In silico proteolysis and analysis of bioactive peptides from sequences of fatty acid desaturase 3 (FAD3) of flaxseed protein

Sapna Langyan, Fatima Nazish Khan, Pranjal Yadav, Alaa Alhazmi, Samy F. Mahmoud, Dalia I. Saleh, Ali Tan Kee Zuan, Ashok Kumar

Division of Germplasm Evaluation, ICAR-National Bureau for Plant Genetic Resources, Pusa, New Delhi, India
Department of Computer Science, Jamia Millia Islamia, New Delhi 110025, India
Division of Plant Physiology, Indian Agricultural Research Institute, Pusa, New Delhi 110012, India
Medical Laboratory Technology Department, Jazan University, Jazan, Saudi Arabia
SMIBES for Consultation in Specialized Medical Laboratories, Jazan University, Jazan, Saudi Arabia
Department of Biotechnology, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
Department of Chemistry, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia
Department of Land Management, Faculty of Agriculture, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

Abstract

Flaxseed (Linum usitatissimum), commonly known as linseed is an oilseed crop, emerging as an important and functional ingredient of food and has been paid more attention due to its nutritional value as well as beneficial effects. It is mainly rich in is α-linolenic acid (ALA, omega-3 fatty acid), fibres and lignans that have potential health benefits in reducing cardiovascular diseases, diabetes, osteoporosis, atherosclerosis, cancer, arthritis, neurological and autoimmune disorders. Due to its richness in omega-3 fatty acid, a group of enzymes known as fatty acid desaturases (FADS) mainly introduce double bonds into fatty acids’ (FAs) hydrocarbon chains that produce unsaturated fatty acids. Fatty acid desaturase 3 (FAD3), the commonest microsomal enzyme of omega-3 fatty acid, synthesizes linolenic acid (C18:3) from linoleic acid located in endoplasmic reticulum (ER) facing towards the cytosol. The emerging field of bioinformatics and large number of databases of bioactive peptides, helps in providing time-saving and efficient method for identification of potential bioactivities of any protein. In this study, 10 unique sequences of FAD3 from flaxseed protein have been used for in silico proteolysis and releasing of various bioactive peptides using three plant proteases, namely ficin, papain and stem bromelain, that are evaluated with the help of BIOPEP database. Overall, 20 biological activities were identified from these proteins. The results showed that FAD3 protein is a potential source of peptides with angiotensin-I-converting enzyme (ACE) inhibitory and dipeptidyl peptidase-IV (DPP-IV) activities, and also various parameters such as $P_A$, $P_B$, AE, W, $B_E$, V and DHt were also calculated. Furthermore, PeptideRanker have been used for screening of novel promising bioactive peptides. Various bioinformatics tools also used to study protein’s physicochemical properties, peptide’s score, toxicity, allergenicity aggregation, water solubility, and drug likeness. The present work suggests that flaxseed protein can be a good source of bioactive peptides for the synthesis of good quality and quantity of oil, and in silico method helps in investigating and production of functional peptides.

1. Introduction

Bioactive peptides majorly derived from food have attracted and increased the interest of researchers from food and health sectors. Currently, there is a large interest in food-derived proteins and their diverse biological activities like angiotensin-I-converting enzyme (ACE) inhibitory, dipeptidyl peptidase-IV (DPP-IV) inhibitory, antibacterial, antioxidative, anticancer, regulating, stimulating, immunomodulating activities, etc. Large
number of studies have been focused on utilizing of food protein (as raw materials) to produce bioactive peptides (Chiribi et al., 2015; Nonongierma et al., 2017; Uraipong and Zhao, 2018; Guo et al., 2020). Amongst researches that have been done previously, the protein digestion is the major challenge for releasing of bioactive peptides, and enzymatic hydrolysis is the most effectively and commonly used method. For the discovery and development of novel bioactive peptides conventional method has been used that includes in vitro/in vivo enzymolysis and also various steps for the separation, purification, and identification of bioactive peptides. However, the conventional method of production of peptides from food is challenging due to time consuming and laborious task and also costly for industrial-scale production (Udenigwe, 2014; Agyey et al., 2016, 2018). With the advancement of bioinformatics and large number of databases of bioactive peptides, in silico approaches that are time-saving and more efficient than conventional methods, have overcome these challenges and been widely utilized for investigation of proteins and peptides bioactive features and also for the production of known and novel sequences of peptides from food proteins (Das et al., 2020). This method is used for investigation of bioactive peptides present in protein sequences and also gave the knowledge of specificities and types of proteases that are used to release sequences of bioactive peptides. For in silico prediction of biological activities of any protein sequence and investigation of release of bioactive peptides using specific proteases, BIOPEP database can be used for initial mining of bioactive peptides, that provided large collection of sequences including bioactive peptides, proteins, sensory peptides and allergenic proteins (Minkiewicz et al., 2008). This database has been previously used for identification and analysis of bioactive peptides from number of sources such as products from plants, animals and seafood like porcine myofibrillar proteins (Keska and Stadnik, 1994). Similarly, the information of undesirable biological features like toxicity and allergenicity are found in endoplasmic reticulum (ER) and plastid. Fatty acid desaturases that produce unsaturated fatty acids. As compared to oil from other sources like corn, soybean, fish oil or marine algae, the oil from flaxseed is rich in n-3 fatty acids. Moreover, it is also rich in mucilage and fibres that help in proper functioning of intestines, and nowadays number of functional products from flaxseed such as mucilage and oils that contain many bioactive substances are available commercially (Oomah, 2003). Therefore, various studies have been reported on peptides that are derived from oilseed proteins have number of bioactive properties such as anti-inflammatory, antioxidative, cholesterol lowering and mineral chelating activities (Wu et al., 2019), however, the industrialization of flaxseed protein and its peptides are yet not achieved due to limiting factor of scale-up production and real-time evaluation of bioactivity (Wu et al., 2019). The aim of the present work is in silico proteolysis and releasing of various bioactive peptides using three plant proteases, namely ficin, papain and stem bromelain from 10 unique sequences of FAD3 from flaxseed protein evaluated with the help of BIOPEP database (Logarušic et al., 2020). Furthermore, PepptideRanker have been used for screening of novel promising bioactive peptides having suitable peptide score. Also, for the discovery and development of novel sequences of bioactive peptides predicting the potency and potential biological properties is important and hence, various parameters such as $\sum A$, $\sum B$, $A_g$, $W$, $V$, and $D$H and $D$ were also calculated. For pharmaceutical and food applications, the information of undesirable biological features like toxicity and allergenicity is very useful, and helps the food technologist to limit the yield of toxic or allergenic peptides. Therefore, various bioinformatics tools also used to study protein's physicochemical properties, peptide's score, toxicity, allergenicity, aggregation, water solubility and drug likeliness. The present work suggests that linseed protein can be a good source of bioactive peptides, and in silico method helps in investigating and production of functional peptides.

2. Materials and methods

The overall methodology is described in the form of flowchart (Fig. 2).

2.1. Selection of protein sequences and enzymes for proteolysis

Ten unique and full-length sequences of FAD3 from flaxseed protein having accession number QT18955, AFK27245, AFK27244, AFK27239, AFJ53097, AFJ53096, AFJ53090, ADV92275,
AFN53698 and ADV92271 were taken from NCBI (https://www.ncbi.nlm.nih.gov/) and UniProt (https://www.uniprot.org/) database for in silico analysis (SS1). Three plant proteases namely ficin (EC 3.4.22.3), papain (EC 3.4.22.2) and stem bromelain (EC 3.4.22.32) from BIOPEP-UWM database (http://www.uwm.edu.pl/biochemia/index.php/pl/biopep) were used for in silico proteolysis and releasing of various bioactive peptides.

2.2. Identification of physiochemical properties of FAD3 sequences of flaxseed protein

With the utilization of online tool called ExPASy's ProtParam (https://web.expasy.org/protparam/), large number of physiochemical properties of the all the FAD3 sequences of flaxseed protein such as molecular weight, theoretical isoelectric point (pI), total number of negatively charged (Asp + Glu) and positively charged (Arg + Lys) residues, amino acid and atomic composition, extinction coefficient, estimated half-life, Grand average of hydropathicity (GRAVY), instability and aliphatic index have been identified (SS2).

2.3. Flaxseed proteins evaluation as bioactive peptides precursor through BIOPEP database

By using the tool "Profiles of potential biological activity" from BIOPEP-UWM database, profiles for flaxseed proteins as bioactive peptides precursor have been extracted that gave the location and type of bioactive fragment present in protein sequences (SS3A for ficin, SS3B for papain, and SS3C for stem bromelain). Also,
the parameter known as frequency of peptides occurrence in a protein \((A)\) has been taken and calculated on the basis of equation:

\[
A = \frac{a}{N}
\]  

(1)

where \(a\) denotes the number of peptides in sequences of protein, \(N\) is the number of residues of amino acid present in the protein sequences. From this, total frequency of all bioactive peptide’s occurrence in the protein \((\sum A)\) was also calculated for all the proteases used (SS4).

And also, the parameter \(B\), estimates the potential of a protein to have a particular biological activity.

\[
B = \frac{\sum_{i} a_{i} k_{i}}{N}
\]  

(2)

where \(a_{i}\) denotes the number of repetitions in the protein sequence of the \(i\)th bioactive property of interest, \(k_{i}\) is the \(i\)th bioactive property concentration of interest corresponding to its half-maximal activity in micromoles per litre, \(k\) is the number of different fragments having bioactive property of interest, and \(N\) is number of residues of amino acid in protein sequences. From this, total frequency of all bioactive peptide’s occurrence in the protein \((\sum B)\) was also calculated for all the proteases used (SS4). We also calculated the sum of monoisotopic (SS5A-C, for ficin, papain and stem bromelain, respectively) and chemical mass (SS6A-C, for ficin, papain and stem bromelain, respectively) of all the bioactive peptides release from FAD3 sequences of flaxseed protein.

2.4. In silico proteolysis and screening of protein sequences

BIOPEP-UWM database has been utilized for in silico proteolysis and three plant proteases namely ficin, papain and stem bromelain were independently used for each flaxseed protein sequences for releasing of various bioactive peptides (Fig. 3). And hence, the frequency and relative frequency of releasing of peptides by specific protease \((AE\) and \(W\), respectively) were calculated based on equations:

\[
AE = \frac{d}{N}
\]  

(3)

and

\[
W = \frac{AE}{A}
\]  

(4)

where \(d\) denotes the number of peptides releasing from the sequences of protein by specific protease and \(N\) is the number of residues of amino acid present in the protein sequences.

Also, the parameters \(B_{E}\) and \(V\) were also calculated (SS8). For the estimation of percent degree of hydrolysis of in silico digestion of peptides, the theoretical degree of hydrolysis \((DHT)\) is generally used by the formula:

\[
DHT = \frac{d}{D} \times 100\%
\]  

(5)

where \(d\) represents the number of peptide bonds that are hydrolyzed and \(D\) is the total number of peptide bonds present in protein’s primary sequence.

Then, the released peptides having only three amino acids were taken to calculate the theoretical bioactivity of peptides using online tool PeptideRanker (http://distildeep.ucd.ie/PeptideRanker/). The results of this ranking were evaluated as score in the form of values ranges between 0 and 1 (poorest and best bioactivity, respectively) (Mooney et al., 2012).

2.5. Toxicity and allergenicity prediction of potential tripeptides

The number of bioactive peptides differentiate with different bioactivities. After the results of in silico hydrolysis the tripeptides having high score of PeptideRanker (i.e., over 0.6) were then evaluated for their toxicity using ToxinPred (http://crdd.osdd.net/raghava/toxinpred/) (Gupta et al., 2013), and also the allergenicity of the peptides can be evaluated by AllergenFP (http://ddg-pharma-fac.net/AllergenFP/) (Dimitrov et al., 2014).

2.6. Aggregation and water solubility prediction of potential tripeptides

Aggregation, one of the universal and mostly used terms for proteins as well as peptides, which generally limits the pharmaceutical and biotechnological applications (Knowles et al., 2007). For this online software such as AGGRESCAN (http://bioinf.uab.es/aggrescan/), has been widely used for the prediction of protein...
and/or peptide aggregation. And also, the water solubility (that mainly influences the peptides absorption, distribution and elimination from the body) with the help of online server, InnovaGen (http://www.innovagen.com/proteomics-tools) that utilizing tool PepCalc (http://pepcalc.com/).

2.7. Drug-likeness evaluation using in silico method

With the help of tool known as SwissADME (http://www.swissadme.ch/index.php#), in silico evaluation of drug-likeness has been done, that mainly uses the properties of ADME (absorption, distribution, metabolism, and excretion) of a compound that estimates and indicates the pharmacokinetics of drugs (Mbarik et al., 2019).

3. Results

3.1. Protein sequence and enzymes selection

Firstly, with the utilization of NCBI and UniProt, ten unique and full-length sequences of FAD3 from flaxseed protein were selected. Three plant proteases namely ficin (EC 3.4.22.3), papain (EC 3.4.22.2) and stem bromelain (EC 3.4.22.32) that were used for in silico proteolysis also selected from BIOPEP-UWM database.

3.2. Physicochemical properties of FAD3 sequences of flaxseed protein

With the utilization of online tool called ExPASY’s ProtParam, large number of physicochemical properties of the all the FAD3 sequences of flaxseed protein. The calculations of different physicochemical parameters of FAD3 sequences with ExPASY’s ProtParam tool are shown (Table 1). The number of amino acid ranges from 391 to 901 in all the 10 FAD3 sequences. The other parameters such as pl value was greater than 7 in all the protein sequences (i.e., basic protein). It was found that the value of instability index was less than 40 (except in AFN53698), and the value of aliphatic index was found to be more than 75 that is extremely high in all the sequences. The value of GRAVY was lies between −2 and +2, suggest that the protein sequences are hydrophobic in nature and rated positively.

3.3. The potential of FAD3 sequences from flaxseed protein as bioactive peptides precursors

For the investigation of potential of flaxseed FAD3 protein sequences as bioactive peptides precursors, a total of ten FAD3 sequences from flaxseed protein in the range of 391 to 901 amino acids were taken and evaluated by “Profiles of potential biological activity” tool available in BIOPEP database. On the basis of information present in BIOPEP database, 20 known bioactivities were found in flaxseed proteins (Fig. 1). Among them, ACE inhibitor, activating ubiquitin-mediated proteolysis, alpha-glucosidase inhibitor, antiinflammatory, antioxidant, antiinflammatory, CaM/PDE inhibitor, dipeptidyl peptide inhibitor III, dipeptidyl peptidase IV inhibitor, immunomodulating, neuropeptide, regulating, renin inhibitor, and stimulating activities presented in all FAD3 sequences after analysing them. The occurrence of total frequency of bioactive peptides was highest in dipeptidyl peptidase IV inhibitor digested with all three plant proteases i.e., ∑A = 7.0007, followed by ACE inhibitor with ∑A = 4.6472, and so on, but the value of ∑B was higher in ACE inhibitor (Table 2).

The calculated sum of monoisotopic mass was found to be highest in DPP-IV inhibitors digested with stem bromelain, for all the FAD3 sequences of flaxseed protein (Fig. 4) and interestingly the similar results have been found for sum of chemical mass for all the sequences taken (Fig. 5).

3.4. In silico proteolysis and screening of flaxseed proteins

Bioactive peptides derived from natural food protein are releasing by the process of enzymolysis for their functions. From these natural food sources number of bioactive peptides were previously generated by using different enzymes (Fu et al., 2016; Lin et al., 2018). In our study, three plant proteases namely ficin, papain and stem bromelain were chosen and applied to the FAD3 sequences of flaxseed protein using tool “Enzyme(s) action” present in BIOPEP database. Hydrolysates having the degree of hydrolysate (DHT) between 35.0384% and 56.1111% were obtained by in silico proteolysis (Table 3). Among the three proteases utilized, stem bromelain gave the highest percentage of DHTs for all FAD3 sequences of flaxseed protein.

We also calculated the parameters such as Ae, W, B, and V of all the protein sequences in this study (Fig. 6). The release frequency (Ae) of was found to be higher in DPP-IV inhibitory peptides released from stem bromelain protease, while the value of relative release frequency (W) was higher in regulating bioactivity released from ficin protease. Tripeptides that were released from in silico proteolysis of FAD3 sequences of flaxseed were further analysed to discover the novel bioactive peptides having specific effect, that were ranked according to scores of best or poor bioactivities (Fig. 7).

3.5. In silico toxicity and allergenicity prediction of potential tripeptides released from FAD3 sequences of flaxseed protein

Large number of bioactive peptides have different biological activities. For in silico prediction of toxicity and allergenicity, tripeptides with higher rank were taken and evaluated. This study

| Table 1 | Physiochemical properties of the all the FAD3 sequences of flaxseed protein. |
|---------|--------------------------------------------------------------------------|
| Accession Number | Number of AA | Molecular weight | Theoretical pI | Formula | Negatively charged residues (Asp + Glu) | Positively charged residues (Arg + Lys) | Extinction coefficients | Instability index | Aliphatic index | GRAVY |
| QT18955 | 392 | 44675.38 | 9.17 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 26 | 35 | 110,030 | 109,780 | 38.85 | 79.31 | −0.169 |
| AKF27245 | 392 | 44752.5 | 9.09 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 27 | 35 | 110,030 | 109,780 | 36.06 | 81.3 | −0.142 |
| AKF27244 | 392 | 44765.5 | 9.09 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 27 | 35 | 110,030 | 109,780 | 37.4 | 81.3 | −0.149 |
| AKF27239 | 392 | 44695.45 | 9.09 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 27 | 35 | 110,030 | 109,780 | 37.7 | 81.56 | −0.139 |
| AFJ53097 | 391 | 44697.34 | 9 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 28 | 35 | 110,030 | 109,780 | 36.72 | 80.28 | −0.169 |
| AFJ53096 | 391 | 44625.28 | 9.09 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 27 | 35 | 110,030 | 109,780 | 36.82 | 80.28 | −0.161 |
| AFJ53090 | 391 | 44651.36 | 9.09 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 27 | 35 | 110,030 | 109,780 | 36.32 | 81.28 | −0.147 |
| ADV92273 | 392 | 44714.36 | 8.31 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 32 | 34 | 110,935 | 110,810 | 33.18 | 85.05 | −0.099 |
| AFN53698 | 901 | 100052.26 | 9.07 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 85 | 100 | 145,690 | 145,190 | 44.42 | 82.14 | −0.214 |
| ADV92271 | 391 | 44677.39 | 9.08 | C_{2038}H_{3072}N_{536}O_{540}S_{14} | 27 | 35 | 111,520 | 111,270 | 36.15 | 81.28 | −0.142 |

S. Langyan, Fatima Nazish Khan, P. Yadava et al. Saudi Journal of Biological Sciences 28 (2021) 5480–5489
was mainly focused on tripeptides with the threshold score of 0.6. The top 27 tripeptides with highest score were GHG, ALP, SGP, QCL, LLP, LWT, IPP, SDF, RWQ, WRS, LPP, WSY, PLP, LWA, VWP, LPM, PPL, WRI, PPP, WRR, LWR, FAP, PWY, PWR, GFL, WPL and PFP, with no previously described bioactivity based on BIOPEP database and literatures, were subjected to in silico prediction of toxicity and allergenicity. The potential toxicity and allergenicity of the flaxseed proteins, as estimated by in silico methods (Fig. 8). The prediction has been given that all the selected tripeptides are nontoxic and are used for further studies in the development and use in producing pharmaceutical products. All tripeptides taken for analysis were considered as promising stimulating, anticancer, antithrombotic, ACE inhibitor, bacterial permease ligand, antioxidative, antibacterial, renin inhibitor, antiinflammatory, dipeptidyl peptidase IV inhibitor, immunostimulating, anti-inflammatory, dipeptidyl peptidase III inhibitor and regulating.

3.6. Potential water solubility and aggregation prediction of released flaxseed tripeptides.

Solubility and aggregation are the universal properties of proteins as well as peptides, which generally limits the pharmaceutical and biotechnological applications. Therefore, for the prediction of water solubility and aggregation, in silico methods were applied. The water solubility and aggregation of potential peptides, predicted by Innovagen's Peptide solubility calculator and AGGRESS-CAN respectively has been determined. As shown in figure, the prediction has been given that among the selected tripeptides, 8 have good solubility in water, while 19 are poorly soluble in water because of large number of hydrophobic residues present in them (Fig. 9). The GRAVY values of flaxseed proteins are related to the hydrophobicity or Boman index.

3.7. In silico evaluation of drug-likeness of peptides

The tool SwissADME has been used for in silico evaluation of drug-likeness, that mainly uses the properties of ADME (absorption, distribution, metabolism, and excretion) of a compound that estimates and indicates the pharmacokinetics of drugs. The drug-likeness of 27 multifunctional tripeptides were predicted and evaluated the ADME and pharmacokinetics properties. We showed the average in-silico drug-likeness and pharmacokinetics of the 27 multifunctional flaxseed tripeptides (Fig. 10).

4. Discussion

In most of the studies it has been noted that bioactive peptides released by in silico proteolysis was much different from the outcomes of experimental enzymolysis (Nongonierma and FitzGerald, 2017; Tu et al., 2018). Although, in silico methods provide an alternative strategy for the identification of novel bioactive peptides, but there also have lots of limitations. The results of previous studies showed that the products obtained from enzymatic hydrolysis could be replace with the degree of hydrolysis that

| Activity                              | \( \Sigma^A \) | \( \Sigma^B \) |
|---------------------------------------|----------------|----------------|
| ACE inhibitor                         | 4.6472         | 0.163724328    |
| activating ubiquitin-mediated proteolysis | 0.0692         |                |
| alpha-glucosidase inhibitor           | 0.2538         | 0.000639102    |
| anti-inflammatory                     | 0.1689         |                |
| antiinflammatory                     | 0.0737         | 5.13E–04       |
| antibacterial                         | 0.0245         |                |
| anticancer                            | 0.0011         |                |
| antioxidative                         | 1.1232         |                |
| antithrombotic                        | 0.0503         | 2.41E–06       |
| bacterial permease ligand             | 0.0734         |                |
| CaMPDE inhibitor                      | 0.076          |                |
| dipeptidyl peptidase III inhibitor    | 1.0506         |                |
| dipeptidyl peptidase IV inhibitor     | 7.0007         | 0.010128318    |
| immunomodulating                      | 0.0711         |                |
| immunostimulating                     | 0.0166         |                |
| neurotropic                           | 0.17           |                |
| opioid agonist                        | 0.0115         |                |
| regulating                            | 0.0918         |                |
| renin inhibitor                       | 0.3186         | 0.002636946    |
| stimulating                           | 0.4803         |                |

Table 2: Total frequency of bioactive peptides \((\Sigma^A \text{ and } \Sigma^B)\).
might affect various factors like features of protein structure, activity of enzymes, temperature, pH, time of hydrolysis and the ratios of enzyme and substrate (Han et al., 2019; Fu et al., 2016; Tu et al., 2018). However, the proteolysis by in silico approaches was more idealistic, digestion was at every specific enzyme’s cutting site, and it was completely done. Besides, in silico proteolysis was according to the recent information of BIOPEP database, which is regularly updated, and the outcomes of analysis might be changed according to new data. The in-silico proteolysis and releasing of various bioactive peptides from FAD3 sequences of flaxseed proteins was done using several bioinformatics software, tools and databases. NCBI database has been used for searching and selecting of 10 unique sequences of FAD3. Three plant proteases, namely ficin, papain and stem bromelain were majorly used for in silico analysis. The higher pI value indicates that the protein sequences are basic in nature, and the structural features of proteins are used to interact electrostatically with lipids that are negatively charged like phospholipids, found in the oil bodies of flaxseed (Huang, 1992). Also, the instability index was less than 40 (except in AFN53698), shows that the protein sequences are stable that will be further evaluated in wet lab experiments for checking the stability of these sequences. Due to high value of aliphatic index, FAD3 sequences are considered as thermostable protein, which means that all sequences are resistant to decay at high temperature. The value of GRAVY suggest that the proteins are hydrophobic in nature and rated positively, and also used their most of the non-polar amino acid residues to interact with and stabilize oil bodies in flaxseed proteins (Tzen et al., 1992). Therefore, during the cellular processes, these proteins decrease the region of linking between non-polar and water molecules, and also utilize the hydrogen bonding between water molecules within cell. The higher value of GRAVY also showing that these flaxseed proteins are more flexible to bind with other proteins. This finding is expected as the FAD3 protein sequences were designed naturally for binding on the lipophilic oil bodies and are less capable to bind to other seed proteins.

For investing the potential of flaxseed FAD3 protein sequences as bioactive peptides precursors, the sequences were evaluated by “Profiles of potential biological activity” tool available in BIOPEP database. Various known bioactivities were found in flaxseed proteins suggesting that these are helpful in different biological activities in oil synthesis in term of quality and quantity. Fatty acid desaturases (FADs) help in the metabolism of fatty acid and introduce double bonds into fatty acids’ (FAs) hydrocarbon chains that produce unsaturated fatty acids, and FAD3, the commonest microsomal enzyme of omega-3 fatty acid, synthesizes linolenic acid (C18:3) from linoleic acid located in endoplasmic reticulum (ER) facing towards the cytosol. Mostly the peptides act as inhibitors against hypertension and various diseases. The ACE and DPP-IV inhibitor were common bioactive peptides in each sequence evaluated by all proteases. ACE regulate the blood pressure by enhancing the active hypertensive hormone production and inactivating vasodilator peptide, serves as the promising target for antihypertensive drugs (Miralles et al., 2018). DPP-IV is linked with incretin degradation and regulating the levels of blood glucose, and also act as inhibitor for the treatment of type 2 diabetes mellitus patients (Juillerat-Jeanneret, 2014). From the last few decades, DPP-IV inhibitors, derived from food protein have been intensively studied (Lacroix and Li-Chan, 2012). Large number of dietary proteins from animal, plants and marine organisms have been utilized for the production of ACE inhibitory hydrolysates (Lee and Hur, 2017). Our study suggested that two bioactive peptides namely DPP-IV and ACE inhibitors are presented largely in all the FAD3 sequences of flaxseed protein. The location and type of bioactive fragment present in protein sequences calculated from BIOPEP-UWM database giving the information to isolate the bioactive peptides in

### Table 3
Degree of hydrolysis (DHT) obtained by in silico proteolysis.

| Accession Number | Ficin (DHT [%]) | Papain (DHT [%]) | Stem Bromelain (DHT [%]) |
|------------------|----------------|-----------------|-------------------------|
| QT18955          | 50.1279        | 35.0384         | 50.3836                 |
| AFK27245         | 50.3836        | 35.5499         | 50.8951                 |
| AFK27244         | 50.3836        | 35.2941         | 50.6394                 |
| AFK27239         | 50.1279        | 35.2941         | 50.6394                 |
| AFJ53097         | 50             | 35.1282         | 50.7692                 |
| AFJ53096         | 50.2564        | 35.3846         | 51.0256                 |
| AFJ53090         | 50             | 35.3846         | 50.7692                 |
| ADV92275         | 48.5934        | 35.0384         | 50.3836                 |
| AFN53698         | 50             | 38.1111         | 56.1111                 |
| ADV92271         | 50             | 35.3846         | 50.7692                 |

### Fig. 5
Sum of chemical mass for proteins digested by all proteases selected, (QT18955, AFK27245, AFK27244, AFK27239, AFJ53097, AFJ53096, AFJ53090, ADV92275, AFN53698 and ADV92271 are accession number of sequences of FAD3 from flaxseed protein).
wet lab experiments for future studies. Various parameters such as $A$, $B$, $A_E$, $W$, $B_E$, $V$ and $D\text{H}t$ were calculated that suggest the frequency of all bioactive peptide’s occurrence in the protein, estimating the potential of a protein to have a particular biological activity, frequency and relative frequency of releasing of peptides by specific protease, and also the estimation of percent degree of hydrolysis of in silico digestion of peptides. The differences in the value of $W$ for different proteases can be suggested by the difference in the value of $D\text{H}t$, number of recognition sites on proteases and the catalytic specificities.

Different enzymes have different potential to release bioactive peptides from proteins, which attribute to their specific cleavage sites (Gomez et al., 2019). In our study, stem bromelain treated FAD3 proteins relatively release higher frequency index of DPP-IV inhibitors as compared to other two protease, while the value of relative release frequency ($W$) was higher in regulating
bioactivity released from ficin protease. This might be due to stem bromelain has the most cutting sites in comparison to other proteases. Actually, there is still need of study in future to identify the unknown peptides with suitable bioactivities. The peptides also screened and ranked according to scores of best or poor bioactivities, to discover the novel bioactive peptides having specific effect. The score assigned to bioactive peptides might be different in when one can apply in wet lab experiments which suggested that they may have other biological functions. In silico toxicity and allergenicity were also important properties for the evaluation of bioactivities of peptides. The property of non-toxin of any peptide was considered as the best and will can take for further studies in the development and use in producing pharmaceutical products. All peptides with peptide score higher than 0.6 were predicted as non-toxin. These findings suggested that plant proteases are more important for the peptides releasing from FAD3 sequences of flaxseed proteins having multifunctional bioactivities. Interestingly, the protease stem bromelain released a greater number of multifunctional peptides from the flaxseed proteins, that might be useful for future studies. Solubility and aggregation are the universal properties of proteins as well as peptides, which generally limits the pharmaceutical and biotechnological applications. Therefore, for the prediction of solubility and aggregation, in silico methods were applied. The water solubility and aggregation of potential peptides, predicted by Innovagen’s Peptide solubility calculator and AGGRESCAN respectively. These kinds of predictions mainly affect the absorption, distribution and elimination of peptides from the body, that will further help in drug designing purposes. The varying degree of solubility of flaxseed peptides demonstrated that they can perform functions in large number of systems, whether they are lipid-soluble, water-soluble or based on emulsion. In silico drug-likeness evaluation was carried with online tool SwissADME, that mainly uses the properties of ADME (absorption, distribution, metabolism, and excretion) of a compound that estimates and indicates the pharmacokinetics properties of drugs.

5. Conclusions

Bioinformatics methods have been widely utilized for studying and releasing the bioactive peptides derived from proteins of food crops in a cost-effective and comprehensive manner. These in silico approaches helps in providing information of peptide conformations, interacting mechanisms of molecules in peptides, and also the improvement of properties of peptides. In this study, in silico evaluation of release of bioactive peptides using three plant proteases (ficin, papain, and stem bromelain) was carried out. On the basis of ten sequences of FAD3 from flaxseed protein, our study suggested that the flaxseed proteins contain number of bioactive peptides with different activities. The results of in silico proteolysis showed that the releasing of ACE and DPP-IV inhibitory peptides are more common in all the plant proteases used, which demonstrates that the flaxseed proteins have play a major role in the generation of peptides with dual functions of ACE and DPP-IV inhibitors. Instead of some limitations in analysing bioactive peptides by in silico approaches, there is enough idea for concluding that FAD3 sequences of flaxseed protein might be the good precursor for bioactive peptides production, and also in silico proteolysis will be the best approach for bioactive peptides preparation. Physicochemical properties such as molecular weight, theoretical isoelectric point (pI), total number of negatively charged (Asp + Glu) and positively charged (Arg + Lys) residues, amino acid and atomic composition, extinction coefficient, estimated half-life, Grand average of hydropathicity (GRAVY), instability and aliphatic index have been identified for all the FAD3 sequences of flaxseed protein. Various other analysis such as toxicity, allergenicity, water solubility, aggregation, and drug likeness have also been done using bioinformatics tools and servers. However, the peptides come from food-derived proteins by enzymatic actions, and therefore these peptides not have undesirable side effects when utilized in clinical management. It was reported that outcomes of in silico analysis involving the discovery and analysing of food-derived bioactive peptides required experimental validation. Thus, future work by our research team will more focus on some of computational analysis such as peptide structure–activity analysis (QSAR analysis), mathematical simulation of hydrolysis and
molecular docking studies of selected peptide released in this study that could overcome the limitations of costly and time-consuming experimental discovery of peptides. We will also focus on the lab-scale synthesis of the specific bioactive peptide (say for example ACE inhibitor) identified through in silico analysis and also the experimental validation of various in silico properties along with in vivo pharmacokinetics studies to confirm the in-silico findings. The analysis of gastrointestinal is also required to study the practical applications of food-derived peptides as bioactive nutrients or health benefits against number of diseases.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was funded by Taif University Researchers Supporting Project number (TURSP-2020/138), Taif University, Taif, Saudi Arabia.

References

Agyei, D., Ongkudon, C.M., Wei, C.Y., Chan, A.S., Danquah, M.K., 2016. Bioprocess challenges to the isolation and purification of bioactive peptides. Food Bioprocess Technol. 9, 244–256. https://doi.org/10.1007/jb.2016.02.003.

Agyei, D., Tsopomo, A., Udengwe, C.C., 2018. Bioinformatics and peptidomics approaches to the discovery and analysis of food-derived bioactive peptides. Anal. Bioanal. Chem. 410 (15), 3463–3472. https://doi.org/10.1007/s00216-018-1097-4.

Belhêt, A.-D., Shavandi, A., Jodja, T., Bürch, T., Teh, S., Mohamed Ahmed, I.A., Al-Jahdali, F.Y., Saedi, P., Belhêt, A.A., 2018. Flaxseed: Composition, detoxification, utilization, and opportunities. Biocatal. Agric. Biotechnol. 13, 129–152. https://doi.org/10.1016/j.bcab.2017.11.017.

Cavazos, A. Gonzalez de Mejia, E., 2013. Identification of bioactive peptides from cereal storage proteins and their potential role in prevention of chronic diseases. Comp. Rev. Food Sci. Food Saf. 12 (34), 364–380. https://doi.org/10.1111/crfs.12133.10141–3143.1207.

Coulombe, Y., Karababa, E., 2007. Some physical properties of flaxseed (Linum usitatissimum L.). J. Food Eng. 78 (3), 1067–1073. https://doi.org/10.1016/j. jfoodeng.2005.12.017.

Dar, A.A., Choudhury, A.R., Kancharla, P.K., Arumugam, N., 2017. The FAD2 gene in chickpea (Cicer arietinum L.). J. Food Eng. 78 (3), 1067–1073. https://doi.org/10.1016/j. jfoodeng.2005.12.017.

Dimitrov, I., Naneva, L., Doytchinova, I., Bangov, I., 2014. AllergenFP: allergenicity assessment of portuguese oyster (Crassostrea angulata) proteins as precursor of dipeptidyl peptidase IV (DPP-IV) inhibitory peptides. Food Chem. 231, 202–211. https://doi.org/10.1016/j.foodchem.2017.07.123.

Dong, Z., Wang, J., Zhang, Y., Ouyang, C., Wang, X., Wang, G., 2019. Biological activities of oilseed proteins using two integrated bioinformatic approaches. Food Res. Int. 115, 283–291. https://doi.org/10.1016/j.foodres.2018.12.015.

Huang, A.H., 1992. Oil bodies and oleosins in seeds. Annu. Rev. Plant Biol. 43 (1), 177–200. https://doi.org/10.1146/annurev.pp.43.060192.091141.