to some potentially high-GI-absorption phytochemicals previously reported to bind to the M^P^29, PLP^29 and spike protein RBD of SARS-CoV-2. Our findings thus suggest that like the natural products derived from edible and medicinal plants, seed storage peptides could also be a natural resource of SARS-CoV-2 inhibitors for combating COVID-19. Notwithstanding, bioinformatic studies have their limitations. For example, such a theoretical or computational approach cannot reveal the relative quantities of the different types of peptides that could be produced upon in vivo GI digestion. Thus, future in vitro and in vivo evaluations are still necessary to validate our findings, focusing on the seven most promising peptides highlighted in this study.

Conclusion

By combining the methods of in silico GI digestion, GI absorption screening, and molecular docking, this study discovered 36 high-GI-absorption peptides with putative anti-COVID-19 potential from the seed storage proteins of five edible seeds. These multi-target peptides could bind to the key binding or catalytic site residues of SARS-CoV-2 spike glycoproteins, M^P^29 and PLP^29. Such peptides could potentially suppress viral attachment and replication concurrently. Quinoa was the most promising among the five seeds, releasing the largest number of multi-target peptides. Thus, its potential as a medicinal food deserves greater attention when designing nutrition approach for curbing COVID-19. Overall, the outcome of this study lays a foundation for future search of multi-target anti-19-COVID-19 peptides from seed storage proteins.

Declaration of Competing Interest

The authors declare no financial or commercial conflict of interest

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.phytotherapy.2020.100016.

References

Agarwal, G., Gabrenia, R., 2020. Antiviral peptides: identification and validation. Int. J. Peptide Res. Therapeut. doi:10.1007/s10929-020-10072-0.
Bao, X., Ma, S., Fu, Y., Wu, J., Zhang, M., 2020. Sensory and structural characterization of umami peptides derived from sunflower seed. Cyt J Food 18, 485–492.
Berman, H.M., Westbrook, J., Feng, Z., Gililand, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res 28, 235–242.
Burley, S.K., Berman, H.M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., Christie, C., Dalekeng, K., Duarte, J.M., Dutta, S., Feng, Z., Ghosh, S., Goodsell, D.S., Green, R.K., Gurunawardana, Y., Guzzenko, D., Hudson, B.P., Kalvo, T., Liang, Y., Lowe, R., Namkoong, H., Peisach, I., Peråkova, I., Prilc, A., Randles, C., Rose, A., Rose, P., Sala, R., Sekhara, M., Shao, C., Tan, L., Tao, Y.-P., Valsavasta, Y., Voigt, M., Westbrook, J., Woo, J., Yang, H., Young, J., Zhuravleva, M., Zardecki, C., 2018. RCSP protein data bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. Nucleic Acids Res 47, D464–D474.
Chai, T.-T., Law, Y.-C., Wong, F.-C., Kim, S.-K., 2017. Enzyme-assisted discovery of antioxidant peptides from edible marine invertebrates: a review. Mar Drugs 15, 42.