Journal of Exploratory Research in Pharmacology **2024** vol. 9(3) | 153–168
DOI: 10.14218/JERP.2023.00030

Original Article

In Silico Assessment of Photosystem I P700 Chlorophyll a Apoprotein A2 (PsaB) from *Chlorella vulgaris* (green microalga) as a Source of Bioactive Peptides

Md Ariful Amin[1], Uzzal Chondra[1,2] and Md Morshedul Alam[1,3*]

[1]*Medical Biotechnology Laboratory, Department of Genetic Engineering and Marine Biotechnology, Bangabandhu Sheikh Mujibur Rahman Maritime University, Dhaka, Bangladesh;* [2]*Department of Microbiology, Gono Bishwabidyalay, Dhaka, Bangladesh;* [3]*Department of Biochemistry and Microbiology, North South University, Dhaka, Bangladesh*

Received: September 21, 2023 | **Revised:** May 08, 2024 | **Accepted:** August 23, 2024 | **Published online:** September 30, 2024

Abstract

Background and objectives: *Chlorella vulgaris* is a green, photosynthetic microalga in the phylum Chlorophyta. The goal of our study was to perform a bioinformatics analysis of Photosystem I P700 chlorophyll a apoprotein A2, one of its photosynthesis-related proteins, and to hunt for potent bioactive peptides.

Methods: To generate peptides and estimate the safety and efficacy of each bioactive peptide, we employed the tools BIOPEP-UWM™, PeptideRanker, DBAASP, and ToxinPred. PepDraw was used to understand the physicochemical properties and primary chemical structures of the selected bioactive peptides.

Results: The liberated peptides exhibit up to 17 distinct bioactivities, as shown by the *in silico* digestion of the protein using several proteolytic enzymes. The peptides with bioactivities are listed as angiotensin-converting enzyme inhibitor, dipeptidyl peptidase IV inhibitor, dipeptidyl peptidase III inhibitor, antioxidative, renin inhibitor, glucose uptake stimulator, neuropeptide regulator (regulating stomach mucosal membrane activity and ion flow), antithrombotic, anti-amnestic, CaMPDE inhibitor, activators of ubiquitin-mediated proteolysis, alpha-glucosidase inhibitor, immunomodulating, calcium-binding, antibacterial, anti-inflammatory, and hypotensive agent. Using the Database of Antimicrobial Activity and Structure of Peptides (DBAASP) prediction method, the antibacterial activity of the released peptides was predicted, highlighting the existence of potent antibacterial peptides. An examination of their physicochemical properties revealed that most peptides are low molecular weight, mildly acidic, and moderately water-soluble. To further establish the non-toxicity profile of the released peptides (sequence length > 3), a ToxinPred analysis was performed, which revealed that most of the peptides are non-toxic. According to the allergenicity analysis, most of the top-ranked peptides are likely non-allergenic.

Conclusions: Thus, our study reveals a less labor-intensive method for discovering new therapeutic targets derived from *C. vulgaris*, which hold both pharmacological and medical significance.

Keywords: *Chlorella vulgaris*; Green microalga; Bioactive peptides; Pharmacological drug; *In silico*; Drug development.

***Correspondence to:** Md Morshedul Alam, Department of Biochemistry and Microbiology, North South University, Bashundhara, Dhaka 1229, Bangladesh. ORCID: https://orcid.org/0000-0003-1564-969X. Tel: +88-01786-063739, Fax: +88-0255668202, E-mail: morshedul.alam@northsouth.edu

How to cite this article: Amin MA, Chondra U, Alam MM. *In Silico* Assessment of Photosystem I P700 Chlorophyll a Apoprotein A2 (PsaB) from *Chlorella vulgaris* (green microalga) as a Source of Bioactive Peptides. *J Explor Res Pharmacol* 2024;9(3):153–168. doi: 10.14218/JERP.2023.00030.

Introduction

Marine water covers around 71% of the Earth's surface, providing humans with access to an abundance of resources. The marine ecosystem consists of mammals, mollusks, bacteria, macroalgae, microalgae, and other creatures. This ecosystem is rich in diverse beneficial bioactive compounds that are useful as ingredients in food and feed, as well as for human health.[1] Because marine algae are a significant source of bioactive chemicals with marine origins and have fewer detrimental effects compared to natural substances found in terrestrial environments, scientists have paid particular attention to them. Marine algae have made significant progress as

a sustainable source of protein due to their ease of production and minimal land occupation compared to terrestrial natural chemicals. Additionally, marine algae-based bioactive peptides show emerging medical promise, attracting researchers to this field of study.[2]

The concept of bioactive peptides is currently being extensively studied in the production of medications, food additives, and other applications. Most bioactive peptides are derived from seaweeds, cereal crops, milk, sea cucumbers, and a few other sources.[2–5] There are also naturally occurring bioactive peptides that serve a variety of functions, including antiviral, antibacterial, anti-inflammatory, anticancer, and antioxidant activities. Due to the greater adaptability of marine organisms in their environment, marine-derived bioactive peptides often have higher pharmacological potential than those from terrestrial sources. Traditional methods for producing bioactive peptides from natural sources include enzymatic hydrolysis, ultrasound, microwave treatment, pulse electric fields, and high hydrostatic pressure-assisted extraction.[6] As a complement to empirical procedures, *in silico* methods can analyze proteins' potential to serve as building blocks for bioactive peptides and predict the functions of specific peptide sequences. In comparison to experimental investigations, they also save more time, energy, and money. Several *in silico* tools (PeptideRanker, DBAASP Antimicrobial Peptide Prediction, ToxinPred) and databases (PEP-UWMTM) are available to identify more precise active peptides, reducing extensive efforts.

Chlorella vulgaris is a green microalga in the genus Chlorella and belongs to the division Chlorophyta. While it is most commonly found in freshwater environments, it can also be found in marine environments.[7] In many countries, it is widely used as a nutritional food supplement. The dry weight output of *C. vulgaris* is significantly higher than that of other microalgae due to its high tolerance to invasive organisms and harsh environments.[8] *C. vulgaris* has a high protein content, ranging from 42 to 58% of its dry weight,[9] and its nutritional quality meets human protein standards.[8] It is also rich in lipids (5–40% dry mass),[9] pigments, including chlorophyll (1–2% of the dry weight),[10] carbohydrates (12–55% dry weight),[11] vitamins, and omega-3 polyunsaturated fatty acids that make *C. vulgaris* an attractive source for dietary supplements and food additives.[9] Furthermore, *C. vulgaris* offers health benefits such as alleviating hyperglycemia and protecting against cancer, oxidative stress, and chronic obstructive pulmonary disease.[12] Despite these benefits, bioactive peptide generation from this species has significant potential due to its specificity for targeted delivery and lower risk of side effects.[2] *C. vulgaris* is a photosynthetic microalga with endogenous Photosystem I (PSI) P700 chlorophyll and an apoprotein A2 encoded by the PsaB gene. Since this species undergoes photosynthesis, PsaB protein expression increases during photosynthesis and electron transport. Generally, the P700 chlorophyll special pair and subsequent electron acceptors are bound by the PsaA/B heterodimer. The central antenna complex in PSI's design collects photons, and an electron transport network converts photonic stimulation into charge separation, transferring electrons from the P700 chlorophyll pair to other electron acceptors, A0, A1, FX, and FB.[13] Additionally, PsaB gene expression is alleviated under metal stress or other environmental stresses, indicating its ubiquitous expression status and importance to the *C. vulgaris* living system.[14] Despite the abundance of PsaB protein in *C. vulgaris*, there are no reports of bioactive peptides generated from PsaB protein in this species. Given the protein's abundance, there is substantial scope to investigate the presence of suitable bioactive peptides *in silico*.

In this study, we investigated *in silico* the possible bioactive peptides from the PsaB protein of *C. vulgaris*, a photosynthetic microalga, using several computational tools. The PsaB protein sequence was obtained from the UniProtKB database and further investigated through PEP-UWM™, PeptideRanker, DBAASP antimicrobial peptide prediction, and ToxinPred tools to identify potential bioactive peptides. Finally, the PepDraw tool was used to obtain the peptide structures of the top-ranked, non-toxic bioactive peptides. Our research deciphered the C. vulgaris PsaB protein as a potential treasure trove of bioactive peptides.

Materials and methods

Retrieval of Photosystem I P700 chlorophyll a apoprotein A2 (PsaB) amino acid sequence

Figure 1 depicts the experimental setup briefly. The amino acid sequence of PsaB was retrieved from UniProtKB (sequence ID: P56342). The protein is part of the light-harvesting complex PS-I and consists of 734 amino acids (Fig. 2).

In silico proteolysis in BIOPEP-UWM: ENZYME(S) ACTION tool

The *in silico* tool allows for the identification of probable peptide fragments from a protein molecule digested by a proteolytic enzyme. Theoretical peptides from PsaB were obtained for further analysis in this study using the BIOPEP-UWM: ENZYME(S) ACTION tool. Peptide sequences were represented using the universal one-letter code for amino acids.[15]

Ranking peptides in PeptideRanker by computing a probability score

PeptideRanker, based on a unique N-to-1 neural network, predicts bioactive peptides (http://distilldeep.ucd.ie/PeptideRanker/).[16] The highest score indicates the most active peptide, while the lowest score indicates the least active. Peptides with unknown bioactivity derived from PsaB were analyzed in PeptideRanker to assess their probability of being bioactive. The score ranged from 0 to 1, and in this study, we set the cutoff score at >0.75 to eliminate false-positive bioactive peptides. The most active peptides were further analyzed.

Prediction of antimicrobial peptides using DBAASP

Theoretical peptides obtained from PsaB were screened in DBAASP (Database of Antimicrobial Activity and Structure of Peptides) to identify potential antimicrobial peptides.[17]

Probability score of bioactive peptide toxicity

The toxicity of bioactive peptides is the major hurdle to their sustainable utilization for functional foods or nutraceuticals. ToxinPred provides tools to design and identify toxic and non-toxic peptides.[18] We used the 'Batch Submission' tool to identify toxic peptides from the pool of theoretical peptides obtained through *in silico* proteolysis of PsaB.

Physicochemical properties and structure of top-ranked bioactive peptides

The physicochemical properties of a protein or enzyme are crucial for its stability and solubility in water or lipids. For synthetic proteins or peptides, stability and dissolution in a living system should be considered before synthesis, leading to the prediction of their physicochemical properties. The physicochemical features (theoretical molecular mass, isoelectric point, hydrophobicity, and extinction coefficient) and the structure of the top-ranked bioac-

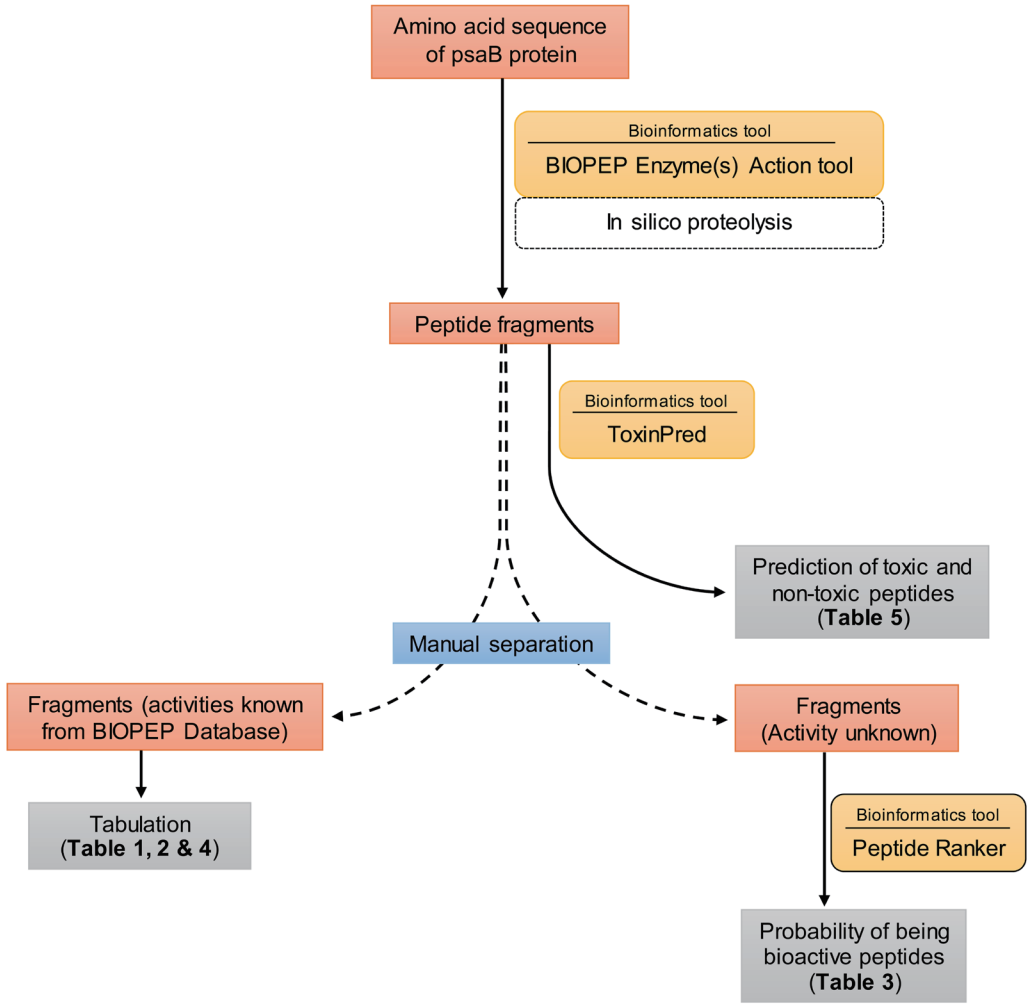

Fig. 1. Experimental setup of the study.

tive peptides were evaluated using the PepDraw (http://pepdraw.com/) tool. This tool draws the primary structure of peptides and calculates the theoretical properties of each peptide.

Prediction of allergenicity of bioactive peptides

Before using a drug or food additive for human health, it is critical to conduct an allergenicity test. In this study, the Allergen FP v.1.0 tool (http://www.ddg-pharmfac.net/AllergenFP) was used to predict the potential allergenicity of the bioactive peptides.[19]

Results

In silico proteolysis of PsaB protein

Numerous peptides derived from the proteolysis of the PsaB protein are illustrated in Figure 3. Among these, peptides with known

```
10         20         30         40         50         60         70         80         90         100
MATKFPKFSQ ALAQDPTTRR LWFGIATAHD FESHDGMTEE RLYQKIFASH FGQLAIIFLW TSGNLFHVAW QGNFEQWVQD PLHIRPIAHA IWDPHFGQAA
110        120        130        140        150        160        170        180        190        200
VEAFTRGGAS GPVNISTSGV YQWWYTIGIR TNQELYVGSI FLLVLAGLFL FAGWLHLQPS FQPALSWFKN AESRLNHHLA GLFGVSSLAW TGHLVHVAIP
210        220        230        240        250        260        270        280        290        300
ESRGQHVGWD NFLTVLPHPA GLTPFFTGNW AAYAENPDSL SQLFGTGEGS GTAILTFLGG FHPQTQSLWL TDMAHHHLAI AVVFILAGHM YRTIFGIGHS
310        320        330        340        350        360        370        380        390        400
MREILEAQTP PSGRLGAGHK GLYDTVNNSL HFQLGLALAS VGTICSLVAQ HMYSLPPYAF LAQDFTTQAS LYTHHQYIAG FIMCGAFAHG AIFFVRDYDP
410        420        430        440        450        460        470        480        490        500
EANRGNVLAR VLDHKEAIIS HLSWVSLFLG FHTLGLYVHN DVVQAFGTPE KQILIEPVFA QWIQAAHGKT VYGFDFLLSS ATSAPSLAGQ SLWLPGWLQG
510        520        530        540        550        560        570        580        590        600
INSDTNSLFL TIGPGDFLVH HAIALGLHTT TLILVKGALD ARGSKLMPDK KDFGYSFPCD GPGRGGTCDI SAWDAFYLAV FWMLNTIGWV TFYFHWKHLG
610        620        630        640        650        660        670        680        690        700
IWQGNVNQFN ESSTYLMGWL RDYLWLNSSQ LINGYNPFGM NSLSVWAWMF LFGHLIYATG FMFLISWRGY WQELIETLAW AHERTPLANL VRWRDKPVAL
710        720        730
SIVQARLVGL THFSVGYVLT YAAFLIASTS GKFG
```

Fig. 2. Peptide sequence of PsaB protein from *Chlorella vulgaris*.

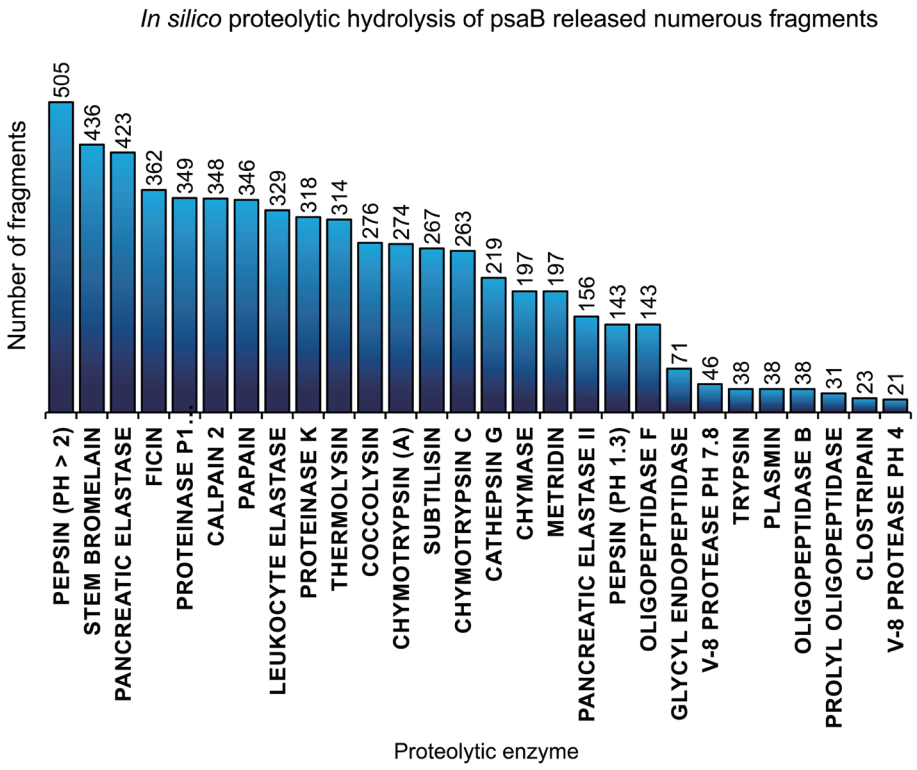

Fig. 3. Number of released peptide fragments for each enzyme.

bioactivities were displayed by searching for active fragments in the BIOPEP-UWM™ database. These peptides are listed with their specific bioactivities in Tables 1 and 2. The bioactive peptides obtained from hydrolysis were categorized into two groups: enzyme inhibitory and antioxidative activities, and regulatory activity. Regarding enzyme inhibitory functions, the generated peptides showed effects such as angiotensin-converting enzyme (ACE) inhibitor, dipeptidyl peptidase IV inhibitor, dipeptidyl peptidase III inhibitor, prolyl endopeptidase inhibitor, renin inhibitor, CaMPDE inhibitor, and alpha-glucosidase inhibitor. On the other hand, regulatory activities included glucose uptake stimulation, neuropeptide regulation (regulating stomach mucosal membrane activity and ion flow), antithrombotic activity, activation of ubiquitin-mediated proteolysis, immunomodulation, calcium binding, antibacterial activity, and anti-inflammatory activity. Thus, our in silico enzymatic hydrolysis generated a significant number of bioactive peptides with diverse pharmacological roles.

Prediction of bioactive peptides by PeptideRanker

Predicting peptides for their bioactivity before obtaining them from bulk protein is important as it reduces costs and time. PeptideRanker helps in this process by reducing both costs and time efforts for investigating the bioactivity of novel peptides. Using PeptideRanker, we ranked the peptides derived from PsaB according to their probability scores. Data shown in Tables 3 and 4 indicate a high probability of these peptides being bioactive. To enhance the likelihood of bioactivity, the cutoff score of >0.75 was set to exclude the possibility of false positives. The top-scoring bioactive peptides generated by chymotrypsin A, pepsin (pH 1.3), proteinase K, pancreatic elastase, thermolysin, chymotrypsin C, papain, ficin, leukocyte elastase, metridin, stem bromelain, calpain 2, pepsin (pH > 2), coccolysin, and subtilisin were PGW (0.98), SWF (0.99), GGF (0.98), WFG (0.99), FWM (0.99), FHW (0.99), AWMF (0.99), VWAWMF (0.98), WMFL (0.99), MGW (0.98), WMF (0.99), WFG (0.99), WWY (0.99), FWM (0.99), and MGW (0.98), respectively. These data suggest a higher bioactivity profile for peptides derived from the PsaB protein of *C.* vulgaris. The top-ranked bioactive peptide structure formulae are mentioned in Figure 4.

Physicochemical characteristics and primary structure of PsaB-derived peptides

As mentioned earlier, predicting the physicochemical properties of peptides is crucial for assessing their stability and dissolution. We evaluated these properties for several bioactive peptides derived from the PsaB protein using PepDraw. The molecular masses of the PsaB-derived peptides ranged from 0.25 to 0.7 kDa, as expected, and most peptides exhibited acidic isoelectric points with moderate to low water solubility (Table 4). Although many of the higher-scored bioactive peptides showed increased hydrophobicity, most were non-allergenic, which suggests potential safety concerns.

Antimicrobial peptides from PsaB

Microbial resistance to drugs is a significant issue for the utilization of antibiotics or other drugs. Therefore, peptides with increased antimicrobial potency could aid in antimicrobial drug development. In our study, we used DBAASP's 'Prediction of general antibacterial activity' tool to identify antimicrobial peptides derived from the proteolytic digestion of PsaB, as listed in Table 5. Bioactive peptides generated from proteolytic digestion mostly consist of tetrapeptides to polypeptides, such as ATKF and IASTSGKF by chymotrypsin A.

Table 1. Theoretical bioactive peptides with enzyme inhibitory and antioxidative activities derived from *in-silico* hydrolysis of PsaB protein

Activity → proteolytic enzyme ↓	ACE inhibitor	Dipeptidyl peptidase IV inhibitor	Antioxidative	Dipeptidyl peptidase III inhibitor	Renin inhibitor	Prolyl endopeptidase inhibitor (Antiamnestic)	CaMPDE inhibitor	Alpha-glucosidase inhibitor (EC 3.2.1.20)
Chymotrypsin A EC 3.4.21.1	IY, GY, AW, GW, AF, GF, GM, GL, GH, SF, AH, PH, TF, DF, IL	AL, SL, GL, VGL, AW, AF, AH, DN, GF, GH, GW, GY, IL, IM, IN, KH, PF, PH, QF, QL, QW, QY, SF, SW, TF, TH, TL, VH, VL, VN	IY, AH, AW, RDY, ISW	TF, GF, PF	QF, SF, TF			
Pepsin (pH 1.3) EC 3.4.23.1	MF, GF, GL, HL, TF, DF, IL, WL	HL, AL, GL, WL, WF, VGL, GF, HF, IL, MF, QL, TF, VL, YF, YL	HL	YF, YL, TF, GF, HL, HF	TF			
Proteinase K (Endopeptidase So) EC 3.4.21.67	GY, GP, AW, GW, RW, RP, AF, GF, GI, GM, GL, HL, GV, AI, SF, KF, HP, TF, AV, TP, DF, QP	GP, TP, RP, HP, EP, NP, QP, HL, AL, SL, GL, AW, AF, AV, DP, GF, GI, GV, GW, GY, HF, HI, HV, HW, KF, QL, QW, RW, SF, SI, SV, SW, TF, TI, TV	HL, RW, AW, RDY	RW, TF, GF, HL, HF, HP	KF, SF, TF	GP	KF	
Pancreatic elastase EC 3.4.21.36	RL, FY, PL, HL, KG, FG, MG, HG, QG, EG, EA, NG, PG, KL, WA, WL, RG	MA, FA, HA, WA, FL, WV, HL, WRG, PL, WL, WT, EG, ES, ET, HI, HS, HT, HV, KG, KT, MG, NG, NL, NT, NV, PG, PS, PV, QA, QG, QL, QS, QT, RG, RL	HL, RDY	RV, HL, FA, FL	FT	PG		EA
Thermolysin EC 3.4.24.27	FQP, LW, LPP, VW, YG, AW, IRP, VG, IG, AG, FG, LG, AR, VE, AH, IEP, LSW, FTTQ, FQ	LW, AW, YT, AG, AH, AS, AT, FQ, IQ, LH, LT, VE, VG, VH, VN, VQ, VS, VT, VW, YG, YS	LH, VHH, AH, AW, VW, LW, YQK	LW, FM, YG	LW			VW, VE
Chymotrypsin C EC 3.4.21.2	RL, IY, GY, HHL, AW, GW, GL, HL, IE, TE, TQ, HP, ASL, TP, IL	TP, HP, FL, HL, AL, SL, GL, VGL, AW, AE, DN, DP, FN, GW, GY, IL, IN, IQ, RL, SW, TE, TL, TQ, VL, VN, VQ	HL, IY, AW, RDY, ISW	HL, HP, FL				
Cathepsin G EC 3.4.21.20	IY, GY, AF, GF, GM, GL, GH, SF, AH, PH, TF, DF, IL, WM, WL	AL, SL, GL, WL, WM, WF, VGL, AF, AH, GF, GH, GY, IL, IM, NH, PH, QL, QY, SF, TF, TH, TL, VH, VL	IY, AH, RDY	TF, GF, WM	SF, TF			
Chymase EC 3.4.21.39	IY, MF, GY, AW, AF, GF, GL, HL, SF, TF, DF, IL	HL, AL, SL, GL, VGL, AW, AF, GF, GY, HF, HW, IL, MF, ML, QL, QW, SF, SW, TF, VL	HL, IY, AW, RDY, ISW	TF, GF, HL, HF	SF, TF			
Papain EC 3.4.22.2	IR, IY, MF, HIR, YW, AY, PL, AW, AF, IF, VG, IG, AG, HL, MG, QG, AI, SG, EG, PG, VR, NF, KF, AR, AH, HP, ASL, DF, DM, IL, WL, QP	HP, QP, HL, AL, VR, PL, WR, WL, WF, AW, YT, AF, AG, AH, AT, AY, EG, HF, HH, HT, HV, IL, IR, KF, KT, MF, MG, NF, NL, NT, NW, PF, PG, PI, QF, QG, QL, QT, VG, VL, YF, YL, YW	HL, HH, VHH, AY, IY, AH, YVL (showed an ORAC-FL value of 0.96µmol Trolox* equivalents per µmol of peptide), IR (Oxygen radical scavenging), AW	YF, YL, HL, HF, HP, PF	IR, KF, QF	PG	IR, KF	
Ficin EC 3.4.22.3	IR, IY, MF, MY, TVY, VK, AF, IF, VG, IG, AG, MG, QG, TG, EG, PG, VR, QK, DG, NF, AR, AH, PH, TF, DY, DF, IL, WL	AL, VR, WR, WK, WL, WF, AF, AG, AH, AS, EG, ES, IL, IR, MF, MG, MY, NF, NH, NL, PG, PH, PK, QG, QH, QL, QS, QY, TF, TG, TH, TL, TR, TS, TY, VG, VH, VK, VL, VS	IY, AH, MY (stimulates expression of the antioxidant defense protein HO-1 in a concentration-depended manner), IR, TY	TF	IR, TF	PG	IR	

(continued)

Table 1. *(continued)*

Activity → proteolytic enzyme ↓	ACE inhibitor	Dipeptidyl peptidase IV inhibitor	Antioxidative	Dipeptidyl peptidase III inhibitor	Renin inhibitor	Prolyl endopeptidase inhibitor (Antiamnestic)	CaMPDE inhibitor	Alpha-glucosidase inhibitor (EC 3.2.1.20)
Leukocyte elastase EC 3.4.21.37	RL, PL, GPV, YA, GI, GA, GL, HL, GS, GV, GT, EA, KL, YV, WA, WL, GHS	MA, FA, HA, GA, WA, FL, WV, HL, GL, PL, WL, WT, YT, ES, ET, GI, GV, HI, HT, HV, NL, NT, PS, QA, QL, QS, QT, RL, YA, YL, YV, GPV	HL	YL, RV, HL, FA, FL	FT, YA, GHS			EA
Metridin EC 3.4.21.3	IY, MF, GY, AW, AF, GF, GL, HL, SF, TF, DF, IL	HL, AL, SL, GL, VGL, AW, AF, GF, GY, HF, HW, IL, MF, ML, QL, QW, SF, SW, TF, VL	HL, IY, AW, RDY, ISW	TF, GF, HL, HF	SF, TF			
Pancreatic elastase II EC 3.4.21.71	GF, GM, GL, HL, TF, DF, IL, WM, WL	HL, AL, GL, WL, WM, WF, VGL, GF, HF, IL, IM, QL, TF, VL, YF, YL	HL	YF, YL, TF, GF, HL, HF, WM	TF			
Stem bromelain EC 3.4.22.32	IR, MF, HIR, YG, PL, IA, YA, IF, IG, HL, KG, MG, HG, QG, EG, EA, PG, NF, KF, KL, DF, YV, IL, WA, WL	MA, HA, IA, WA, WV, HL, PL, WR, WL, WT, WF, YT, EG, ES, HF, HS, HT, HV, IL, IR, KF, KG, KT, MF, MG, NF, NL, NR, NT, NV, PF, PG, PS, PV, QA, QG, QL, QS, QT, YA, YF, YG, YL, YS, YV,	HL, IR	YF, YL, HL, HF, PF, YG	IR, KF, NR, YA	PG	IR, KF	EA
Calpain 2 EC 3.4.22.53	IR, IY, HIR, AY, PL, AW, VK, AF, AP, IF, VG, IG, AG, HL, FG, MG, AI, SG, EG, PG, VR, IFG, AR, AH, PP, HK, FNE, AFL, DM, IL, WL, ST, DFG	PP, AP, FL, HL, AL, SL, VR, PL, WR, WL, AW, YT, AE, AF, AG, AH, AT, AY, EG, HF, HT, IL, IN, IR, MG, MN, NL, NQ, NT, NW, PG, PI, PV, SK, VG, VK, VL, YI, YL, YQ	HL, VHH, AY, IY, AH, EL, YVL, IR, AW	YL, HL, HK, HF, FL, YI	IR	PG	IR	PP
Proteinase P1 (lactocepin) EC 3.4.21.96	HIR, AW, IA, GW, AG, AI, SG, EA, SF, KF, AR, IE, EK, TF, AV, FNE, YV, ST	FA, IA, WV, EK, AIAV, WI, AW, AG, AV, GW, HF, HV, KF, NN, SF, SK, SV, TF, TV, YV	NHH, AW, RDY	TF, HF, FA	KF, SF, TF		KF	EA
Pepsin (pH > 2) EC 3.4.23.1	RL, IY, VF, VY, HHL, PL, IRP, VK, IA, IF, VG, IG, HL, HG, SG, PG, SF, IE, VE, PT, HK, IL, WA, WM, WL, RG, ST	VA, PA, HA, IA, WA, HL, SL, PL, WL, WQ, WM, WT, WF, HD, HE, HF, HT, IL, IM, IN, IQ, PF, PG, PK, PT, RG, RL, SF, SK, SW, VE, VF, VG, VH, VK, VL, VN, VQ, VT, VY, WD	HL, IY, PHF, VY	HL, HK, HF, PE, PF, WM, VY	SF	PG		VE, PE
Coccolysin EC 3.4.24.30	FQP, LW, LPP, YG, AW, IRP, IG, AG, FG, LG, AR, AH, LSW, FTTQ, AV, IVQ, FQ, YV, AVV	LW, AW, YT, AG, AH, AS, AT, AV, FQ, IQ, LH, LT, LV, YG, YS, YV	LH, AH, AW, LW, YQK	LW, FM, YG	LW			
Subtilisin EC 3.4.21.62	RL, IY, VF, MF, VW, VY, GY, GP, AW, AF, GF, IF, GL, HL, KL, AR, IEP, TF, DF, IL	GP, HL, AL, GL, VGL, AW, AF, AS, ES, GF, GY, HF, HW, IL, MF, ML, QL, QW, RL, TF, TS, TY, VF, VH, VL, VS, VW, VY	HL, IY, TY, VY, AW, VW, RDY	TF, GF, HL, HF, VY	TF	GP		VW

ACE, angiotensin converting enzyme; CaMPDE, calmodulin (CaM)-dependent cyclic nucleotide phosphodiesterase.

Table 2. Bioactive peptides with regulatory activities derived from PsaB protein

Activity→ proteolytic enzyme ↓	Glucose uptake stimulating	Neuropeptide	Regulating	Antithrombotic	Activating ubiquitin-mediated proteolysis	Immunomodulating	Binding	Antibacterial	Anti-inflammatory
Papain EC 3.4.22.2	VL, IL	YL (Anxiolytic)	PG (regulating the stomach mucosal membrane activity)	PG			YDT (calcium binding peptide)	YVL (active against mainly gram-positive bacteria)	YW
Calpain 2 EC 3.4.22.53	VL, IL	YL (Anxiolytic)	PG, SL	PG			YDT (calcium binding peptide)	YVL (active against mainly gram-positive bacteria)	
Stem bromelain EC 3.4.22.32	VL, IL	YL (Anxiolytic)	PG	PG	WA	YG (enhancing protein biosynthesis in lymphocytes)	YDT (calcium binding peptide)		
Chymotrypsin A EC 3.4.21.1	VL, IL		SL (Regulator of phospho-glycerate kinase activity)						
Pepsin (pH 1.3) EC 3.4.23.1	VL, IL	YL (Anxiolytic)							
Proteinase K (Endopeptidase So) EC 3.4.21.67			GP, SL	GP					
Pancreatic elastase EC 3.4.21.36			PG	PG	WA				
Thermolysin EC 3.4.24.27						YG	YDT (calcium binding peptide)		
Chymotrypsin C EC 3.4.21.2	VL, IL		SL						
Cathepsin G EC 3.4.21.20	VL, IL		SL						
Chymase EC 3.4.21.39	VL, IL		SL						
Ficin EC 3.4.22.3	VL, IL		DY (ion flow regulating), PG	PG					
Leukocyte elastase EC 3.4.21.37		YL (Anxiolytic)			WA		YDT (calcium binding peptide)		
Metridin EC 3.4.21.3	VL, IL		SL						
Pancreatic elastase II EC 3.4.21.71	VL, IL	YL (Anxiolytic)							
Pepsin (pH > 2) EC 3.4.23.1	VL, IL		PG, SL	PG	WA				
Coccolysin EC 3.4.24.30	LV					YG			
Subtilisin EC 3.4.21.62	VL, IL		GP	GP					

Table 3. Scoring of peptides derived from PsaB (by proteolytic digestion) as a prediction of having bioactivity

Chymotrypsin A		*Proteinase K*		*Thermolysin*	
0.987911	PGW	0.840052	QGNF	0.922013	FQP
0.987345	GGF	0.833322	AAF	0.902121	FPK
0.959298	AGW	0.819503	SAW	0.897906	FS
0.958541	CGAF	0.804969	DAF	0.877191	FGQ
0.954702	GIW	0.793307	QAF	0.859479	FTGNW
0.953169	GM	0.793186	CDGP	0.845373	IMCG
0.894831	PKF	0.78476	AQW	0.844668	FPCDGPGRGGTCD
0.881794	TPF	0.775374	DNF	0.833345	FHWKH
0.868141	PPY	*Pancreatic elastase*		0.828049	LPP
0.852204	VGW	0.996278	WFG	0.78476	AQW
0.849992	AIW	0.994262	WMFL	0.779545	ISWRG
0.848096	GAIF	0.994192	FWML	0.766312	FEQW
0.844471	QPSF	0.994063	FMFL	0.758608	IWDPH
0.842881	PCDGPGRGGTCDISAW	0.977738	FDFL	0.755836	AWD
0.833322	AAF	0.969838	QWWY	*Chymotrypsin C*	
0.832896	IAGF	0.967512	FFV	0.990882	FHW
0.804969	DAF	0.967425	WRG	0.989795	FGM
0.78476	AQW	0.963529	WDPHFG	0.982697	FAGW
0.76345	VRW	0.958912	WDNFL	0.971483	GFDFL
0.752534	DGM	0.958072	PFFT	0.954702	GIW
Pepsin pH 1.3		0.950178	MCG	0.939625	SFP
0.99088	SWF	0.948084	FI	0.928594	AFL
0.987345	GGF	0.939404	FPCDG	0.912685	GDFL
0.984773	WML	0.933871	NPFG	0.910107	AVFW
0.979473	PGWL	0.904268	HFG	0.899536	FGHL
0.979247	MGWL	0.904195	WQG	0.890489	GGFHP
0.978146	SVWAWMF	0.889906	DFL	0.839976	FFTGN
0.948718	AGWL	0.869072	FHWKHL	0.830136	ATGFM
0.94657	PPYAF	0.868141	PPY	0.822086	AAFL
0.894831	PKF	0.842785	KFG	0.816879	FHVAW
0.881794	TPF	0.81261	FQPA	0.796958	IAGFIM
0.876307	GYSF	0.790112	QWI	0.793186	CDGP
0.856383	IMCGAF	0.751998	NWA	0.789885	AIIFL
0.844471	QPSF	*Thermolysin*		0.76345	VRW
0.838569	PCDGPGRGGTCDISAWDAF	0.997525	FWM	0.758123	TFL
0.795121	INGYNPF	0.972599	AWM	0.757343	ISW
0.788788	PHPAGL	0.97045	LPGW	*Papain*	
0.766454	AHGAIF	0.968639	YQWW	0.992532	AWMF
Proteinase K		0.965053	LMGW	0.99088	SWF
0.987345	GGF	0.959298	AGW	0.984773	WML
0.959298	AGW	0.952832	FH	0.961534	WDNF
0.958541	CGAF	0.951958	IGW	0.937045	AIWDPHF
0.95682	AGF	0.925995	FGH	0.92937	YNPF
0.941533	GDF	0.922094	FD	0.897085	QWWYT

(continued)

Table 3. *(continued)*

Papain		*Leukocyte elastase*		*Stem bromelain*	
0.894831	PKF	0.899536	FGHL	0.904195	WQG
0.868141	PPY	0.894015	FPCDGPGRGGT	0.894831	PKF
0.845373	IMCG	0.865062	YQWWYT	0.865062	YQWWYT
0.844471	QPSF	0.854452	GWV	0.845373	IMCG
0.839011	YSF	0.836456	FGT	0.782785	IIF
0.833341	AIF	0.81261	FQPA	0.752918	PCDG
0.805497	SIF	0.806335	GNWA	0.751998	NWA
0.790112	QWI	0.796542	MCGA	*Calpain 2*	
0.755836	AWD	0.796185	WDPHFGQA	0.996278	WFG
0.752918	PCDG	0.790112	QWI	0.994063	FMFL
0.750945	ISWR	0.787708	FGV	0.988038	AWMFL
Ficin		0.77598	GKFG	0.977738	FDFL
0.986284	VWAWMF	0.775024	GRL	0.962274	FYFHWK
0.984773	WML	0.75667	WRGYWQEL	0.960563	AVFWML
0.969838	QWWY	*Metridin*		0.959178	FIMCG
0.961534	WDNF	0.989792	MGW	0.958912	WDNFL
0.951874	NPF	0.987911	PGW	0.958072	PFFT
0.940814	AWDAF	0.987345	GGF	0.947248	WWYT
0.881794	TPF	0.959298	AGW	0.937842	SWFK
0.868141	PPY	0.954702	GIW	0.935844	AIWDPHFG
0.845373	IMCG	0.951874	NPF	0.934192	AFG
0.838591	WVTF	0.896783	DPHF	0.928594	AFL
0.833341	AIF	0.894831	PKF	0.919895	IFG
0.833322	AAF	0.881794	TPF	0.919474	YNPFG
0.760915	AIWDPH	0.868141	PPY	0.910435	DFG
0.752918	PCDG	0.856383	IMCGAF	0.902121	FPK
Leukocyte elastase		0.844471	QPSF	0.901482	SFPCDG
0.994262	WMFL	0.842881		0.890464	AFYL
0.994192	FWML	0.840052	QGNF	0.889906	DFL
0.991137	GFMFL	0.833322	AAF	0.878963	SWR
0.981018	WFGI	0.832896	IAGF	0.868141	PPY
0.980505	GWL	0.804969	DAF	0.843778	SIFL
0.979473	PGWL	0.788788	PHPAGL	0.812713	WT
0.979247	MGWL	0.78476	AQW	0.808892	YWQ
0.967512	FFV	0.775374	DNF	0.801775	AIFFVR
0.95973	YGFDFL	0.766454	AHGAIF	0.794483	SHFG
0.959629	GWDNFL	0.76345	VRW	0.789885	AIIFL
0.958072	PFFT	0.757343	ISW	0.755836	AWD
0.950547	FYL	*Stem bromelain*		*Pepsin pH>2*	
0.948084	FI	0.997633	WMF	0.992227	WWY
0.945267	GPGDFL	0.984773	WML	0.99088	SWF
0.943742	GFI	0.961534	WDNF	0.938016	PHF
0.930864	FYFHWKHL	0.941151	IWDPHF	0.93391	SW
0.920052	FGI	0.92937	YNPF	0.931922	CG

(continued)

Table 3. *(continued)*

Pepsin pH>2		*Subtilisin*	
0.920402	PSF	0.989792	MGW
0.868141	PPY	0.987911	PGW
0.80286	SHF	0.987345	GGF
0.779545	ISWRG	0.959298	AGW
Coccolysin		0.954702	GIW
0.997525	FWM	0.951874	NPF
0.972599	AWM	0.896783	DPHF
0.97045	LPGW	0.894831	PKF
0.968639	YQWW	0.883411	GKF
0.965053	LMGW	0.881794	TPF
0.959298	AGW	0.868141	PPY
0.925995	FGH	0.856383	IMCGAF
0.922013	FQP	0.852204	VGW
0.902121	FPK	0.840052	QGNF
0.877191	FGQ	0.833322	AAF
0.859479	FTGNW	0.832896	IAGF
0.845373	IMCG	0.804969	DAF
0.844668	FPCDGPGRGGTCD	0.788788	PHPAGL
0.833345	FHWKH	0.78476	AQW
0.828049	LPP	0.775374	DNF
0.78476	AQW	0.775024	GRL
0.779545	ISWRG	0.766454	AHGAIF
0.758608	IWDPH	0.76345	VRW
0.755836	AWD		

Cut off score > 0.75.

Toxicity profile of all peptides

In silico digestion of PsaB by various proteolytic enzymes generates numerous peptide fragments. It is crucial to determine the toxicity level of each peptide to ensure safety. The toxicity profile of these peptides is shown in Tables 6. Data indicate that the peptides generated by proteolytic hydrolysis of PsaB were mostly non-toxic. These findings suggest the increased possibility of using the PsaB-derived peptides.

Allergenicity prediction of bioactive peptides

The frequency of food allergies is increasing, highlighting the need to assess the allergenicity of food additives or drugs beforehand. In this study, some top-ranked peptides generated by *in silico* proteolysis were analyzed for allergenicity. According to Table 4, a few peptides were found to be potential allergens, including WML, FWML, FAGW, QWWY, GFMFL, MWG, WDNF, FMFL, AWMFL, and PHF. However, the other top-ranked peptides were not allergenic. Although these few bioactive peptides exhibited allergenic properties, they were also non-toxic. This necessitates further wet lab studies, including cell line and *in vivo* mouse studies, particularly for these selected peptides.

Discussion

This study focuses on the potential of obtaining bioactive peptides from *C. vulgaris*, a marine or freshwater microalga commonly used as a food supplement. In addition to its food applications, *C. vulgaris* is also considered a promising candidate for bioremediation and biofuel production due to its rapid growth rate.[20] As a unicellular photosynthetic microalga,[21] *C. vulgaris* produces a variety of proteins and enzymes to capture photons from sunlight, supporting its fast growth. Photosystem I P700 chlorophyll and apoprotein A2 (PsaB) bind P700, playing a crucial role in photosynthesis. Besides binding P700, this protein has several other important functions, such as binding 4 iron–4 sulfur clusters, facilitating electron transfer, and binding magnesium ions.[22] In this study, PsaB has been targeted as a parent protein for *in silico* analysis to identify bioactive peptides. It has been previously shown that Rubisco, a key enzyme for CO_2 fixation, can be an excellent source of bioactive peptides.[5,23]

PsaB, a crucial protein of the P700 (photosystem I), contains many bioactive peptides within its amino acid sequence. These active peptide fragments can be released by digesting the protein with various proteolytic enzymes (Tables 1 and 2). At least 17 different bioactivities of the peptides have been identified. In addition to these active peptides, proteolytic digestion of PsaB released numerous peptide fragments (Fig. 3). Proteolytic enzymes such as pepsin (pH > 2), stem bromelain, pancreatic elastase, ficin, proteinase P1, calpain 2, and papain released the highest number of peptides from PsaB. Although many peptides obtained from *in silico* proteolysis had unknown biological functions, their potential could not be ruled out. The peptides shown in Tables 3 and 4, and in Figure 4, were ranked by PeptideRanker, indicating that these peptides derived from PsaB have high probabilities of being bioactive. Therefore, these peptides warrant further analysis in both dry and wet laboratories to explore their potential interactions with biomolecules. Based on wet lab data, further docking analysis will reveal the precise interaction patterns of each peptide with corresponding enzymes. Additionally, several peptides should be tested in laboratories to assess their antimicrobial potency, as suggested by DBAASP's antimicrobial activity prediction algorithms (Table 5). Antibiotics are valuable but limited resources in the fight against infectious diseases. However, the rapid increase in antimicrobial resistance poses a global public health threat. In this context, the development and modification of antimicrobial peptides may offer a solution, as developing resistance against antimicrobial peptides would be a slower and costlier process for microbes.[24] From the toxicity profile of the peptide pool, we observed that most peptides generated from PsaB digestion were non-toxic (Table 6). Therefore, these peptides are likely safe for use as medicine or dietary supplements.

As demonstrated in Tables 1 and 2, PsaB can be a potent source of bioactive peptides with pharmaceutical value. ACE inhibitors and renin inhibitors are commonly used to treat high blood pressure and cardiovascular disease.[25] Furthermore, peptides with ACE inhibitory activity may be useful in treating SARS-CoV-2 infection.[5,26] Peptides that inhibit DPP-IV and stimulate glucose uptake can be used to treat type 2 diabetes.[2,27,28] Alpha-glucosidase inhibitors can lower blood glucose levels by delaying carbohydrate digestion. Other peptides with antioxidative, anxiolytic, prolyl endopeptidase inhibitory, CaMPDE inhibitory, and antibacterial activities may be beneficial for treating aging and cancer, anxiety, neurodegenerative diseases, erectile dysfunction, and bacterial infections, respectively. Additionally, our study identified calcium-binding peptides, DPP-III inhibitors, regulators of phosphoglycer-

Table 4. Theoretical physicochemical properties and allergenicity prediction of some top-ranked bioactive peptides generated by proteolytic hydrolysis

Enzymes	Peptide sequence	Length	Mass (g/mol)	Isoelectric point (pI)	Hydrophobicity (Kcal mol^{-1})	Extinction coefficient ($M^{-1}cm^{-1}$)	Allergenicity Prediction
Chymotrypsin A	PGW	3	358.1637	5.75	+7.10	5,500	Probable non-allergen
	GGF	3	325.1093	5.13	+7.32	0	Probable non-allergen
	AGW	3	332.1481	5.71	+7.46	5,500	Probable non-allergen
Pepsin (pH 1.3)	SWF	3	438.1898	5.39	+4.56	5,500	Probable non-allergen
	GCF	3	325.1093	5.13	+7.32	0	Probable non-allergen
	WML	3	448.2138	5.53	+3.89	5,500	Probable allergen
Proteinase K	GGF	3	279.1216	5.47	+8.49	0	Probable non-allergen
	AGW	3	332.1481	5.71	+7.46	5,500	Probable non-allergen
	CGAF	4	396.1263	5.13	+7.82	0	Probable non-allergen
Pancreatic elastase	WFG	3	408.1793	5.55	+5.25	5,500	Probable non-allergen
	WMFL	4	595.2820	5.53	+2.18	5,500	Probable non-allergen
	FWML	4	595.2820	5.52	+2.18	5,500	Probable allergen
Thermolysin	FWM	3	482.1982	5.35	+3.43	5,500	Probable non-allergen
	AWM	3	406.1670	5.41	+5.64	5,500	Probable non-allergen
	LPGW	4	471.2475	5.69	+5.85	5,500	Probable non-allergen
Chymotrypsin C	FHW	3	488.2167	7.68	+6.43	5,500	Probable non-allergen
	FGM	3	353.1405	5.35	+6.67	0	Probable non-allergen
	FAGW	4	479.2163	5.62	+5.75	5,500	Probable allergen
Papain	AWMF	4	553.2352	5.47	+3.93	5,500	Probable non-allergen
	SWF	3	438.1898	5.39	+4.56	5,500	Probable non-allergen
	WML	3	448.2138	5.53	+3.89	5,500	Probable allergen
Ficin	VWAWMF	6	838.3825	5.46	+1.38	11,000	Probable non-allergen
	WML	3	448.2138	5.53	+3.89	5,500	Probable allergen
	QWWY	4	681.2903	5.37	+3.78	12,490	Probable allergen
Leukocyte elastase	WMFL	4	595.2820	5.53	+2.18	5,500	Probable non-allergen
	FWML	4	595.2820	5.52	+2.18	5,500	Probable allergen
	GFMFL	5	613.2925	5.58	+3.71	0	Probable allergen
Metridin	MWG	3	392.1514	5.53	+6.29	5,500	Probable allergen
	PGW	3	358.1637	5.75	+7.10	5,500	Probable non-allergen
	GGF	3	279.1216	5.47	+8.49	0	Probable non-allergen
Stem bromelain	WMF	3	482.1982	5.42	+3.43	5,500	Probable non-allergen
	WML	3	448.2138	5.53	+3.89	5,500	Probable allergen
	WDNF	4	580.2275	3.05	+8.59	5,500	Probable allergen
Calpain 2	WFG	3	408.1793	5.55	+5.25	5,500	Probable non-allergen
	FMFL	4	556.2711	5.52	+2.56	0	Probable allergen
	AWMFL	5	666.3190	5.59	+2.68	5,500	Probable allergen
Papsin (pH > 2)	WWY	3	553.2319	5.41	+301	12,490	Probable non-allergen
	SWF	3	438.1898	5.39	+4.56	5,500	Probable non-allergen
	PHF	3	399.1902	8.32	+8.66	0	Probable allergen
Coccolysin	FWM	3	482.1982	5.35	+3.43	5,500	Probable non-allergen
	AWM	3	406.1670	5.41	+5.64	5,500	Probable non-allergen
	LPGW	4	471.2475	5.69	+5.85	5,500	Probable non-allergen
Subtilisin	MGW	3	392.1514	5.61	+6.29	5,500	Probable non-allergen
	PGW	3	358.1637	5.75	+7.10	5,500	Probable non-allergen
	GGF	3	279.1216	5.47	+8.49	0	Probable non-allergen

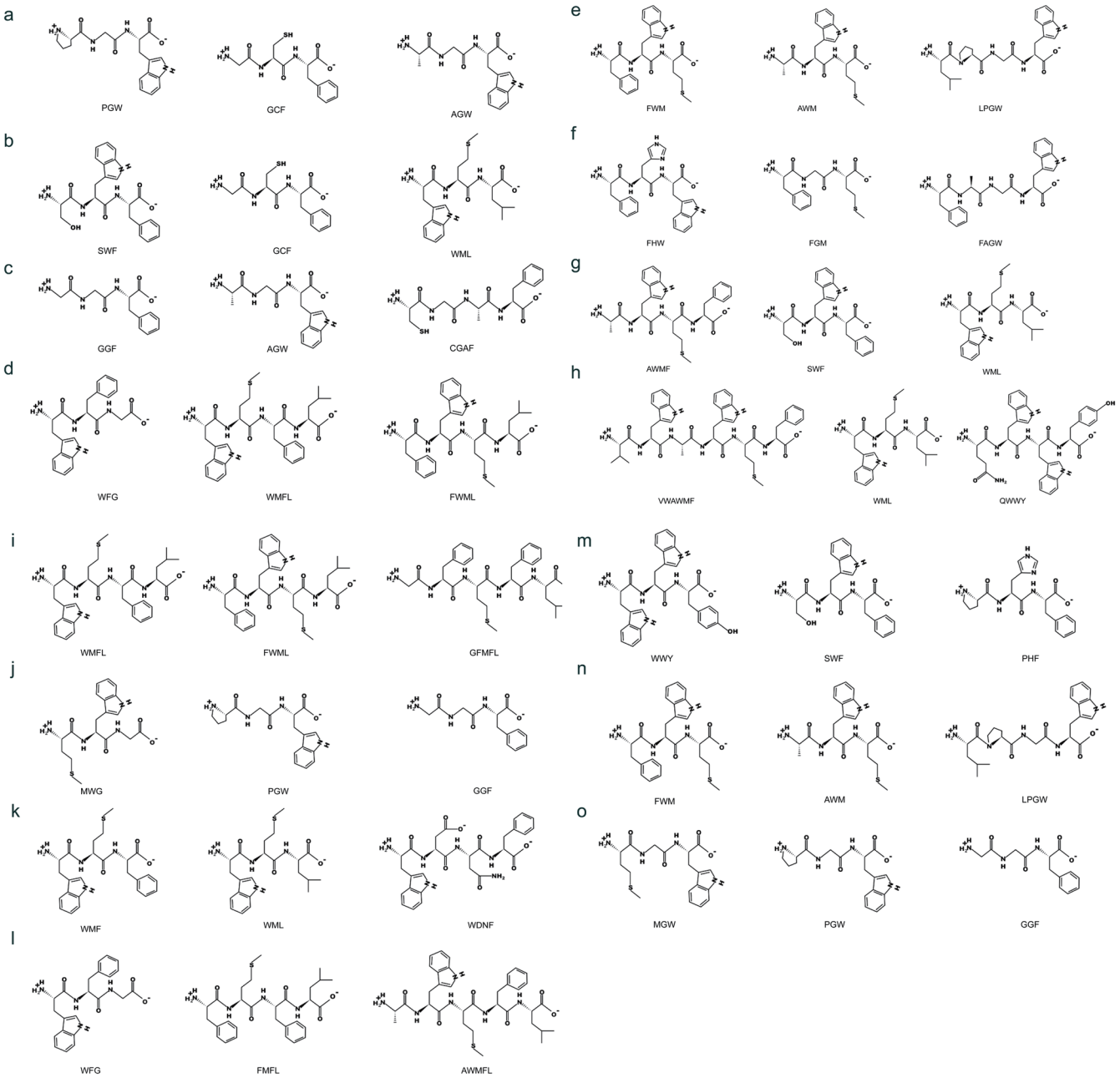

Fig. 4. Structural formulae of top-ranked bioactive peptides generated by proteolytic hydrolysis. Primary structures of peptides generated from various proteolytic enzymes: a. Chymotrypsin A, b. Pepsin (pH 1.), c. Proteinase K, d. Pancreatic elastase, e. Thermolysin, f. Chymotrypsin C, g. Papain, h. Ficin, i. Leukocyte elastase, j. Metridin, k. Stem bromelain, l. Calpain 2, m. Pepsin (pH > 2), n. Coccolysin, o. Subtilisin.

ate kinase and stomach mucosal membrane, antithrombotic peptides, activators of ubiquitin-mediated proteolysis, and peptides that enhance protein biosynthesis in lymphocytes, all of which have significant roles and importance in the medical field.

A single bioactive peptide can possess multiple bioactivities, which can be of great interest for treating patients with multiple diseases such as diabetes, hypertension, and erectile dysfunction.[5] Our study identified a substantial number of peptides with multiple bioactivities. For example, PG, a dipeptide with five different bioactivities (ACE inhibitor, DPP-IV inhibitor, regulator, antiamnestic, and antithrombotic), can be obtained by digesting PsaB with various proteolytic enzymes such as papain, ficin, and calpain 2 (Tables 1 and 2). In the physicochemical properties analysis, we found that most peptides derived from PsaB protein are low molecular weight, mildly acidic, and moderately to poorly water soluble (Table 4; Fig. 4). Our findings are consistent with previous research, which found that most bioactive peptides have a low molecular weight profile.[29]

Toxicity level detection or prediction is a crucial step before developing any drug or food additive.[18,30] In this study, the proteolytic

Table 5. Antimicrobial peptides from PsaB

Proteolytic enzyme	Peptides with predicted antimicrobial activity
Chymotrypsin A	ATKF, QKIF, IRPIAH, RTIF, ARVL, VKGAL, IASTSGKF
Pepsin pH 1.3	MATKF, YQKIF, KNAESRL, AGHMYRTIF, ARVL, VKGAL, HWKHL, IASTSGKF
Proteinase K (Endopeptidase So)	ATKF, KNAESRL
Pancreatic elastage	KFPKFS, WFKNA, RWRDKPV
Thermolysin	–
Chymotrypsin C	ATKFP, KFSQ, KIFASHFGQ, RTIFGIGHSM, ARVL, VKGAL
Papain	HWKHL, HMYR, QKIF, HSMR
Ficin	–
Leukocyte elastase	KFPKFS, WFKNA, GHMYRT, RWRDKPV, GHKGL, YQKI
Metridin	MATKF, RTIF, QKIF, HIRPIAHAIW, IASTSGKF, ARVL, VKGAL, KNAESRL
Stem bromelain	HWKHL, HMYR, YQKIF
Calpain 2	SWFK, HMYR
Pepsin pH > 2	–
Coccolysin	LVKG, LVRWRDKPV
Subtilisin	MATKF, QKIF, HIRPIAHAIW, RTIF, VKGAL
V-8 protease (glutamyl endopeptidase) pH 7.8	MATKFPKFSQALAQD, RLYQKIFASHFGQLAIIFLWTSGNLFHVAWQGNFE, AQTPPSGRLGAGHKGLYD
Trypsin	LYQK, TPLANLVR
Prolyl oligopeptidase	MATKFP, ALSWFKNAESRLNHHLAGLFGVSSLAWTGHLVHVAIP, LANLVRWRDKP
V-8 protease pH 4	RLYQKIFASHFGQLAIIFLWTSGNLFHVAWQGNFE
Plasmin	LYQK, TPLANLVR
Cathepsin G	ATKF, QKIF, IRPIAH, KNAESRL, RTIF, ARVL, VKGAL, IASTSGKF
Clostripain	MATKFPKFSQALAQDPTTR, TPLANLVR
Pancreatic elastage II	ATKF, YQKIF, KNAESRL, YRTIF, ARVL, VKGAL, HWKHL, IASTSGKF
Glycyl endopeptidase	MATKFPKFSQALAQDPTTRRLWFG, WLHLQPSFQPALSWFKNAESRLNHHLAG, HMYRTIFG
Proteinase P1	MATKF, KIFA, AGHMYRTIFGIGH

enzymes used were primarily derived from plant and animal sources and are commonly utilized in many food-processing industries. It is known that peptides with low molecular weights are generally non-toxic and less allergenic compared with their native proteins.[31,32] From the ToxinPred analysis, it was observed that most of the studied peptides are non-toxic, except for FPCDGPGRGGTCD and FPCDGPGRGGTCD (SVM scores < 0) (Table 6). The primary components of non-toxic peptides are V, T, R, Q, M, L, K, I, F, and A, and most bioactive peptides contain these non-toxic amino acid components according to our findings. As a result, these peptides can be considered safe potential functional ingredients, though further *in vitro* and *in vivo* testing is required to prove the safety concern.

Human health issues are driving an increased demand for allergenicity safety concerns. Most allergens are animal and plant-based proteins. Food allergens affect approximately 1–2% of adults and 8% of children.[32] The European Food Safety Authority also encourages the prediction of the probable allergenicity of food proteins through *in silico* approaches.[33] Furthermore, proteolytic hydrolysis by pepsin may result in the elimination of linear epitopes, which is a primary concern for allergenicity.[29,34,35] Only a few of the top-ranked peptides in our study were allergenic (Table 4), according to the AllergenFP v.1.0 tool.[19] The allergenicity of the selected peptides needs to be clarified through *in vitro* and *in vivo* approaches.[33] Additionally, we found that very few of the bioactive peptides are toxic or allergic (Tables 4 and 6), which would raise concerns about the compatibility of purifying these peptides. Generally, bioactive peptides are manufactured using various techniques such as microwave-assisted extraction, chemical hydrolysis, organic synthesis, and enzyme hydrolysis. The targeted peptides are then obtained through further purification methods such as gel filtration, ultrafiltration, size exclusion chromatography, ion-exchange column chromatography, reversed-phase high-performance liquid chromatography, and so on.[36] Thus, after confirming bioactivity *in vitro* and *in vivo* assays, further purification processes could be applied to separate the toxic from the non-toxic or allergenic from the non-allergenic bioactive peptides.

In this study, we observed that some bioactive peptides possess poor water solubility due to slightly increased hydrophobicity (Table 4). This could present a barrier to absorption in the human body, as the body contains a lot of water. Therefore, water solubility is an important factor when considering functional foods, food

Table 6. Toxicity profile of the bioactive peptides generated by various proteolytic enzymes

Proteolytic enzyme	Peptides	SVM score	Prediction
Chymotrypsin A	All	−ve	Non-toxin
Pepsin (pH 1.3)	All	−ve	Non-toxin
Proteinase K	All	−ve	Non-toxin
Pancreatic elastase	All	−ve	Non-toxin
Thermolysin	FPCDGPGRGGTCD	+0.06	Toxin
Chymotrypsin C	All	−ve	Non-toxin
Cathepsin G	All	−ve	Non-toxin
Papain	All	−ve	Non-toxin
Ficin	All	−ve	Non-toxin
Leukocyte elastase	All	−ve	Non-toxin
Metridin	All	−ve	Non-toxin
Pancreatic elastase II	All	−ve	Non-toxin
Stem bromelain	All	−ve	Non-toxin
Calpain 2	All	−ve	Non-toxin
Proteinase P1	All	−ve	Non-toxin
Pepsin (pH > 2)	All	−ve	Non-toxin
Coccolysin	FPCDGPGRGGTCD	+0.06	Toxin
Subtilisin	All	−ve	Non-toxin

SVM, support vector machine (a binary classifier that computes the probabilities of possible outcomes for samples in X within a range of negative (−ve) to positive (+ve) values. If the output of the scoring function is negative, then the input is classified as non-toxic.).

ingredients, or drug development. Allergenicity and toxicity concerns are also significant. A previous *in silico* study observed that water-soluble peptides are not always non-allergenic, while poor water-soluble peptides may be non-allergenic.[37] In our study, we noted that some poor water-soluble peptides might affect absorption efficiency in the human body. Regarding toxicity and allergenicity, most of these poor water-soluble peptides are non-toxic and non-allergenic (Tables 4 and 6). These issues will be further examined by assessing the impact of these peptides in mammalian cell lines and *in vivo* models. Additionally, for better absorption of these peptides, nanoparticle-based or solid dispersion-based approaches could be applied,[38] which are commonly used in therapies for poorly water-soluble drugs.

This study focused solely on the *in silico* aspect of the process (Fig. 5). The theoretical bioactive peptides should be produced and validated in a wet lab with the appropriate proteolytic enzymes, enzyme concentrations, substrate concentrations, optimal pH, and temperature. The limitation of this study is that it only involves an *in silico* approach to identify novel bioactive peptides from this species. The potential bioactive peptides should be tested in wet lab-based experiments using both *in transfecto* and *in vivo* models to clarify their physiological relevance.

Future directions

Finding out of functional peptides has got great attention to all. In this study, we found lots of functional peptides from C. vulgaris having no toxicity and allergenicity that is promising to pharmaceutical industries to produce those peptides for human welfare. Through investigating the in vivo studies, the real finding in biological system will come in front. So that, few more realistic experiments should be done to satisfy the proposed goal.

Conclusions

C. vulgaris is a widely cultivated green microalga, with annual gross production of *Chlorella* biomass exceeding 2000 tons in 2005. It is extensively used in food and as a supplement in many countries. This study highlights the importance of *C. vulgaris* in pharmaceutical and medical contexts and encourages the pharmaceutical industry to explore the production and commercialization of bioactive peptides as commercial products.

Acknowledgments

We thank the authorities of the corresponding online bioinformatics tools.

Funding

This study has not received any funding.

Conflict of interest

All authors declared no conflict of interest in this study.

Author contributions

Conceiving the idea, designing the research, supervising the re-

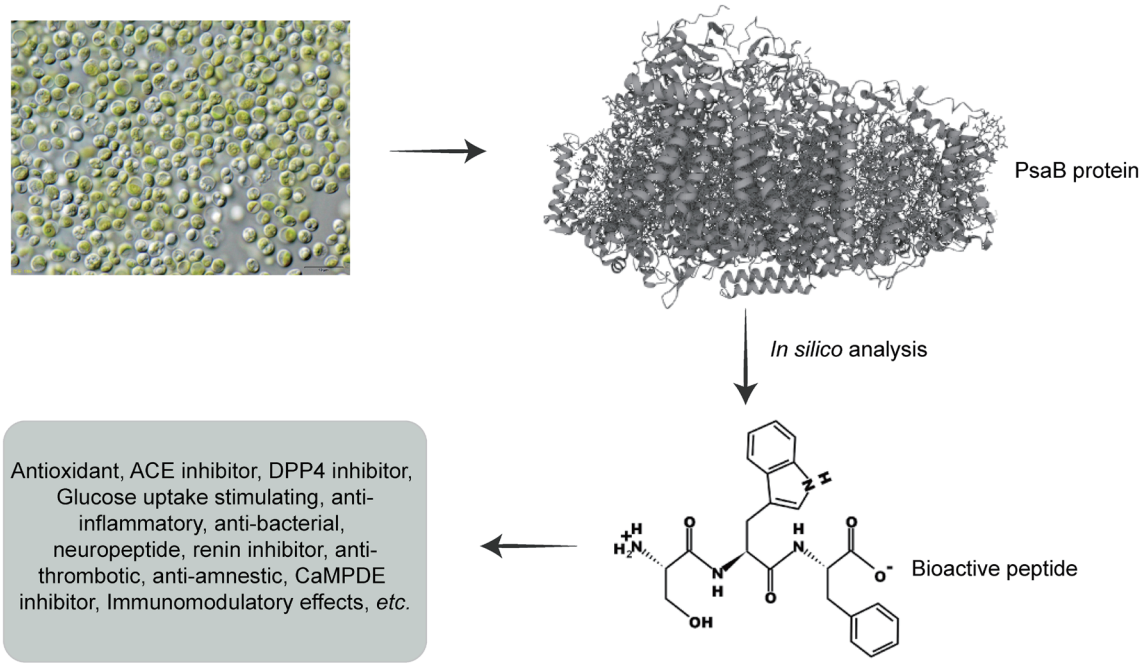

Fig. 5. *Chlorella vulgaris* with bioactive peptides and their pharmacological relevance.

search, analyzing the data, writing, editing, and revising the manuscript (MMA), conducting the research, analyzing the data, writing the initial draft of the manuscript (MAA); and editing the manuscript (UC).

Data sharing statement

No additional data are available.

References

[1] Morshedul A. Therapeutic potential of marine bioactive compounds against SARS-CoV2 infection. CPQ Medicine 2020;11(1):1–18.
[2] Morshedul A. Prospect of marine bioactive peptides as DPP4 inhibitor. Oceanogr Fish Open Access J 2022;14(5):555896. doi:10.19080/OFOAJ.2022.14.555896.
[3] Admassu H, Gasmalla MAA, Yang R, Zhao W. Bioactive Peptides Derived from Seaweed Protein and Their Health Benefits: Antihypertensive, Antioxidant, and Antidiabetic Properties. J Food Sci 2018;83(1):6–16. doi:10.1111/1750-3841.14011, PMID:29227526.
[4] Udenigwe CC, Gong M, Wu S. In silico analysis of the large and small subunits of cereal RuBisCO as precursors of cryptic bioactive peptides. Process Biochem 2013;48(11):1794–1799. doi:10.1016/j.procbio.2013.08.013.
[5] Amin MA, Chondra U, Mostafa E, Alam MM. Green seaweed Ulva lactuca, a potential source of bioactive peptides revealed by in silico analysis. Informatics Med Unlocked 2022;33:101099. doi:10.1016/j.imu.2022.101099.
[6] Cermeño M, Kleekayai T, Amigo-Benavent M, Harnedy-Rothwell P, FitzGerald RJ. Current knowledge on the extraction, purification, identification, and validation of bioactive peptides from seaweed. Electrophoresis 2020;41(20):1694–1717. doi:10.1002/elps.202000153.
[7] Darienko T, Rad-Menéndez C, Campbell C, Pröschold T. Are there any true marine Chlorella species? Molecular phylogenetic assessment and ecology of marine Chlorella-like organisms, including a description of Droopiella gen. nov. Syst Biodivers 2019;17(8):811–829. doi:10.1080/14772000.2019.1690597, PMID:32256217.
[8] Safi C, Zebib B, Merah O, Pontalier PY, Vaca-Garica C. Morphology, composition, production, processing, and applications of Chlorella vulgaris: A review. Renewable Sustainable Energy Rev 2014;35:265–278. doi:10.1016/j.rser.2014.04.007.
[9] Seyfabadi J, Ramezanpour Z, Khoeyi ZA. Protein, fatty acid, and pigment content of Chlorella vulgaris under different light regimes. J Appl Phycol 2011;23(4):721–726. doi:10.1007/s10811-010-9569-8.
[10] de-Bashan LE, Bashan Y, Moreno M, Lebsky VK, Bustillos JJ. Increased pigment and lipid content, lipid variety, and cell and population size of the microalgae Chlorella spp. when co-immobilized in alginate beads with the microalgae-growth-promoting bacterium Azospirillum brasilense. Can J Microbiol 2002;48(6):514–521. doi:10.1139/w02-051, PMID:12166678.
[11] Choix FJ, de-Bashan LE, Bashan Y. Enhanced accumulation of starch and total carbohydrates in alginate-immobilized Chlorella spp. induced by Azospirillum brasilense: I. Autotrophic conditions. Enzyme Microb Technol 2012;51(5):294–299. doi:10.1016/j.enzmictec.2012.07.013, PMID:22975128.
[12] Panahi Y, Darvishi B, Jowzi N, Beiraghdar F, Sahebkar A. Chlorella vulgaris: A Multifunctional Dietary Supplement with Diverse Medicinal Properties. Curr Pharm Des 2016;22(2):164–173. doi:10.2174/1381612822666151112145226, PMID:26561078.
[13] Wakasugi T, Nagai T, Kapoor M, Sugita M, Ito M, Ito S, *et al*. Complete nucleotide sequence of the chloroplast genome from the green alga Chlorella vulgaris: the existence of genes possibly involved in chloroplast division. Proc Natl Acad Sci USA 1997;94(11):5967–5972. doi:10.1073/pnas.94.11.5967, PMID:9159184.
[14] Senturk T, Batir MB, Camli C, Yildiz S. Evaluating the essential and non-essential metal remediation efficacy of Chlorella vulgaris, and photosynthetic gene expression level changes during the process. Aqua Res 2019;2(3):134–142. doi:10.3153/AR19011.
[15] Minkiewicz P, Iwaniak A, Darewicz M. BIOPEP-UWM Database of Bioactive Peptides: Current Opportunities. Int J Mol Sci 2019;20(23):5978. doi:10.3390/ijms20235978, PMID:31783634.
[16] Mooney C, Haslam NJ, Pollastri G, Shields DC. Towards the improved discovery and design of functional peptides: common features of diverse classes permit generalized prediction of bioactivity. PLoS One 2012;7(10):e45012. doi:10.1371/journal.pone.0045012,

PMID:23056189.
[17] Pirtskhalava M, Amstrong AA, Grigolava M, Chubinidze M, Alimbarashvili E, Vishnepolsky B, *et al*. DBAASP v3: Database of antimicrobial/cytotoxic activity and structure of peptides as a resource for development of new therapeutics. Nucleic Acids Res 2021;49(D1):D288–D297. doi:10.1093/nar/gkaa991, PMID:33151284.
[18] Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Open Source Drug Discovery Consortium, *et al*. In silico approach for predicting toxicity of peptides and proteins. PLoS One 2013;8(9):e73957. doi:10.1371/journal.pone.0073957, PMID:24058508.
[19] Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v.2—a server for in silico prediction of allergens. J Mol Model 2014;20(6):2278. doi:10.1007/s00894-014-2278-5, PMID:24878803.
[20] Masojídek J, Torzillo G. Mass Cultivation of Freshwater Microalgae. Reference Module in Earth Systems and Environmental Sciences 2014;AmsterdamElsevierdoi:10.1016/B978-0-12-409548-9.09373-8.
[21] Ramaraj R, Unpaprom Y, Dussadee N. Cultivation of Green Microalga , Chlorella Vulgaris for Biogas Purification. Int J New Technol Res 2016;2(3):117–122.
[22] UniProt. KB-P56342 (PSAB_CHLVU). Online: UniProt. 2023.
[23] Kandemir-Cavas C, Pérez-Sanchez H, Mert-Ozupek N, Cavas L. In Silico Analysis of Bioactive Peptides in Invasive Sea Grass Halophila stipulacea. Cells 2019;8(6):557. doi:10.3390/cells8060557, PMID:31181665.
[24] Marr AK, Gooderham WJ, Hancock RE. Antibacterial peptides for therapeutic use: obstacles and realistic outlook. Curr Opin Pharmacol 2006;6(5):468–72. doi:10.1016/j.coph.2006.04.006, PMID:16890021.
[25] Knott L. ACE Inhibitors. Leeds: Patient. 2022.
[26] Rossi GP, Sanga V, Barton M. Potential harmful effects of discontinuing ACE-inhibitors and ARBs in COVID-19 patients. Elife 2020;9:e57278. doi:10.7554/eLife.57278, PMID:32250244.
[27] Kęska P, Stadnik J, Bąk O, Borowski P. Meat Proteins as Dipeptidyl Peptidase IV Inhibitors and Glucose Uptake Stimulating Peptides for the Management of a Type 2 Diabetes Mellitus In Silico Study. Nutrients 2019;11(10):2537. doi:10.3390/nu11102537, PMID:31640215.
[28] Lee DS, Lee ES, Alam MM, Jang JH, Lee HS, Oh H, *et al*. Soluble DPP-4 up-regulates toll-like receptors and augments inflammatory reactions, which are ameliorated by vildagliptin or mannose-6-phosphate. Metabolism 2016;65(2):89–101. doi:10.1016/j.metabol.2015.10.002, PMID:26773932.
[29] Pal GK, Suresh PV. Comparative assessment of physico-chemical characteristics and fibril formation capacity of thermostable carp scales collagen. Mater Sci Eng C Mater Biol Appl 2017;70(Pt 1):32–40. doi:10.1016/j.msec.2016.08.047, PMID:27770898.
[30] Fu Y, Young JF, Løkke MM, Lametsch R, Aluko RE, Therkildsen M. Revalorisation of Bovine Collagen as a Potential Precursor of Angiotensin I-Converting Enzyme (ACE) Inhibitory Peptides Based on in Silico and in Vitro Protein Digestions. J Functional Foods 2016;24:196–206.
[31] Schaafsma G. Safety of protein hydrolysates, fractions thereof and bioactive peptides in human nutrition. Eur J Clin Nutr 2009;63(10):1161–1168. doi:10.1038/ejcn.2009.56, PMID:19623200.
[32] Lee SY, Song EJ, Kim KE, Ahn DH. Others Effect of Heat and Autoclave on Allergenicity of Porcine Serum Albumin. Food Sci Biotechnol 2011;20:455–459. doi:10.1007/s10068-011-0063-0.
[33] Lafarga T, Wilm M, Wynne K, Hayes M. Bioactive Hydrolysates from Bovine Blood Globulins: Generation, Characterisation, and in Silico Prediction of Toxicity and Allergenicity. J Functional Foods 2016;24:142–155. doi:10.1016/j.jff.2016.03.031.
[34] Pal GK, Suresh PV. Physico-chemical characteristics and fibril-forming capacity of carp swim bladder collagens and exploration of their potential bioactive peptides by in silico approaches. Int J Biol Macromol 2017;101:304–313. doi:10.1016/j.ijbiomac.2017.03.061, PMID:28315441.
[35] Sabadin IS, Villas-Boas MB, De Lima Zollner R, Netto FM. Effect of Combined Treatment of Hydrolysis and Polymerization with Transglutaminase on β-lactoglobulin Antigenicity. Eur Food Res Technol 2012;235:801–809. doi:10.1007/s00217-012-1802-z.
[36] Wang X, Yu H, Xing R, Li P. Characterization, Preparation, and Purification of Marine Bioactive Peptides. Biomed Res Int 2017;2017:9746720. doi:10.1155/2017/9746720, PMID:28761878.
[37] Pooja K, Rani S, Prakash B. In silico approaches towards the exploration of rice bran proteins-derived angiotensin-I-converting enzyme inhibitory peptides. Int J Food Properties 2017;20(2):2178–2191.
[38] Kakran M, Li L. Muller RH. Overcoming the challenge of poor drug solubility. Pharm Eng 2012;32(4):1–7.