Microchemical Journal 203 (2024) 110872

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Extensive characterization and antioxidant function prediction of endogenous peptides derived from *Polygonatum cyrtonema* Hua

Yanyan Zhang [a], Haixia Li [b], Peizi Liu [b], Keyi Chen [b], Shengjun Ma [a,*], Wei Cai [b,*]

[a] School of Food Science and Pharmacy, Xinjiang Agricultural University, Urumqi, Xinjiang, 830000, PR China
[b] School of Pharmacy, Hunan University of Medicine, Huaihua, Hunan, 418000, PR China

ARTICLE INFO

Keywords:
Polygonatum cyrtonema Hua
Endogenous peptides
Differential analysis
Nano-LC-Q Exactive Plus Quadrupole-Orbitrap MS
Molecular docking

ABSTRACT

Polygonatum cyrtonema Hua (*P. cyrtonema*), a medicinal and edible, has a long-standing tradition of treating respiratory diseases, regulating immune function, and enhancing liver health. Although polysaccharides, saponins, and flavonoids have been identified as its major bioactive components, the presence and properties of bioactive peptides in *P. cyrtonema* are relatively understudied. This study utilized an alkali extraction-acid precipitation method to extract crude protein from *P. cyrtonema*. Subsequently, Nano-LC-Q Exactive MS was utilized to characterize successfully the sequences and abundances of endogenous peptides from different sources. A total of 2571 M1(endogenous) peptides and 2122 M2(crude protein) peptides were identified using a combination of database searches and *de novo* strategies, with significantly higher confidence levels observed for M2 peptides identified through the database search approach. Further analysis of subsequent overall differences reveals that, based on the chemical properties and bioinformatics analyses of peptides from different sources, the M1 peptide primarily consists of 4–6 amino acids (62.1 %), while the M2 peptide is mainly composed of more than 8 amino acids (60.8 %). Molecular docking technology was then applied to verify the antioxidant prediction results and to evaluate the differences in antioxidant potential between M1 peptides and M2 peptides, with seven of the top ten ranked peptides in terms of docking energy being M1 peptides. These findings notably advance our understanding of endogenous peptides in *P. cyrtonema* and provide valuable reference information for identifying potential candidate bioactive compounds. This information will support future comprehensive validation of the biological activities of these peptides in *P. cyrtonema* and contribute to drug discovery research.

1. Introduction

Polygonatum cyrtonema Hua (*P. cyrtonema*) is a perennial herbaceous plant belonging to the Liliales family [1]. *P. cyrtonema* is primarily distributed across the temperate regions of the Northern Hemisphere and thrives in shaded, moist environments, such as forests understory and shrublands [2]. Historical records reveal that *P. cyrtonema* has been utilized for medicinal purposes for nearly two thousand years, with primary medicinal properties attributed to its rhizomes [2]. There are valued for their effectiveness in treating symptoms of dizziness, fatigue, and respiratory illnesses [3]. Additionally, *P. cyrtonema* is commonly used as a food ingredient in soups, porridge, and brewing alcohol [4]. Previous investigated shown that polysaccharides [5,6], saponins [7], and flavonoids [8]were the main active components of *P. cyrtonema*.

Bioactive peptides (BPs) represent a class of macromolecular functional components with extensive bioactivities in medicinal plants [9,10]. Based on their source, BPs can be classified as exogenous or endogenous peptides. Exogenous peptides are obtained method such as enzymatic hydrolysis and microbial fermentation, while endogenous peptides are produced by proteases in biological or plant organisms and can be extracted using solvents [11–13]. Exogenous peptides often receive more attention due to the optimized methods that yield fixed and optimal results for preparation of target components [13]. Additionally, these peptides, derived from parent proteins, are generally more abundant, facilitating achieving more accurate characterization through mass spectrometry. However, endogenous peptides, compatible with the plant's biological system, offer higher safety and can serve as biological markers for specific plant species [14–16]. Despite this, the sequence and abundance of endogenous peptides in medicinal plants, such as *P. cyrtonema*, are largely unknown. The preparation of

* Corresponding authors.
E-mail addresses: zhangyanyan9527@163.com (Y. Zhang), 13684563626@163.com (H. Li), 18522650390@163.com (P. Liu), 19983213436@163.com (K. Chen), wlmqmsj@xjau.edu.cn (S. Ma), 20120941161@bucm.edu.cn (W. Cai).

https://doi.org/10.1016/j.microc.2024.110872
Received 14 February 2024; Received in revised form 21 May 2024; Accepted 28 May 2024
Available online 29 May 2024

exogenous peptides typically involves extraction crude proteins through methods like alkali extraction and acid precipitation. However, these technique are often criticized for potentially to inducing protein or BPs inactivation and denaturation, which could impact subsequent experimentations involving enzymatic hydrolysis of parent proteins to generate exogenous BPs [17,18].

The primary methods for analyzing peptide sequences from mass spectrometry data are currently database search and *de novo* sequencing [13]. The database search method relies on simulating enzymatic cleavage at user-specified cleavage sites within selected proteins to generate virtual peptide fragments [12,19–21]. These theoretical peptide spectra are then compared against each experimental MS/MS spectrum, with peptide scores generated using scoring algorithms [22]. Peptides with the highest scores are considered most likely to be correct matches. Database search results are generally reliable when the database includes comprehensive sequences of the corresponding proteins under study. However, research on proteins from traditional Chinese medicine (TCM), especially botanical TCM, is relatively scarce, and peptide variants along with unexpected degradation products may emerge during experiments [23]. The *de novo* strategy fill this gap by not relying on protein databases and instead infers peptide sequences directly from the mass differences of ion fragment peaks obtained from secondary mass analyzers [12]. This method provides more complete and accurate peptide sequence information. However, the credibility of identification results diminishes when ion fragments are lost [24,25]. With the advancement of high-resolution mass spectrometers, *de novo* sequencing is increasing becoming a viable alternative for peptide identification.

The study began with the extraction of crude proteins from *P. cyrtonema* using an alkaline extraction followed by acid precipitation. Analysis was carried out using nano-LC-Q Exactive Plus Quadrupole Orbitrap MS, coupled with PEAK Studio software, to identifying two types of peptides [26]. The subsequent data analysis employed both database searching and *de novo* sequencing approaches, yielding 2571 M1 peptides and 2122 M2 peptides in crude proteins. Differential analysis based on the chemical properties of the peptides-such as molecular weight (MW), peptide length, charge state, average isoelectric point, grand average of hydropathy (GRAVY) value and retention times-was performed. Bioinformatics prediction tools were utilized to select peptides with the highest antioxidative potential for molecular docking [27]. The potential for antioxidative activity was inferred from the binding energies. The chemical properties and structural characteristics of the peptides with the highest antioxidant potential were analyzed to determine consistency with previous research findings. Using antioxidative potential as the selection criterion, two types of peptides were screened, primarily differing in peptide length and MW: M1 peptides, predominantly consisting of segments with 4–6 amino acids (62.1 %), and M2 peptides, mainly comprising peptides with more than 8 amino acids (60.8 %). Molecular docking energy was utilized to validate the predicted antioxidative potential of peptide segments, revealing that the anti-oxidative potential of M1 peptides surpassed that of M2 peptides. This experiment, exploring for the first time the impact of a peptide preparation scheme on endogenous peptides through differential analysis, provides valuable insights and scientific basis for future preparation and screening of BPs from *P. cyrtonema*.

2. Materials and method

2.1. Chemicals and reagents

P. cyrtonema was purchased from Hunan Bestcome Traditional medicine co., Ltd, dried at 60℃, and subsequently sealed for use. Petroleum ether (boiling range:60-90℃) was purchased from Tianjin Aopusheng Chemical Co., Ltd (China). Sodium hydroxide was obtained from Sinopharm Chemical Reagent Co., Ltd (China). Hydrochloric acid was obtained from Chengdu Jinshan Chemical Reagent Co., Ltd (China). High-performance liquid chromatography (HPLC)-grade acetonitrile (A998-4) was obtained from Fisher scientific (China). LC-MS-grade formic acid (28905) was purchased from Thermo Scientific (China), HPLC-grade Trifluoroacetic acid (T818781) was obtained from Shanghai Macklin Biochemical Technology Co., Ltd (China). ZipTip C18 micro chromatographic column (ZTC18S096) was obtained from MerckMillipore (China) and ultra-pure water was prepared using ultra-pure water system (KZ-ZDX-30L, Kezhi, Shanghai Kezhi Environmental Protection Equipment Co., Ltd, China).

2.2. Extraction of crude protein from Polygonatum cyrtonema Hua

Dried *P. cyrtonema* was ground (FSJ-1000C, Lingsum, Yongkang Hongtaiyang Electromechanical Co., Ltd, Zhejiang, China), and sieved through a 0.178 mm mesh. The collected powder underwent defatting treatment using petroleum ether (1:3 w/v) for three rounds of 5 min each, with continuous stirring until the petroleum ether became clear. After filtration (SHZ-III, Jinye, Shanghai Yarong Biochemical Instrument Factory, Shanghai, China), the defatted *P. cyrtonema* powder was air-dried. For extraction, the defatted powder was mixed with pure water (1:22 w/v) at 40℃. Ultrasonic (KQ-300DE, Shumei, Kunshan ultrasonic instrument Co., Ltd, Jiangsu, China) assistance (100 W) and pH adjustment to 11 using a 1 M sodium hydroxide solution were applied during the 40-minute extraction process. The obtained extract was centrifuged (L530 R, Cence, Hunan Xiangyi Experimental Instrument Development Co., Ltd, Hunan, China) at 4000 rpm for 25 min to collect the supernatant. Adjusting the pH of the supernatant to 3 with 1 M hydrochloric acid, the mixture was left to stand for 1 h to allow the crude protein to precipitate. The precipitate was obtained by centrifuging at 4000 rpm for 10 min and removing the supernatant. The precipitate was collected and freeze-dried (LGJ-10C, Foring, Foring technology development (Beijing) Co., Ltd, Beijing, China) into powder for sample preparation of Nano-LC-Q Exactive Plus Quadrupole-Orbitrap analysis.

2.3. Sample desalting and preparation

50 mg of defatted *P. cyrtonema* powder (M1) and crude protein powder(M2) were dissolved in 200 μL of 0.1 % TFA. The sample solution was then desalted using ZipTip C18 microcolumns; the column was rinsed 10 times with 50 μL of 60 % ACN containing 0.1 % TFA, followed by 10 rinses with 10 μL of 0.1 % TFA. Subsequently, the sample solution was aspirated and dispensed 20 times, and the ZipTip C18 micro-column was rinsed 5 times with 10 μL of 0.1 % TFA. Finally, the peptides were eluted into an EP tube using 10 μL of 60 % ACN containing 0.1 % TFA, and then subjected to vacuum freeze-drying. The freeze-dried peptide was reconstituted in 20 μL of 5 % ACN solution containing 0.1 % TFA, followed by vigorous vortexing (VORTEX-5, Kylin-Bell, Haimen Kylin-Bell lab instruments Co., Ltd, Jiangsu, China). The solution was then centrifuged at 13500 rpm for 20 min at 4℃ (5424R, Eppendorf, Eppendorf (Shanghai) International Trade Co., Ltd, Shanghai, China), and the supernatant was used for Nano-LC-Q Exactive Plus Quadrupole-Orbitrap MS identification.

2.4. Nano-LC-Q Exactive plus Quadrupole-Orbitrap MS analysis

A Thermo Scientific Q Exactive Plus Quadrupole-Orbitrap MS (Thermo Fisher Scientific, Bremen, Germany) system was used to analyze peptides derived from *P. cyrtonema* powder and crude protein powder. This system was coupled online with a Thermo Scientific EASY-nLC 1200 system (LC140, Thermo Scientific, Bremen, Germany), and a self-packed C18 column (75 μm ID, 150 mm length), using Acclaim PepMap RSLC C18 as packing material (2 μm, 100 Å; Nanoviper). The LC-MS analysis Conditions were set as follows: a linear gradient elution was employed using solvent A (0.1 % formic acid) and solvent B (80 % ACN containing 0.1 % formic acid) with a stepped gradient program: 0–3 min, 3 % B, 400 nL/min; 3–7 min, 3–8 % B, 400 nL/min; 7–46 min,

Table 1
Protein sources of M1 peptides.

Accession	−10lgP	Coverage (%)	Description
Q8L568|Q8L568 9ASPA	96.93	50.00 %	Mannose/sialic acid-binding lectin OS = *Polygonatum cyrtonema*
Q5EER7|Q5EER7 9ASPA	61.15	14.47 %	Mannose/sialic acid-binding lectin OS = *Polygonatum roseum*
R4QNU6|R4QNU6 9ASPA	22.05	2.83 %	NAD(P)H-quinone oxidoreductastsubunit 5, chloroplastic (Fragment) OS = *Polygonatum stenophyllum*
AOA8E5N7T4| AOA8E5N7T4 9ASPA	4.06	1.70 %	Photosystem Il D2 protein OS = *Polgonatum hunanense*
AOA7S8F9L2| AOA7S8F9L2 9ASPA	3.20	1.57 %	Abiotic stress-related transcription factor OS = *Polvaonatum cyrtonema*
Q9M653|RIPT POLML	3.08	1.00 %	Ribosome-inactivating protein PMRIt OS = *Polyaonatum multiflorum*

Table 2
Protein sources of M2 peptides.

Accession	−10lgP	Coverage (%)	Description
Q8L568|Q8L568 9ASPA	210.00	50.00 %	Mannose/sialic acid-binding lectin OS = *Polygonatum cyrtonema*
O24274|O24274 POLML	136.45	30.63 %	Mannose-specific lectin OS = *Polygonatum multiflorum*
AOAOM1RFB4| AOAOM1RFB4 9ASPA	2.65	16.19 %	Protein TIC 214 OS = *Polygonatum cirrhifolium*
AOA678Q5P8| AOA678Q5P8 9ASPA	1.02	9.12 %	Maturase K (Fragment) OS = *Polygonatum arisanense*
AOA7M3V8K1| IAOA7M3V8K1 9ASPA	7.75	2.35 %	Protein TIC214 OS = *Polygonatum cirrhifolium*
O3LFE3|O3LFE3 9ASPA	3.33	2.26 %	Cytochrome *c* oxidase subunit 3 OS = *Polygonatum hookeri*

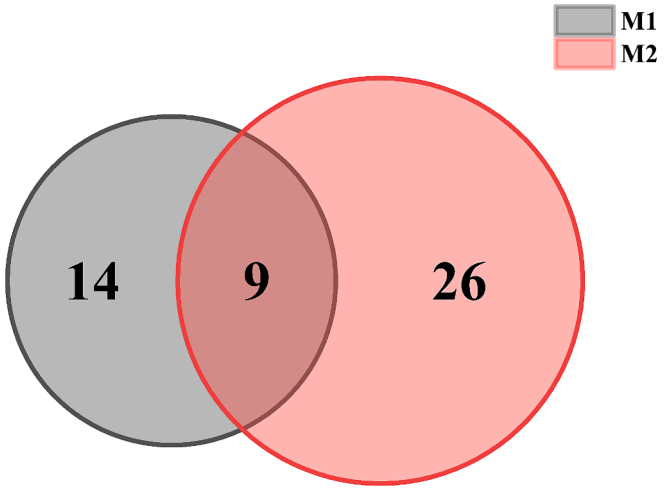

Fig. 1. Comparison between M1 peptides and M2 peptides.

8–32 % B, 400 nL/min; 46–51 min, 32–44 % B, 400 nL/min; 51–60 min, 44–99 % B, 400 nL/min; 60–65 min, 99–3 % B, 400 nL/min. The total run time was 65 min per injection with a sample injection volume of 8 μL. A Thermo Scientific Q Exactive Plus Quadrupole-Orbitrap (Thermo Fisher Scientific, Bremen, Germany) system was employed for MS and MS/MS data acquisition and conditions for it were set as follows: the precursor ion was scanned using full scanning mode in mass range at 350–1550 *m/z*, with a resolution of 120,000. The AGC target was set at $4e^5$, and the maximum injection time was 50 ms. Fragmentation was performed using data dependent acquisition (DDA) on the top 20 ions based on precursor ion intensity. The MS/MS resolution was set at 30000, with an AGC target of $2e^5$ and a maximum injection time of 100 ms. The normalized collision energy (NCE) was set at 32 %.

2.5. Data analysis

The raw files obtained from mass spectrometry were imported into PEAKS studio 11 (11.5 bulid 20230821) for peptide identification, utilizing both database and *de novo* methods. For the database search, the UniProt database with taxonomy ID 16195 was used, with a mass tolerance of 15 ppm for precursor ions and 0.05 Da for fragment ions. No enzyme was specified for protein cleavage. Variable modifications were selected for acetylation (N-terminal, delta mass + 42.01) and oxidation (delta mass + 15.99) for both the database search and Auto *De Novo* sequencing. The "Deep Learning Boost" option was selected to employ enhanced deep learning algorithms for improved peptides identification and peptide-spectra match (PSM) accuracy during the database search process. Information including protein source, amino acid sequence, MW, charge, retention time, −10lgP score, and average local confidence (ALC) score, among other parameters, were exported. Results from *de novo* sequencing were filtered for an ALC score greater than 60 %. Venn diagrams and heat maps were generated using GraphPad Prism (9.5.0) and Origin Pro 2021 (9.8.0.200), respectively, to analyze the differences in the data obtained.

2.6. Bioinformatics analysis

The isoelectric point and charge of the peptides were calculated using IPC 2.0, an online tool (https://isoelectric.org/index.html). The Gravy score was predicted using the GRAVY CALCULATOR (https://www.gravy-calculator.de/index.php). The full name of GRAVY is Grand Average of Hydropathy. The GRAVY score of a peptide considers the hydrophilicity and hydrophobicity of each amino acid residue, specifically the side chains. This score is calculated by summing the hydropathy values of all amino acid residues and dividing it by the total number of residues [28]. Positive values indicate polar peptides, whereas negative values indicate nonpolar peptides [29,30].

2.7. Peptide function prediction analysis

Bioactivity of peptides was predicted using various databases, including BIOPEP-UWM (https://biochemia.uwm.edu.pl/biopep) [31], Peptide Ranker (https://distilldeep.ucd.ie/PeptideRanker) [32], AnOxPePred1.0 (https://services.healthtech.dtu.dk/services/AnOxPePred-1.0) [33]. The screening of peptides exhibiting potential antioxidant activity was initially conducted using the BIOPEP-UWM database, which comprises 4651 BPs. Subsequently, AnOxPePred 1.0 was used to evaluate the quality of their antioxidant activities based on parameters such as free radical scavenging rate and ion chelation ability. The convolutional neural network method was employed for prediction and scoring, with the top 100 scoring peptides further assessed on the peptide ranker website to identify those with scores exceeding 0.6 for further sequence divergence analysis. The Peptide ranker predict the likelihood of peptide bioactivity, as the probability approaches unity, the peptide is more likely to exhibit bioactivity in all peptides which submit to server. These tools utilize quantitative structure–activity relationship (QSAR) regression models to predict peptide bioactivity [30].

2.8. Molecular docking analysis

The semi-flexible binding energy between Keap1 and the peptides was calculated using the molecular docking software AutoDock 4.2 [34].

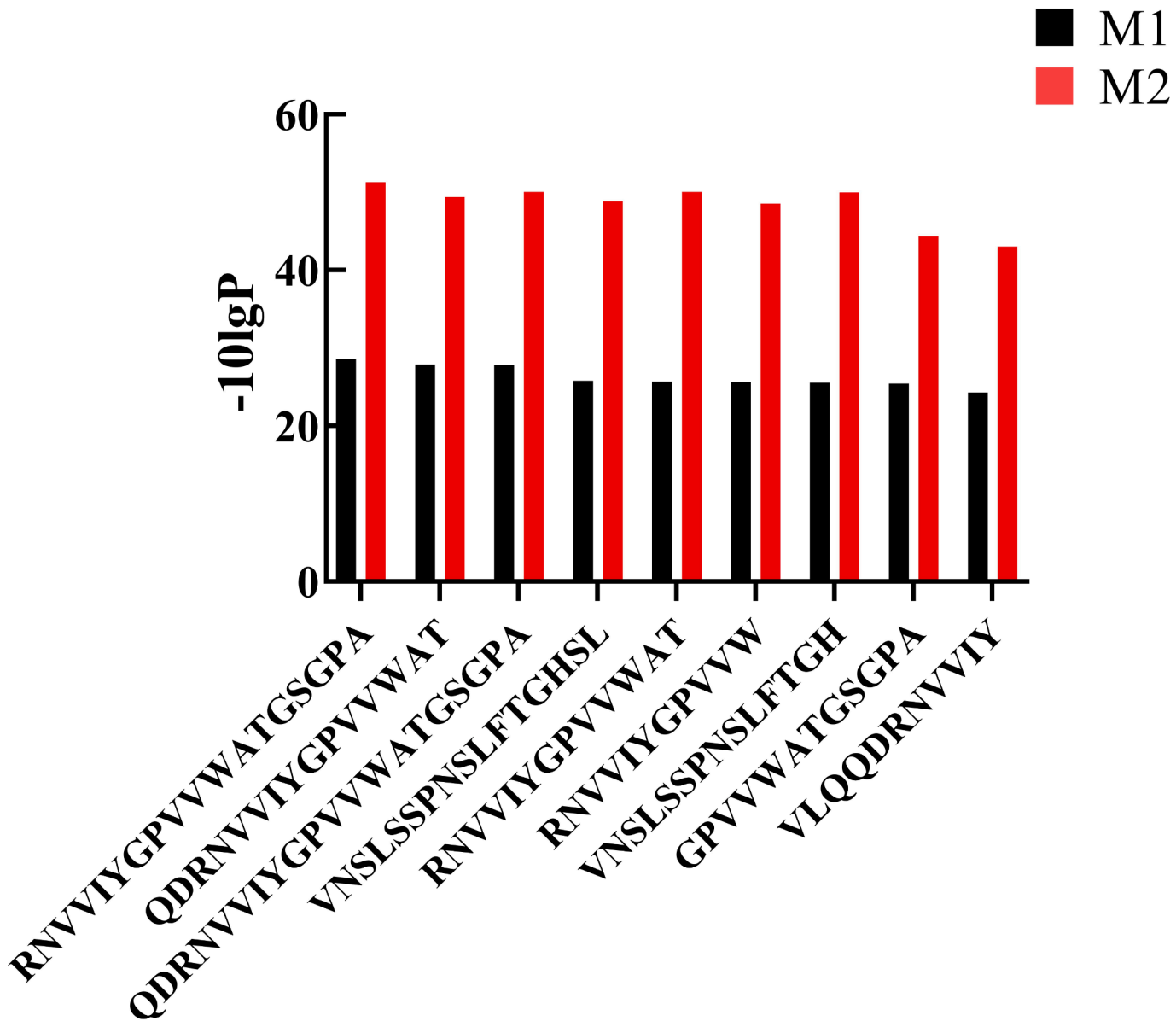

Fig. 2. Comparison of −10lgP score in identical 9 peptides between M1 peptides and M2 peptides.

Table 3
Comparison of area in identical 9 peptides between M1 peptides and M2 peptides.

Identical peptides	Source	Area
RNVVIYGPVVWATGSGPA	M1	2.64×10^7
	M2	8.30×10^7
QDRNVVIYGPVVWAT	M1	1.08×10^5
	M2	8.47×10^5
QDRNVVIYGPVVWATGSGPA	M1	2.92×10^5
	M2	1.32×10^6
VNSLSSPNSLFTGHSL	M1	3.63×10^6
	M2	1.22×10^6
RNVVIYGPVVWAT	M1	1.11×10^7
	M2	4.98×10^6
RNVVIYGPVVW	M1	1.03×10^6
	M2	1.17×10^6
VNSLSSPNSLFTGH	M1	2.33×10^6
	M2	7.46×10^6
GPVVWATGSGPA	M1	5.31×10^6
	M2	2.67×10^6
VLQQDRNVVIY	M1	3.56×10^5
	M2	3.84×10^6

Prior to the docking, the pdb format file of Keap1 (PDB ID: 2FUL) was obtained from the RCSB Protein Database (https://www.rcsb.org). Keap1 served as the receptor protein for docking with peptides [30]. Initially, the three-dimensional structure of the peptide was created using ChemDraw 20.0, followed by energy minimization in using Chem 3D 20.0, and saved in pdb format. After importing the three-dimensional structures of peptides and receptor proteins into Autodock tool 1.5.7, hydrogen atoms were added and the protein types were set as “ligand” and “Macromolecule”, respectively. The structures were thensaved in pdbqt format, and the grid box spacing was set to 0.375 Å to generate the gpf file [35]. The binding energies were optimized using the Lamarckian genetic algorithm 4.2, and a dpf file was generated for molecular docking [36]. Upon completion of docking, and “conformation play by ranked energy” was selected to display the molecular docking results and establish hydrogen bonds to obtain docking energies base on the resulting dlg file. The docking results in pdbqt format were converted to pdb format using Open Babel 3.3.1, and the pdb file was opened using PyMOL 2.5.2 to generate a three-dimensional docking plot [37]. Finally, a two-dimensional interaction diagram between the ligand and the receptor was generated using Discovery Studio 24.1.0, for the analysis of

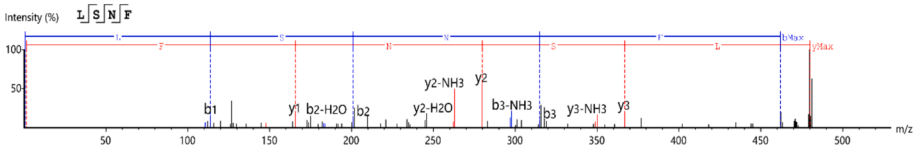

Fig. 3. Peptide "LSNF" MS/MS spectrogram.

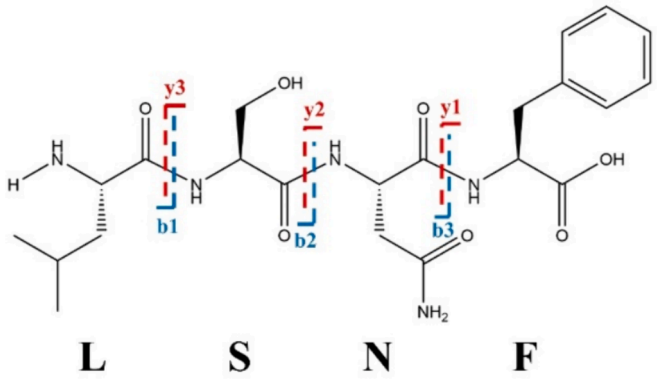

Fig. 4. Peptide "LSNF" b ion and y ion fragmentation picture.

Table 4
Peptide "LSNF" y ion mass spectrometry data and identification results.

	Amino Acid Sequence	Calculated value	Measured value
y1	−F	166.086	166.09
y2	−NF	280.129	280.130
y3	−SNF	367.161	367.160
Total	LSNF	480.245	480.2461

Table 5
Peptide "LSNF" b ion mass spectrometry data and identification results.

	Amino Acid Sequence	Calculated value	Measured value
b1	L-	114.092	114.100
b2	LS-	201.124	201.120
b3	LSN-	315.167	315.170
Total	LSNF	480.245	480.2461

Table 6
De novo sequencing results of M1 peptides.

Peptide sequence	ALC (%)	z	*m/z*	Mass	RT (min)	ppm	Mode
LSNF	99.8	1	480.24597	479.238	19.84	1.5	HCD
P (+42.01) PAF	98.7	2	237.12341	472.2322	24.53	0.2	HCD
FRAP	98.5	2	245.64354	489.27	13.59	5.3	HCD
H (+42.01) LHH	98.3	2	293.14923	584.2819	17.10	3.4	HCD
K (+42.01) VNDPF	97.8	1	761.38312	760.3755	36.65	0.4	HCD
LRFSR	97.8	3	226.80705	677.3973	13.23	3	HCD
YVFP	97.8	2	263.13968	524.2635	18.45	2.6	HCD
H (+42.01) VLF	97.5	2	279.15973	556.3009	23.54	7.2	HCD
METNY	97.5	1	657.25208	656.2476	33.57	−4.2	HCD
M (+42.01) DLRVF	97.4	2	411.71359	821.4105	55.39	2.5	HCD

hydrogen bonding and hydrophobic interactions.

3. Results

3.1. Database search identification results

Database searching is currently one of the two primary methods for peptide identification from MS/MS data. It involves generating a corresponding peptide list from the MS/MS data against a user-specified protein database, and scoring the peptides based on the fragmentation spectra [38,39]. A total of 57 peptides were identified in M1 through database searching, and 58 peptides were identified in M2. The −10lgP scores exported from PEAKS studio database searching are primarily associated with linear discriminant function (LDF), which can be used to assess the quality of peptide spectrum matching [39]. Thus, higher −10lgP scores reflect more credible peptide identification outcomes. It is worth noting that when the peptide dataset is less than 100, a threshold of 20 is typically considered appropriate for −10lgP [39]. Significantly, the M1 peptides were predominantly sourced from 45 protein sequences spanning 6 protein groups, whereas the M2 peptides originated from 123 protein sequences within 6 groups. The two protein groups exhibiting the highest coverage for both M1 and M2 were identified as accession Q8L568_9ASPA from *Polygonatum cyrtonema*, with a coverage rate of 50 %, consistent with the source of M2. However, the −10lgP for the M2 peptides was significantly higher than those of the M1 peptides. The other major protein groups included *Polygonatum roseum, Polgonatum hunanense, Polyaonatum multiflorum, Polygonatum hookeri, Polygonatum cirrhifolium, Polygonatum arisanense,* and *Polygonatum stenophyllum*. Table 1 and Table 2 show a detailed overview of the protein sources for both M1 and M2 of *P. cyrtonema*. To augment the reliability of analytical outcomes in database research, a favorable filtering criterion was adopted, requiring a threshold of $-10lgP \geq 20$ and a minimum peak area of 10^4. Notably, such filtering resulted in a comprehensive identification of 23 M1 peptides and 37 M2 peptides. Upon comparing the both sets of peptides; it was observed that 9 peptides showed an identical alignment in terms of amino acid sequence. In particular, the M2 peptides displayed a high abundance of M1 peptides, reflected by their high −10lgP scores and increased peak areas. These significant findings strongly imply an augmented extent of peptide diversity and abundance. The results of Venn diagram comparison between M1 peptides and M2 peptides of *P. cyrtonema*, as well as the alignment of 9 identical peptides, are shown in Fig. 1, Fig. 2 and Table 3, respectively.

3.2. De novo identification results

De novo sequencing is a powerful methodology that bypasses reliance on protein database references, instead enabling the direct inference of amino acid sequence from experimental MS/MS spectra fragment ion analysis [40]. This approach may facilitate identification of peptide sequences not captured within the existing protein databases, thereby considerably augmenting the pool of identified peptide sequences relative to database search techniques. The number of M1 and M2 peptides (Automated Spectrum Confidence Level (ALC) > 50 %) obtained through *de novo* sequencing were significantly higher than those identified through database search, totaling 2514 and 2064 peptides, respectively. Common fragmentation ions are classified into −a, −b, −c,

Table 7
De novo sequencing results of M2 peptides.

Peptide sequence	ALC (%)	z	*m/z*	Mass	RT (min)	ppm	Mode
ETGSWNHEVEAAVPAGRLF	99.9	2	1035.50635	2068.9966	50.9549	1.5	HCD
SVHLF	99.6	2	301.66922	601.3224	27.1034	0.2	HCD
ETGSWNHEVEAAVPAGRL	99.3	2	961.97314	1921.9282	43.0266	5.3	HCD
LDNF	98.9	1	508.24261	507.2329	20.8217	3.4	HCD
LLGF	98.8	1	449.27744	448.2686	38.084	0.4	HCD
FTSF	98.8	1	501.23581	500.2271	27.8402	3	HCD
SHVKLER	98.6	3	290.17236	867.4926	10.1659	2.6	HCD
LADL	98.5	1	431.25095	430.2427	20.7011	7.2	HCD
FSLDF	98.3	1	628.2984	627.2904	48.8823	−4.2	HCD
GRLF	98.2	2	246.65031	491.2856	18.2975	2.5	HCD

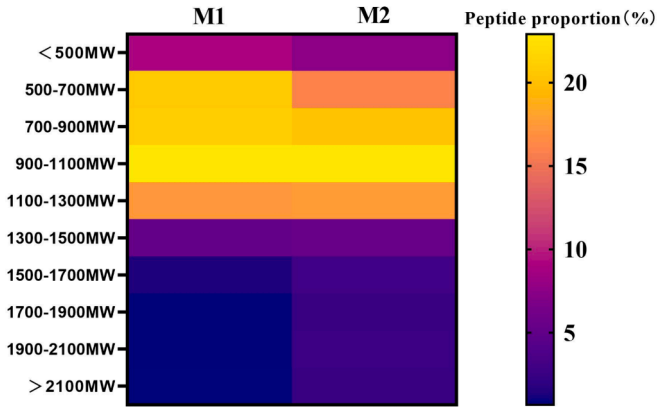

Fig. 5. The heat map of comparison between the two sources of peptides.

−x, −y, and −z based on their different cleavage positions in the peptide backbone. It is worth noting that the −b and −y ions, which are the primary ions generated by CID and HCD fragmentation methods, differ by one amino acid residue adjacent to each [41]. Therefore, these ions are mainly used in *de novo* peptide sequencing. In the *de novo* sequencing of M1 peptides in *P. cyrtonema*, the peptide sequence "LSNF" with the highest confidence level (ALC = 99.8 %) has a MW of 480.2461 and a y1 ion of 166.09. The corresponding amino acid residue has a MW of 165.09 (y1-H), indicating that it is a phenylalanine(F). By applying the same method to the y2-y1 and y3-y2 values of 114.04 and 87.03, respectively, their amino acid residue MW are determined to be 132.04 (y2-y1 + H_2O) and 87.03 + 18 = 105.03 (y3-y2 + H_2O), corresponding to tyrosine(N) and serine(S), respectively. Finally, by subtracting the y3 ion from the total MW value of 113.084, the amino acid residue MWs are determined to be 113.084 + H_2O = 131.173, corresponding to is determined to be leucine(L), confirming the sequence as LSNF. Similar to the above-mentioned method, the amino acid sequence can be inferred through the use of b ions. The measured and calculated values of the b and y ions for the peptide sequence LSNF, as well as its MS/MS spectrogram and ion fragmentation picture, can be found in Fig. 3, Fig. 4, Table 4, and Table 5. In addition, the partial results of *de novo* sequencing of M1 peptides and M2 peptides of *P. cyrtonema* can be observed in Table 6 and Table 7, showing the top 10 ranking based on ALC.

3.3. Differential analysis

3.3.1. Peptides overall differential analysis

Based on *De novo* sequenced peptide sequences, the study assessed the comprehensive differences between the M1 and M2 peptides, examining their MW, PI, GRAVY value, charge at pH 7.4, and retention time. The MWs of M1 peptides ranged from 294 to 3239 Da, averaging of 906.57 Da, while M2 peptides ranged from 278 to 3009 Da, averaging of 1002.76 Da. Predominantly, both types of peptides fell within the 500–1300 MW range, constituting 81.9 % and 76.6 % of the sample, respectively, indicating medium to long peptide lengths. However, M2 peptide, particularly within the 500–700 MW interval, display a distinct violet hue, suggesting a relatively lower abundance of peptides compared to M1 peptides within this mass range. The heat map of MW comparison between the two sources of peptides is shown in Fig. 5. The average PI distribution ranged from 2.8 to 12.0 for M1 peptides and from 2.7 to 12.0 for M2 peptides, with a higher proportion of M1 peptides within the 5.5–6.0 Ip range. In term of charges in pH 7.4, M2 peptides exhibited greater variability across a wide range compared to M1 peptides, with charges ranging from − 6 to 4.9 for M2, and − 4 to 3.9 for M1. The GRAVY value ranged from − 3.6 to 3.2 for M1 and from − 3.0 to 3.7 for M2. The retention time demonstrate a concentration of peptides towards the first half of the elution time. Collectively, based on these criteria, minimal disparities were observed between the peptides derived from the two sources. The comparison of axis distribution diagram average of PI, peptides charge states in pH 7.4, peptides GRAVY value and retention times from different peptides source are shown in Fig. 6, respectively.

3.3.2. Peptide differential analysis based on antioxidant bioactivity prediction

De novo sequencing results with ALC score greater than 60 % were utilized to predict biological activity. The bioactivities of peptides predicted by the BIOPEP-UWM database search mainly included DPP-IV inhibition, ACE inhibition, antioxidant activity, dipeptidyl peptidase III inhibitor and renin inhibitor. Specifically, the percentage for M1 peptides were 40.0 %, 31.2 %, 8.1 %, 6.5 % and 2.9 %, respectively, while for M2 peptides were 39.9 %, 32.0 %, 5.4 %, 5.5 % and 2.4 %. A comparison chart of predicted proportions of bioactivity of peptides from different sources are shown in Fig. 7. *De novo* sequencing result with an ALC score greater than60% were scored to filter out peptides, prioritizing those with top 100 AnOxPePred 1.0 scores and peptide ranker scores greater than 0.6. This resulted in 66 M1 peptides and 56 M2 peptides being selected, with partial results displayed in Tables 8 and 9, and comprehensive results in Tables S1 and S2. Analysis of sequence variations in antioxidant peptides revealed that M2 peptides (21), with peptide ranker scores exceeding 0.6, exhibited significantly higher capabilities for free radical scavenging and metal ion binding compared to M1 peptides (10). The Venn diagram of predicted proportions of bioactivity of peptides from different sources is shown in Fig. 8. "Structure-activity relationship" represents a fundamental attribute of BPs, wherein variances in structural attributes potentially give rise to disparate bioactivities. Notably, the presence of multi-BPs s suggests the possibility of shared structural characteristics, including shorter peptide length, lower MW, and an increased content of hydrophobic amino acids, among other factors [13,32,42]. Therefore, an analysis and comparison of the amino acid count, type, and arrangement were conducted on the selected peptide from different sources that potentially possess antioxidant bioactivity. The length and MW of peptides frequently influence the bioactivity of peptides, likely due to the fact that shorter peptides are more capable of fully exposing their active

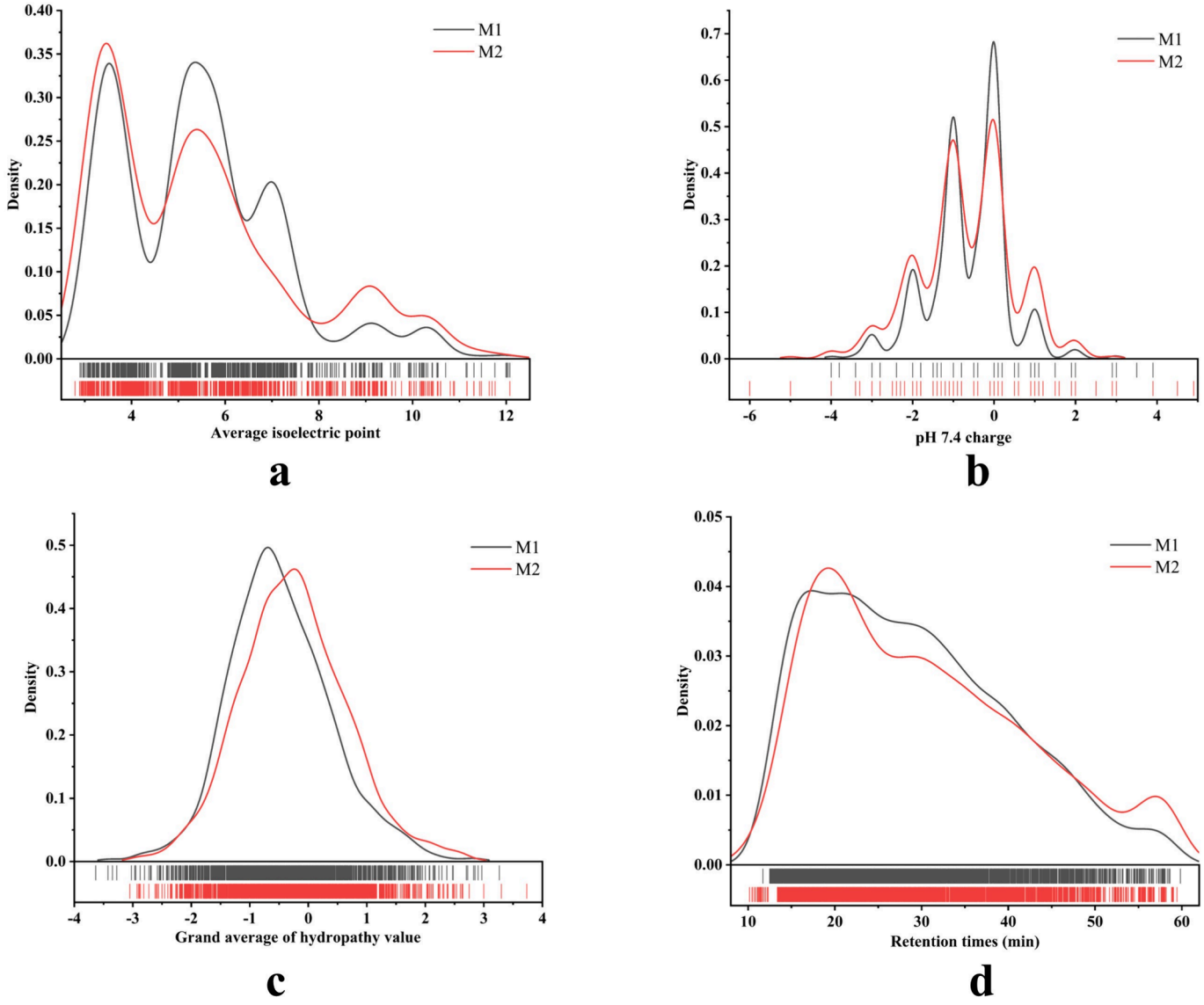

Fig. 6. Comparison of axis distribution diagram from different peptide source a: average isoelectric point; b: charges state in pH 7.4; c: GRAVY value d: retention times.

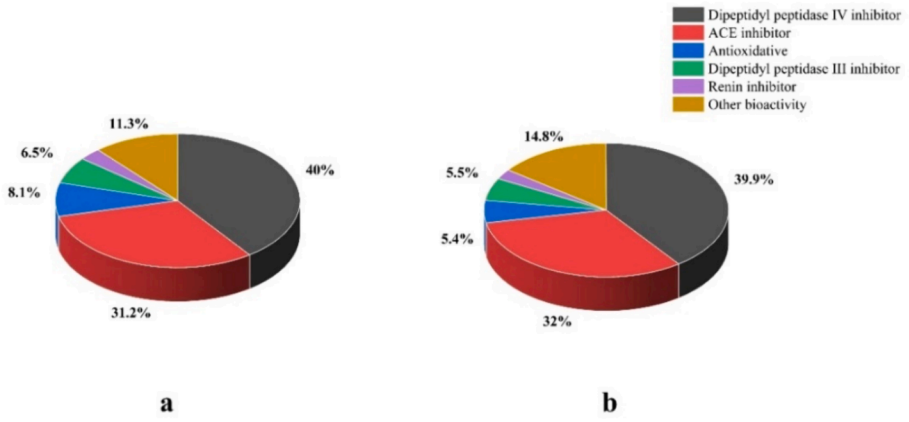

Fig. 7. Comparison chart of predicted proportions of bioactivity of peptides from different sources a: M1 peptides; b: M2 peptides.

Table 8
Prediction and screening results of M1 peptides antioxidant activity.

Peptide	Peptide ranker score	Amino acid counts	Molecular weight (Da)
FCMGF	0.994231	5	603.75174
YFWC	0.992844	4	617.71984
YWWTWDF	0.980126	7	1103.20114
HFGWY	0.971507	5	708.77404
HMYF	0.945458	4	596.70154
YFFEYF	0.941898	6	915.01254
HFYMTWHF	0.937693	8	1168.33754
AWGY	0.922049	4	495.53514
SGWY	0.917596	4	511.53454
MYGGGF	0.915297	6	630.71614

Table 9
Prediction and screening results of M2 peptides antioxidant activity.

Peptide	Peptide ranker score	Amino acid counts	Molecular weight (Da)
YHCCFPPF	0.984394	8	1013.19654
VWDPFHGWPF	0.981128	10	1287.44244
EWGGF	0.94254	5	594.62434
QNCGAYFF	0.930461	8	949.04844
YFGGGCPF	0.929305	8	846.95564
LWGGHLF	0.927747	7	828.96874
FHGSYFDYF	0.923859	9	1182.25684
FELPF	0.922217	5	651.76004
LLGF	0.914277	4	448.56254
YHCCFPPF	0.984394	8	1013.19654

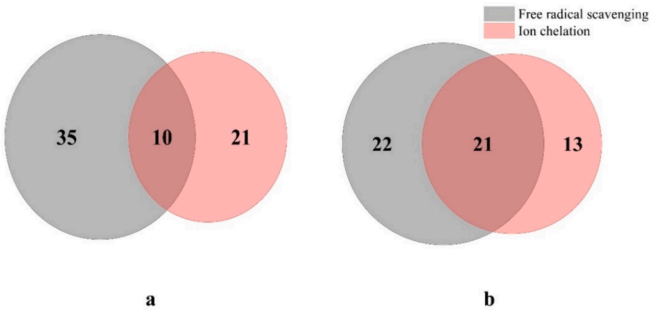

Fig. 8. The Venn diagram of predicted proportions of bioactivity of peptides from different sources a: M1 peptides; b: M2 peptides.

amino acid sites, and are also apt to more readily traverse biological membranes in order to initiate their effects via binding reactions [43–46]. Significant differences were observed in the length and MW of peptides from two different sources. M1 peptides are primarily comprised of 4–6 amino acids, accounting for 62.1 %, while M2 peptides are predominantly composed of 8–10 and greater than 10 amino acids, accounting for 60.8 %. Notably, peptides longer than 10 amino acids make up 27 % of the total. The average MW were 810.73 Da for M1 peptide and 1052.53 Da for M2 peptide, showing statistically significant difference. The comparison of number of amino acids in peptides from different sources is shown in Fig. 9. Secondly, amino acid composition is closely with peptides functionality. The content of hydrophobic amino acids and aromatic amino acids are generally enhance antioxidant activity [47]. As shown in Table 10, there were no significant differences in the proportions and distribution of hydrophobic amino acids and aromatic amino acids from different sources. Apart from amino acid composition, the positional arrangement of amino acids within peptides also serves as a critical determinant of peptide bioactivity. For instance, the hydrophobicity of amino acids positioned at the N-terminus or C-terminus of peptides is known to significantly influence their antioxidant capacity, ACE inhibitory activity, and anti-inflammatory ability [46,48,49]. As shown in Table 11, the proportion of peptides ending with hydrophobic amino acids was 76.2 % and 91.0 %, respectively. The significant difference in these data is primarily attributed to the considerably higher proportion of M2 peptides (30.3 %) with hydrophobic termini on both ends compared to M1 peptides (15.1 %).

3.3.3. Molecular docking

Aerobic biological cells are frequently subjected to internal metabolism, respiratory reactions, and environment oxidizing and electrophilic damage [50]. The transcriptional induction of antioxidant and cytoprotective enzymes in mammalian cells is mediated by several key signaling components and regulatory events [51]. Kelch-like ECH-associated protein 1 (KEAP1) features multiple stress sensors and deactivation mechanisms, which collectively accommodate a variety of cellular inputs. These input range from oxidative stress and cellular metabolites to disrupted autophagy, there by regulating the nuclear factor erythroid 2-related factor 2 (Nrf2) activity [52]. The accumulation of Nrf2 is critical for the expression of a series of phase II detoxifying enzymes such as superoxide dismutase (SOD) within the cell [53]. Therefore, external molecules (such as peptides) capable of binding to Keap1 can inhibit the Keap1-Nrf2 interaction, enhancing the cell's resistance to oxidative stress [36]. Using Autodock 4.2 to simulate the docking of antioxidant peptides with the Keap1-Nrf2 receptor protein and to calculate the binding energy, a higher absolute value of the binding energy indicates that a stronger likelihood of the peptide binding to the receptor protein. The top 10 binding energy results can be found in Table 12. 2D and 3D diagrams of peptide molecular docking are shown in Fig. 10. YFWC forms nine hydrogen bonds with Leu557, Ile416, Val512, Ile559, Val369, and forms carbon hydrogen bond with Gly605, and forms PI alkyl bonds with Arg605, Ala336, Val465, Cys368, Ala607, and forms PI lone pair with Leu365 with binding energy of − 10.82 kcal/mol. FELPF forms ten hydrogen bonds with Ile416, Ile559, Val418, Val465, Val606, Gly367 and forms carbon hydrogen bond with Gly605, and forms PI alkyl bonds with Ala466, Ala467, Ala556, Arg415, and forms alkyl bond with Cys513, and forms pi-sigma bonds with Gly364 and Gly419, with binding energy of − 9.67 kcal/mol. SGWY forms nine hydrogen bonds with Ile416, Ile559, Val418, Val465, Val606, Gly367, Leu557, Arg415 and forms carbon hydrogen bonds with Gly367, Val465 and forms pi-donor hydrogen bonds with Ile559, and forms pi-sigma bond with Val606, and forms pi-alkyl bonds with Ala556, Arg415 with binding energy of − 9.24 kcal/mol. HMYF forms eight hydrogen bonds with Ile559, Val512, Val465, Val463, Leu557, and forms carbon hydrogen bonds with Gly511, Gly558 and forms pi-donor hydrogen bonds with Val418, and forms pi-sigma bond with Ala366, and forms alkyl bonds with Ala607, Cys368, Val369 with binding energy of − 9.19 kcal/mol. AWGY forms eight hydrogen bonds with Val416, Val418, Val465, Val463, Val604, Val606, Leu365, Ile416 and forms carbon hydrogen bond with Ile559 and forms pi-Alkyl bond with Arg415, with binding energy of − 9.19 kcal/mol. It is worth noting that among these five potential antioxidant peptides, the hydrophobic amino acids and aromatic amino acids of the YFWC and AWGY accounted for 100 % of the residues, while the other three peptides accounted for over 75 %. This demonstrates that the presence of hydrophobic amino acid and aromatic amino acid residues is beneficial for the interaction between the peptide and Keap1 protein, leading to the formation of a more stable complex with the receptor protein and interfering with the Keap1-Nrf2 PPI, thus possessing strong antioxidant potential [54,55].

4. Discussion

Research on endogenous BPs in medicinal plants has been limited over the past decade due to their relatively low abundance and challenges associated with directly extracting high-purity peptide monomers [12,13]. However, the lack of attention has not impeded advancement of endogenous peptides in other areas. For instance, over 100 endogenous

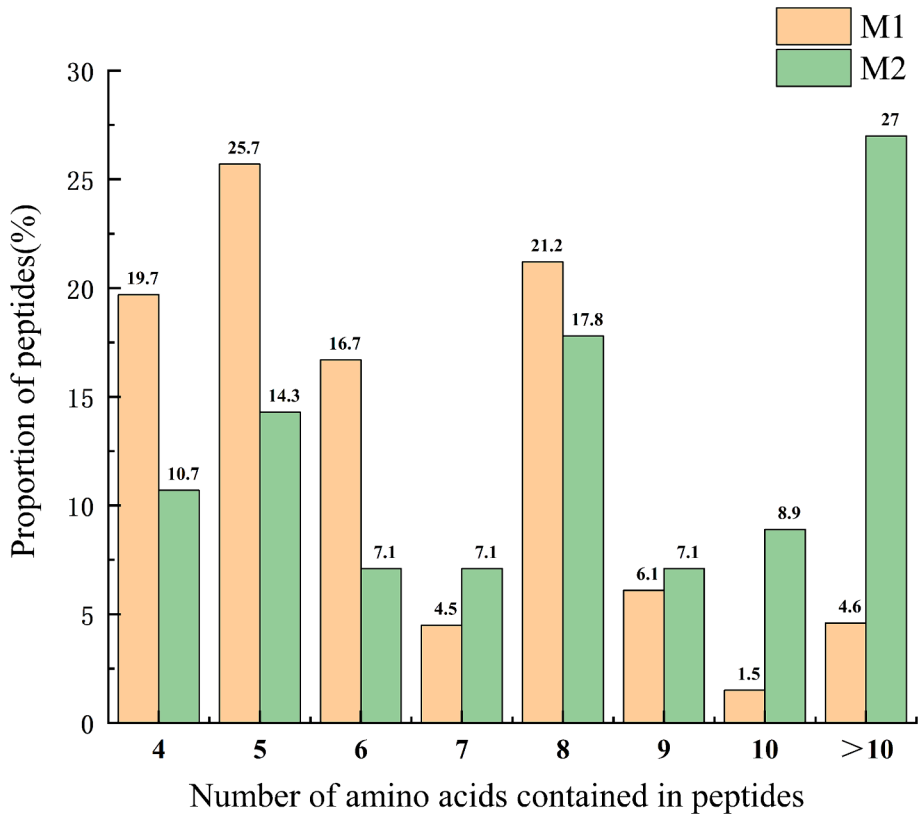

Fig. 9. Comparison of number of amino acids in peptides from different sources.

Table 10
Comparison of amino acid composition of peptides from different sources.

Amino acids	M1 peptides	M2 peptides
Trp (Hydrophobic/Aromatic)	23	22
Phe (Hydrophobic/Aromatic)	61	73
Val (Hydrophobic)	6	29
Leu (Hydrophobic)	14	46
Ala (Hydrophobic)	31	20
Met (Hydrophobic)	12	8
Tyr (Aromatic)	55	45
Total	202	243
Proportion	51.6 %	49.1 %

Table 11
Comparison of sequence of amino acid in peptides from different sources.

	M1 Peptides	M2 peptides
Hydrophobic amino acids at both ends	15.1 %	30.3 %
Hydrophobic amino acids at single end	61.1 %	60.7 %
Total	76.2 %	91.0 %

Table 12
Top 10 peptide binding energy results.

Peptide	Peptide ranker	Amino acid counts	Molecular weight	Binding Energy (kcal/mol)
YFWC	0.992844	4	617.71984	−10.82
FELPF	0.922217	5	651.76004	−9.67
SGWY	0.917596	4	511.53454	−9.24
HMYF	0.945458	4	596.70154	−9.19
AWGY	0.922049	4	495.53514	−9.19
LLGF	0.914277	4	448.56254	−9.07
EWGGF	0.94254	5	594.62434	−8.02
MYGGGF	0.915297	6	630.71614	−7.82
FCMGF	0.994231	5	603.75174	−6.67
HFGWY	0.971507	5	708.77404	−1.96

peptides have been identified in mammalian systems, with their underlying mechanisms have been leveraged to develop therapeutic drugs [56–58]. In this experiment, Peaks DB was used to initially identify and screen 23 M1 peptides and 37 M2 peptides, setting a threshold of −10lgP > 20. These findings contrast with those presented by Ye et al.[14], who reported 125 peptides having been identified in *Panax ginseng* C. A. Meyer. The discrepancy might be attributed to the insufficient research on TCM proteins, resulting in a fewer of protein entries in the database, which subsequently limits matching of mass spectrometry data. The identification from the M2 peptides dataset (123 entires) showed a significant increase compared to that from the M1 peptides (45 entires). Furthermore, the comparison of nine identical peptides identified from two different sources showed higher −10lgP values and peak areas overall, indicating that the crude protein extracted through alkali and acid precipitation method releases more peptides from *P. cyrtonema* compared to the herb itself. Moreover, the M1 peptides remained relatively intact even under extreme pH conditions. Furthermore, the increase in protein concentration of sample had a positive impact on improving the confidence of peptide mass spectrometry identification. Through *de novo* sequencing, a total of 2,515 M1 peptides and 2,065 M2 peptides were obtained from *P. cyrtonema*. This is in contrast to the 183 endogenous peptides (ALC > 50 %) identified by Ye et al [14] from *Panax ginseng* C. A. Meyer, the 799 endogenous peptides identified by Bai et al [59] from *Bombyx batryticatus*, and the 145 endougenous peptides identified by Liu et al [60] from Isatidis Radix through a combination of *de novo* sequencing and database searching. Considering these figures, *P. cyrtonema* emerges as a plant with significant potential in the field of BPs. The amino acid sequences were confirmed to be identical to those deduced through manual calculations using ion difference. Regarding the overall analysis of peptide variances, the

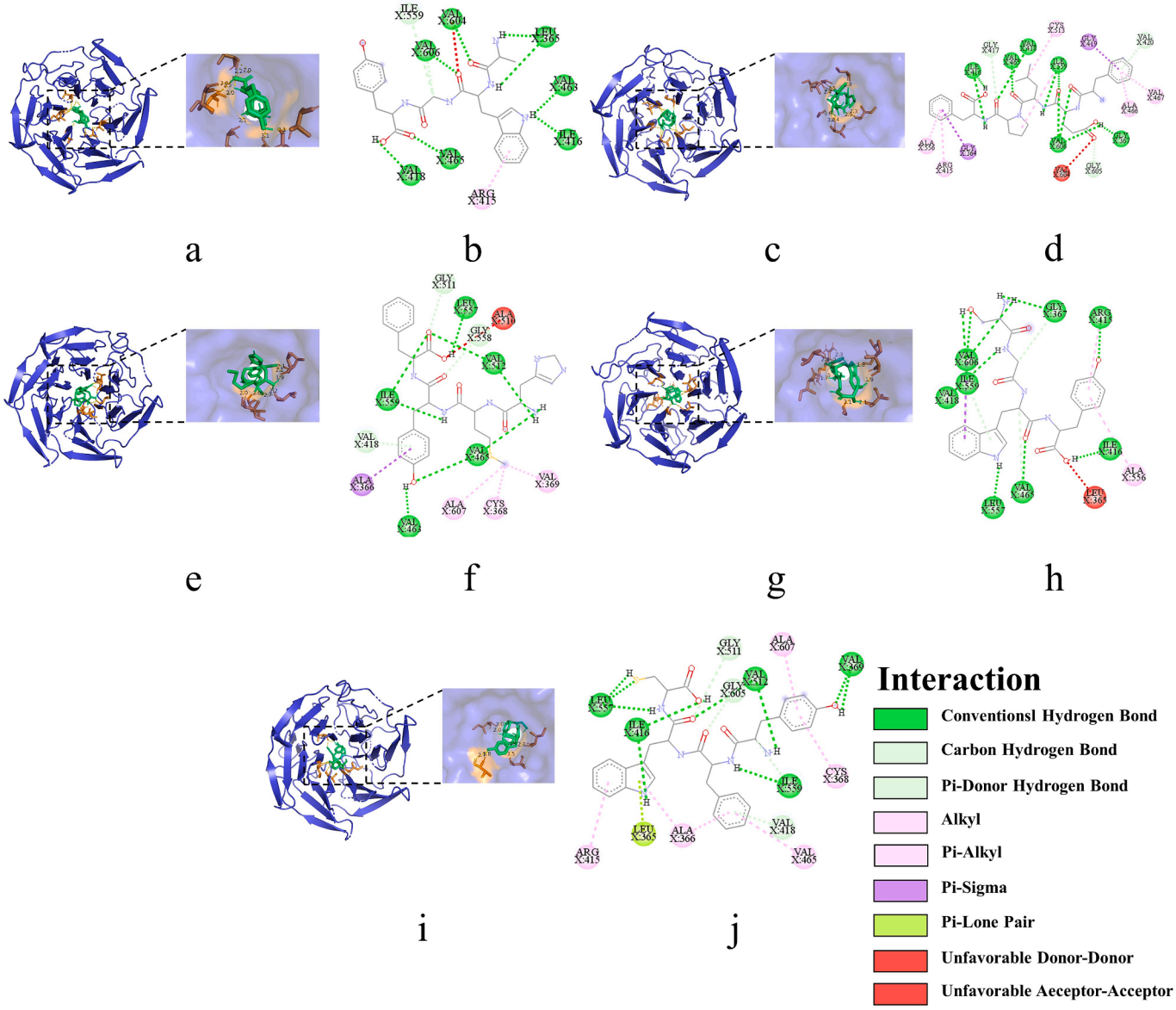

Fig. 10. 2D and 3D diagrams of peptide molecule docking a,b: YFWC; c,d: FELPF; e,f: SGWY g,h: HMYF; i,j: AWGY.

differences in peptides from two different sources are not significantly distinct. It is notable that the average molecular weight (MW) of the M2 peptides is slightly higher than that of the M1 peptides, likely due to the alkali and acid precipitation process hydrolyzing a small amount of short peptides into free amino acids. Additionally, the proportion of M2 peptides with an PI between 6–8 is significantly lower than that of M1 peptides, which predominantly have PIs greater than 8. The proportion of M2 peptides uncharged at pH 7.4 is also significantly higher than that of M1 peptides [13,61–63]. In the analysis based on predicted antioxidant activity, selected peptides exhibited similar patterns to the overall differential analysis [46,64]. Notably, the average MWs of the M2 peptide (1052.53 Da), selected through antioxidant activity prediction, is significantly higher than that of the M1 peptide (810.73 Da). This is attributed to M2 peptide primarily consisting of medium-length peptides with more than 8 amino acids, while the M1 peptide is mainly composed of short peptides with 4–6 amino acids. The longer M2 peptides are predicted to have higher dual-function antioxidant potential (free radical scavenging and ion chelation) than M1 peptides [65]. Regarding molecular docking, of the top 10 peptides with the highest absolute binding energy, seven were derived from M1 peptides, indicating that peptide with antioxidant potential are generally found in short to medium lengths, consistent with previous research findings [46].

5. Conclusion

Overall, this experiment employed a Nano-LC-Q Exactive Plus Quadrupole-Orbitrap MS in combination with PEAK studio software to successfully identified and analyzed 2571 M1 peptides of *P. cyrtonema*, as well as 2122 M2 peptides obtained through an alkali extraction-acid precipitation process. The overall differences were assessed based on the peptide sequences obtained, with a focus on antioxidant activity prediction, and their antioxidant potential was further predicted through bioinformatics analysis and molecular docking. The results indicate that *P. cyrtonema* is a pharmaceutical resource rich in endogenous peptides. Furthermore, increasing the protein concentration of the samples through alkali extraction and acid precipitation significantly contributes to enhancing the credibility of peptide identification. Additionally, compared to the M1 peptide, the M2 peptide, due to its longer peptide sequence, has the potential to contain a greater variety of antioxidative segments, thus demonstrating multifaceted antioxidative potential.

However, it is noteworthy that molecular docking results indicate a greater antioxidative potential for the M1 peptide compared to the M2 peptide. This demonstrates that MWs remains the primary factor influencing the antioxidative potential of peptides. Moreover, it is essential to enzymatically hydrolyze the M2 peptide into lower MWs peptides to expose more antioxidative groups, which is currently the mainstream method for preparing antioxidative peptides in research. Therefore, in the future, it is crucial to enhance the detection resolution of mass spectrometers to reduce ion loss for improved accuracy in detecting endogenous peptides in natural products. Furthermore, the development or improvement of peptide bioactivity prediction accuracy is essential for exploring structure–activity relationships and identifying potential antioxidative peptides. Simultaneously, the development of more targeted methods for preparing bioactive peptides is also a key area for advancement. Lastly, thorough in *vivo* and in *vitro* biological experiments are necessary to validate the predicted antioxidative activity.

CRediT authorship contribution statement

Yanyan Zhang: Writing – original draft, Visualization, Supervision, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Haixia Li:** Formal analysis, Data curation. **Peizi Liu:** Data curation. **Keyi Chen:** Data curation. **Shengjun Ma:** Supervision, Conceptualization. **Wei Cai:** Writing – review & editing, Project administration, Methodology, Conceptualization.

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.

Data availability

Data will be made available on request.

Acknowledgments

This project was financially supported by the Science and Technology Innovation Program of Hunan Province (no. 2022RC1228), Hunan Province Social Science Innovation Research Base (Ethnic medicine and ethnic culture research base).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.microc.2024.110872.

References

[1] L. Li, K. Thakur, B.-Y. Liao, J.-G. Zhang, Z.-J. Wei, Antioxidant and antimicrobial potential of polysaccharides sequentially extracted from Polygonatum cyrtonema Hua, Int. J. Biol. Macromol. 114 (2018) 317–323, https://doi.org/10.1016/j.ijbiomac.2018.03.121.

[2] The genus Polygonatum, A review of ethnopharmacology, phytochemistry and pharmacology, J. Ethnopharmacol. 214 (2018) 274–291, https://doi.org/10.1016/j.jep.2017.12.006.

[3] F. Liu, Y. Liu, Y. Meng, M. Yang, K. He, Structure of polysaccharide from Polygonatum cyrtonema Hua and the antiherpetic activity of its hydrolyzed fragments, Antiviral Res. 63 (2004) 183–189, https://doi.org/10.1016/j.antiviral.2004.04.006.

[4] Saboon, Y. Bibi, M. Arshad, S. Sabir, M.S. Amjad, E. Ahmed, S.K. Chaudhari, Pharmacology and biochemistry of Polygonatum verticillatum: A review, Journal of Coastal Life medicine 4 (2016) 406–415. https://doi.org/10.12980/jclm.4.2016J5-228.

[5] L.-L. Su, X. Li, Z.-J. Guo, X.-Y. Xiao, P. Chen, J.-B. Zhang, C.-Q. Mao, D. Ji, J. Mao, B. Gao, T.-L. Lu, Effects of different steaming times on the composition, structure and immune activity of Polygonatum Polysaccharide, J. Ethnopharmacol. 310 (2023) 116351, https://doi.org/10.1016/j.jep.2023.116351.

[6] C. Wang, D. Peng, J. Zhu, D. Zhao, Y. Shi, S. Zhang, K. Ma, J. Wu, L. Huang, Transcriptome analysis of Polygonatum cyrtonema Hua: identification of genes involved in polysaccharide biosynthesis, Plant Methods 15 (2019) 65, https://doi.org/10.1186/s13007-019-0441-9.

[7] G. Liu, S. Feng, M. Sui, B. Chen, P. Sun, Extraction and identification of steroidal saponins from Polygonatum cyrtonema Hua using natural deep eutectic solvent-synergistic quartz sand-assisted extraction method, J. Sep. Sci. 46 (2023) e2200823.

[8] W. Wang, X. Dabu, J. He, H. Yang, S. Yang, J. Chen, W. Fan, G. Zhang, J. Cai, H. Ai, M. Hai, Polygonatone H, a new homoisoflavanone with cytotoxicity from Polygonatum Cyrtonema Hua, Nat. Prod. Res. 33 (2019) 1727–1733, https://doi.org/10.1080/14786419.2018.1434645.

[9] P.J. Sreelekshmi, V. Devika, L.S. Aiswarya, S.R. Jeevan, K. Ramanunni, P.B. Nair, S. Sadanandan, Recent Advances in Bioactive Peptides as Functional Food for Health Promotions and Medicinal Applications, Protein Pept. Lett. 30 (2023) 626–639, https://doi.org/10.2174/0929866530666230706104923.

[10] S. Meena, B. Kanthaliya, A. Joshi, F. Khan, J. Arora, Biologia futura: medicinal plants-derived bioactive peptides in functional perspective—a review, BIOLOGIA FUTURA 71 (2020) 195–208, https://doi.org/10.1007/s42977-020-00042-4.

[11] A.C. Sabbione, S.M. Ibañez, E.N. Martínez, M.C. Añón, A.A. Scilingo, Antithrombotic and Antioxidant Activity of Amaranth Hydrolysate Obtained by Activation of an Endogenous Protease, Plant Foods Hum. Nutr. 71 (2016) 174–182, https://doi.org/10.1007/s11130-016-0540-y.

[12] S. Piovesana, A.L. Capriotti, C. Cavaliere, G. La Barbera, C.M. Montone, R. Zenezini Chiozzi, A. Laganà, Recent trends and analytical challenges in plant bioactive peptide separation, identification and validation, Anal Bioanal Chem 410 (2018) 3425–3444. https://doi.org/10.1007/s00216-018-0852-x.

[13] Y. Zhang, L. Liu, M. Zhang, S. Li, J. Wu, Q. Sun, S. Ma, W. Cai, The Research Progress of Bioactive Peptides Derived from Traditional Natural Products in China, Molecules 28 (2023) 6421, https://doi.org/10.3390/molecules28176421.

[14] X. Ye, N. Zhao, X. Yu, X. Han, H. Gao, X. Zhang, Extensive characterization of peptides from *Panax ginseng* C. A. Meyer using mass spectrometric approach, Proteomics 16 (2016) 2788–2791, https://doi.org/10.1002/pmic.201600183.

[15] F.A. Ortiz-Morea, D.V. Savatin, W. Dejonghe, R. Kumar, Y. Luo, M. Adamowski, J. Van den Begin, K. Dressano, G. Pereira de Oliveira, X. Zhao, Q. Lu, A. Madder, J. Friml, D. Scherer de Moura, E. Russinova, Danger-associated peptide signaling in Arabidopsis requires clathrin, Proc Natl Acad Sci U S A 113 (2016) 11028–11033, https://doi.org/10.1073/pnas.1605588113.

[16] X. Dou, D. Yan, S. Liu, N. Gao, Z. Ma, Z. Shi, N. Dong, A. Shan, Host Defense Peptides in Nutrition and Diseases: A Contributor of Immunology Modulation, J. Agric. Food Chem. (2023), https://doi.org/10.1021/acs.jafc.2c08522.

[17] M. Kumar, M. Tomar, J. Potkule, R. Verma, S. Punia, A. Mahapatra, T. Belwal, A. Dahuja, S. Joshi, M.K. Berwal, V. Satankar, A.G. Bhoite, R. Amarowicz, C. Kaur, J.F. Kennedy, Advances in the plant protein extraction: Mechanism and recommendations, Food Hydrocoll. 115 (2021) 106595, https://doi.org/10.1016/j.foodhyd.2021.106595.

[18] Q. Cui, X. Ni, L. Zeng, Z. Tu, J. Li, K. Sun, X. Chen, X. Li, Optimization of Protein Extraction and Decoloration Conditions for Tea Residues, Hortic. Plant J. 3 (2017) 172–176, https://doi.org/10.1016/j.hpj.2017.06.003.

[19] K. Fields, T.J. Falla, K. Rodan, L. Bush, Bioactive peptides: signaling the future, J. Cosmet. Dermatol. 8 (2009) 8–13, https://doi.org/10.1111/j.1473-2165.2009.00416.x.

[20] R. Craig, R.C. Beavis, TANDEM: matching proteins with tandem mass spectra, Bioinformatics 20 (2004) 1466–1467, https://doi.org/10.1093/bioinformatics/bth092.

[21] J.K. Eng, B.C. Searle, K.R. Clauser, D.L. Tabb, A face in the crowd: recognizing peptides through database search, Mol Cell Proteomics 10 (2011) R111.009522. https://doi.org/10.1074/mcp.R111.009522.

[22] A.I. Nesvizhskii, Protein Identification by Tandem Mass Spectrometry and Sequence Database Searching, in: Mass Spectrometry Data Analysis in Proteomics, Humana Press, New Jersey, 2006: pp. 87–120. https://doi.org/10.1385/1-59745-275-0:87.

[23] J. Doellinger, L. Schaade, A. Nitsche, Comparison of the Cowpox Virus and Vaccinia Virus Mature Virion Proteome: Analysis of the Species- and Strain-Specific Proteome, PLoS One 10 (2015) e0141527.

[24] R.S. Johnson, M.T. Davis, J.A. Taylor, S.D. Patterson, Informatics for protein identification by mass spectrometry, Methods 35 (2005) 223–236, https://doi.org/10.1016/j.ymeth.2004.08.014.

[25] D. Kumar, A.K. Yadav, D. Dash, Choosing an Optimal Database for Protein Identification from Tandem Mass Spectrometry Data, in: S. Keerthikumar, S. Mathivanan (Eds.), Proteome Bioinformatics, Springer New York, New York, NY, 2017: pp. 17–29. https://doi.org/10.1007/978-1-4939-6740-7_3.

[26] L. Zhong, L. Zhu, Z.-W. Cai, Mass Spectrometry-based Proteomics and Glycoproteomics in COVID-19 Biomarkers Identification: A Mini-review, J. Anal. Test. 5 (2021) 298–313, https://doi.org/10.1007/s41664-021-00197-6.

[27] T.F. Rakotondrabe, M.-X. Fan, Y.-L. Zhang, M.-Q. Guo, Simultaneous Screening and Analysis of Anti-inflammatory and Antiproliferative Compounds from Euphorbia maculata Combining Bio-affinity Ultrafiltration with Multiple Drug Targets, Journal of Analysis and Testing 6 (2022), https://doi.org/10.1007/s41664-022-00225-z.

[28] M. Depke, S. Michalik, A. Rabe, K. Surmann, L. Brinkmann, N. Jehmlich, J. Bernhardt, M. Hecker, B. Wollscheid, Z. Sun, R.L. Moritz, U. Völker, F. Schmidt, A peptide resource for the analysis of Staphylococcus aureus in host-pathogen interaction studies, Proteomics 15 (2015) 3648–3661, https://doi.org/10.1002/pmic.201500091.

[29] V. Vijayakumar, A.N. Guerrero, N. Davey, C.B. Lebrilla, D.C. Shields, N. Khaldi, EnzymePredictor: a tool for predicting and visualizing enzymatic cleavages of

digested proteins, J. Proteome Res. 11 (2012) 6056–6065, https://doi.org/10.1021/pr300721f.
[30] J. Ning, M. Li, W. Chen, M. Yang, J. Chen, X. Luo, X. Yue, Characterization and biological function analysis of endogenous peptides derived from donkey colostrum proteins, Food Funct. 14 (2023) 8261–8275, https://doi.org/10.1039/D3FO01703F.
[31] P. Minkiewicz, A. Iwaniak, M. Darewicz, BIOPEP-UWM Database of Bioactive Peptides: Current Opportunities, Int. J. Mol. Sci. 20 (2019) 5978, https://doi.org/10.3390/ijms20235978.
[32] C. Mooney, N.J. Haslam, G. Pollastri, D.C. Shields, Towards the Improved Discovery and Design of Functional Peptides: Common Features of Diverse Classes Permit Generalized Prediction of Bioactivity, PLoS One 7 (2012) e45012.
[33] T.H. Olsen, B. Yesiltas, F.I. Marin, M. Pertseva, P.J. García-Moreno, S. Gregersen, M.T. Overgaard, C. Jacobsen, O. Lund, E.B. Hansen, P. Marcatili, AnOxPePred: using deep learning for the prediction of antioxidative properties of peptides, Sci. Rep. 10 (2020) 21471, https://doi.org/10.1038/s41598-020-78319-w.
[34] J. Han, S. Tang, Y. Li, W. Bao, H. Wan, C. Lu, J. Zhou, Y. Li, L. Cheong, X. Su, In silico analysis and in vivo tests of the tuna dark muscle hydrolysate anti-oxidation effect, RSC Adv. 8 (2018) 14109–14119, https://doi.org/10.1039/c8ra00889b.
[35] J. Zhang, M. Li, G. Zhang, Y. Tian, F. Kong, S. Xiong, S. Zhao, D. Jia, A. Manyande, H. Du, Identification of novel antioxidant peptides from snakehead (Channa argus) soup generated during gastrointestinal digestion and insights into the anti-oxidation mechanisms, Food Chem. 337 (2021) 127921, https://doi.org/10.1016/j.foodchem.2020.127921.
[36] X. Hu, J. Liu, J. Li, Y. Song, S. Chen, S. Zhou, X. Yang, Preparation, purification, and identification of novel antioxidant peptides derived from Gracilariopsis lemaneiformis protein hydrolysates, Front. Nutr. 9 (2022) 971419, https://doi.org/10.3389/fnut.2022.971419.
[37] N.M. O'Boyle, C. Morley, G.R. Hutchison, Pybel: a Python wrapper for the OpenBabel cheminformatics toolkit, Chem. Cent. J. 2 (2008) 5, https://doi.org/10.1186/1752-153X-2-5.
[38] F.L. Nii Adoquaye Acquaye, A. Kertesz-Farkas, W.S. Noble, Efficient Indexing of Peptides for Database Search Using Tide, J. Proteome Res. 22 (2023) 577–584, https://doi.org/10.1021/acs.jproteome.2c00617.
[39] J. Zhang, L. Xin, B. Shan, W. Chen, M. Xie, D. Yuen, W. Zhang, Z. Zhang, G.A. Lajoie, B. Ma, PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification, Mol Cell Proteomics 11 (2012) M111.010587. https://doi.org/10.1074/mcp.M111.010587.
[40] T. Muth, B.Y. Renard, Evaluating de novo sequencing in proteomics: already an accurate alternative to database-driven peptide identification? Brief. Bioinform. 19 (2018) 954–970, https://doi.org/10.1093/bib/bbx033.
[41] Y. Yan, A.J. Kusalik, F.-X. Wu, De novo peptide sequencing using CID and HCD spectra pairs, Proteomics 16 (2016) 2615–2624, https://doi.org/10.1002/pmic.201500251.
[42] C. Acquah, Y.W. Chan, S. Pan, D. Agyei, C.C. Udenigwe, Structure-informed separation of bioactive peptides, J. Food Biochem. 43 (2019) e12765.
[43] S.K. Ns, N. Ra, J. r,, Purification and identification of antioxidant peptides from the skin protein hydrolysate of two marine fishes, horse mackerel (Magalaspis cordyla) and croaker (Otolithes ruber), Amino Acids 42 (2012), https://doi.org/10.1007/s00726-011-0858-6.
[44] R. Natesh, S.L.U. Schwager, E.D. Sturrock, K.R. Acharya, Crystal structure of the human angiotensin-converting enzyme-lisinopril complex, Nature 421 (2003) 551–554, https://doi.org/10.1038/nature01370.
[45] B.H. Sarmadi, A. Ismail, Antioxidative peptides from food proteins: a review, Peptides 31 (2010) 1949–1956, https://doi.org/10.1016/j.peptides.2010.06.020.
[46] T.-B. Zou, T.-P. He, H.-B. Li, H.-W. Tang, E.-Q. Xia, The Structure-Activity Relationship of the Antioxidant Peptides from Natural Proteins, Molecules 21 (2016) 72, https://doi.org/10.3390/molecules21010072.
[47] S.C. Cheison, Z. Wang, S.-Y. Xu, Preparation of whey protein hydrolysates using a single- and two-stage enzymatic membrane reactor and their immunological and antioxidant properties: characterization by multivariate data analysis, J. Agric. Food Chem. 55 (2007) 3896–3904, https://doi.org/10.1021/jf062936i.
[48] S. Guha, K. Majumder, Structural-features of food-derived bioactive peptides with anti-inflammatory activity: A brief review, J. Food Biochem. 43 (2019) e12531, https://doi.org/10.1111/jfbc.12531.
[49] Y. Li, J. Yu, Research Progress in Structure-Activity Relationship of Bioactive Peptides, J. Med. Food 18 (2015) 147–156, https://doi.org/10.1089/jmf.2014.0028.
[50] L. Baird, M. Yamamoto, The Molecular Mechanisms Regulating the KEAP1-NRF2 Pathway, Mol. Cell Biol. 40 (2020) e00099–e00120, https://doi.org/10.1128/MCB.00099-20.
[51] F. Hong, K.R. Sekhar, M.L. Freeman, D.C. Liebler, Specific patterns of electrophile adduction trigger Keap1 ubiquitination and Nrf2 activation, J. Biol. Chem. 280 (2005) 31768–31775, https://doi.org/10.1074/jbc.M503346200.
[52] S. Liu, J. Pi, Q. Zhang, Signal amplification in the KEAP1-NRF2-ARE antioxidant response pathway, Redox Biol. 54 (2022) 102389, https://doi.org/10.1016/j.redox.2022.102389.
[53] D.D. Zhang, M. Hannink, Distinct cysteine residues in Keap1 are required for Keap1-dependent ubiquitination of Nrf2 and for stabilization of Nrf2 by chemopreventive agents and oxidative stress, Mol Cell Biol 23 (2003) 8137–8151, https://doi.org/10.1128/MCB.23.22.8137-8151.2003.
[54] K. Wang, L. Han, H. Hong, J. Pan, H. Liu, Y. Luo, Purification and identification of novel antioxidant peptides from silver carp muscle hydrolysate after simulated gastrointestinal digestion and transepithelial transport, Food Chem. 342 (2021) 128275, https://doi.org/10.1016/j.foodchem.2020.128275.
[55] S. He, Y. Zhang, H. Sun, M. Du, J. Qiu, M. Tang, X. Sun, B. Zhu, Antioxidative Peptides from Proteolytic Hydrolysates of False Abalone (Volutharpa ampullacea perryi): Characterization, Identification, and Molecular Docking, Mar Drugs 17 (2019) 116, https://doi.org/10.3390/md17020116.
[56] L.D. Fricker, J. Lim, H. Pan, F.-Y. Che, Peptidomics: identification and quantification of endogenous peptides in neuroendocrine tissues, Mass Spectrom. Rev. 25 (2006) 327–344, https://doi.org/10.1002/mas.20079.
[57] K.D.B. Anapindi, E.V. Romanova, J.W. Checco, J.V. Sweedler, Mass Spectrometry Approaches Empowering Neuropeptide Discovery and Therapeutics, Pharmacol. Rev. 74 (2022) 662–679, https://doi.org/10.1124/pharmrev.121.000423.
[58] J.W. Checco, Identifying and Measuring Endogenous Peptides through Peptidomics, ACS Chem. Nerosci. 14 (2023) 3728–3731, https://doi.org/10.1021/acschemneuro.3c00546.
[59] Y. Bai, Q. Zhao, M. He, X. Ye, X. Zhang, Extensive characterization and differential analysis of endogenous peptides from Bombyx batryticatus using mass spectrometric approach, J. Pharm. Biomed. Anal. 163 (2019) 78–87, https://doi.org/10.1016/j.jpba.2018.09.033.
[60] X. Liu, C. Zhang, L. He, M. Lin, Study on Identification and Activity Virtual Screening of Endogenous Peptides from Isatidis Radix Based on Nano LC-MS / MS, Journal of Chinese Medicinal Materials 44 (2021) 1149–1154, https://doi.org/10.13863/j.issn1001-4454.2021.05.021.
[61] C. Zhang, J.P.M. Sanders, T.T. Xiao, M.E. Bruins, How Does Alkali Aid Protein Extraction in Green Tea Leaf Residue: A Basis for Integrated Biorefinery of Leaves, PLoS One 10 (2015) e0133046.
[62] F. Zhu, J. Cao, Y. Song, P. Yu, E. Su, Plant Protein-Derived Active Peptides: A Comprehensive Review, J. Agric. Food Chem. 71 (2023) 20479–20499, https://doi.org/10.1021/acs.jafc.3c06882.
[63] Y.W. Sari, W.J. Mulder, J.P.M. Sanders, M.E. Bruins, Towards plant protein refinery: Review on protein extraction using alkali and potential enzymatic assistance, Biotechnol. J. 10 (2015) 1138–1157, https://doi.org/10.1002/biot.201400569.
[64] R.J. Elias, S.S. Kellerby, E.A. Decker, Antioxidant Activity of Proteins and Peptides, Crit. Rev. Food Sci. Nutr. 48 (2008) 430–441, https://doi.org/10.1080/10408390701425615.
[65] Y. Zhang, Y. Li, T. Ren, P. Xiao, J.-A. Duan, Novel and efficient techniques in the discovery of antioxidant peptides, Crit. Rev. Food Sci. Nutr. (2023) 1–15, https://doi.org/10.1080/10408398.2023.2245052.