Bioinformatics, 38(12), 2022, 3275–3280
https://doi.org/10.1093/bioinformatics/btac323
Advance Access Publication Date: 13 May 2022
Original Paper

Databases and ontologies

DFBP: a comprehensive database of food-derived bioactive peptides for peptidomics research

Dongya Qin[1], Weichen Bo[1], Xin Zheng[1], Youjin Hao[2], Bo Li [2,*], Jie Zheng[3,*] and Guizhao Liang [1,*]

[1]Key Laboratory of Biorheological Science and Technology, Ministry of Education, Bioengineering College, Chongqing University, Chongqing 400044, China, [2]College of Life Sciences, Chongqing Normal University, Chingqing 401331, China and [3]Department of Chemical, Biomolecular, and Corrosion Engineering, The University of Akron, Akron, OH 44325, USA

*To whom correspondence should be addressed.

Associate Editor: Zhiyong Lu

Received on October 21, 2021; revised on March 31, 2022; editorial decision on May 7, 2022; accepted on May 9, 2022

Abstract

Motivation: Food-derived bioactive peptides (FBPs) have demonstrated their significance in pharmaceuticals, diets and nutraceuticals, benefiting public health and global ecology. While significant efforts have been made to discover FBPs and to elucidate the underlying bioactivity mechanisms, there is lack of a systemic study of sequence–structure–activity relationship of FBPs in a large dataset.

Results: Here, we construct a database of food-derived bioactive peptides (DFBP), containing a total of 6276 peptide entries in 31 types from different sources. Further, we develop a series of analysis tools for function discovery/repurposing, traceability, multifunctional bioactive exploration and physiochemical property assessment of peptides. Finally, we apply this database and data-mining techniques to discover new FBPs as potential drugs for cardiovascular diseases. The DFBP serves as a useful platform for not only the fundamental understanding of sequence–structure–activity of FBPs but also the design, discovery, and repurposing of peptide-based drugs, vaccines, materials and food ingredients.

Availability and implementation: DFBP service can be accessed freely via http://www.cqudfbp.net/. All data are incorporated into the article and its online supplementary material.

Contact: gzliang@cqu.edu.cn or zhengj@uakron.edu or libcell@cqnu.edu.cn

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Food-derived bioactive peptides (FBPs), derived from edible cereal proteins (Gong *et al.*, 2020), vegetables (Chan-Zapata *et al.*, 2020), milk (Nongonierma and FitzGerald, 2016), animal by-products (Bechaux *et al.*, 2019), seafood (Pavlicevic *et al.*, 2020) and microorganisms (Zhou *et al.*, 2020), have demonstrated their significant impacts on a wide variety of biomedical, nutraceutical and pharmaceutical applications, including targeted drugs, biological materials, functional foods and antibacterial ingredients (Barati *et al.*, 2020; Hajfathalian *et al.*, 2018; Nongonierma and FitzGerald, 2017). Apart from common merits (e.g. rich sources, easy absorption, production at low cost, sequence diversity and hypoallergenicity), FBPs of 2–20 natural amino acids also possess diverse biological activities such as anti-hypertensive, angiotensin-converting enzyme (ACE) inhibitory, antioxidant, immunomodulatory, anti-tumoral, anti-microbial, opiate-like, hypocholesterolemic and antithrombotic actions (Bo *et al.*, 2021; Chai *et al.*, 2020; Hamley, 2017; Rafiq *et al.*, 2021; Singh *et al.*, 2021). More importantly, recent studies (Bhullar *et al.*, 2021; Lammi and Arnoldi, 2021; Schütz *et al.*, 2020) have shown that FBPs can be further repurposed as potential drugs to reduce the SARS-CoV-2-induced oxidative stress and inflammation by targeting angiotensin-converting enzyme 2 receptor, viral spike protein, protease furin, transmembrane serine protease 2, cathepsin L and renin-angiotensin-aldosterone system members.

Literature search of keywords of 'food and bioactive peptide' or 'food and functional peptide' in PubMed, Web of Science and Elsevier ScienceDirect databases leads to a total of 12 469, 35 043 and 1976 records, respectively (Supplementary Fig. S1). Despite biological significance, some fundamental questions about the sequence–structure–property (activity) relationship of FBPs still remain to be answered: For example, what are the critical sequence and structural characteristics for FBPs? Statistically, are there sequences in FBPs that have preferential bindings to specific structures of targets? Do FBPs with similar sequences have similar biological functions? What are the targets and rules of action of multifunctional bioactive peptide (MBPs)?

With the advent of big-data technologies in many research fields, a number of research efforts have been made to construct different

bioactive peptide (BP) databases, including BIOPEP-UWM (Minkiewicz *et al.*, 2019), APD3 (Wang *et al.*, 2016), AHTPDB (Kumar *et al.*, 2015), MBPDB (Nielsen *et al.*, 2017), DADP (Novković *et al.*, 2012) CancerPPD (Tyagi *et al.*, 2015), BioPepDB (Li *et al.*, 2018) and DPL (Wang *et al.*, 2020). However, the existing database needs to be improved in the following aspects: (i) lack of FBP database; (ii) the functional classification, physicochemical properties, biological information, literature records of FBP are insufficiently annotated; (iii) data mining and chart visualization are lacking; (iv) MBPs are not included. In addition, some heterogeneous data from multiple experiments at different conditions and labs are not clearly presented. Some datasets even contain erroneous or inconsistent entries, for example, the parameters on four peptides (VV-hemorphin-7, LVV-hemorphin-7, LVV-hemorphin-5 and VV-hemorphin-5) extracted from the same original articles are inconsistent. This fact raises a key question on how to gather a reliable benchmarking database in combination with computational techniques to better assess the component–structure–property–performance relationship of FBPs.

Motivated by the high impacts of FBPs and current limitations of BP databases, here we constructed a web-based database of food-derived bioactive peptides (DFBP) containing a total of 3997 specific FBPs and 21 249 food-borne proteins, with multi-data-mining modules and chemoinformatics analysis of (i) intelligent searching for different types, sequences and structures of FBPs, (ii) repurposing new bioactivity of existing FBPs, (iii) clarifying the sequence and activity relationship of FBPs, (iv) discovering the functional/sequential relationship between FBPs and food-derived proteins and (v) deriving physicochemical parameters of FBPs from bioinformatics models. Collectively, this database provides a more comprehensive and specific platform for the screening, discovery and design of FBPs in a wide sequence, structure and bioactivity space.

2 Materials and methods

2.1 Data collections of FBPs and food-derived proteins

A total of 6276 FBP entries, as collected from 1063 literature sources, were grouped into 31 categories in terms of their functions/activities (Supplementary Table S1). To this end, different keywords, e.g. 'ACE inhibitory peptide', 'anticancer peptide', 'antioxidative peptide', were used to search references in three databases of Web of Science, SpringerLink and ScienceDirect (Supplementary Table S2). Then, general information about peptide sequences, food sources, preparation methods, physical and chemical characteristics (such as functional activity, toxicity, bitterness, stability) and references were screened, checked and integrated manually. Additional information retrieval was required by looking into the references if they are not available directly. Finally, FBPs with all necessary information from literature were used to construct our database with different modules. (Main selection criteria for FBPs is summarized in Supplementary Note S1.)

Further, to study the relationship between proteins and FBPs, a total of 21 249 food-derived proteins from 60 different species sources were retrieved from database UniProt (https://www.uniprot.org/) and grouped into five categories: terrestrial animals (no milk source), terrestrial plants, milk sources, marine sources and microorganism sources.

2.2 Annotation of food-derived bioactive peptides and proteins

In DFBP, (i) each FBP was encoded with different sequential/structural/functional features; (ii) some physicochemical properties and multifunctional properties, if not experimentally available, are calculated/predicted by different tools in DFBP of Multi-cross Analysis, Peptide Calculator, BPP-Tool, EHP-Tool and HotSpot Search; (iii) SMILES structure and toxicity of peptides are obtained by the online tool PepSMI and ToxinPred, respectively; (iv) Information on FBPs with the same sequence but from different sources are listed in the 'Link-sources' for comparison; (v) Other miscellaneous information is also provided in the 'Additional information' section; (vi) Relationships between FBPs, target proteins and related diseases were established (FBP–target protein–disease) via DFBP tools and UniProt (see Supplementary Table S3 for details).

For food-derived proteins, each protein is encoded with (i) protein name, organism, length, sequence and UniProtKB were cited from UniProt; (ii) types and cumulative amounts of BPs in proteins; (iii) enzymatic hydrolysis characteristics (see Supplementary Table S4 for details). Additional information (e.g. Species classification, FBP list, function difference, statistics view and peptide traceability), if necessary, can also be collected through database search and comparison, while some analysis results can be directly visualized by GUI.

2.3 Data structure design and web interface of DFBP

The construction process of DFBP was divided into five steps: (i) Reference data collection from Web of science, SpringerLink, ScienceDirect, while food-derived proteins from UniProt; (ii) Integration and supplementation of peptide and protein properties; (iii) Development of FBP and protein analysis tools; (iv) Statistics of FBPs and proteins and related references; (v) Web page construction (Supplementary Fig. S2).

All data and results were stored in a MySQL-based database on the Tencent Cloud server. The web page was constructed using HTML, CSS3, JavaScript, while the visualization of dynamic charts was realized by HTML5 Canvas and eCharts. Moreover, six JAVA-based tools were developed for realizing a wide variety of bioinformatics and chemoinformatics analyses. As a user-friendly operating platform, DFBP provides detailed online instructions, responses and error messages during the operation. In addition, DFBP was equipped with a complete background management system, which can periodically perform data updates, backups, restores and website optimization.

3 Results

3.1 Overview of data architecture and web interface of DFBP

The DFBP collects 31 types of FBPs from 1063 references, leading to a total of 6276 entries for 3997 unique sequences (Supplementary Table S1). The data in DFBP showed that FBPs often possess multiple activities and can be obtained from abundant sources in large quantities, especially for ACE inhibitory peptides (1956), antioxidant peptides (1028), antibacterial peptides (476), antihypertensive peptides (418) and dipeptidyl peptidase IV enzyme (DPP IV) inhibitory peptides (244) (Supplementary Fig. S3).

To display the complete information of FBPs, we integrated about 30 types of characteristics for each peptide derived from experimental data points, predicted data points and external links (Supplementary Table S3). These characteristic data can be used for the better understanding of physicochemical properties of FBPs and guiding the setup and use of bioinformatics models.

Meanwhile, food-derived proteins including FBPs in 60 species were traced back by analyzing 10 types of attributes for 21 249 proteins (Supplementary Table S5). DFBP not only presents different level information to describe the sequence-function characteristics of FBPs and the relationship between FBPs and their precursor proteins, but also provides six application tools, including HotSpot Search, Enzymatic Hydrolysis Prediction (EHP-Tool), the Bitterness of Peptide Prediction (BPP-Tool), Amino Acid Structural Descriptor (AASD-Tool), Peptide Calculator and Multi-cross Analysis for MBPs (Supplementary Fig. S2D).

The main web interface for DFBP contain all menus of Home, Browse, Tools, Statistics, Download and Help, with a wide variety of search, analysis, display and predictive functions for revealing the essential physicochemical properties, sequence characteristics and multifunctional relationship of FBPs, as well as sequence–activity relationship of FBPs and food-borne proteins (Supplementary Fig. S4).

3.2 Rapid search of sequences with desirable functions

DFBP can be used for a rapid search and identification of sequences with desirable functions. Two search liberties of 'Peptide Search' and 'Protein Search' with three different search engines of 'Quick Search', 'Advanced Search' and 'Multifunctional Peptide Search' are available in DFBP (see Supplementary Fig. S5 for details).

3.3 Six application tools for the discovery and repurposing of FBPs

Apart from offering basic information of FBPs and food-derived proteins, DFBP also provides a collection of Java-based analysis tools to analyze the sequence–activity relationship of FBPs, screen food-derived proteins rich in FBPs, calculate physiochemical properties, predict the bitterness, simulate hydrolysis of proteins and generate feature vectors (Supplementary Table S6). Of note, all these tools in the DFBP are interdependent, in which mutual assistance and cross-check between different tools are allowed to obtain more accurate and complement information for any given demand. As a typical example, the peptide fragments derived from enzymatic hydrolysis can be predicted whether they possess bitterness or not, followed by additional validation via fragment traceability, multifunction observation and physicochemical attributes.

3.4 Quantitative distribution and targeted activity of MBPs

A total of 3997 peptides of 31 types as collected from different sources are mainly analyzed to obtain their sequence-function relationship using a 'Multi-cross Analysis' tool. It can be seen from Figure 1A, except for Celiac disease peptides (the peptide involving in eliciting diarrhea), other 29 types of FBPs presented two or more functional activities. Particularly, ACE inhibitory peptides, antihypertensive peptides, antioxidant peptides and DPP IV-inhibitory peptides exhibited the highest population of MBPs, as evidenced by 649 (39.33%), 327 (89.84%), 302 (30.08%) and 113 (47.08%) peptides, respectively.

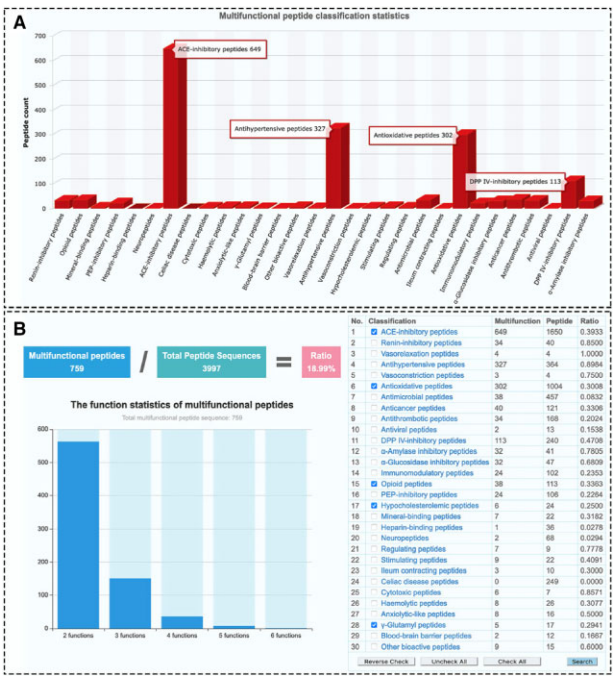

Fig. 1. Quantitative distribution and characterization of MBPs from 30 types of FBPs using 'Multi-cross Analysis'. (**A**) The number of MBPs in different types of FBPs. (**B**) The number of functions, fragments and the corresponding proportion of MBPs (The number of MBPs in a category/the total number of FBPs in this category)

Through 'Multi-cross Analysis' page, 759 of 3997 peptides were identified as MBPs, accounting for 18.99% of all peptides and the largest collection of MBPs among the existing databases (Fig. 1B). These MBPs are mainly distributed in the categories of ACE-inhibitory, Antihypertensive, Antioxidative, DPP IV-inhibitory, Anticancer, Antimicrobial, Opioid, Antithrombotic and Renin inhibitory, where they generally have 2–4 functions, and 8 peptide fragments have 5 functions (Fig. 1B). For instance, peptide YP (DFBPMUFU0106) exhibited six different bioactivities, including ACE inhibitory activity (IC_{50} = 120 μM) (Yamada *et al.*, 2013), anti-hypertensive activity (SBP: -32.1 mmHg, Dose: 2.0 mg/kg) (Yamamoto *et al.*, 1999), DPPH free radical scavenging activity (EC_{50} > 5 mM) (Nongonierma and FitzGerald, 2014), DPP IV competitive inhibitory activity (IC_{50} = 658.1 μM) (Nongonierma and FitzGerald, 2013), α-glucosidase inhibitory activity (IC_{50} = 16.8 mM) (Matsui *et al.*, 1999) and opioid agonist activity (Lottspeich *et al.*, 1980). Our DFBP enables to explore the relationship between their sequences and activities of MBPs, to guide the discovery, design and repurposing of multifunctional and multi-target drug molecules, food active components or peptide-based materials.

3.5 Bidirectional traceability of food-derived proteins and FBPs

Based on the protein data currently reviewed and identified by UniProt, 'Statistics→Proteins' allows to analyze the sequence coding features of FBPs encoded in 21 249 food-borne proteins. As shown in Figure 2A, in the 60 species of food-derived proteins, terrestrial

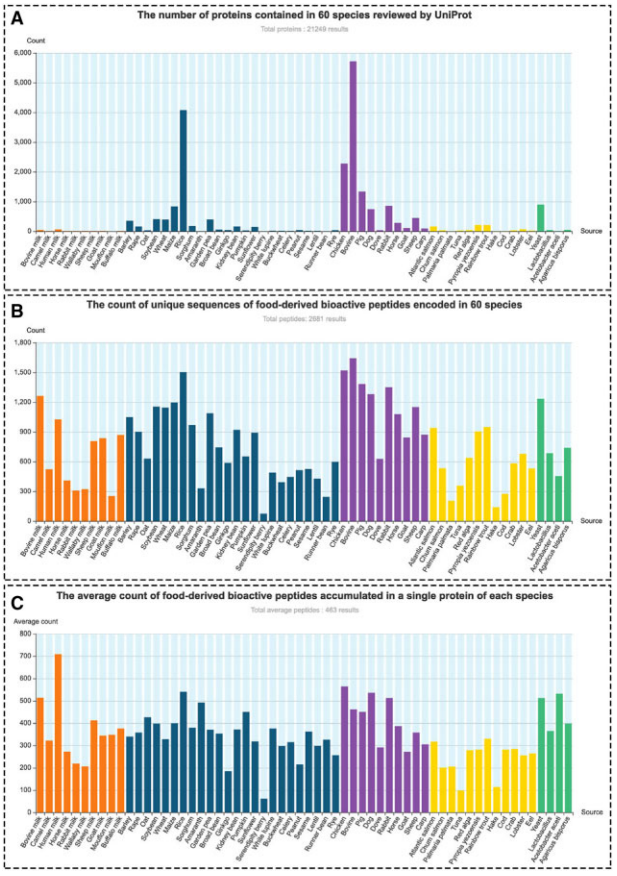

Fig. 2. The quantitative distribution characteristics between food-borne proteins and their FBP coding sequences. (**A**) The number of proteins contained in 60 species reviewed by UniProt. (**B**) The count of unique sequences of FBPs encoded in different species. (**C**) The average count of FBPs accumulated in a single protein of each species. The orange, blue, purple, yellow and green bars represent milk sources, terrestrial plants, terrestrial animals (no milk source), marine sources and microorganisms, respectively (A color version of this figure appears in the online version of this article)

Table 1. Overall performance comparison of DFBP database with other popular BP databases[a]

Feature	Database								
	DFBP	BIOPEP-UWM	APD3	AHTPDB	MBPDB	DADP	CancerPPD	BioPepDB	DPL
Entries	6276	4456	3324	6000	944	2571	3491	4807	1011
Unique sequences	3997	–	3324	1700	606	1923	624	3371	1011
Included BP types	31	43	25	1	8	2	1	10	10
MBPs	759	–	–	–	–	–	–	–	–
BP properties	30	12	10	10	13	21	16	10	10
Retrievable properties	12	8	15	15	3	9	16	5	–
Included FBPs	√	√		√	√			√	
Protein-BP analysis & statistics	√	√						√	
Sequence feature analysis & statistics	√		√						
Bitter prediction	√	√							
Mimic enzymatic hydrolysis	√	√							
Physicochemical property analysis	√	√							
Target protein & disease	√						√		
Advanced & Multiple searching	√			√		√	√		
MBP screening	√								
Category list & Visualizations	√		√				√	√	
Heterogeneous data comparison	√								
BP-protein bidirectional traceability	√	√							

[a]Character '-' refers to the unclear count in this cell for some reasons. The retrieval time is March 29, 2022.

plants (11 878) and animals (7410) are the main sources, especially the number of proteins reviewed in beef (5713) and rice (4080) is more significant. In addition, milk sources (170), marine sources (801) and microorganisms (990) have gradually become important protein sources for FBPs. Figure 2B displays that these reviewed proteins derived from terrestrial animal and plant coded more unique FBP sequences, especially FBPs encoded in beef (1642), chicken (1520), rice (1503), pig (1383), rabbit (1350), dog (1281), bovine milk (1263) and yeast (1235) were prominent. Moreover, human milk (709), chicken (564), rice (540), dog (536), *Acetobacter aceti* (532), rabbit (513) and yeast (513) coded more FBP fragments (Fig. 2C), suggesting that these reviewed proteins may have great potential for generating biopeptides.

In DFBP, we also traced the source of FBPs to view the sequence distribution of each FBP in different precursor proteins using the 'Precursor proteins' list or the 'HotSpot Search' tool. As a well-studied ACE-inhibitory peptide IPP (DFBPACEI1197), it is not only presented in the ovine, bovine and human milk, but also widely distributed in 1435 protein sequences (Supplementary Fig. S6). This indicates that DFBP realizes the bidirectional traceability from both food-derived proteins and BPs, i.e. DFBP can not only trace the original protein sources to identify a peptide sequence of interest and their occurring frequency, but also analyze the types and amounts of FBPs in each protein. This helps for designing targeted peptide experiments or tracking down the protein origin of peptides.

4 Discussion

4.1 The comparison among different peptide-based databases

Compared with existing BP databases, our DFBP with a web interface mainly focuses on physicochemical properties and the sequence–activity relationship of FBPs by developing a wide variety of online data-driven analysis tools. This web interface of the database allows to provide a fast and reliable way to not only query peptide sequences and return the respective sequence/structure/function information about their precursor and connected proteins, but also design and repurpose BPs. To better assess the overall performance of our DFBP, we conducted a systematic comparison between DFBP and other listed BP databases (e.g. BIOPEP-UWM, APD3, AHTPDB, MBPDB, DADP, CancerPPD, BioPepDB and DPL.). As shown in Table 1, DFBP is the only one to achieve 12 criteria simultaneously: the largest number of FBPs (6276 entries and 3997 unique sequences), abundant tools for characterizing sequence characteristics and activity, experimental tracking and comparison of same and similar sequences, physicochemical properties, traceability between peptides and proteins, MBPs, bitter characteristics, mimic enzymatic hydrolysis, cytotoxicity, target proteins, the associated diseases and multiple search tools. Therefore, based on the above peptide and protein entries, as well as application tools and multiple search engines, researchers can acquire more peptides by searching DFBP than searching existing databases (see Supplementary Table S7 for examples). Clearly, DFBP database outperforms the widely used public databases in terms of large-scale, fast, reliable data acquisition, tracking, analysis and predication.

4.2 The sequence–activity relationship of FBPs

It is critical and interesting to examine the sequence-bioactivity of different FBPs (Chen *et al.*, 1998; Li and Li, 2013). To this end, we analyzed the position specific properties (e.g. amino acid composition, two terminals, molecular weight and physicochemical properties) of 12 types of FBPs (sample size > 100) (Supplementary Table S8). The compositional propensities of the 20 amino acids in FBPs showed that (i) Pro (8.89–29.32%), Leu (7.7–12.25%) and Gly (7.14–10.76%) had the highest occurrence in 10, 9 and 7 types of FBPs; (ii) N-terminal had the strong preference for two hydrophobic amino acids of Val (7.68–11.32%) and Ile (7.84–15.09%) and a non-sidechain amino acid Gly (7.14–25.62%), while C-terminal favored Pro (7.67–30.19%), Phe (7.03–15.26%), Lys (8.36–23.21%), Trp (7.55–8.24%) and Cys (14.66%); (iii) in addition, the residue distributions at both N-terminal and C-terminals are mainly non-polar Leu (7.7–16.25%) and Ala (7.52–11.25%), polar Tyr (7.97–66.37%) and Gln (39.36–39.76%), and basic Arg (7.08–20.83%); and (iv) FBPs possessed the average length of 2–10 residues and molecular mass of <1500 Da. Particularly, antimicrobial peptides, Celiac disease peptides and anticancer peptides had the relatively longer length and higher molecular mass than other peptides.

Considering that 5 of 12 types of FBPs (i.e. Antioxidative, ACE-inhibitory, Antihypertensive, Anticancer and DPP IV-inhibitory peptides) are pathologically associated with some chronic diseases (e.g.

oxidative aging, hypertension, cancer and diabetes) (Nongonierma and FitzGerald, 2016; Sabbatino *et al.*, 2021), we further examined the sequence propensity of these 5 types of FBPs, whose average lengths/molecular weights were 2–9 and 180–1100 Da, respectively. The data showed that non-polar Leu (8.26–16.25%), Gly (6.08–25.62%) and Ala (4.13–7.57%) were the three residues with the highest sequence propensity at N-terminals, while non-polar Leu (7.7–15.83%), polar Tyr (4.96–11.81%), basic Lys (5.77–15.7%) had more prominent sequence propensity at C-terminals (Supplementary Fig. S7). Hydrophobic Leu (8.63–12.25%), Gly (7.14–9.33%) and Pro (4.69–19.25%) were found in all 5 types of peptides (Supplementary Fig. S8).

To link sequence propensity to their bioactivity, several studies have shown that the length and terminal residues of FBPs are critical for their bioactivity (As proof-of-examples are summarized in Supplementary Note S2.). Collective evidence indicates that subtle change of residue types, positions and propensities will somehow significantly impact on their bioactivity, which provides a large room for the rational design and repurposing of highly effective BPs. Moreover, it is worth mentioning that based on the sequence characteristics and physicochemical parameters of antioxidative peptides, we are utilizing the deep-learning methods to develop a predictor for antioxidative peptides [Antioxidative peptide predictor (AnOxPP), http://www.cqudfbp.net/AnOxPP/index.jsp, unpublished]. This greatly saves time and materials resources for achieving the large-scale screening and development of antioxidative peptides.

4.3 Exploring the food-derived MBPs for potential treatment of cardiovascular diseases

Hypertension, atherosclerosis, ageing, insulin resistance and hyperlipidaemia are the major risk factors for Cardiovascular diseases (CVDs), which were mainly caused by oxidative stress and the imbalance of a variety of metabolic enzymes including ACE, renin, 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGCoAR), α-amylase, DPP IV, α-glucosidase and thrombin (Dabravolski *et al.*, 2021; Halliwell, 1994; Johansen *et al.*, 2005; Lammi *et al.*, 2019; Singh *et al.*, 2021). The current drug molecules mainly regulate these target proteins; for example, Captopril lowered blood pressure by inhibiting ACE (Posvar and Sedman, 1991); β-sitosterol exhibited the hypolipidemic effect by inhibiting HMGCoAR to limit cholesterol synthesis (Babu and Jayaraman, 2020); and Sitagliptin and alogliptin reduced blood sugar by inhibiting DPP IV to increase endogenous GLP-1 and insulin levels (Argyrakopoulou and Doupis, 2009). As a multifactorial comprehensive disease, CVD involves multiple metabolic pathways and the regulation of a single target protein in the pathway usually has a limited effect, so conventional CVD treatment generally requires the patients to take multiple drugs, which causes some potential adverse effects in the long-term.

Fortunately, MBPs not only have lower negative side effects and rich sources, but also are able to impart more than one physiological outcome by affecting different target proteins, so MBPs are considered as potential safer agents for CVD treatment (Gu and Wu, 2016; Hajfathalian *et al.*, 2018; Lammi *et al.*, 2019; Patil *et al.*, 2015). Multiple studies have shown that these MBPs can modulate multiple CVD-related pathway proteins (Supplementary Table S9). For example, the soybean-derived peptides (IAVPTGVA, IAVPGVEA and LPYP) exhibited an inhibition on HMGCoAR and DPP IV (Lammi *et al.*, 2019); the *Takifugu flavidus*-derived peptide PPLLFAAL exhibited a more potent antihypertensive effect than *captopril* by inhibiting ACE (Su *et al.*, 2021); the β-lactoglobulin-derived peptide IIAEK exhibited hypocholesterolemic, ACE-inhibitory, antioxidative and DPP IV-inhibitory activities; as a novel cholesterol-lowering peptide, IIAEK exhibited a stronger cholesterol-lowering activity than β-sitosterol (Nagaoka *et al.*, 2001; Power-Grant *et al.*, 2014); and the γ-glutamyl dipeptides (γ-EM, γ-EL, γ-EF, γ-EW and γ-EY), as competitive inhibitors of dipeptidyl peptidase-IV, were considered as potential functional ingredients in the type 2 diabetic diet (Yang *et al.*, 2018).

Motivated by the potential use of less-explored MBPs, we used 'Multi-cross Analysis' to analyze 10 kinds of FBPs, including antioxidant, ACE inhibitory, renin inhibitory, antihypertensive, α-amylase inhibitory, α-glucosidase inhibitory, DPP IV-inhibitory, antithrombotic, hypocholesterolemic and γ-gultamyl peptides with potential bioactivities related to CVDs (Supplementary Table S9). Among a total of 645 MBPs, there were 513, 107, 21 and 4 MBPs with 2, 3, 4 and 5 functions being identified as CVD-related MBPs in food-derived proteins. Of particular concern is that the four MBPs, RVPSL (DFBPMUFU0062), QIGLF (DFBPMUFU0063), GP (DFBPMUFU0152) and YP (DFBPMUFU0106) were identified with five bioactivities (Supplementary Table S10). Moreover, these four MBPs can indeed act on two or more target proteins in the above pathways simultaneously: (i) RVPSL with ACE-inhibitory, antihypertensive and anticoagulation activities; (ii) QIGLF with ACE-inhibitory, DPPH radical scavenging and anticoagulation activities; (iii) GP with ACE-inhibitory, antihypertensive, DPPH radical scavenging and anticoagulation activities; and (iv) YP with ACE-inhibitory, antihypertensive, DPPH radical scavenging, DPP IV-inhibitory and α-Glucosidase inhibitory activities (Supplementary Table S11). Therefore, these MBPs could serve as drug components with multi-targets, high activity and low toxicity functions for a potential prevention/treatment for multifactorial CVDs (Gu and Wu, 2016).

5 Conclusion

In conclusion, we developed DFBP with a web interface, consisting of 3997 FBPs in 31 categories. The database serves as a reliable data-driven platform to provide (i) sequence-based information for physicochemical properties, toxicity properties, biological pathways, multifunctional activity, enzymatic stability, traceability distribution and length/mass/composition features, (ii) a wide variety of analysis and data-mining tools for sequence query, hotspot identification, MBP discovery, proteolysis prediction and physiochemical property assessment, (iii) discovery and repurposing of FBPs via sequence design, engineering and optimization for different bioapplications, (iv) a user-friendly web interface for open data access, with up-to-date information monthly http://www.cqudfbp.net/.

Funding

This work was supported by the National Natural Science Foundation of China [32172196, 31771975 and 31571782 to G.L.]; NSF [2107619] and Faculty Research Fellowship from the University of Akron (to J.Z.); and the Natural Science Foundation of Chongqing, China [CSTC2019JCYJ-MSXMX0527 to B.L.].

Conflict of Interest: The authors declare no competing financial interest.

References

Argyrakopoulou,G. and Doupis,J. (2009) DPP4 inhibitors: from sitagliptin monotherapy to the new alogliptin-pioglitazone combination therapy. *Adv. Ther.*, **26**, 272–280.

Babu,S. and Jayaraman,S. (2020) An update on β-sitosterol: a potential herbal nutraceutical for diabetic management. *Biomed. Pharmacother.*, **131**, 110702.

Barati,M. *et al.* (2020) Techniques, perspectives, and challenges of bioactive peptide generation: a comprehensive systematic review. *Compr. Rev. Food Sci. Food Saf.*, **19**, 1488–1520.

Bechaux,J. *et al.* (2019) A comprehensive review of bioactive peptides obtained from animal byproducts and their applications. *Food Funct.*, **10**, 6244–6266.

Bhullar,K.S. *et al.* (2021) Translating bioactive peptides for COVID-19 therapy. *Eur. J. Pharmacol.*, **890**, 173661.

Bo,W. *et al.* (2021) Application of quantitative structure-activity relationship to food-derived peptides: methods, situations, challenges and prospects. *Trends Food Sci. Technol.*, **114**, 176–188.

Chai,K.F. *et al.* (2020) Bioactive peptides from food fermentation: a comprehensive review of their sources, bioactivities, applications, and future development. *Compr. Rev. Food Sci. Food Saf.*, **19**, 3825–3885.

Chan-Zapata,I. *et al.* (2020) Proteins and peptides from vegetable food sources as therapeutic adjuvants for the type 2 diabetes mellitus. *Crit. Rev. Food Sci. Nutr.*, **2020**, 1–10.

Chen,H.M. *et al.* (1998) Antioxidative properties of histidine-containing peptides designed from peptide fragments found in the digests of a soybean protein. *J. Agric. Food Chem.*, **46**, 49–53.

Dabravolski,S.A. *et al.* (2021) The role of mitochondria-derived peptides in cardiovascular diseases and their potential as therapeutic targets. *Int. J. Mol. Sci.*, **22**, 8770.

Gong,X. *et al.* (2020) Prospects of cereal protein-derived bioactive peptides: sources, bioactivities diversity, and production. *Crit. Rev. Food Sci. Nutr.*, **62**, 1–17.

Gu,Y.C. and Wu,J.P. (2016) The potential of antioxidative and anti-inflammatory peptides in reducing the risk of cardiovascular diseases. *Curr. Opin. Food Sci.*, **8**, 25–32.

Hajfathalian,M. *et al.* (2018) Peptides: production, bioactivity, functionality, and applications. *Crit. Rev. Food Sci. Nutr.*, **58**, 3097–3129.

Halliwell,B. (1994) Free radicals, antioxidants, and human disease: curiosity, cause, or consequence? *Lancet*, **344**, 721–724.

Hamley,I.W. (2017) Small bioactive peptides for biomaterials design and therapeutics. *Chem. Rev.*, **117**, 14015–14041.

Johansen,J.S. *et al.* (2005) Oxidative stress and the use of antioxidants in diabetes: linking basic science to clinical practice. *Cardiovasc. Diabetol.*, **4**, 5.

Kumar,R. *et al.* (2015) AHTPDB: a comprehensive platform for analysis and presentation of antihypertensive peptides. *Nucleic Acids Res.*, **43**, D956–962.

Lammi,C. *et al.* (2019) Multifunctional peptides for the prevention of cardiovascular disease: a new concept in the area of bioactive food-derived peptides. *J. Funct. Foods*, **55**, 135–145.

Lammi,C. and Arnoldi,A. (2021) Food-derived antioxidants and COVID-19. *J. Food Biochem.*, **45**, e13557.

Li,Y.W. and Li,B. (2013) Characterization of structure-antioxidant activity relationship of peptides in free radical systems using QSAR models: key sequence positions and their amino acid properties. *J. Theor. Biol.*, **318**, 29–43.

Li,Q. *et al.* (2018) BioPepDB: an integrated data platform for food-derived bioactive peptides. *Int. J. Food Sci. Nutr.*, **69**, 963–968.

Lottspeich,F. *et al.* (1980) Novel opioid peptides derived from casein (beta-casomorphins). III. Synthetic peptides corresponding to components from bovine casein peptone. *Hoppe Seylers Z. Physiol. Chem.*, **361**, 1835–1839.

Matsui,T. *et al.* (1999) Isolation and identification of peptidic alpha-glucosidase inhibitors derived from sardine muscle hydrolyzate. *Z. Naturforsch. C. J. Biosci.*, **54**, 259–263.

Minkiewicz,P. *et al.* (2019) BIOPEP-UWM database of bioactive peptides: current opportunities. *Int. J. Mol. Sci.*, **20**, 5978.

Nagaoka,S. *et al.* (2001) Identification of novel hypocholesterolemic peptides derived from bovine milk beta-lactoglobulin. *Biochem. Biophys. Res. Commun.*, **281**, 11–17.

Nielsen,S.D. *et al.* (2017) Milk bioactive peptide database: a comprehensive database of milk protein-derived bioactive peptides and novel visualization. *Food Chem.*, **232**, 673–682.

Nongonierma,A.B. and FitzGerald,R.J. (2013) Inhibition of dipeptidyl peptidase IV (DPP-IV) by proline containing casein-derived peptides. *J. Funct. Foods*, **5**, 1909–1917.

Nongonierma,A.B. and FitzGerald,R.J. (2014) Susceptibility of milk protein-derived peptides to dipeptidyl peptidase IV (DPP-IV) hydrolysis. *Food Chem.*, **145**, 845–852.

Nongonierma,A.B. and FitzGerald,R.J. (2016) Strategies for the discovery, identification and validation of milk protein-derived bioactive peptides. *Trends Food Sci. Technol.*, **50**, 26–43.

Nongonierma,A.B. and FitzGerald,R.J. (2017) Strategies for the discovery and identification of food protein-derived biologically active peptides. *Trends Food Sci. Technol.*, **69**, 289–305.

Novković,M. *et al.* (2012) DADP: the database of anuran defense peptides. *Bioinformatics*, **28**, 1406–1407.

Patil,P. *et al.* (2015) Food protein-derived bioactive peptides in management of type 2 diabetes. *Eur. J. Nutr.*, **54**, 863–880.

Pavlicevic,M. *et al.* (2020) Marine bioactive peptides-an overview of generation, structure and application with a focus on food sources. *Mar. Drugs*, **18**, 424.

Posvar,E.L. and Sedman,A.J. (1991) ACE inhibitors in the elderly. *Angiology*, **42**, 387–396.

Power-Grant,O. *et al.* (2014) Selective enrichment of bioactive properties during ultrafiltration of a tryptic digest of β-lactoglobulin. *J. Funct. Foods*, **9**, 38–47.

Rafiq,S. *et al.* (2021) Functional role of bioactive peptides with special reference to cheeses. *Int. J. Dairy Technol.*, **74**, 1–16.

Sabbatino,F. *et al.* (2021) Molecules and mechanisms to overcome oxidative stress inducing cardiovascular disease in cancer patients. *Life (Basel)*, **11**, 105.

Schütz,D. *et al.* (2020) Peptide and peptide-based inhibitors of SARS-CoV-2 entry. *Adv. Drug Deliv. Rev.*, **167**, 47–65.

Singh,B.P. *et al.* (2021) Bioactive peptides in the management of lifestyle-related diseases: current trends and future perspectives. *Crit. Rev. Food Sci. Nutr.*, 1–14.

Su,Y. *et al.* (2021) A novel angiotensin-I-converting enzyme (ACE) inhibitory peptide from *Takifugu flavidus*. *Mar. Drugs*, **19**, 651.

Tyagi,A. *et al.* (2015) CancerPPD: a database of anticancer peptides and proteins. *Nucleic Acids Res.*, **43**, D837–843.

Wang,G. *et al.* (2016) APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res.*, **44**, D1087–1093.

Wang,F. *et al.* (2020) DPL: a comprehensive database on sequences, structures, sources and functions of peptide ligands. *Database (Oxford)*, **2020**, baaa089.

Yamada,A. *et al.* (2013) Novel angiotensin I-converting enzyme inhibitory peptide derived from bovine casein. *Food Chem.*, **141**, 3781–3789.

Yamamoto,N. *et al.* (1999) Purification and characterization of an antihypertensive peptide from a yogurt-like product fermented by *Lactobacillus helveticus* CPN4. *J. Dairy Sci.*, **82**, 1388–1393.

Yang,J. *et al.* (2018) γ-Glu-Met synthesised using a bacterial glutaminase as a potential inhibitor of dipeptidyl peptidase IV. *Int. J. Food Sci. Technol.*, **53**, 1166–1175.

Zhou,J. *et al.* (2020) A review on mushroom-derived bioactive peptides: preparation and biological activities. *Food Res. Int.*, **134**, 109230.