UNIVERSITY OF LEEDS

This is a repository copy of *Identification of angiotensin converting enzyme and dipeptidyl peptidase-IV inhibitory peptides derived from oilseed proteins using two integrated bioinformatic approaches*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/140835/

Version: Accepted Version

Article:
Han, R, Maycock, J, Murray, BS et al. (1 more author) (2019) Identification of angiotensin converting enzyme and dipeptidyl peptidase-IV inhibitory peptides derived from oilseed proteins using two integrated bioinformatic approaches. Food Research International, 115. pp. 283-291. ISSN 0963-9969

https://doi.org/10.1016/j.foodres.2018.12.015

© 2018 Published by Elsevier Ltd. Licensed under the Creative Commons Attribution-Non Commercial No Derivatives 4.0 International License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Identification of Angiotensin Converting Enzyme and Dipeptidyl Peptidase-IV inhibitory peptides derived from oilseed proteins using two integrated bioinformatic approaches

Ruixian Han, Joanne Maycock, Brent S. Murray, Christine Boesch*

School of Food Science and Nutrition, University of Leeds, LS2 9JT, Leeds, UK

* corresponding author: School of Food Science and Nutrition, University of Leeds, Woodhouse Lane, LS2 9JT, Leeds, UK; phone: +44 113 3430268; email: c.bosch@leeds.ac.uk

Running title: Bioactive peptide prediction from oilseed protein

Abbreviations: ACE, Angiotensin-converting enzyme; DPP-IV, and dipeptidyl peptidase-IV

Abstract

Angiotensin-converting enzyme (ACE) and dipeptidyl peptidase-IV (DPP-IV) play critical roles in the development of hypertension and type 2 diabetes, respectively. Inhibiting the ACE and DPP-IV activity using peptides has become part of new therapeutic strategies for supporting medicinal treatment of both diseases. In this study, oilseed proteins, including soybean, flaxseed, rapeseed, sunflower and sesame are evaluated for the possibility of generating ACE and DPP-IV inhibitory peptides using different integrated bioinformatic approaches (UniProt knowledgebase, ProtParam, BLAST, BIOPEP, PeptideRanker, Pepsite2 and ToxinPred), and three bovine proteins (β-lactoglobulin, β-casein and κ-casein) as comparisons. Compared with bovine proteins, the potency indices of ACE and DPP-IV inhibitory peptides, calculated using the BIOPEP database, suggest that oilseed proteins may be considered as good precursors of ACE inhibitory peptides but generate a relative lower yield of DPP-IV inhibitory peptides following subtilisin, pepsin (pH=1.3) or pepsin (pH>2) hydrolysis. Average scores aligned using PeptideRanker confirmed oilseeds proteins as significant potential sources of bioactive peptides: over 105 peptides scored over 0.8. Pepsite2 predicted that these peptides would largely bind via Gln281, His353, Lys511, His513, Tyr520 and Tyr523 of ACE to give the inhibition, while Trp629 is the predominant binding site of peptides in reducing DPP-IV activity. All peptides were capable of inhibiting ACE and DPP-IV whilst 65 of these 105 peptides are not currently recorded in BIOPEP database. In conclusion, our *in silico* study demonstrates that oilseed proteins could be considered as good precursors of ACE and DPP-IV inhibitory peptides as well as so far unexplored peptides that potentially have roles in ACE and DPP-IV inhibition and beyond.

Keywords: Angiotensin-converting enzyme, dipeptidyl peptidase-IV, bioactive peptides, hypertension, diabetes, bioinformatics, *in silico* analysis, oilseed proteins

1. Introduction

Defatted oilseeds, such as flaxseed, rapeseed, sunflower and sesame, are by-products from the food industry and currently used as livestock feed or waste. In the last two decades these under-utilised materials have gained growing interest, due to their high protein content, which in turn makes them a potential low-cost source of bioactive peptides. Peptides derived from some oilseed proteins have demonstrated a wide range of bioactive properties including: antioxidative (He, Girgih, Malomo, Ju, & Aluko, 2013), mineral chelating (Megías et al., 2008), anti-inflammatory (Udenigwe, Lu, Han, Hou, & Aluko, 2009) and cholesterol lowering (Cho, Juillerat, & Lee, 2007) activities. These activities have been linked to beneficial health outcomes and peptides could be applied as value-added components in functional foods, nutraceuticals and pharmaceuticals (Hartmann & Meisel, 2007; Korhonen & Pihlanto, 2003; Li-Chan, 2015; Udenigwe & Aluko, 2012).

Among several health benefits of bioactive peptides, inhibition of Angiotensin-Converting Enzyme (ACE) and Dipeptidyl Peptidase-IV (DPP-IV) have been well documented, as well as their synergistic effects in combination with synthetic drugs (Guang & Phillips, 2009; Marczak et al., 2003; Nongonierma & FitzGerald, 2015; Patil, Mandal, Tomar, & Anand, 2015). In the human body, ACE, is associated with elevated blood pressure, by cleaving a dipeptide (HL) from the decapeptide angiotensin I to form a potent vasoconstrictor, angiotensin II. Further, ACE inhibits and degrades bradykinin, a potent vasodilator (Bénéteau-Burnat & Baudin, 1991). DPP-IV is a metabolic serine peptidase which is widely distributed in almost all human tissues, causing the degradation and inactivation of glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP), both are incretin hormones and responsible for stimulating the secretion of insulin (Lambeir, Durinx, Scharpé, & De Meester, 2003). Therefore, inhibiting ACE and DPP-IV activities have become two major therapeutic targets for the management of hypertension and type 2 diabetes mellitus, directly leading to significant reduction in blood pressure and blood glucose levels (Hansson et al., 1999; Kieffer, McIntosh, & Pederson, 1995), respectively.

Proteomic approaches are widely applied for exploring and evaluating the biological activities of bioactive peptides and are traditionally comprised of several steps: 1) manually quantifying the protein content; 2) extracting the protein isolate; 3) hydrolysing protein with the aid of one or more proteases to release the bioactive peptides; 4) fractionating and purifying the protein hydrolysates; 5) identifying the amino acid sequences of peptides; 6) synthesizing the identified peptides and confirming the bioactive properties (Carrasco-Castilla, Hernández-Álvarez, Jiménez-Martínez, Gutiérrez-López, & Dávila-Ortiz, 2012; Dupont, 2017; Sánchez-Rivera, Martínez-Maqueda, Cruz-Huerta, Miralles, & Recio, 2014). These methods are time-consuming and expensive, and yield low amounts of targeted peptides, which limits their further appraisal with respect to *in vivo* studies.

Complementary to traditional proteomic approaches, database-aided bioinformatic evaluations (*in silico*) are suggested as a potential cost-effective tool to screen and theoretically predict the potency of specific protein sequences as precursors for ACE and DPP-IV inhibitors. The release of peptide fractions can be predicted through simulation of enzymatic hydrolysis of identified protein sequences based on protease cleavage specificities, which allow evaluation of the hydrolysis capability of enzymes and gastrointestinal digestive tolerance of the peptides. Such screening delivers information on the potential production of potent bioactive peptides and can highlight novel bioactive peptides for further chemical or recombinant DNA synthesis (Udenigwe, 2014). Peptide sequences with ACE and DPP-IV inhibiting activities have already been extensively explored and identified in the literature, as well being included in appropriate databases; this therefore greatly improves the accuracy and reliability of *in silico* screening for these particular peptide activities. In addition, a number of *in silico* studies have already been published with the purpose of predicting other biological activities of peptides derived from food material including milk (Vukic et al., 2017), deer skin (Jin, Yan, Yu, & Qi, 2015), rice (Pooja, Rani, & Prakash, 2017), crude barley (Gangopadhyay et al., 2016), green algae Caulerpa (Agirbasli & Cavas, 2017) and cumin (Siow & Gan, 2016). These studies suggest that integrated bioinformatic

evaluations are effective in predicting the peptides released from the parent proteins. However, to the best of our knowledge, the comparison of two *in silico* predicting approaches for potency of precursor proteins generating ACE and DPP-IV inhibitors is missing. Thus, the aims of this study are: (a) to screen peptides released from precursor proteins using BIOPEP and (b) to rank the peptides using PeptideRanker, together with predicting the binding sites of promising peptides to ACE and DPP-IV using Pepsite2.

2. Methods

2.1 Protein sequences

In this study, ten storage proteins from five oilseeds sources, flaxseed (*Linum usitatissimum* Q8LPD4), rapeseed (*Brassica napus* P17333, P11090), sunflower (*Helianthus annuus* P19084, P15461), sesame (*Sesamum indicum* Q9XHP1, Q9XHP0) and soybean (*Glycine max* P04347, P11827, P13916), were selected and assessed, together with three bovine proteins (*BOS Taurus* P02754, P02666, P02666) for comparison, to investigate their potency as precursors of ACE and DPP-IV inhibitory peptides. An overview of the integrated bioinformatic approach is presented in Figure 1.

All sequence information was retrieved from UniProt Knowledgebase (http://www.uniprot.org/) (Table 1). The specific oilseed crops chosen were selected based on the protein levels in their seeds as well as the availability of their amino acid sequences in databases. The bovine caseins β-casein and κ-casein, plus the principal bovine whey protein β-lactoglobulin, are all well-documented as abundant resources of ACE and DPP-IV inhibitory peptides (Maruyama & Suzuki, 1982; Nongonierma & FitzGerald, 2013; Pihlanto-Leppälä, Koskinen, Piilola, Tupasela, & KORHONEN, 2000; Silveira, Martínez-Maqueda, Recio, & Hernández-Ledesma, 2013). ProtParam (https://web.expasy.org/protparam/) was used to count the amino acid percentage in the selected proteins (Gasteiger et al., 2005). In addition, BLAST server (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to evaluate significant similarities (homologies) of the 13 protein sequences (Papadopoulos & Agarwala, 2007).

2.2 *In silico* hydrolysis

The peptide bond cleavage was simulated according to the specificity of the enzymes subtilisin (EC 3.4.21.62), pepsin (pH = 1.3) (EC 3.4.23.1) and pepsin (pH >2) (EC 3.4.23.1) using the BIOPEP 'Enzyme(s) action' tool. The peptide profiles generated were identified based on the information recorded in BIOPEP database (currently 3669 peptides). Two parameters, frequency index (A) and potency index (B) for generating ACE and DPP-IV inhibitory peptides were calculated using Eq (A.1) and Eq (A.2), respectively (Minkiewicz, Dziuba, Iwaniak, Dziuba, & Darewicz, 2008).

2.3 The potential of peptide profiles exerting biological activities

All the peptide profiles generated via *in silico* digestion were evaluated for their likelihood of being bioactive using PeptideRanker. This tool assigns a score to each peptide, within the range 0 (poorest activity) to 1 (most promising activity). The peptides with score > 0.8 were described as 'promising bioactive peptides ' and subsequently subjected to toxicity prediction using ToxinPred (http://crdd.osdd.net/raghava/toxinpred/) and further binding site prediction (Gupta et al., 2013; Mooney, Haslam, Pollastri, & Shields, 2012). In addition, an average score of all the predicted peptides generated from oilseed and bovine proteins was calculated via Eq (A.3) to assess overall protein promise of releasing bioactive peptides.

2.4 Predicting ACE and DPP-IV binding sites within the generated peptides

The interactions between the peptides and the targeted enzyme were predicted using Pepsite2 (http://pepsite2.russelllab.org/) (Trabuco, Lise, Petsalaki, & Russell, 2012). The three-dimensional structures of human DPP-IV (PDB code: 1NU6) and ACE (PDB code: 1O8A) were obtained from Protein Data Bank (PDB) (https://www.rcsb.org/). Colour scales were applied in this study to evaluate the predictions, which are 1) Red colour refers to highly significant; 2) Yellow colour means moderately significant; 3) White colour is considered as no significance (Trabuco et al., 2012). For each peptide,

only the prediction with the lowest p-value was selected. In addition, sequences comprising more than ten amino acid residues were ignored (the maximum length accepted by this database).

2.5 Potential gastrointestinal digestive tolerance of peptides

The bioavailability of peptides *in vivo* is also determined by their survival during digestion. Peptide profiles with a score over 0.8 (aligned by PeptideRanker) were evaluated for their tolerance against the cleavage of pepsin (pH > 2.0, EC 3.4.23.1), trypsin (EC 3.4.21.4) and chymotrypsin (3.4.21.2), using the 'Enzyme(s) action' tool obtained from BIOPEP database (http://www.uwm.edu.pl/biochemia/index.php/en/biopep).

3. Results and Discussion

3.1 Amino acid compositions and homology of the oilseed proteins

All sequences underwent homology analysis using BLAST with the method 'compositional matrix adjustment'. The distributions of alignment scores among selected proteins conducted in pairs are shown in Table 2. Out of 78 pairs, 7 gave high scores (> 200), meaning that these pairs show similar molecular features and therefore the peptides derived from them might also be expected to have similar sequences and biological activities (Pooja et al., 2017). Table 3 summarises the details, including the identities, positives, gaps, and bit scores, of protein sequences with alignment scores over 200. 76% identities and 83% positives were found between the α-chain and α'-chain of β-conglycinin, both from soybean. In addition, 11S globulin (sunflower), cruciferin (rapeseed), 11S globulin (sesame) and glycinin (soybean) displayed similarity. This finding agrees with the previous research on related proteins - for example, Chang and Alli (2012) suggested there are approximately 30% similar amino acid residues between legumin (chickpea) and 12S globulin (oat).

Quantitative structure-activity relationship (QSAR) studies, e.g., Lafarga, O'Connor, and Hayes (2014) have shown how 7 amino acid residues - Gly, Ile, Leu, Phe, Pro, Trp

and Tyr, are present at high frequencies in ACE inhibitory peptides. In addition, Ala, Gly, Pro and Tyr play key roles in the composition of DPP-IV inhibitory peptides (Jin et al., 2015; Lacroix & Li-Chan, 2012). Therefore, protein sequences containing high concentrations of the above amino acid residues are expected to be promising sources of ACE or DPP-IV inhibitory peptides. The number and percentage of specific amino acid residues were counted via ProtParam (Table 4). The α chain and α' chain of β-conglycinin (soybean) demonstrated similar percentages of these specific amino acids. In addition, cruciferin (rapeseed) has a similar amino acid composition to 11S globulin storage protein (sunflower) and 11S Globulin Seed storage protein (sesame), but differences could be found in their Ala, Ile and Pro contents when compared with glycinin (soybean). Other sequences showed major differences when compared with each other.

Obviously, the variations in amino acid content and sequence between oilseed and milk proteins contributes to the differences in bioactive properties of peptides generated.

3.2 Stage I: Screening the ACE and DPP-IV inhibitory peptides

Oilseed and bovine protein sequences were analysed using BIOPEP database. Meanwhile, the bacterial protease subtilisin and the human gastric enzyme pepsin, were selected as enzymes for protein hydrolysis. Udenigwe (2016) suggested that pepsin exerts a narrower specificity when the pH of the medium is below 2. Therefore, two pH conditions, pH =1.3 and pH > 2, were selected for pepsin hydrolysis. Only peptides with short amino acid sequences, such as di- and tri-peptides, contribute to the major part of peptide bioactivity (Iwaniak & Dziuba, 2009), exerting antithrombotic, antiamnestic, antioxidative, hypotensive and ubiquitin-mediated proteolysis (Supplementary Table 1S). However, here only ACE and DPP-IV inhibitory activity were investigated.

Table 5 presents the frequency index of ACE and DPP-IV inhibitory peptides generated from 13 protein sequences. Apart from pepsin (pH = 1.3)-treated napin (rapeseed), pepsin (pH = 1.3)-treated 2S storage protein (sunflower), pepsin (pH = 1.3)-treated 2S

storage protein (sesame) and pepsin (pH > 2)-treated napin (rapeseed), the frequency index of ACE inhibitory peptide was lower than that of DPP-IV inhibitory peptides generated from the same sequences. In addition, pepsin (pH > 2) gave the highest frequency index among the three enzymes, followed by subtilisin, with pepsin at pH 1.3 being lowest. Among all the pepsin (pH > 2)-treated oilseed proteins, napin, derived from rapeseed (A 0.0883), exerted the highest frequency index of ACE inhibition, only κ-casein being higher (A 0.0947) when compared with the three milk proteins. Regarding the frequency index of DPP-IV inhibition, cruciferin (rapeseed) gave the highest value (A 0.1127) amongst the oilseed proteins, but this was lower than for all 3 milk proteins (A between 0.1180 and 0.1518).

High predicted frequency values of ACE and DPP-IV inhibition do not directly translate from the precursor protein to a good source of ACE and DPP-IV inhibitors. The value of IC_{50} of each active peptide should be used to adjust the frequency to get the potency index (μM^{-1}).

Pepsin (pH > 2)-treated napin showed the highest potency index (B 0.00622135 μM^{-1}) of ACE inhibitor amongst all the proteins investigated (Table 6). With regards to DPP-IV inhibition, pepsin (pH > 2)- treated milk proteins gave more promising values than oilseed proteins: 0.00032434μM^{-1} (β-lactoglobulin), 0.00030789μM^{-1} (β-casein) and 0.00026140μM^{-1} (κ-casein), whilst the most promising amongst the oilseed proteins was pepsin (pH > 2)-treated napin (B 0.00023281μM^{-1}). Thus, bovine milk proteins might be a more promising source of DPP-IV inhibitors than oilseed proteins in general. In comparison to animal peptide data, plant protein-derived peptide sequence availability is limited which may have an impact on the outcome of prediction analysis and therefore contribute to underestimation of frequency and potency indices of plant proteins. To be able to predict unrecorded ACE and DPP-IV inhibitor candidate peptides potentially obtainable from oilseed and milk proteins, the PeptideRanker application was used together with Pepsite2 (Stage II - see Figure 1 and below).

The frequency and potency indices among all proteins vary notably, even though sequences possessing significant similarity (aligned score > 200 via BLAST). Lafarga

et al. (2014) also highlighted that the peptides derived from one 'parent' protein might not always be generated from highly similar proteins.

3.3 Stage II: Predicting ACE and DPP-IV inhibitory peptides

Entire peptide profiles from *in silico* hydrolysis are provided with scores using PeptideRanker (Supplementary Table 2S). A threshold of 0.8 was set in order to reduce the number of false positives Mooney et al. (2012) and the resulting numbers of peptides for each source are shown in Figure 2. Unlike the results of Udenigwe, Gong, and Wu (2013), the numbers of bioactive peptides did not always appear to be strongly dependent on the type of enzyme, or correlate with the frequency index calculated using the BIOPEP database. In most cases, pepsin (pH > 2) gave the highest number bioactive peptides, except for colinin (flaxseed) and glycinin (soybean). Subtilisin treatment predicts the same number of bioactive peptides in 2S storage protein (sunflower) as in the α-chain and α'-chain of β-conglycinin (soybean) and β-lactoglobulin (bovine). For the other proteins, pepsin (pH = 1.3) gave the lowest number of bioactive peptides, correlating with the trends in the frequency index of the proteins with different enzymes. In addition, the highest numbers of bioactive peptides were predicted from oilseed proteins compared to milk proteins. However, the total numbers of peptide fragments are remarkably different for each protein sequence. Therefore, the average PeptideRanker score for all the sequences was calculated (see Table 7). Pepsin (pH > 2) gave the highest average score, pepsin (pH = 1.3) the lowest. This tendency is the same as the influence of enzyme on frequency Index of ACE and DPP-IV inhibitor peptides. For the pepsin (pH>2)-treated proteins, colinin exerted the highest average score (0.2678), while the lowest was for κ-casein (0.1972). Thus, the oilseed proteins (0.2103 – 0.2678) might have equal or even better release of bioactive peptides compared to β-lactoglobulin (0.2406), β-casein (0.2260) or κ-casein (0.1972).

Peptides with a score > 0.80 via PeptideRanker suggest high bioactive possibilities. However, their biological activity still needs be explored via Pepsite2. Remembering that this tool ignores peptides with > 10 amino acid residues, 89 oilseed peptides and

16 milk protein peptides were finally investigated (Supplementary Table 3S). (In addition, 10 peptides were removed because their sequences were identical in the oilseed- and milk-derived systems). Studies investigating the binding models of ACE and DPP-IV inhibitors are available but the binding sites for different inhibitors are not always the same (Table 8). Nevertheless, the important amino acids in ACE binding are summarized as Glu162, Gln281, His353, Ala354, His383, Glu384, His387, Glu411, Lys511, Phe512, His513, Val518, Tyr520 and Tyr523 in the study by Ngoh and Gan (2017) whilst the predominant amino acids of DPP-IV binding have been described as Arg125, Glu205, Glu206, Val207, Ser209, Phe357, Arg358, Tyr547, Gly(Trp)629, Ser630, Tyr631, Gly633, Val656, Trp659, Tyr662, Tyr666, Asp708, Asn710, Val711and His740. These are slightly different from the ones summarized by Mudgil, Kamal, Yuen, and Maqsood (2018), who do not mention Arg356, Glu403, Val404 and Tyr585 and who modelled DPP-IV (PDB code: 4A5S) forming complexes with the inhibitor, whereas the one used in our study is human DPP-IV (PDB Code: 1NU6).

Table 9 enumerates the 105 peptides binding to the amino acids presented in Table 8. Gln281, His353, Lys511, His513, Tyr520 and Tyr523 are major binding sites of these peptides predicting high ACE inhibiting activity, whilst only W629 is frequently bound by these peptides to exert the DPP-IV inhibition. PF, TF, VF, SF, PSF, MKF, KF, IPF, IF, HF, CF, NF and PM are considered as promising ACE inhibitory peptides, while MW, AW, WF, AF, MKF, KF, QCAW, HWL, WA, IPF, MAPF, WM, IF, ACQCL, PQNIPPL and VYPF could be considered as promising DPP-IV inhibitory peptides, due to all their predicted binding sites being at the critical amino acid in ACE and DPP-IV. In addition, ACF shows the highest p-value (0.05557) for predicting ACE binding sites. Regarding DPP-IV, the highest p-value is 0.06617, coming from the dipeptide IF. This means all the candidates could be considered to interact with both ACE and DPP-IV. Comparison of the sequences recorded in the BIOPEP databases revealed that out of these, 105 peptides are unrecorded in this database (Supplementary Table 4S). The toxicity of the peptides was analysed using ToxinPred as suggested by Gupta et al. (2013) but no toxic properties could be found.

315 The stability of peptides in the gastrointestinal (GI) tract (gastric phase + intestinal
316 phase) determines their availability and in vivo efficacy (You et al., 2010). In order to
317 predict whether the GI tract could be a barrier for effectivity of oral administration of
318 peptides, we used *in silico* GI digestion (mainly, pepsin (pH>2), trypsin and
319 chymotrypsin) applied to the peptides with high PeptideRanker scores. Out of the 105
320 peptides, only 21 were predicted as stable on exposure to these enzymes during
321 digestion (PG, VCPF, PF, PL, VF, SF, PSF, SPF, CL, VPF, IPF, IF, CG, CY, SPM, CF,
322 PF, CM, PM, VPPF and IPPL). In this case, the most promising peptides exerted a low
323 oral bioavailability, which is similar to the finding of Udenigwe and Fogliano (2017)
324 which is indicating that peptides may need to be protected by appropriate encapsulation
325 techniques as recently suggested (Mohan, Rajendran, He, Bazinet, & Udenigwe, 2015).
326
327 3.4 Limitations
328 The predicted generation of peptides *in silico* is based on the specificity of enzymes
329 and the primary structure of precursor proteins. However, peptide generation during
330 actual enzymatic hydrolysis is influenced by different factors such as accessibility of
331 individual peptide bonds through the target enzyme, the presence of enzyme inhibitors
332 and protein structural features such as tertiary structures that will impact on the
333 hydrolysis outcome (Agyei, Tsopmo, & Udenigwe, 2018; Nishinari, Fang, Guo, &
334 Phillips, 2014). In addition, frequency and potency indices of protein sequences are
335 calculated basing on the current knowledge recorded in BIOPEP databases. Therefore,
336 with more peptide sequences added to databases, changes in the frequency and potency
337 indices will undoubtedly occur (BIOPEP database). Then, the bioinformatic tools
338 involved in this study only consider the amino acid sequences of peptides to predict the
339 toxicity (ToxinPred); to align the score (PeptideRanker) and to predict the binding site
340 of ACE and DPP-IV (Pepsite 2). Furthermore, the binding of promising peptides at the
341 pertinent site of the targeted enzyme is not necessarily correlated with their inhibiting
342 activity, due to the fact that the interaction may not be associated with the targeted
343 activity (Pepsite 2) (Li-Chan, 2015). Finally, total protein isolates contain different

protein sequences, therefore, limiting our considerations to storage proteins is lacking to represent the total biological activity of a protein source.

4. Conclusions

Our results, based on selected amino acid sequences from different protein sources hydrolyzed *in silico* using subtilisin and pepsin, indicate that oilseed proteins may be good sources for bioactive peptides, in particular for ACE inhibitory peptides, compared to bovine milk proteins. Further studies are highly warranted to validate the predictions, in particular to confirm the presence and activity of peptides that are currently not described and to establish their overall relevance for enzyme inhibition and beneficial health properties *in vitro* and *in vivo*.

Declarations of interest

None

References

Agirbasli, Z., & Cavas, L. (2017). In silico evaluation of bioactive peptides from the green algae Caulerpa. *Journal of Applied Phycology, 29* (3), 1635-1646.

Agyei, D., Tsopmo, A., & Udenigwe, C. C. (2018). Bioinformatics and peptidomics approaches to the discovery and analysis of food-derived bioactive peptides. *Analytical and bioanalytical chemistry*, 1-10.

Allen, R., Cohen, E., Haar, R. V., Adams, C., Ma, D., Nessler, C., & Thomas, T. (1987). Sequence and expression of a gene encoding an albumin storage protein in sunflower. *Molecular and General Genetics MGG, 210* (2), 211-218.

Allen, R. D., Nessler, C. L., & Thomas, T. L. (1985). Developmental expression of sunflower 11S storage protein genes. *Plant molecular biology, 5* (3), 165-173.

Bénéteau-Burnat, B., & Baudin, B. (1991). Angiotensin-converting enzyme: clinical applications and laboratory investigations on serum and other biological fluids. *Critical reviews in clinical laboratory sciences, 28* (5-6), 337-356.

Berger, J. P., SinhaRoy, R., Pocai, A., Kelly, T. M., Scapin, G., Gao, Y. D., Pryor, K. A. D., Wu, J. K., Eiermann, G. J., & Xu, S. S. (2018). A comparative study of the binding properties, dipeptidyl peptidase‐4 (DPP‐4) inhibitory activity and glucose‐lowering efficacy of the DPP‐4 inhibitors alogliptin, linagliptin, saxagliptin, sitagliptin and vildagliptin in mice. *Endocrinology, Diabetes & Metabolism, 1* (1), e00002.

Carrasco-Castilla, J., Hernández-Álvarez, A. J., Jiménez-Martínez, C., Gutiérrez-López, G. F., & Dávila-Ortiz, G. (2012). Use of proteomics and peptidomics methods in food bioactive peptide science and engineering. *Food Engineering Reviews, 4* (4), 224-243.

Chang, Y.-W., & Alli, I. (2012). In silico assessment: Suggested homology of chickpea (Cicer arietinum L.) legumin and prediction of ACE-inhibitory peptides from chickpea proteins using BLAST and BIOPEP analyses. *Food Research International, 49* (1), 477-486.

Cho, S.-J., Juillerat, M. A., & Lee, C.-H. (2007). Cholesterol lowering mechanism of soybean protein hydrolysate. *Journal of agricultural and food chemistry, 55* (26), 10599-10604.

Chung, M., Lei, B., & Li-Chan, E. (2005). Isolation and structural characterization of the major protein fraction from NorMan flaxseed (Linum usitatissimum L.). *Food chemistry, 90* (1-2), 271-279.

Dalgleish, D. G. (2011). On the structural models of bovine casein micelles—review and possible improvements. *Soft matter, 7* (6), 2265-2272.

Dupont, D. (2017). Peptidomic as a tool for assessing protein digestion. *Current Opinion in Food Science, 16*, 53-58.

Engel, M., Hoffmann, T., Wagner, L., Wermann, M., Heiser, U., Kiefersauer, R., Huber, R., Bode, W., Demuth, H.-U., & Brandstetter, H. (2003). The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism. *Proceedings of the National Academy of Sciences, 100* (9), 5063-5068.

Ericson, M., Rödin, J., Lenman, M., Glimelius, K., Josefsson, L.-G., & Rask, L.

(1986). Structure of the rapeseed 1.7 S storage protein, napin, and its precursor. *Journal of Biological Chemistry, 261* (31), 14576-14581.

Fujiwara, T., Hirai, M. Y., Chino, M., Komeda, Y., & Naito, S. (1992). Effects of sulfur nutrition on expression of the soybean seed storage protein genes in transgenic petunia. *Plant Physiology, 99* (1), 263-268.

Gangopadhyay, N., Wynne, K., O'Connor, P., Gallagher, E., Brunton, N. P., Rai, D. K., & Hayes, M. (2016). In silico and in vitro analyses of the angiotensin-I converting enzyme inhibitory activity of hydrolysates generated from crude barley (Hordeum vulgare) protein concentrates. *Food chemistry, 203*, 367-374.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook* (pp. 571-607): Springer.

Guang, C., & Phillips, R. D. (2009). Plant food-derived angiotensin I converting enzyme inhibitory peptides. *Journal of Agricultural and Food Chemistry, 57* (12), 5113-5120.

Gueguen, J., Bollecker, S., Schwenke, K. D., & Raab, B. (1990). Effect of succinylation on some physicochemical and functional properties of the 12S storage protein from rapeseed (Brassica napus L.). *Journal of agricultural and food chemistry, 38* (1), 61-69.

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Raghava, G. P., & Consortium, O. S. D. D. (2013). In silico approach for predicting toxicity of peptides and proteins. *PLoS One, 8* (9), e73957.

Hansson, L., Lindholm, L. H., Niskanen, L., Lanke, J., Hedner, T., Niklason, A., Luomanmäki, K., Dahlöf, B., de Faire, U., & Mörlin, C. (1999). Effect of angiotensin-converting-enzyme inhibition compared with conventional therapy on cardiovascular morbidity and mortality in hypertension: the Captopril Prevention Project (CAPPP) randomised trial. *The Lancet, 353* (9153), 611-616.

Hartmann, R., & Meisel, H. (2007). Food-derived peptides with biological activity: from research to food applications. *Current opinion in biotechnology, 18* (2), 163-169.

He, R., Girgih, A. T., Malomo, S. A., Ju, X., & Aluko, R. E. (2013). Antioxidant activities of enzymatic rapeseed protein hydrolysates and the membrane ultrafiltration fractions. *Journal of Functional Foods, 5* (1), 219-227.

Iwaniak, A., & Dziuba, J. (2009). Animal and plant proteins as precursors of peptides with ACE Inhibitory Activity—An in silico strategy of protein evaluation. *Food Technology and Biotechnology, 31* (4), 441.

Jin, Y., Yan, J., Yu, Y., & Qi, Y. (2015). Screening and identification of DPP-IV inhibitory peptides from deer skin hydrolysates by an integrated approach of LC－MS/MS and in silico analysis. *Journal of Functional Foods, 18*, 344-357.

Kieffer, T. J., McIntosh, C., & Pederson, R. A. (1995). Degradation of glucose-dependent insulinotropic polypeptide and truncated glucagon-like peptide 1 in vitro and in vivo by dipeptidyl peptidase IV. *Endocrinology, 136* (8), 3585-3596.

Korhonen, H., & Pihlanto, A. (2003). Food-derived bioactive peptides-opportunities

for designing future foods. *Current pharmaceutical design, 9* (16), 1297-1308.
Lacroix, I. M., & Li-Chan, E. C. (2012). Evaluation of the potential of dietary proteins as precursors of dipeptidyl peptidase (DPP)-IV inhibitors by an in silico approach. *Journal of Functional Foods, 4* (2), 403-422.
Lafarga, T., O'Connor, P., & Hayes, M. (2014). Identification of novel dipeptidyl peptidase-IV and angiotensin-I-converting enzyme inhibitory peptides from meat proteins using in silico analysis. *Peptides, 59*, 53-62.
Lambeir, A.-M., Durinx, C., Scharpé, S., & De Meester, I. (2003). Dipeptidyl-peptidase IV from bench to bedside: an update on structural properties, functions, and clinical aspects of the enzyme DPP IV. *Critical reviews in clinical laboratory sciences, 40* (3), 209-294.
Li-Chan, E. C. (2015). Bioactive peptides and protein hydrolysates: research trends and challenges for application as nutraceuticals and functional food ingredients. *Current Opinion in Food Science, 1*, 28-37.
Madureira, A. R., Pereira, C. I., Gomes, A. M., Pintado, M. E., & Malcata, F. X. (2007). Bovine whey proteins–overview on their main biological properties. *Food Research International, 40* (10), 1197-1211.
Marczak, E. D., Usui, H., Fujita, H., Yang, Y., Yokoo, M., Lipkowski, A. W., & Yoshikawa, M. (2003). New antihypertensive peptides isolated from rapeseed. *Peptides, 24* (6), 791-798.
Maruyama, S., & Suzuki, H. (1982). A peptide inhibitor of angiotensin I converting enzyme in the tryptic hydrolysate of casein. *Agricultural and Biological Chemistry, 46* (5), 1393-1394.
Masuyer, G., Schwager, S. L., Sturrock, E. D., Isaac, R. E., & Acharya, K. R. (2012). Molecular recognition and regulation of human angiotensin-I converting enzyme (ACE) activity by natural inhibitory peptides. *Scientific reports, 2*, 717.
Megías, C., Pedroche, J., Yust, M. M., Girón-Calle, J., Alaiz, M., Millán, F., & Vioque, J. (2008). Production of copper-chelating peptides after hydrolysis of sunflower proteins with pepsin and pancreatin. *LWT-Food Science and Technology, 41* (10), 1973-1977.
Meinke, D., Chen, J., & Beachy, R. (1981). Expression of storage-protein genes during soybean seed development. *Planta, 153* (2), 130-139.
Metzler, W. J., Yanchunas, J., Weigelt, C., Kish, K., Klei, H. E., Xie, D., Zhang, Y., Corbett, M., Tamura, J. K., & He, B. (2008). Involvement of DPP-IV catalytic residues in enzyme–saxagliptin complex formation. *Protein Science, 17* (2), 240-250.
Minkiewicz, P., Dziuba, J., Iwaniak, A., Dziuba, M., & Darewicz, M. (2008). BIOPEP database and other programs for processing bioactive peptide sequences. *Journal of AOAC International, 91* (4), 965-980.
Mohan, A., Rajendran, S. R., He, Q. S., Bazinet, L., & Udenigwe, C. C. (2015). Encapsulation of food protein hydrolysates and peptides: a review. *RSC Advances, 5* (97), 79270-79278.
Mooney, C., Haslam, N. J., Pollastri, G., & Shields, D. C. (2012). Towards the improved discovery and design of functional peptides: common features of

493 diverse classes permit generalized prediction of bioactivity. *PloS one, 7* (10), e45012.
494
495 Mudgil, P., Kamal, H., Yuen, G. C., & Maqsood, S. (2018). Characterization and identification of novel antidiabetic and anti-obesity peptides from camel milk protein hydrolysates. *Food chemistry, 259*, 46-54.
496
497
498 Nabeno, M., Akahoshi, F., Kishida, H., Miyaguchi, I., Tanaka, Y., Ishii, S., & Kadowaki, T. (2013). A comparative study of the binding modes of recently launched dipeptidyl peptidase IV inhibitors in the active site. *Biochemical and biophysical research communications, 434* (2), 191-196.
499
500
501
502 Natesh, R., Schwager, S. L., Sturrock, E. D., & Acharya, K. R. (2003). Crystal structure of the human angiotensin-converting enzyme–lisinopril complex. *Nature, 421* (6922), 551.
503
504
505 Ngoh, Y.-Y., & Gan, C.-Y. (2017). Identification of Pinto bean peptides with inhibitory effects on α-amylase and angiotensin converting enzyme (ACE) activities using an integrated bioinformatics-assisted approach. *Food Chemistry*.
506
507
508
509 Nishinari, K., Fang, Y., Guo, S., & Phillips, G. (2014). Soy proteins: A review on composition, aggregation and emulsification. *Food hydrocolloids, 39*, 301-318.
510
511 Nongonierma, A. B., & FitzGerald, R. J. (2013). Inhibition of dipeptidyl peptidase IV (DPP-IV) by proline containing casein-derived peptides. *Journal of functional foods, 5* (4), 1909-1917.
512
513
514 Nongonierma, A. B., & FitzGerald, R. J. (2015). Investigation of the potential of hemp, pea, rice and soy protein hydrolysates as a source of dipeptidyl peptidase IV (DPP-IV) inhibitory peptides. *Food Digestion: Research and Current Opinion, 6* (1-3), 19-29.
515
516
517
518 Orruno, E., & Morgan, M. (2007). Purification and characterisation of the 7S globulin storage protein from sesame (Sesamum indicum L.). *Food Chemistry, 100* (3), 926-934.
519
520
521 Papadopoulos, J. S., & Agarwala, R. (2007). COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics, 23* (9), 1073-1079.
522
523 Patil, P., Mandal, S., Tomar, S. K., & Anand, S. (2015). Food protein-derived bioactive peptides in management of type 2 diabetes. *European journal of nutrition, 54* (6), 863-880.
524
525
526 Pihlanto-Leppälä, A., Koskinen, P., Piilola, K., Tupasela, T., & KORHONEN, H. (2000). Angiotensin I-converting enzyme inhibitory properties of whey protein digests: concentration and characterization of active peptides. *Journal of Dairy Research, 67* (1), 53-64.
527
528
529
530 Pooja, K., Rani, S., & Prakash, B. (2017). In silico approaches towards the exploration of rice bran proteins-derived angiotensin-I-converting enzyme inhibitory peptides. *International Journal of Food Properties*, 1-14.
531
532
533 Priyanto, A. D., Doerksen, R. J., Chang, C.-I., Sung, W.-C., Widjanarko, S. B., Kusnadi, J., Lin, Y.-C., Wang, T.-C., & Hsu, J.-L. (2015). Screening, discovery, and characterization of angiotensin-I converting enzyme inhibitory peptides derived from proteolytic hydrolysate of bitter melon seed proteins.
534
535
536

Journal of proteomics, 128, 424-435.

Sánchez-Rivera, L., Martínez-Maqueda, D., Cruz-Huerta, E., Miralles, B., & Recio, I. (2014). Peptidomics for discovery, bioavailability and monitoring of dairy bioactive peptides. *Food research international, 63*, 170-181.

Silveira, S. T., Martínez-Maqueda, D., Recio, I., & Hernández-Ledesma, B. (2013). Dipeptidyl peptidase-IV inhibitory peptides generated by tryptic hydrolysis of a whey protein concentrate rich in β-lactoglobulin. *Food chemistry, 141* (2), 1072-1077.

Siow, H.-L., & Gan, C.-Y. (2016). Extraction, identification, and structure–activity relationship of antioxidative and α-amylase inhibitory peptides from cumin seeds (Cuminum cyminum). *Journal of Functional Foods, 22*, 1-12.

Tai, S. S., Wu, L. S., Chen, E. C., & Tzen, J. T. (1999). Molecular cloning of 11S globulin and 2S albumin, the two major seed storage proteins in sesame. *Journal of Agricultural and Food Chemistry, 47* (12), 4932-4938.

Trabuco, L. G., Lise, S., Petsalaki, E., & Russell, R. B. (2012). PepSite: prediction of peptide-binding sites from protein surfaces. *Nucleic acids research, 40* (W1), W423-W427.

Truksa, M., MacKenzie, S. L., & Qiu, X. (2003). Molecular analysis of flax 2S storage protein conlinin and seed specific activity of its promoter. *Plant Physiology and Biochemistry, 41* (2), 141-147.

Udenigwe, C. C. (2014). Bioinformatics approaches, prospects and challenges of food bioactive peptide research. *Trends in Food Science & Technology, 36* (2), 137-143.

Udenigwe, C. C. (2016). Towards rice bran protein utilization: In silico insight on the role of oryzacystatins in biologically-active peptide production. *Food chemistry, 191*, 135-138.

Udenigwe, C. C., & Aluko, R. E. (2012). Food protein-derived bioactive peptides: production, processing, and potential health benefits. *Journal of Food Science, 77* (1).

Udenigwe, C. C., Gong, M., & Wu, S. (2013). In silico analysis of the large and small subunits of cereal RuBisCO as precursors of cryptic bioactive peptides. *Process Biochemistry, 48* (11), 1794-1799.

Udenigwe, C. C., Lu, Y.-L., Han, C.-H., Hou, W.-C., & Aluko, R. E. (2009). Flaxseed protein-derived peptide fractions: Antioxidant properties and inhibition of lipopolysaccharide-induced nitric oxide production in murine macrophages. *Food Chemistry, 116* (1), 277-284.

Vukic, V. R., Vukic, D. V., Milanovic, S. D., Ilicic, M. D., Kanuric, K. G., & Johnson, M. S. (2017). In silico identification of milk antihypertensive di- and tripeptides involved in angiotensin I-converting enzyme inhibitory activity. *Nutrition Research, 46*, 22-30.

Wang, X., Wu, S., Xu, D., Xie, D., & Guo, H. (2011). Inhibitor and substrate binding by angiotensin-converting enzyme: quantum mechanical/molecular mechanical molecular dynamics studies. *Journal of chemical information and modeling, 51* (5), 1074-1082.

581

582

583

584

585

586 **Table 1.** Overview on oilseed and bovine protein sequences used for bioinformatic
587 analyses
588

Protein	UniProtKB Code	Length	References
Flaxseed			
Linin	-	-	(Chung, Lei, & Li-Chan, 2005; Truksa, MacKenzie, & Qiu, 2003)
Conlinin	Q8LPD4	169	
Rapeseed			
Napin	P17333	180	(Ericson et al., 1986; Gueguen, Bollecker, Schwenke, & Raab, 1990)
Cruciferin	P11090	488	
Sunflower			
11S globulin seed storage protein	P19084	493	(R. Allen et al., 1987; R. D. Allen, Nessler, & Thomas, 1985)
2S seed storage protein	P15461	295	
Sesame			
2S seed protein protein	Q9XHP1	148	(Orruno & Morgan, 2007; Tai, Wu, Chen, & Tzen, 1999)
11S globulin seed storage protein	Q9XHP0	459	
Soybean			
Glycinin	P04347	516	(Fujiwara, Hirai, Chino, Komeda, & Naito, 1992; Meinke, Chen, & Beachy, 1981)
β-conglycinin, α' chain	P11827	639	
β-conglycinin, α chain	P13916	605	
Bovine			
β-lactoglobulin	P02754	178	(Dalgleish, 2011; Madureira, Pereira, Gomes, Pintado, & Malcata, 2007)
β-casein	P02666	224	
κ-casein	P02668	190	

589
590
591
592
593

Table 2. Distribution of alignment scores for 13 proteins sequences. (**BLAST**)

Alignment scores	< 40	40 – 50	50 - 80	80 - 200	> 200
Number of groups	65	4	2	0	7

High value of alignment score indicates high homology

Table 3. Summary of protein sequences with alignment scores over 200. (**BLAST**)

	Identities	Positives	Gaps	Scores
P19084 vs P11090	189/494(38%)	265/494(53%)	87/494(17%)	304
P11090 vs Q9XHP0	150/442(34%)	251/442(56%)	35/442(7%)	277
P11090 vs P04347	167/520(32%)	258/520(49%)	116/520(22%)	254
P19084 vs Q9XHP0	161/471(34%)	257/471(54%)	49/471(10%)	275
Q9XHP0 vs P11090	146/442(33%)	246/442(55%)	35/442(7%)	271
P04347 vs Q9XHP0	149/495(30%)	249/495(50%)	83/495(16%)	252
P13916 vs P11827	495/649(76%)	540/649(83%)	54/649(8%)	900

P19084: Sunflower, 11S Globulin seed storage protein; **P11090**: Rapeseed, Cruciferin; **Q9XHP0**: Sesame, 11S Globulin seed protein; **P04347**: Soybean, Glycinin; **P13916**: Soybean, β-conglycinin, α'-chain; **P11827**: Soybean, β-conglycinin, α-chain; **Identities**: same amino acid residues in same position; **Positives**: amino acid residues with the similar molecular characteristics in same position; **Gaps**: Spaces for deletions and insertions; **Scores**: reflects homology.

612 **Table 4.** Number and percentage of amino acid residues frequently discovered in ACE
613 and DPP-IV inhibitory peptides present in oilseed and bovine proteins (**ProtParam**)
614

Protein	Ala	Gly	Ile	Leu	Phe	Pro	Trp	Tyr
Flaxseed								
2S Conlinin	10	19	8	9	7	2	3	3
%	5.90%	11.20%	4.70%	5.30%	4.10%	1.20%	1.80%	1.80%
Rapeseed								
Napin	12	9	6	15	9	15	2	3
%	6.70%	5.00%	3.30%	8.30%	5.00%	8.30%	1.10%	1.70%
Cruciferin	33	47	22	45	22	25	5	10
%	6.80%	9.60%	4.50%	9.20%	4.50%	5.10%	1.00%	2.00%
Sunflower								
11S Globulin Seed Storage Protein	38	35	24	37	25	22	8	5
%	7.70%	7.10%	4.90%	7.50%	5.10%	4.50%	1.60%	1.00%
2S Seed Storage Protein	14	17	16	15	10	13	0	1
%	4.70%	5.80%	5.40%	5.10%	3.40%	4.40%	0%	0.30%
Sesame								
2S Seed Storage Protein	12	6	2	7	6	3	2	4
%	8.10%	4.10%	1.40%	4.70%	4.10%	2.00%	1.40%	2.70%
11S Globulin Seed Storage Protein	34	33	20	35	18	19	4	12
%	7.40%	7.20%	4.40%	7.60%	3.90%	4.10%	0.90%	2.60%
Soybean								
Glycinin	20	41	17	41	18	38	4	15
%	3.90%	7.90%	3.30%	7.90%	3.50%	7.40%	0.80%	2.90%
β-conglycinin, α' chain	28	31	29	49	32	35	3	15
%	4.40%	4.90%	4.50%	7.70%	5.00%	5.50%	0.50%	2.30%
β-conglycinin, αchain	28	26	31	54	30	40	2	15
%	4.60%	4.30%	5.10%	8.90%	5.00%	6.60%	0.30%	2.50%
Bovine								
β-lactoglobulin	19	5	10	27	4	8	2	4
%	10.70%	2.80%	5.60%	15.20%	2.20%	4.50%	1.10%	2.20%
β-casein	9	5	11	27	9	35	1	4
%	4.00%	2.20%	4.90%	12.10%	4.00%	15.60%	0.40%	1.80%
κ-casein	16	3	13	13	7	21	1	9
%	8.40%	1.60%	6.80%	6.80%	3.70%	11.10%	0.50%	4.70%

615
616
617
618

Table 5. Frequency indices of ACE and DPP-IV inhibitory peptides generated *in silico* from oilseed and bovine proteins using enzymatic hydrolysis with subtilisin and pepsin (pH 1.3 and pH>2) (**BIOPEP**)

	Subtilisin		Pepsin (pH 1.3)		Pepsin (pH>2)	
	ACE inhibitor	DPP-IV inhibitor	ACE inhibitor	DPP-IV inhibitor	ACE inhibitor	DPP-IV inhibitor
Flaxseed						
2S Conlinin	0.0296	0.0355	n/a	n/a	0.0533	0.1006
Rapeseed						
Napin	0.0389	0.0444	0.0111	0.0056	0.0883	0.0883
Curciferin	0.0246	0.0574	0.0164	0.0184	0.084	0.1127
Sunflower						
2S Seed Storage Protein	0.0102	0.0271	0.0068	0.0068	0.0644	0.0949
11S Globulin Seed Storage Protein	0.0243	0.0446	0.0101	0.0142	0.0649	0.0852
Sesame						
2S Seed Storage Protein	n/a	0.0405	0.0068	n/a	0.0743	0.0946
11S Globulin Seed Storage Protein	0.0194	0.0367	0.0043	0.013	0.0475	0.0907
Soybean						
Glycinin	0.031	0.0581	0.0078	0.0116	0.0523	0.0969
β-conglycinin, α' chain	0.0203	0.0423	0.0063	0.0125	0.0673	0.1095
β-conglycinin, α chain	0.0198	0.0347	0.0099	0.0165	0.0793	0.1124
Bovine						
β-lactoglobulin	0.0169	0.0337	0.0056	0.0225	0.0562	0.118
β-casein	0.0268	0.0938	0.0268	0.0357	0.067	0.1518
κ-casein	0.0158	0.0789	0.0053	0.0211	0.0947	0.1421

625 **Table 6.** Potency indices (μM^{-1}) of ACE and DPP-IV inhibitory peptides generated from
626 oilseeds and bovine proteins (**BIOPEP)**
627

	Subtilisin		**Pepsin (pH 1.3)**		**Pepsin (pH>2)**	
	ACE inhibitor	DPP-IV inhibitor	ACE inhibitor	DPP-IV inhibitor	ACE inhibitor	DPP-IV inhibitor
Flaxseed						
2S Conlinin	0.0014197	0.0000060	n/a	n/a	0.0019097	0.0000375
Rapeseed						
Napin	0.0030663	0.0000009	0.0000445	n/a	0.0062213	0.0002050
Curciferin	0.0039421	0.0001985	0.0002079	0.0000031	0.0038966	0.0002328
Sunflower						
2S Seed Storage Protein	0.0004810	0.0000497	0.0001125	0.0000038	0.0010453	0.0000876
11S Globulin Seed Storage Protein	0.0015792	0.0000274	0.0002115	n/a	0.0016763	0.0000688
Sesame						
2S Seed Storage Protein	n/a	0.0001826	0.0009543	n/a	0.000873	0.0002536
11S Globulin Seed Storage Protein	0.0009832	0.0000153	0.0000940	0.0000050	0.0016192	0.0000818
Soybean						
Glycinin	0.0029164	0.0001214	0.0003297	0.0000015	0.0035738	0.0001779
β-conglycinin, α' chain	0.0004772	0.0000653	0.0000549	n/a	0.0015270	0.0001141
β-conglycinin, α chain	0.0002859	0.0000453	0.0002303	0.0000013	0.0016910	0.0000831
Bovine						
β-lactoglobulin	0.0014124	0.0001646	0.0000431	0.0000887	0.0016405	0.0003243
β-casein	0.0007556	0.0002810	0.0006728	0.0000410	0.0010948	0.0003079
κ-casein	0.0006648	0.0002997	0.0000404	0.0000060	0.0046858	0.0002614

628

Table 7. Average scores of fragments released from oilseeds and bovine proteins (**PeptideRanker**)

Protein sequences	Average score		
	Subtilisin	Pepsin (pH 1.3)	Pepsin (pH > 2)
Flaxseed - Colinin	0.0936	0.0498	0.2678
Rapeseed - Napin	0.1628	0.0731	0.2476
Rapeseed - Cruciferin	0.1095	0.0622	0.2553
Sunflower - 2S seed storage protein	0.0520	0.0393	0.2053
Sunflower - 11S globulin seed storage protein G3	0.0946	0.0582	0.2538
Sesame - 2S seed storage protein	0.0771	0.0391	0.2673
Sasame - 11S globulin seed stroage protein	0.0902	0.0480	0.2381
Soybean - Glycinin	0.0981	0.0515	0.2299
Soybean -β-conglycinin, α'-chain	0.0940	0.0650	0.2103
Soybean -β-conglycinin, α-chain	0.1015	0.0748	0.2137
Bovine -β-lactoglobulin	0.1068	0.0699	0.2406
Bovine -β-casein	0.1236	0.0701	0.2260
Bovine -κ-casein	0.0975	0.0581	0.1972

637 **Table 8.** Summary of important active sites or binding sites in ACE and DPP-IV *(Homo*
638 *sapiens*)
639

Important sites	References
Angiotensin-converting enzyme	
Binding sites of lisinopril in ACE: Glu384, Val518, Glu162, Lys511 and Tyr520	(Natesh, Schwager, Sturrock, & Acharya, 2003) (Priyanto et al., 2015)
Active sites: Glu384, Ala354, Glu162, His353,	
Active sites of ACE binding by Lisinopril and Enalaprilat: Glu162, His353, Ala354, Glu384, His387, Glu411, His383, Tyr523, Tyr520, Lys511	(Wang, Wu, Xu, Xie, & Guo, 2011)
Important binding sites of two natural ACE inhibitory peptides: Gln281, Tyr520, Lys511, Tyr523, His353, Ala354, His513, His353, Ala354, Phe512, Glu384	(Masuyer, Schwager, Sturrock, Isaac, & Acharya, 2012)
Dipeptidyl-peptidase IV	
Active sites: Ser630, Asp708, His740, Gly629, Gly633, Tyr631, Glu205 and Glu206	(Lambeir et al., 2003)
Important binding sites of commercial DPP-IV inhibitors: Ser630, Tyr666, Tyr547 (vildapliptin and saxagliptin); Tyr547 and Trp629 (alogliptin and linagliptin); Asn710 (sitagliptin and teneligliptin); Glu205 and Glu206 play a key role in DPP-IV inhibiting activities for all the DPP-IV inhibitor **Active sites:** Val207, Ser209, Phe357, Arg358	(Nabeno et al., 2013)
Binding sites of saxagliptin in DPP-IV: Val711, Val656, Tyr662, Tyr666, Trp659, Tyr547, Asn710, Glu205, Glu206, Tyr 547and Arg125	(Metzler et al., 2008)
DPP-IV inhibitors binding sites: Ser630, Glu205, Glu206, Arg125, Phe357, Tyr 547, Arg125, Trp629	(Berger et al., 2018)
Active sites: Ser630, His740, Asp708; Tyr547; Tyr666; Tyr662; Val711; Val656; Trp659; Arg125; Asn710; Glu205; Glu206 and Arg358	(Engel et al., 2003)

640
641
642
643

644

645 **Table 9.** The frequency of 105 peptides binding to the amino acids detained in Table 8
646 (**Pepsite2**)
647

Binding sites in ACE	Number of peptides	Binding sites in DPP-IV	Number of peptides
E(Glu)162	0	R(Arg)125	0
Q(Gln)281	103	E(Glu)205	3
H(His)353	102	E(Glu)206	5
A(Ala)354	0	V(Val)207	0
H(His)383	57	S(Ser)209	0
E(Glu)384	36	F(Phe)357	12
H(His)387	27	R(Arg)358	0
E(Glu)411	54	Y(Tyr)547	33
K(Lys)511	86	W(Trp)629	93
F(Phe)512	0	S(Ser)630	47
H(His)513	102	Y(Tyr)631	9
V(Val)518	0	G(Gly)633	0
Y(Tyr)520	104	V(Val)656	2
Y(Tyr)523	102	W(Trp)659	0
		Y(Tyr)662	7
		Y(Tyr)666	25
		D(Asp)708	0
		N(Asn)710	2
		V(Val)711	2
		H(His)740	5

648
649

List of figures

Figure 1. Bioinformatic methodology applied for screening and predicting ACE and DPP-IV inhibitory peptides from oilseed and bovine proteins.

Figure 2. Peptide profiles generated from *in silico* hydrolysis of thirteen proteins which demonstrated scores over 0.8

List of Equations

$$A = \frac{a}{N} \quad \textbf{Eq (A.1)}$$

A: the frequency index of targeted peptides released from precursor proteins
a: the number of targeted peptides
N: the total number of amino acid residues

$$\mathbf{B} = \frac{\sum_{i=1}^{K} \frac{a_i}{IC_{50}}}{N} \quad \textbf{Eq (A.2)}$$

B: the potency index of targeted biological activity
a_i: the number of repetitions of peptides with same amino acid residues released from precursor proteins
IC_{50}: the concentration of bioactive peptides needed to inhibit half-maximal enzyme activity
K: the number of different peptides with targeted biological activities
N: the total number of amino acid residues

$$Average\ score = \frac{\sum_{i=1}^{k} b_i}{N} \quad \textbf{Eq (A.3)}$$

b_i: the score of peptides
k: the number of peptides released from precursor protein
N: the total number of amino acid residues