Food Research 7 (2) : 248 - 261 (April 2023)
Journal homepage: https://www.myfoodresearch.com

FOOD RESEARCH

FULL PAPER

Anchovy's protein as a potential precursor of Angiotensin-I Converting Enzyme (ACE) inhibitory peptide and Dipeptidyl Peptidase-IV (DPP-IV) inhibitory peptide by an *in silico* approach

Kari, N.M., Azmi, S.H., Tuan Zainazor, T.C., Zamri, A.I. and *Ahmad, F.

Faculty of Fisheries and Food Science Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

Article history:
Received: 3 October 2021
Received in revised form: 21 November 2021
Accepted: 2 March 2022
Available Online: 18 April 2023

Keywords:
Anchovy,
Bioactive Peptides,
in silico,
BIOPEP,
Angiotensin-I Converting Enzyme (ACE),
Dipeptidyl Peptidase-IV (DPP-IV)

DOI:
https://doi.org/10.26656/fr.2017.7(2).792

Abstract

Protein from fish is known as the precursor of biologically active peptides that exert various health benefits such as antihypertensive, antitumor, and immunomodulatory properties. Hence, this study aimed to perform an extraction of anchovy, LC-MS/MS analysis and *in silico* evaluation of the major proteins in anchovies as potential precursors of biologically active peptides in addition to determining whether such peptides can be released by selected proteolytic enzymes. Anchovy was subjected to protein extraction followed by total soluble protein concentration determination using Bradford assay. The sample was subjected to in-solution trypsin digestion, which was then analysed by liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS). A bioinformatic approach by PEAKS Studio was used to identify the protein. A total of four anchovy's proteins which are myosin light chain 1, V(D)J recombination-activating protein, 60 kDa chaperonin and heat shock protein 90AA1 were identified. Then, the identified proteins were subjected to *in silico* approach using the BIOPEP database. The biological potential of the theoretically released angiotensin-I converting enzyme (ACE) inhibitory peptides and dipeptidyl peptidase (DPP)-IV inhibitory peptides were predicted by determining the frequency of occurrence of fragments with a given activity. 60 kDa chaperonin and heat shock protein 90AA1 predicted the highest number of biological activities for ACE inhibitory peptides (284 and 264 fragments) and DPP-IV inhibitory peptides (395 and 409 fragments). The most promising enzyme for the generation of bioactive peptides from anchovy protein was anticipated to be pepsin (pH > 2), which theoretically released a high number of DPP-IV inhibitory peptides and ACE inhibitory peptides through the action of *in silico* proteolysis. Overall, this work highlighted that anchovy protein could be a promising precursor of bioactive peptides that have ACE and DPP-IV inhibitory activities for developing functional food or nutraceutical products.

1. Introduction

Anchovies are the pelagic species that are widely found throughout the world's oceans. It is a small fish belonging to the Engraulidae family, which includes 140 species in 16 genera (Thienchai and Chaiyanan, 2012). *Encrasicholina* and *Stolephorus* are two genera that are habitually found in Malaysian coastal waters (DOF, 2019). The main commercial species of anchovy in Malaysia are *Encrasicholina heteroloba, Encrasicholina punctifer, Stolephorus commersonii* and *Stolephorus andhraensis* (Froese and Pauly, 2018). The saltwater fish can be identified by its silvery blue-green back appearance and can grow up to 20 cm (Shiriskar *et al.*, 2010). In the oceans and main seas such as the Atlantic, Indian and Pacific Oceans, anchovy populations can be found in most temperate and productive coastal areas (Thienchai and Chaiyanan, 2012; Checkley *et al.,* 2017).

Approximately 1.7 million tonnes of fishery production in 2017 were contributed by Malaysia's fisheries sector whereas pelagic marine fish contributed to 556, 342 tonnes of Malaysian fishery capture in 2019. Anchovies account for 6% of RM 2.4 billion in the income generated by the Malaysian fisheries sub-sector (FAO, 2019). Fish is one of the main sources of protein diet. In Southeast Asia, anchovies are diversely added in dishes, whether in fresh form or dried form. In addition, anchovies are the main ingredient in fish sauce, which is a food product that is commonly eaten in the Southeast Asia region (Tanasupawat and Visessanguan, 2014).

*Corresponding author.
Email: *fisal@umt.edu.my*

eISSN: 2550-2166
Published by Rynnye Lyan Resources

Anchovies have been traditionally regarded as food items in Peninsular Malaysia, and they are still consumed in dry, fermented, or raw form (Ahmad *et al.*, 2018).

According to Palani Kumar *et al.* (2014), the protein composition of live-weight whole fish is approximately 20 to 30% and may differ according to feeding conditions, sex and stage of maturity. Generally, the content of protein in fish muscle is between 12.2 to 21.79% (Pyz-łukasik, 2020). Fish muscle is composed of myofibrillar protein, sarcoplasmic protein, connective tissue, stroma protein, polypeptide, nucleotide and non-protein nitrogen compounds (Jim *et al.,* 2017). Myofibrillar proteins are mainly composed of actin and myosin, which are responsible for muscle contraction (Le Gouic *et al.*, 2018). Food protein is known as the precursor of biologically active peptides, apart from many other functions that can be obtained from various sources such as dairy products, marine and land animals, plants and cyanobacteria (Minkiewicz *et al.,* 2011).

A bioactive peptide is defined as an active protein fragment that remains inactive as long as it were intact with parent proteins (Hernández-Ledesma *et al.,* 2014). Usually, the process of enzymatic hydrolysis of the parent protein will release the bioactive peptide, in addition to other processes such as cooking, ripening and fermentation (Daliri *et al.,* 2017). Proteases with a broad specificity of action are used to produce bioactive peptides with specific activities. These enzymes are commonly extracted from vegetable, animal and microbial tissues such as bromelain, papain, pepsin, chymotrypsin, proteinase K and subtilisin (Jakubczyk *et al.,* 2020). Bioactive peptides have shown diverse functions in physiological effects based on their actions such as antimicrobial, antithrombotic, antihypertensive, opioid, immunomodulatory, mineral binding, and antioxidative (Sánchez and Vázquez, 2017).

The identification of bioactive peptides by conventional method is tedious, and costly, and leads to a lower yield of isolated bioactive peptides and loss of the potential bioactivities. Thus, computational simulation tools or *in silico* methods that can predict the release of a bioactive peptide can be used as an alternative to the traditional method (Pooja *et al.,* 2017). Current available online tools are Peptide Cutter (*in silico* peptide digestion tools), PeptideLocator and PeptideRanker (bioactive peptide prediction tools), mMass and BIOPEP (an online database that combines *in silico* digestion and bioactive peptide prediction) (Anekthanakul *et al.,* 2018). In this paper, the BIOPEP database has been selected as an online tool to predict the biologically active peptides in anchovies. BIOPEP is a database program that predicts the profiles of the potential biological activity of protein fragments as well as the prediction of bonds that are susceptible to hydrolysis by an enzyme in the protein chain (Minkiewicz *et al.,* 2008).

A study by Darewicz *et al.* (2014) has identified peptides with angiotensin I-converting enzyme inhibitory (ACE) activity in products of salmon protein hydrolysis/ digestion using the BIOPEP database and mass spectrometry (MS) approach. Panjaitan *et al.* (2018) conducted a study where major bioactive peptides predicted from dried giant grouper roe using proteomic analysis and in *silico* approach are angiotensin-I converting enzyme (ACE-I) and dipeptidyl peptidase-IV (DPP-IV). To the best of our knowledge, little information is available regarding the sequences of bioactive peptides released from anchovy proteins and the selection of optimal proteinases for producing different bioactive peptides. Hence, the study aimed to identify the major proteins contained in selected anchovy using a proteomics approach using LC/MS-MS, perform an *in silico* evaluation of the major proteins in anchovies as potential precursors of biologically active peptides (BIOPEP) and determine whether such peptides can be released by selected proteolytic enzymes. Such data can be useful in further research for deciphering the profile of protein sequences and could be helpful in the production of specific bioactive peptides with required biological functions.

2. Materials and methods

2.1 Materials

Anchovies from the species *Encrasicholina devisi* were selected for this study. The samples were purchased at Pantai Tok Bali, Terengganu and transported to the laboratory. The anchovies were stored at -80^{o}C until use. All chemicals used in this research were of analytical grade and electrophoresis grade. For further analysis, the anchovies were thawed and minced into smaller pieces beforehand.

2.2 Protein extraction

Protein from anchovy was extracted by using the precipitation approach, which is TCA/Acetone precipitation extraction method as described by Isaacson *et al.* (2006) with a slight modification. Approximately 2 g of anchovy was suspended in a buffer consisting of 50 mM Tris-HCl (pH 8.5), 100 mM KCl, 5 mM EDTA and 2% (w/v) Mercaptoethanol. The solution was then homogenized for 1 min. This was followed by centrifuging the sample at 11, 000 rpm for 15 mins at 4^{o}C. The supernatant was collected and precipitated with TCA/acetone overnight at -20^{o}C followed by

https://doi.org/10.26656/fr.2017.7(2).792

centrifugation for 15 mins at 4°C. Acetone was discarded and the pellet was re-suspended in TCA/acetone for 1 hr at -20°C and centrifuged to collect the final pellet. The pellet was then air-dried at room temperature for 5 mins and solubilized with solubilization buffer for further analysis.

2.3 Determination of total soluble protein concentration by Bradford assay

Bradford assay was selected as a method to determine the concentration of protein in the solution of extracted protein from anchovy as described in Bradford (1976) and modified by Shukla (2015). The sample that contains a blank, a protein standard and the unknown concentration of the solution to be assayed was mixed with the Bradford reagent. The procedure was performed at room temperature where the absorbance was measured at 595 nm. Then, a standard curve of absorbance vs. protein concentration of each standard was plotted. The concentration of the unknown protein sample was determined by comparing the Net A_{595} values against the standard curve.

2.4 In-solution trypsin digestion

In-solution trypsin digestion was done according to Kwan and Ismail (2018) where the protein samples were evaporated and re-suspend in 6 M urea, 100 mM Tris buffer at 10 mg/mL. 200 mM dithiothreitol (DTT) was added to each fraction and incubated at room temperature for 1 hr. After that, 200 mM of iodoacetamide was added and kept at room temperature for 1 h, followed by 20 µL of 200 mM DTT again. Next, 775 µL of water was added to dilute the samples. Digestion was carried out by introducing 20 µg of bovine trypsin to each sample and incubating overnight at 37°C. The digestion was stopped the next day by adjusting the pH of the buffer to pH <6 using concentrated acetic acid. The digested samples were then concentrated to less than 20 µL each.

2.5 LC-MS/MS analysis

Liquid chromatography mass spectrometry/mass spectrometry (LC-MS/MS) analysis was performed according to Kwan *et al.* (2016). For this analysis, LTQ-Orbitrap Velos Pro mass spectrometer coupled with Easy -nLC II nano liquid chromatography system was used. An aliquot of 100 µL of 0.1% formic acid in deionized water was added to each of the peptides and then filtered using the 0.45um regenerated cellulose (RC) membrane syringe filter (Sartorius AG, Goettingen, Germany). Easy column C18 was used (100 mm, 0.75 mm i. d 3 µm; Thermoscientific USA) as the analytical column while Easy Column C18 (20 mm, 0.1 mm, 5 µm; Thermo Scientific USA) was used as pre-column at flow rate of 3 µL/mm for 15 µL. The analytical column was equilibrated at flow rate of 0.3 µL/min for 4 µL. Then, 3 µL of prepared vial samples were injected at a flow rate of 0.3 µL/min. Running buffer of 0.1 % formic acid in deionized water and 0.1% of formic acid in acetonitrile were used. Sample eluent was sprayed into a mass spectrometer at 220°C capillary temperature and a 2.1 kV source. A full scan mass analysis from mz 300-2000 was used to detect protein and peptide at a resolving power of 60,000 with a data dependent MS/MS analysis (ITMS) triggered by the eight most abundant ions from a parent mass list of predicted peptide, with rejection or unassigned charge states. The fragmentation technique used was collision induced dissociation (CID) with a collision energy of 35. Each sample was done in replicates.

2.6 LC-MS/MS data analysis using PEAKS Studio

The LC-MS/MS raw data were further proceeded to an analysis by using PEAKS Studio software (Bioinfor Inc., CA, USA) to produce a predicted peptide sequence according to Kwan *et al.* (2016). The de novo sequencing and database matching analysis was done by PEAKS Studio Version 7.5 (Bioinformatics Solution, Waterloo, Canada). The database matching used was the UniProt database (UniProt Consortium, 2019). The carbamidomethylation and methionine oxidation were set as fixed modifications and the maximum missed cleavage was set at 2. 0.1 Da is the value set for parent mass and precursor mass tolerance. For protein acceptance, a false detection rate (FDR) <0.1% and a significant score (-10logP) for protein >20 were used. Minimum unique peptide and maximum variable post-translational modification were set at 1 and 4 subsequently. The fixed PTMs in the parameter chosen were phosphorylation, methylation, oxidation, hydroxylation, biotinylation, acetylation, sulfation, amidation, myristoylation, carboxylation, ubiquitination and farnesylation while the average local confidence (ALC) was set at >15%.

2.7 In silico assessment of anchovy proteins using BIOPEP database

2.7.1 Potential biological activity profile

The protein sequences of identified proteins and their accession ID were obtained from UniProt database (http://www.uniprot.org/) in the FASTA format. BIOPEP database (http://www.uwm.edu.pl/biochemia/index.php/en/biopep) was used to predict the profiles of potential biological activities of selected anchovy proteins using the "profiles of potential biological activity" option in the database (Minkiewicz *et al.*, 2008). Data from BIOPEP database were imported into Microsoft Excel 2016, where the predicted bioactive

https://doi.org/10.26656/fr.2017.7(2).792

peptides and activities were sorted, analyzed and tabulated into tables and figures.

Through this option, BIOPEP ID, name of peptides, the potential activity of the peptide, number of peptides and location of bioactive peptide in protein sequences were acquired. Meanwhile, the frequency of occurrence of fragments with given activity (A) in the selected protein was taken as the evaluation parameter and calculated based on the equation:

$$A = a/ N$$

Where a = number of bioactive peptides and N = total number of amino acid (AA) residues in the protein chain. In addition, the total frequency of bioactive fragments (∑A), in each four sequences was also calculated.

2.7.2 In silico proteolysis

The sequences of anchovy proteins were subjected to *in silico* proteolysis by using BIOPEP's 'enzyme/s action' tool. A total of 33 enzymes were selected during *in silico* proteolysis which are chymotrypsin A, trypsin, pepsin (pH 1.3), proteinase K, pancreatic elastase, oligopeptidase V-8, protease (pH 4), thermolysin, chymotrypsin C, plasmin, cathepsin, clostripain, chymase, papain, ficin, leukocyte elastase, metridin, thrombin, pancreatic elastase II, stem bromelain, glutamyl endopeptidases II, oligopeptidase B, calpain 2, glycyl endopeptidase, oligopeptidase F, proteinase p1, xaa pro dipeptidase, pepsin (pH>2), coccolysin, subtilisin, chymosin, ginger protease and V-8 protease (pH 7.8). Predicted degree of hydrolysis (DH %) for each enzyme used was measured. The efficiency of the released bioactive fragments was measured by the frequency of release of the peptide with given activity by selected enzymes (A_E) and also the relative frequency for the release of peptides with given activity by a selected enzyme (W).

$$A_E = \frac{d}{N} \quad W = \frac{A_E}{A}$$

Where d is the number of fragments with given activity in the protein sequence, which may be released by enzymes and N is the number of amino acid residues in the protein chains.

3. Result and discussion

3.1 Total soluble protein concentration and protein yield of selected anchovy sample.

Table 1 shows the total soluble protein concentration and yield of the anchovy sample (*Encrasicholina devisi*). The total soluble protein concentration obtained from Bradford assay determination of the anchovy was 18.44 ±0.73 mg/mL in content, whereas the protein yield from the extraction was reported as 20.28 mg/g. The amount of total soluble protein of the anchovy protein was relatively low compared to the previous study reported Ceruso *et al.* (2015) where the total soluble protein concentration of *M. galloprovincialis* obtained was 23.33 mg/ml. This may be due to poor compatibility between the extraction method and the sample of anchovy. Thus, future research can focus on improving the extraction technique to increase the yields of protein extraction.

Table 1. Total soluble protein concentration and protein yield of anchovy sample extracted using TCA/Acetone precipitation method

Sample	Protein concentration (mg/mL)	Protein yield (mg/g)
Encrasicholina devisi	18.44±0.73	20.28

3.2 Protein identification through LC-MS/MS analysis and data analysis using PEAKS Studio

Protein identification using LC-MS/MS and PEAKS Studio managed to make identification of 97 proteins in total, which is based on the registered sequence in the UniprotKB library database. Table 2 shows a list of 20 proteins out of 97, which were listed based on the confidence of the results according to the -10lgP from the LC-MS/MS and coverage (%) to match proteins in the database. There were several major proteins identified, which were myosin light chain that has the highest coverage (8-21%), followed by V(D)J recombination-activating protein (6-10%), HSP90AA1 (3%), 60 kDA chaperonin (2%) and rhodopsin (1%).

Identification of myosin light chain as the highest confidence was found to comply with previous studies, which had shown myofibrillar proteins make up to 77% of total protein in anchovy (*Stolephorus* sp.), consisting of actin and myosin (Dewi, 2002). Choi *et al*. (2004) also proved the presence of a myosin heavy chain, along with actin and tropomyosin in the myofibrillar protein of anchovy (*Engraulis japonicus*) using SDS-PAGE. The proteins were also identified from various species including *Engraulis* sp., *Amazonsprattus* sp., *Thryssa* sp., *Coila* sp., *Anchoa* sp. and *Anchoviella* sp.

However, among the identified proteins, only four major proteins were chosen for *in silico* evaluation (Table 3). The selected protein was myosin light chain 1 (accession ID: Q9IB21), V(D)J recombination activating protein 1 (accession ID: A0A2R4GAA0), 60 kDa chaperonin (accession ID: A0A2H4PU29) and heat shock protein 90AA1 (A0A3G1CVR4). These proteins were selected based on their composition as components of muscle and structural protein from the anchovy body. For example, myosin is known as the myofibrillar protein that constitutes 65-75% of total muscle protein in

https://doi.org/10.26656/fr.2017.7(2).792

Table 2. Protein identification by LC-MS/MS

Protein ID	Protein Name	-10lgP	Coverage (%)	Species
Q90W53	Myosin light chain 3 OS	186.74	21	*Engraulis japonicus*
Q9IB20	Myosin light chain 2 OS	118.36	20	*Engraulis japonicus*
Q9IB21	Myosin light chain 1 OS	114.93	8	*Engraulis japonicus*
H6W0N4	V(D)J recombination-activating protein 2 (Fragment) OS	45.51	7	*Amazonsprattus scintilla OX*
A0A2R4G9P6	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	10	*Thryssa Kammalensis OX*
A0A2R4G9V6	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	10	*Coilia dussumieri OX*
H6W0A9	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	7	*Anchoa filifera OX*
H6W0B4	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	7	*Engraulis mordax OX*
H6W0C3	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	7	*Anchoa Parva OX*
H6W0D5	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	7	*Anchoa Walkeri OX*
H6W0A7	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	7	*Engraulis encrasicolus OX*
H6W0C4	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	7	*Anchoa delicatissima OX*
H6W0C8	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	7	*Anchovia macrolepidota OX*
H6W0D3	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	7	*Anchoviella balboae OX*
A0A2R4G9R3	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	6	*Engraulis australis OX*
A0A2R4G9S0	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	6	*Anchoa hepsetus OX*
A0A2R4GAA0	V(D)J recombination-activating protein 1 (Fragment) OS	45.17	6	*Thryssa spinidens OX*
A0A2H4PU29	60 kDa chaperonin OS	43.01	2	*Coila nasus OX*
A0A3G1CVR4	HSP90AA1 OS	24.69	3	*Coila nasus OX*
A9QUF3	Rhodopsin (Fragment) OS	5	1	*Engraulis encrasicolus OX*

Table 3. Selected proteins identified from LC-MS/MS analysis and PEAKS Studio database.

Protein ID	Coverage (%)	Average Mass	Protein description
68512	8	21592	Myosin light chain 1 OS = Engraulis japonicas; OX = 42892 GN = mlc1; PE = 2 SV = 1
32563	6	57376	-- V(D)J recombination-activating protein 1 (Fragment) OS = Thryssa spinidens; OX = 2137661 PE = 3 SV = 1
32287	5	61138	60 kDa chaperonin OS = Coilia nasus OX = 365059 PE = 2 SV = 1
68379	3	85550	HSP90AA1 OS = Coilia nasus OX = 365059 GN = HSP90AA1 PE = 2 SV = 1

OS = Organism name, OX = Organism identifier, GN = Gene name, PE = Protein existence, SV = Sequence version.

fish, along with other minor proteins (Medina and Pazos, 2010).

In addition, the selection of the proteins for *in silico* evaluation was based on their molecular mass where higher masses were preferred. This preference was implemented since proteins with higher molecular mass, could probably have a higher probability of producing bioactive peptides after enzyme treatment rather than lower molecular weight proteins. Table 3 shows protein ID, protein group, percentage of coverage, average mass and description of the protein selected. A similar study has been done by Huang *et al.* (2015) where a total of 7 tilapia proteins were identified by proteomic techniques in which LC-MS/MS analysis revealed structural proteins from tilapia co-products using proteomic approaches.

3.3 In silico assessment of anchovy proteins using BIOPEP database

3.3.1 Profile of the selected protein

Table 4 lists the identified proteins with their

https://doi.org/10.26656/fr.2017.7(2).792

Table 4. Name, accession number, sequences, amino acid residue and molecular mass of anchovy protein used in *in silico* analysis

Name of protein	Accession number (UniProt)	Sequences	Amino acid residues	Molecular mass (kDa)
Myosin light chain 1	Q9IB21	MAPKKDAKPAAKAAPAAKKEEPKPAEPAPEPPKPAAP TVDLSAVKLDFTPDQIEDFREAFSLFDRLGDNKVAYNQ IADIMRALGQNPTNKEVKLILSDPSPEDMEKKRIEFEQF LPMLQTVINNPNKAGFEDYVEGLRVFDKEGNGTVMG AELRIVLSTLGEKMNEKEIDALMAGGQEDENGCVNYET	195	21.59
V(D)J recombination-activating protein 1	A0A2R4GAA0	SGRQIFQPLHTLRTAEKELLPGFHAFEWQPALKAVSSS WDVGILDGLSGWTSSVEDVPADTIARRFRYDVALVSA LKDLEEDILEGLREHSLDDSSVTGFSVTIKESCDGMGD VSEKHGSGPVVPEKAVRFSFTIMSVSAHLEGEEASVNIF QEIKPNSELSCKPLCLMFVDESDHETLTAILGPVVAERN AMKESRLILSLAGLPRSFRFHFRGTGYDEKMVREMEGL EASGSTYICTLCDSTRAEASHNMVLHSITRSHQENLER YELWRTNPFSESADELRDRVKGVSAKPFMETQPTLDA LHCDIGNATEFYKIFQDEIGEMHARKHVPSREERRQWR AALDKQLRKKMKLKAVMRMNGNYARRLMTAEAVEV VCELVPSEERRQALRELMALYLQMKPVWRSTWPARE CPDQLCRYSFNSQRFAELLSSTFKYRYDGKITNYLHKT LAHVPEIVERDGSIGAWASEGNESGNKLFRRFRKMNA RQSKSFELEDVLKHH	503	57.34
60 kDa chaperonin	A0A2H4PU29	MFRLPTLMRQVRPVCRALAPHLTRAYAKDVKFGADA RALMLQGVDLLADAVAVTMGPKGRTVIIEQSWGSPKV TKDGVTVAKSIDLKDKYKNIGAKLVQDVANNTNEEAG DGTTTATVLARAIAKEGFDTISKGANPFEIRRGVMLAV ETVISELKKLSKPVTTPEEIAQVATISANGDSEVGTIISN AMKKVGRKGVITVKDGKTLHDELEIIEGLKFDRGYISP YFINTAKGQKCEFQDAYLLLSEKKISSVQSIVPALEIAN QHRKPLVIIAEDVDGEALSTLVLNRLKVGLQVVAVKA PGFGDNRKNMLKDMAVASGGAVFGDEAIALALEDIQP HDFGKVGEVQVTKDDTLILRGRGDPAAIEKRISEIAEQL DNTTSDYEKEKLNERLAKLSDGVAVLKVGGTSDVEVN EKKDRVTDALNATRAAVEEGIVPGGGCALLRCIPALDT IVPANADQKIGVDIIRRALRVPAMTIAKNAGVEGSLVV EKILQLQGVMGYDAMQGEYVNMVEKGIIDPTKVVRT ALLDAAGVASLLSTAEAVVTEIPKEEKEGGMPGGMGG MGGMGGMGGGMF	576	61.14
Heat shock protein 90AA1	A0A3G1CVR4	MPEPQDQNMDEEAETFAFQAEIAQLMSLIINTFYSNKEI FLRELISNSSDALDKIRYESLTDPTKLDDGKDLKIDIIPN KAERTLTLIDTGIGMTKADLINNLGTIAKSGTKAFMEA LQAGADISMIGQFGVGFYSAYLVAERVTVITKHNDDE QYAWESSAGGSFTVKVDTSEPMGRGTKVILHMKEDQI EYLEDRRVKEIVKKHSQFIGYPITLHVEKEREKEVSDDE AEEEEDKEKEKEKDKEKDGEEEEEKDKDKPEIEDVGS DEDEDHDHDHDKCGDKKKKKKKIKEKYIDQEELNKT KPLWTRNPDDITNEEYGEFYKSLTNDWEDHLAIKHFSV EGQLEFRALLFVPRRAPFDLFENKKKKNNIKLYVRRVF IMDNCDELIPEYLNFIRGVVDSEDLPLNISREMLQQSKIL KVIRKNLVKKCLELFTELSEDKDNCKKLYEQFSKNIKL GIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYVARM KDTQKHIYYITGESKDQVANSAFVERLRKAGLEVIYMI EPIDEYCVQQLKEFEGKNLVSVTKEGLELPEDEDEKKK QEEKKSQFENLCKIMKDILEKKVEKVTVSNRLVSSPCCI VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHL EINPDHPIVQTLRQKAEADKNDKSVKDLVILLFETALLS SGFSLDDPQTHSNRIYRMIKLGLGIDEDDLSTDDTSAAP VEDMPPLEGDDDTSRMEEVD	742	85.55

https://doi.org/10.26656/fr.2017.7(2).792

characteristics (Uniprot accession number, sequences, number of amino acids residue and molecular mass). After protein identification using LC-MS/MS analysis and data analysis using PEAKS Studio, four proteins were selected which were myosin light chain 1 (accession ID: Q9IB21), V(D)J recombination activating protein 1 (accession ID: A0A2R4GAA0), 60 kDa chaperonin (accession ID: A0A2H4PU29) and heat shock protein 90AA1 (A0A3G1CVR4). The sequences of amino acids of the proteins were obtained from Uniprot database. As shown in Table 4, the number of amino acids residue in the proteins ranged from 195 to 742 and molecular weights were between 21.59 and 85.55 kDa. Heat shock protein90AA1 contains the highest number of amino acids (742) and a molecular weight of 85.55 kDa.

Myosin light chain 1 (MLC1), also known as myosin essential light chain, is a vital structural component of the actomyosin cross-bridge that also helps with muscle contraction and force development (Hernandez *et al.,* 2021). There are two main types of essential light chains which are short and long isoforms. The N-terminus of the later isoform contain an extended sequence that comprises the repeats of Pro and Ala residues, as well as positively charged amino acids (Nieznańska *et al.,* 2002). A study by Darewicz *et al.* (2016) had shown that ACE inhibitory and antioxidant peptides were the most predominant fragments of proteins in other subunits of myosin, which is a myosin heavy chain, after *in silico* evaluation.

The RAG complex is a multi-protein complex that mediates the DNA cleavage process during V(D)J recombination and comprises a catalytic component. V (D)J recombination deploys a wide repertoire of immunoglobulin and T-cell receptor genes of V and T-lymphocytes development via rearrangement of various V (variable), D (diversity), and J (joining) gene segments (UniProt Consortium, 2019). The adaptive immune response highly depends on V(D)J recombination actions. The human immune system is weakened when it is not present. It causes chromosomal translocations and B- and T-cell malignancies when it is not appropriately regulated (Market and Papavasiliou, 2003).

Moreover, 60 kDa chaperonin is known as a ubiquitous family of sequence-related molecular chaperones that consist of oligomeric proteins with a subunit mass of about 60 kDa that are required for protein folding in both normal and stressed situations (Levy-Rimler *et al.*, 2001). 60 kDa heat shock protein is the other name for 60 kDa chaperonin which it is usually found in eubacteria, mitochondria and chloroplast. It has the ability to act as intercellular signals in various biological effects such as immunity and inflammation (Maguire *et al*., 2002; UniProt Consortium, 2019).

Heat shock proteins (Hsps), sometimes referred to as stress proteins or extrinsic chaperones, are a group of highly preserved proteins found in all living creatures (Shi *et al*., 2016). Meanwhile, heat shock protein 90 (HSP90) is an important part of the stress response's protective mechanism. Signal transmission, cell cycle control, genomic silencing, and protein trafficking are all aided by HSP90s (Park and Kwak, 2014). Protein folding, damaged protein repair, nascent protein transport, and immune presentation are a few of the biological roles of heat shock proteins (Shi *et al*., 2016). Heat shock protein 90 has been reported in the liver, ovary, muscle, brain, and stomach tissue of *Penaeus monodon*, which is a species of black tiger shrimp (Jiang *et al.,* 2009).

3.3.2 Profile of the potential biological activity and A parameter of selected anchovy proteins

The profile of the potential biological activity for each selected protein was measured (Table 5). It is a parameter assessment for the quantitative prediction of protein sequences as a precursor to bioactive peptides (Iwaniak *et al.,* 2005). The number of bioactive peptides predicted to be released from proteins were generated by BIOPEP database. As of 27 June 2021, 4325 peptides functioned in 56 bioactivities have been collected in the database. The biological activities present in Table 5 include ACE inhibitor, DPP-IV inhibitor, DPP-III inhibitor, antioxidative, neuropeptide, regulating, antithrombotic, bacterial permease ligand, inhibitor, renin inhibitor, stimulating, activating ubiquitin mediated proteolysis, immuno-stimulating, antiamnestic, anti-inflammatory, hypolipidemic, alpha-glucosidase Inhibitor, CaMPDE inhibitor, HMG-CoA reductase inhibitor, embryotoxic, immunomodulating and celiac toxic.

A total of 22 predicted biological activities were found in the selected anchovy proteins. From the list, 60 kDa chaperonin predicted the highest biological activities which were 21 activities, whereas myosin light chain 1 predicted the lowest value with 14 activities. 60 kDa chaperonin and heat shock protein 90AA1 generated the highest number of predicted biological activities for ACE inhibitory peptides (284 and 264 fragments) and DPP-IV inhibitory peptides (395 and 409 fragments) respectively. Meanwhile myosin light chain 1 turned out to be the lowest source of ACE-inhibitory peptides and DPP-IV inhibitory peptides, which were 93 and 125 fragments respectively.

The frequency of occurrence (A) of potential

https://doi.org/10.26656/fr.2017.7(2).792

Table 5. Total number of potential bioactive peptides from anchovy proteins predicted using BIOPEP-*Profiles of potential biological activity* tool

Protein (ID)	Number of activities	Number of peptides				
		ACE inhibitor	DPP-IV inhibitor	DPP- III inhibitor	Anti-oxidative	Other activities*
Myosin light chain 1	14	93	125	15	10	28
V(D)J recombination-	16	202	288	49	42	71
60 kDa chaperonin	21	284	395	34	30	121
Heat shock protein	16	264	409	54	59	144
Total number of potential biological	22	-	-			-

*Other activities include Neuropeptide, Regulating, Antithrombotic, Bacterial Permease Ligand, Inhibitor, renin inhibitor, stimulating, Activating Ubiquitin Mediated Proteolysis, Immuno-stimulating, Antiamnestic, Anti-inflammatory, Hypolipidemic, Alpha Glucosidase Inhibitor, CaMPDE inhibitor, HMG-CoA reductase inhibitor, Embryotoxic, Immunomodulating, Celiac toxic.

bioactive peptides for the identified proteins was summarized in Table 6. According to Minkiewicz *et al.* (2011), the frequency of occurrence of bioactive peptide (A) indicates the potential for the bioactivity of the protein where the higher value of parameter A, the higher the probability of bioactive peptide to be released. In the case of anchovy proteins, DPP-IV inhibitory peptides (A value ranged from 0.5647 to 0.7017) were the highest compared to ACE inhibitory peptides (0.3639 to 0.4931), followed by DPP III-inhibitory peptides (0.0590 to 0.0944) and antioxidative (0.0513-0.0855).

However, DPP-IV and ACE inhibitory peptides were the two dominant frequencies of occurrence of bioactive peptides. This theoretical value suggested the protein anchovy protein contains the high potential for DPP-IV inhibitory and ACE inhibitory bioactive peptides activity. Among the selected protein, myosin light chain 1 and 60 kDa chaperonin generated the highest A value for DPP-IV inhibitory peptide and ACE inhibitory peptide as shown in Table 6. The result from Tables 5 and 6 shows that the profile of biological activities of protein and its frequency of occurrence of fragments with a given activity proved the potency of anchovy as a source of bioactive peptides.

3.3.3 In silico proteolysis of anchovy protein for the production of peptides

The enzyme action tools of BIOPEP database allow the prediction of bioactive peptides released by the action of different enzymes. Table 7 summarized the peptide fragment that was released during simulated proteolysis of anchovy protein. The predictions of theoretical sequences were tentatively cleaved by 33 proteinases, but only 5 proteinases showed the ability to generate relatively numerous ACE inhibitory peptides and DPP-IV inhibitory peptides. Simulated proteolysis of the four selected proteins using five proteases (ficin, calpain 2, pancreatic elastase, stem bromelain and pepsin (pH>2) have generated between 9 to 51 dipeptide and tripeptide bioactive fragments of ACE inhibitory peptide. In the meantime, for DPP-IV inhibitory peptide, the dipeptide and tripeptide that were produced were between 9 to 79 fragments. Among the five proteases, pepsin (pH>2) had initially predicted that the enzyme was able to release the highest amount of ACE inhibitory peptides and DPP-IV inhibitory compared to other proteases.

A previous study by Huang *et al.* (2015) has theoretically determined that protein isolated from tilapia skin and frame generated abundant ACE inhibitory peptides through BIOPEP database analysis. The finding was consistent with Lin *et al.* (2017) whereby tilapia by-products exhibited great ACE inhibitory activities hydrolyzed by pepsin *in vitro*. According to Shevchenko *et al.* (2007), trypsin enzyme was commonly used for the release of bioactive peptides in protein identification, while pepsin (pH>2) was able to hydrolyze the proteins to release a higher amount of bioactive peptides, compared to other enzymes (Panjaitan *et al.*, 2018).

Table 8 summarized the values of parameters

Table 6. The frequency of occurrence of bioactive fragment in anchovy protein chain (A)

Proteins	UniProt Accession no	∑A	Frequency of occurrence of bioactive fragment in protein chain (A)			
			ACE	DPP IV Inhibitor	DPP III Inhibitor	Antioxidative
Myosin light chain 1	Q9IB21	1.390	0.4769	0.6410	0.0769	0.0513
V(D)J recombination-activating protein 1	A0A2H4PU29	1.324	0.4155	0.5805	0.0944	0.0855
60 kDa chaperonin	A0A3G1CVR4	1.477	0.4931	0.6858	0.0590	0.0521
Heat shock protein 90AA1	A0A2R4GAA0	1.279	0.3639	0.5647	0.0728	0.0809

https://doi.org/10.26656/fr.2017.7(2).792

https://doi.org/10.26656/fr.2017.7(2).792

Table 7. Fragments of ACE-Inhibitory and DPP-IV inhibitory peptides predicted to be released by five proteases using BIOPEP enzyme action tool

Proteins	Enzyme used	ACE- inhibitory peptides	DPP IV-inhibitory peptides
Myosin Light Chain 1	Ficin	[9] VF, VAY, AG, EG, NG, EK, DF, IL, AEL	[9] EK, AL, AG, DR, EG, IL, NG, TL, VF
	Calpain 2	[11] AF, AP, AG (2), EG, NG, NK, EK, IL, AEL, ST	[13] FL, EK, AL(2), SL, AF, AG, AG(2), EG, IL, NG, SV
	Pancreatic elastase	[9] DA, MG, EG, NG, NY, KL (2), KA, PT	[12] MA(2), KA, PA, EG, ET, MG, NG, NY, PT, QT, RI
	Stem bromelain	[10] MG, EG, EA, NG, KL (2), KA, PT, DL, IL	[11] MA (2), KA, PA, DR, EG, IL, MG, NG, PT
	Pepsin (pH>2)	[15] RL, VF, VK (2), IA, RA, PPK, IE (2), VE, PT (2), IL, ST, VM	[28] VA, PA (5), IA, RA, SL, IL, IM (2), IN, PK (2), PM, PN, PT (2), RL, VD, VE, VF, VK (2), VL, VM, VN
V(D)J Recombination Activating Protein	Ficin	[29] MF, PR, PL, VK, AF, IF, IG, AG, MG, TG, EG (2), PG, DG (3), NY, NK, AR (2), AH (2), EK, IL (2), AEL, ER, TF	[36] EK, AL, PL, WR, AF, AG, AH (2), DR, EG (3), ES (3), IL (2), MF, MG, MK, NY, PF, PG, QL, QS, TF, TG, TL (2), TR, TY, VK, VS (3)
	Calpain 2	[41] PR, YL, PL (2), AW, VK, VP, (3), AG, MG, HG, SG (5), EG (2), PG, DG (2), NY, NK, AR (5), PT, HK, AV, IL (2), AEL, ST (4), ER	[48] VP, WP, AL (7), SL (2), PL (2), WR (2), WT, AW, AE (2), AG, AT, AV, DR, EG (2), EH, FR (3), HH, HT, IL (2), IM, MG, MK (2), MN, NE, NY, PG, PN, PT, SK, SV, VK, YL, YR
	Pancreatic elastase	[23] RL, RY, RA, HL, KG, DA, MG, EG (3), EA (3), PG, NKL, DG (2), NY (2), KA (2), EV, WA	[32] MA, KA (2), PA, RA, WA, HL, WRS, WRT, WT, EG (3), ES, EV, HS, HT, HV, KG, KI (2), KS, MG, NA, NY (2), PG, PS (2), PV (2), RL.
	Stem bromelain	[30] MF, PR, YL, IA, IG, HL, KG, DA, MG, EG (3), EA (3), PG, NKL, DG (2), KA (2), EV, IL (3), WA, EF, ER (3)	[40] MA, KA (2), PA, HA, IA, WA, HL, WR (2), WT, DR, EG (3), ES, EV, HF, HS, HT, HV, IL (3), KG, KS, MF, MG, MR, NA (2), PG, PS (2), PV (2), QA, QS, YL, YS
	Pepsin (pH>2)	[42] RY (3), RF (2), PL, PL (2), VK, IA, RA, IF (3), VG, IG (3), HL, HG, SG (5), PG, SF (3), VE (3), PT, HK, IL (4), WA, RG, ST (2), VM.	[53] VA (2), PA, HA (2), IA, RA, WA, HL, SL, WRT, WRA, PL (2), WQ, WT, HE, HF, HH, HS, HT, IL (4), IM, PF (2), PG, PN, PT, RG, RK (3), RM, RN, SF (3), SK, VD, VE (3), VG, VK, VL (2), VM, VN, VT (2)
60 kDa chaperonin	Ficin	[33] MF (2), AY, PL, VG (4), IG, MG (4), WG, QG (2), EG (2), QK, DG (4), AR, EY, EK (3), DY, DF, IL (2), VVR	[42] APG, EK (3), AL (4), PL, WG, AY, DR (2), EG (2), EY, IL (2), MF (2), MG (4), ML, MR, NR, PK (2), PY, QG (2), QL, TL (2), TR, TS, VG (4), VL

https://doi.org/10.26656/fr.2017.7(2).792

Table 7 (Cont.). Fragments of ACE-Inhibitory and DPP-IV inhibitory peptides predicted to be released by five proteases using BIOPEP enzyme action tool

Proteins	Enzyme used	ACE- inhibitory peptides	DPP IV-inhibitory peptides
60 kDa chaperonin	Calpain 2	[43] MF, AY, PL, VK, VP, VG (4), IG, AG (3), FG (2), MG (5), AI (2), SG, EG (2), VR, DG (3), YK, AR (2), EI, PT, EK, AV (2), AVL, DM, IL (2), ST (2)	[69] VP, VV, APG, EK, AL, AL (10), SL (2), VR, PL, AD (2), AE, AG (3), AT (3), AV (2), AY, DP, DR, EG (2), EI, HR, II, IL (2), MF, MG (5), ML, MR, NR, PK, PT, SK (2), VG (4), VI, VK, VL (2), VQ, VT (4), YD, YI, YK
	Pancreatic elastase	[48] MF, RA (4) , KG (2), FG (2), DA (5), MG (6), WG, QG (2), EG (2), EA (2), NG, PG (2), DG (3), KL (2), KA, EY, EI (4), EV (3), PT, DY, RG (2)	[56] KA, PA (4), RA (4), WG, EG (2), EI (2), ET, EV (3), EY, KG (2), KS, KT, KV (4), MF, MG (6), ML (2) , NA (3), NG, NT, PG (2), PT, PY, QG (2), QL, QS, QV (3), RG (2).
	Stem bromelain	[42] MF (2), YL, IA (3), YA, KG (3), DA (4), MG (6), WG, QG (2), EG, EA (2), NG, PG (2), DG (3), KF (2) , KL (2), KA, EV (3), PT, IL	[56] KA, PA (3), IA (3), WG, DR, EG, ET, EV (3), IL, KF (2), KG (3), KS, KT, KV (4), MF (2), MG (6), ML (2), MR, NA (3), NG, NR, PG (2), PT, PV, QG (2), QL, QS, QV (4), YA, YL
	Pepsin (pH>2)	[51] RL (3), VF, VRP, IPA, PL, VK (3), IA (7), RA (4), VG (5), IG (2), SG, PG (2), IE (3), VE (6), PT (2), IL (2), RG (3), ST (2), VM (2).	[79] VA (8), PA, IPA, IA (7), RA (4), SL (2), PL, HD, IL (2), IN, IQ, PF, PG (2), PK, PT (2), RG (3), RK (2), RL (3), SK, VD (3), VE (6), VF, VG (5), VK (3), VL (3), VM (2), VN (2), VQ (3), VT (7).
Heat shock protein 90AA1	Ficin	[41] IR (2), IY (2), VF, AY, PL (2), VK (3), AF (3), VG, IG (2), AG (2), EG (3), VR, QK, DG, NF, NK (2), EK (7), DY, IL, EF (3), ER	[52] EK (7), AL, VR, PL (2), AF (3), AG (2), AS, AY, EG (3), ES (2), IH, IL, IR (2), MK (2), NF, NL (2), NR (2), QF (3), QL, TK (3), TL (2), TS, TY, VF, VG, VK (3), VS (2)
	Calpain 2	[39] IR, YG, FY (2) , PL (2), VK (3), VG, IG, AG (3), HL, FG, MG, SG (3), EG (2), DG, NK, AR, EV, EK (7), YE, IL, ST (3), ER	[52] HL, EK (7), AL (5), SL (5), PL (2), WT (2), AE, AG (3), DN, DQ (2), EG (2), EV, HF, IL, IR, MG, MK, NL (2), NR, SK (3), VG, VK (3), VT (2), YE, YG, YM
	Pancreatic elastase	[22] RY (2), FY (2), PL, DA, MG, EG (2), EA, KL (4), KA (2), EY, EI (2), EV, RG (2).	[39] MA, KA (2), FA, FL, PL, WT (2), EG (2), EI (2), ES (2), ET, EV, EY, HS, HV, KI (2), KS (2), KV (3), MG, MI (2), NT, PI, PV, QA (2), QL (2), QT, RG (2).
	Stem bromelain	[29] IR (2), PR, YL, YG, PL, IA, IG (2), DA, MG, EG (3), EA, NF, KL, KA (3), EV, YV, IL (2), EF (2), ER (3).	[44] KA (3), IA, PL, WT (2), EG (3), ES, ET (2), EV, HS, HV, IL (2), IR (2), KS, KV (3), MG, NF, NR (2), PF, PV, QA (3), QF (3), QL (2), QT, YG, YL, YS (2), YV
	Pepsin (pH>2)	[50] RL (2), RY, IY (3), VF, PL (2), VK (5), IA (2), RA, IF, VG (2), IG (3), HL (2), SG (2), SF, RR, IE (3), VE (5), PT, PQ (2), HP, IL (4), RG, ST (3), PPL	[75] VA (3), HP, IA (2), RA, HL (2), SL (4), PL (2), PPL, WT (2), WE (2), HD (3), HF, IL (4), IM (3), IN (3), PF, PM, PQ (2), PT, RG, RK (2), RL (2), RM (2), RN, RR, SF, SK (3), VD (3), VE (5), VF, VG (2), VK (5), VQ (2) , VS, VT (4).

Table 8. The values of parameters describing the predicted efficiency of release of bioactive fragments from selected anchovy proteins by *in silico* hydrolysis

Protein	Enzymes	DH_t (%)	ACE Inhibitor		DPP IV Inhibitor	
			A_E	W	A_E	W
Myosin light chain 1 (Q9IB21)	Ficin	35.0515	0.0462	0.0969	0.0462	0.0721
	Calpain 2	42.2680	0.0564	0.1183	0.0667	0.1041
	Pancreatic elastase	41.7526	0.0462	0.0969	0.0615	0.0959
	Stem bromelain	47.7835	0.0513	0.1076	0.0564	0.0880
	Pepsin (pH>2)	77.8357	0.0769	0.1612	0.1436	0.2240
V(D)J Recombination-Activating Protein (A0A2H4PU29)	Ficin	47.4104	0.0577	0.2020	0.0716	0.1233
	Calpain 2	47.2112	0.0815	0.1961	0.0954	0.1643
	Pancreatic elastase	48.0080	0.0457	0.1100	0.0636	0.1096
	Stem bromelain	54.7809	0.0596	0.1434	0.0795	0.1370
	Pepsin (pH>2)	65.9363	0.0835	0.2010	0.1054	0.1816
60 kDa chaperonin (A0A3G1CVR4)	Ficin	39.1304	0.0573	0.1393	0.0729	0.1063
	Calpain 2	49.2174	0.0764	0.1549	0.1198	0.1747
	Pancreatic elastase	57.7391	0.0833	0.1689	0.0972	0.1417
	Stem bromelain	55.8261	0.0729	0.1795	0.0972	0.1417
	Pepsin (pH>2)	73.9130	0.0885	0.1478	0.1372	0.2001
Heat shock protein 90AA1 (A0A2R4GAA0)	Ficin	42.5101	0.0553	0.1605	0.0701	0.1241
	Calpain 2	43.0499	0.0526	0.1445	0.0701	0.1241
	Pancreatic elastase	43.5897	0.0296	0.0813	0.0526	0.0931
	Stem bromelain	40.8907	0.0391	0.1074	0.0593	0.1050
	Pepsin (pH>2)	76.3036	0.0674	0.1852	0.1011	0.1790

describing the predicted efficiency of the release of bioactive fragments from selected anchovy proteins by *in silico* proteolysis. From Table 8, the degree of hydrolysis (DH%) for the anchovy proteins that generated ACE inhibitory peptides and DPP-IV inhibitory peptides were between 35.0515% and 77.83576%. Among the five enzymes, pepsin (pH >2) gave the highest percentage of DH value for four anchovies proteins. This indicated the high ability of hydrolysis of pepsin (pH>2) compared to other enzymes.

The two most critical parameters that could be used to determine the possibility of bioactive peptides being generated from protein sequences by proteases are AE and W values (Minkiewicz *et al.*, 2011). Based on the AE and W value of each protein, pepsin (pH>2) marked its highest value (AE value between 0.0674-0.885) and (W value between 0.1478-0.2010). Parameter W showed the probability of fragments with a specific activity being released from a protein with a specified frequency of occurrence. The probability and "effectiveness" of the release of bioactive fragments from intact protein sequences were influenced by a higher value of parameter W (Borawska-Dziadkiewicz *et al.*, 2021).

4. Conclusion

The combination of proteomic technique and *in silico* approach can identify proteins from anchovy through LC-MS/MS and predict the potential bioactive peptide from these protein sequences using the BIOPEP database. Moreover, through *in silico* hydrolysis, the release of fragments with given activity can be predicted by a selection of suitable protease. Among the proteins, 60 kDa chaperonin and heat shock protein 90AA1 were the highest potential sources of bioactive fragments exhibiting all types of activity, specifically ACE inhibitory and DPP-IV inhibitory peptides. Meanwhile, pepsin (pH>2) released the highest number of ACE inhibitory peptides and DPP-IV inhibitory peptides compared to other proteases. This study will provide a database and guidelines for further research. Therefore, coupled with appropriate techniques of protein extraction and identification, these anchovy proteins can be accepted to high value-added products or bioactive peptides ingredients that can be used in the food, cosmetic and biomedical industries.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This work was funded by the Fundamental Research Grant Scheme (FRGS/1/2018/WAB09/UMT/03/2). The authors would like to extend their gratitude to everyone who helped in the article writing process.

https://doi.org/10.26656/fr.2017.7(2).792

References

Ahmad, F., Ayub, M.N.A, Mohamad, S.N. and Ibrahim, S. (2018). Proximate and amino acid compositions of selected dried anchovies (whole and processed) from Pangkor Island, Malaysia. *Malaysian Fisheries Journal*, 17, 39–50.

Anekthanakul, K., Hongsthong, A., Senachak, J. and Ruengjitchatchawalya, M. (2018). SpirPep: An *in silico* digestion-based platform to assist bioactive peptides discovery from a genome-wide database. *BMC Bioinformatics*, 19(1), 1–10. https://doi.org/10.1186/s12859-018-2143-0

Borawska-Dziadkiewicz, J., Darewicz, M. and Tarczyńska, A.S. (2021). Properties of peptides released from salmon and carp via simulated human-like gastrointestinal digestion described applying quantitative parameters. *PloS ONE*, 16(8), 1–23. https://doi.org/10.1371/journal.pone.0255969

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical *Biochemistry*, 7(1), 248-254. https://doi.org/10.1016/0003-2697(76)90527-3

Ceruso, M., Chirollo, C., Boccia, F., Smaldone, G., Marrone, R. and Pepe, T. (2015). Evaluation and comparison of four protein extraction protocols for mono- and two-dimensional electrophoresis in *Mytilus galloprovincialis*. *Italian Journal of Food Safety*, 4(3), 179–181. https://doi.org/10.4081/ijfs.2015.4534

Checkley, D.M., Asch, R.G. and Rykaczewski, R.R. (2017). Climate, anchovy, and sardine. *Annual Review of Marine Science*, 9(1), 469–493. https://doi.org/10.1146/annurev-marine-122414-033819

Choi, Y.J., Heu, M.S., Kim, H.R. and Pyeun, J.H. (2004). Properties of proteases responsible for degradation of muscle proteins during anchovy sauce fermentation. *Development of Food Sciences*, 42 (41), 425-439. https://doi.org/10.1016/S0167-4501(04)80042-6

Daliri, E.B.M., Oh, D.H. and Lee, B.H. (2017). Bioactive peptides. *Foods*, 6(5), 32. https://doi.org/10.3390/foods6050032

Darewicz, M., Borawska, J. and Pliszka, M. (2016). Carp proteins as a source of bioactive peptides – an *in silico* approach. *Czech Academy of Agricultural Sciences*, 2, 111–117. https://doi.org/10.17221/300/2015-CJFS

Darewicz, M., Borawska, J. Vegarud, G.E., Minkiewicz, P. and Iwaniak, A. (2014). Angiotensin I-converting enzyme (ACE) inhibitory activity and ACE inhibitory peptides of salmon (*Salmo salar*) protein hydrolysates obtained by human and porcine gastrointestinal enzymes. *International Journal of Molecular Sciences*, 15(8), 14077–14101. https://doi.org/10.3390/ijms150814077

Department of Fisheries Malaysia (DOF). (2019). Perangkaan Perikanan Tahunan 2019. Retrieved on July 15, 2021 from DOF website: http://www.dof.gov.my/index.php/pages/view/4046

Dewi, E.N. (2002). Chemical Analysis During the processing of dried salted anchovy. *Journal of Coastal Development*, 5(2), 55-65.

Food and Agriculture Organization of the United Nations (FAO). (2019). Fishery and Aquaculture Country Profiles. Malaysia. Country Profile Fact Sheets. Retrieved on July 15, 2021 from FAO website: http://www.fao.org/fishery/countrysector/naso_malaysia/en

Froese, R. and Pauly, D. (2020). FishBase. World Wide Web electronic publication. Retrieved on June 26, 2021 from Fish Base website: https://www.fishbase.se/search.php

Hernández-Ledesma, B., García-Nebot, M.J., Fernández-Tomé, S., Amigo, L. and Recio, I. (2014). Dairy protein hydrolysates: Peptides for health benefits. *International Dairy Journal*, 38(2), 82–100. https://doi.org/10.1016/j.idairyj.2013.11.004

Hernandez, O.M., Jones, M., Guzman, G. and Szczesna-cordary, D. (2006). Myosin essential light chain in health and disease. *American Journal of Physiology-Heart and Circulatory Physiology*, 292(4), H1643-H1654. https://doi.org/10.1152/ajpheart.00931.2006

Huang, B.B., Lin, H.C. and Chang, Y.W. (2015). Analysis of proteins and potential bioactive peptides from tilapia (*Oreochromis* spp.) processing co-products using proteomic techniques coupled with BIOPEP database. *Journal of Functional Foods*, 19 (Part A), 629–640. https://doi.org/10.1016/j.jff.2015.09.065

Isaacson, T., Damasceno, C.M.B., Saravanan, R.S., He, Y., Catalá, C., Saladié, M. and Rose, J.K.C. (2006). Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nature Protocols*, 1(2), 769–774. https://doi.org/10.1038/nprot.2006.102

Iwaniak, A., Dziuba, J. and Niklewicz, M. (2005). The BIOPEP database - A tool for the *in silico* method of classification of food proteins as the source of peptides with antihypertensive activity. *Acta Alimentaria*, 34(4), 417–425. https://doi.org/10.1556/AAlim.34.2005.4.9

Jakubczyk, A., Karas, M., Rybczynska-Tkaczyk, K., Zielinska, E. and Zielinski, D. (2020). Current trends

of bioactive peptides - New sources and therapeutic effect. *Foods*, 9(7), 846. https://doi.org/10.3390/foods9070846

Jiang, S., Qiu, L., Zhou, F., Huang, J., Guo, Y. and Yang, K. (2009). Molecular cloning and expression analysis of a heat shock protein (Hsp90) gene from black tiger shrimp (*Penaeus monodon*). *Molecular Biology Reports*, 36(1), 127–134. https://doi.org/10.1007/s11033-007-9160-9

Jim, F., Garamumhango, P. and Musara, C. (2017). Comparative analysis of nutritional value of *Oreochromis niloticus* (Linnaeus), nile tilapia, meat from three different ecosystems. *Journal of Food Quality*, 2017, 6714347. https://doi.org/10.1155/2017/6714347

Kwan, S.H., Baie, S. and Ismail, M.N. (2016). Profiling of proteins and post translational modifications of *Channa striatus* dried meat. *Current Proteomics*, 13 (1), 9–19. https://doi.org/10.2174/1570164613666160413123846

Kwan, S.H. and Ismail, M.N. (2018). Identification of the potential bio-active proteins associated with wound healing properties in snakehead fish (*Channa striata*) mucus. *Current Proteomics*, 15(4), 299–312. https://doi.org/10.2174/1570164615666180717143418

Le Gouic, A.V., Harnedy, P.A., FitzGerald, R.J., Mérillon, J.M. and Ramawat, K.G. (2018). Bioactive peptides from fish protein by-products. In Merillon, J.M. and Ramawat, K. (Eds). Bioactive Molecules in Food, p. 1-35. Cham, Netherland: Springer International Publishing. https://doi.org/10.1007/978-3-319-54528-8_29-1

Levy-Rimler, G., Viitanen, P., Weiss, C., Sharkia, R., Greenberg, A., Niv, A., Lustig, A., Delarea, Y. and Azem, A. (2001). The effect of nucleotides and mitochondrial chaperonin 10 on the structure and chaperone activity of mitochondrial chaperonin 60. *European Journal of Biochemistry*, 268(12), 3465–3472. https://doi.org/10.1046/j.1432-1327.2001.02243.x

Lin, H., Alashi, A.M., Aluko, R.E. and Pan, B.S. (2017). Antihypertensive properties of tilapia (O*reochromis spp.*) frame and skin enzymatic protein hydrolysates. *Food and Nutrition Research*, 61(1), 1–11. https://doi.org/10.1080/16546628.2017.1391666

Maguire, M., Coates, A.R.M. and Henderson, B. (2002). Chaperonin 60 Unfolds its Secrets of Cellular Communication. *Cell Stress and Chaperones,* 7(4), 317–329. https://doi.org/10.1379/1466-1268(2002)007<0317:CUISOC>2.0.CO;2

Market, E. and Papavasiliou, F.N. (2003). V (D) J recombination and the evolution of the adaptive immune system. *PLOS Biology*, 1(1), 24–27. https://doi.org/10.1371/journal.pbio.0000016

Medina, I. and Pazos, M. (2010). Oxidation and protection of fish. In Decker, E., Elias, R. and McClements, D.J. (Eds). Oxidation in Foods and Beverages and Antioxidant Applications, p. 91–120. Cambridge, United Kingdom: Woodhead Publishing Limited. https://doi.org/10.1533/9780857090331.1.91

Minkiewicz, P., Dziuba, J. and Michalska, J. (2011). Bovine meat proteins as potential precursors of biologically active peptides - A computational study based on the BIOPEP database. *Food Science and Technology International*, 17(1), 39–45. https://doi.org/10.1177/1082013210368461

Minkiewicz, P., Dziuba, J., Iwaniak, A., Dziuba, M. and Darewicz, M. (2008). BIOPEP database and other programs for processing bioactive peptide sequences. *Journal of AOAC International*, 91(4), 965–980. https://doi.org/10.1093/jaoac/91.4.965

Nieznańska, H., Nieznański, K. and Stêpkowski, D. (2002). The effects of the Interaction of myosin essential light chain isoforms with actin in skeletal muscles. *Acta Biochimica Polonica*, 49(3), 709–719. https://doi.org/10.18388/abp.2002_3780

Palani Kumar, Ruba Annathai, A., Jeya Shakila, R. and Shanmugan, S. (2014). Proximate and major mineral composition of 23 medium sized marine fin fishes landed in the Thoothukudi Coast of India. *Journal of Nutrition and Food Sciences*, 4(1), 259. https://doi.org/10.4172/2155-9600.1000259

Panjaitan, F.C.A., Gomez, H.L.R. and Chang, Y.W. (2018). *In silico* analysis of bioactive peptides released from Giant Grouper (*Epinephelus lanceolatus*) roe proteins identified by proteomics approach. *Molecules*, 23(11), 2910. https://doi.org/10.3390/molecules23112910

Park, K. and Kwak, I.S. (2014). Characterize and gene expression of heat shock protein 90 in marine crab *Charybdis japonica* following bisphenol A and 4-nonylphenol exposures. *Environmental Health and Toxicology*, 29(2), e2014002. https://doi.org/10.5620/eht.2014.29.e2014002

Pooja, K., Rani, S. and Prakash, B. (2017). *In silico* approaches towards the exploration of rice bran proteins-derived angiotensin-I-converting enzyme inhibitory peptides. *International Journal of Food Properties*, 20(2), 2178–2191. https://doi.org/10.1080/10942912.2017.1368552

Pyz-łukasik, R., Chałabis-Mazurek, A. and Gondek, M. (2020). Basic and functional nutrients in the muscles

of fish: A review. *International Journal of Food Properties*, 23(1), 1941–1950. https://doi.org/10.1080/10942912.2020.1828457

Sánchez, A. and Vázquez, A. (2017). Bioactive peptides: A review. *Food Quality and Safety*, 1(1), 29–46. https://doi.org/10.1093/fqs/fyx006

Shevchenko, A., Tomas, H., Havliš, J., Olsen, J.V. and Mann, M. (2007). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, 1(6), 2856–2860. https://doi.org/10.1038/nprot.2006.468

Shi, J., Fu, M., Zhao, C., Zhou, F. and Yang, Q. (2016). Characterization and function analysis of Hsp60 and Hsp10 under different acute stresses in black tiger shrimp (*Penaeus monodon)*. *Cell Stress and Chaperones*, 21(2), 295–312. https://doi.org/10.1007/s12192-015-0660-6

Shiriskar, D.A., Khedkar, G.D. and Sudhakara, N.S. (2010). Preparation of boiled and dried products from anchovies (*Stolephorus* sp.) and studies on quality changes during storage. *Journal of Food Processing and Preservation*, 34(SUPPL. 1), 73–86. https://doi.org/10.1111/j.1745-4549.2008.00280.x

Shukla, S. (2015). A Bradford multiplexing method for protein estimation in fermented foods: Soy sauces. *Bangladesh Journal of Pharmacology*, 11(1), 24–27. https://doi.org/10.3329/bjp.v11i1.25796

Tanasupawat, S. and Visessanguan, W. (2014). Fish Fermentation. In Boziaris, I.S. (Ed). Seafood Processing: Technology, Quality and Safety, p. 177–207. Chichester, United Kingdom: John Wiley and Sons Limited. https://doi.org/10.1002/9781118346174.ch8

Thienchai, S. and Chaiyanan, S. (2012). Anchovy paste production from lactic acid fermentation of anchovy fish found in the Gulf of Thailand. *Asia-Pacific Journal of Science and Technology*, 17(4), 556–564.

UniProt Consortium. (2019). UniProt: the universal protein knowledge. Nucleic acids research, 49(D1). Retrieved on August 21, 2021 from UniProt website: https://www.uniprot.org/

