Proceedings of the 36th European Peptide Symposium
Michal Lebl (Editor)
European Peptide Society, 2022 *https://doi.org/10.17952/36EPS.2022.279*

Bioactive Peptides from Salmon Collagen: An *in silico* Approach

Constanza Cárdenas[1], Tanya Román[1,2], Claudio Alvarez[3,4], Paula Santana[5], and Fanny Guzmán[1]

[1]Nucleo de Biotecnología Curauma. Pontificia Universidad Católica de Valparaíso. Valparaíso, Chile.; [2]Programa de Doctorado en Biotecnología. Pontificia Universidad Católica de Valparaíso y Universidad Federico Santa María. Valparaíso, Chile; [3]Laboratorio de Fisiología y Genética Marina (FIGEMA), Centro de Estudios Avanzados en Zonas Áridas (CEAZA), Coquimbo, Chile; [4]Facultad de Ciencias del Mar, Universidad Católica del Norte, Coquimbo, Chile; [5]Facultad de Ingeniería, Instituto de Ciencias Químicas Aplicadas, Universidad Autónoma de Chile, Santiago, Chile

Introduction

Bioprospecting is a source for finding new bioactive compounds, and today there are several sources to achieve this search. Collagen is one of the most abundant proteins in animal tissues, almost 30% of total protein [1], and in the case of salmon processing by-products, such as scales and skin, is the most abundant protein. Although about 30 types of collagens are reported, type I is the most abundant and the main one in the by-products of salmon processing [1,2].

Collagen is a target for applications in many fields such as the cosmetic industry, food additives, tissue regeneration, bioactive peptides, etc [3-8]. In this report, we work with the reported sequences of Atlantic salmon and rainbow trout collagen, doing an *in silico* hydrolysis with Alcalase (subtilisin) and characterizing the peptides obtained by a set of descriptors and physicochemical properties obtained with the Peptide package [5] implemented in Rstudio [6], and comparing them with other reported peptides of known activities. As a result, it is possible to classify the peptides from *Salmo salar* and *Onchorhynchus mykiss* collagen by including them in clusters of sequences related with specific activities. This methodology allows us to reduce the search space and to select sequences with potential activity that could be synthesized and tested. Additionally, this approach can be used as a way to add value to the by-products from the salmon industry [11,12].

Results and Discussion

In this work we chose proteins from collagen type I, which is the main type of collagen present in the byproducts of fish processing (Table 1). These proteins were hydrolyzed in silico, through the BIOPEP server [13], by using the enzyme Alcalase, resulting in a total of 477 unique peptides.

Table 1. Salmo salar and Oncorhynchus mykiss collagen proteins used for this study.

Chain	*Source*	*Uniprot ID*	*Length*	*# Unique peptides*
Alpha 1	*Salmo salar*	A0A1S3RF89	544	125
	Oncorhynchus mykiss	Q910C0	1449	70
Alpha 2	*Salmo salar*	A0A1S3Q205	1356	132
	Oncorhynchus mykiss	O93484	1356	83
Alpha 3	*Oncorhynchus mykiss*	O93486	678	67

In parallel, peptides reported in databases with different activities were selected, as shown in Table 2.

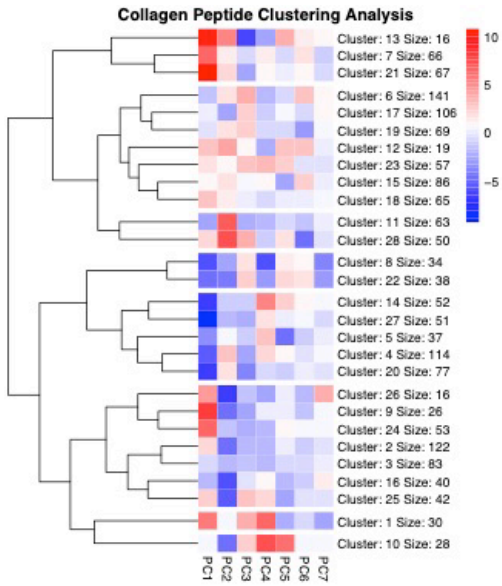
Table 2. Number of peptides from databases with different activities, used in this study.

Activity	*# Peptides*
Antioxidant	230
Antifreezing	214
Cell Penetrating Peptides	228
Antiparasitic	403
Insecticidal	58
Collagen peptides	477
Total	2781

Peptide sequences were characterized in Rstudio [6], with the Peptides package [5], by initially using 71 descriptors listed below:

- aIndex: Aliphatic index defined as the relative volume occupied by the side chains of the aliphatic amino acids (Ala, Ile, Leu y Val).
- blosumIndices: Indices derived from the physicochemical properties contained in the BLOSUM62 matrices. It includes 10 indices from BLOSUM1 to BLOSUM10
- boman: Boman index, the potential for protein interaction
- Charge: Theoretical net charge at pH = 7 and using pKscale = EMBOSS.
- crucianiProperties: Use the scaled principal component scores that summarize a broad set of descriptors calculated based on the interaction of each amino acid residue with several chemical groups (or "probes"), such as charged ions, methyl, hydroxyl groups, and so forth; PP1: related to polarity; PP2: related to hydrophobicity; PP3: related to formation of hydrogen bonds
- fasgaiVectors: The FASGAI vectors (Factor Analysis Scales of Generalized Amino Acid Information) are a set of amino acid descriptors, that reflects hydrophobicity, alpha and turn propensities, bulky properties, compositional characteristics, local flexibility, and electronic properties, that can be utilized to represent the sequence structural features of peptides or protein motifs, F1 to F6
- Hydrophobicity: Calculates the GRAVY index or average hydrophobicity based on the Kidera scale.
- kideraFactors: The Kidera Factors were originally derived by applying multivariate analysis to 188 physical properties of the 20 amino acids and using dimension reduction techniques. This function calculates the average of the ten Kidera factors for a protein sequence. KF1 to KF10.
- mswhimScores: Molecular Surface-Weighted Holistic Invariant Molecular. MS-WHIM scores were derived from 36 electrostatic potential properties derived from the three dimensional structure of the 20 natural amino acids.
- pI: Isoelectric point with pKscale = EMBOSS
- protFP: The ProtFP descriptor set was constructed from a large initial selection of indices obtained from the AAindex database for all 20 naturally occurring amino acids. From ProtFP1 to ProtFP8
- stScales: ST-scales were proposed by Yang et al, taking 827 properties into account which are mainly constitutional, topological, geometrical, hydrophobic, electronic, and steric properties of a total set of 167

The 71 descriptors were transformed by using principal component analysis, selecting the first seven Principal Components (PC) to perform a Kmeans clustering analysis with pheatmap [14] implemented in Rstudio. As result we obtained the clusters showed in Figure 1.



Cluster	Antiox	AFP	CPP	Antip	Insect	Other	S_A1	O_A1	S_A2	O_A2	O_A3	Total
13			5	5			2		2	1	1	16
7	4		40	17	1		1		2		1	66
21	4		55	7		1						67
6	4	94	1	10	20	2	5	2	1	1	1	141
17	9	48	10	20	1	4	1	1	5	4	3	106
19	14	6	2	14	2	9	6	3	6	3	4	69
12	2	1	3	1		2	3	1	3	1	2	19
23	34			8	1	1	6	3	2	1	1	57
15	6			45	27	1	2	1	2	1	1	86
18	4	1	3	42	5	6	3				1	65
11	2					2	10	13	16	12	8	63
28	21			7		1	6	3	3	3	6	50
8	2	24	1			4	2	1				34
22	1	33		1		1			1	1		38
14	8						12	6	14	6	6	52
27	3						9	4	17	13	5	51
5	15			6			6	2	5		3	37
4	17						25	16	29	16	11	114
20	4	2					21	12	16	16	6	77
26	6		1	9								16
9	3		7	14					1		1	26
24	1		25	26							1	53
2	2		57	62		1						122
3	7		6	65			2		2	1		83
16	12		1	22		1	2		1	1		40
25	15	4	1	14				1	2		5	42
1	17			5		3	1	1	1	2		30
10	13		10	3	1				1			28
Total	230	213	228	403	58	39	125	70	132	83	67	1648

Fig. 1. Clustering analysis of the 2781 peptides characterized with seven PC. Left: dendrogram of the kmean clustering with 28 clusters.Rigth: clusters members according the activity, Antioxidant, Antifreezing (AFP), Cell penetrating peptides (CPP), Antiparasitic (Antip), Insecticidal (Insect), others, and the fish collagen peptides S_A1,O_A1 for alpha 1 chain, S_A2, O_A2 for alpha chain 2 for salmon and oncorhynchus, and O_A3 for oncorhynchus alpha 3 chain.

From these results it is possible to establish groups of similarity of the peptides obtained from the *in silico* collagen hydrolysis with the reported bioactive peptides. We found 24 peptides associated with antioxidant activity, clusters 6 and 17 (highlighted in blue in Figure 1). 278 peptides with potential antioxidant activity in clusters 14, 27, 5, 4 and 20 (highlighted in orange in Figure 1) and three different groups of clusters of peptides with potential antiparasitic activity, the first group of 10 peptides of clusters 13 and 7, the second of 56 peptides of clusters 19, 12, 23, 15 and 18, and the third with 20 peptides of the clusters 9, 24, 3, 16 and 25.

This procedure constitutes the first filter in the search for bioactive peptides, and the projection of the work will be the synthesis of these potentially active peptides to be tested in *in vitro* assays against the parasite *Caligus rogercresseyi*. Another projection is to use byproducts from fish processing to obtain hydrolysates, and be able to isolate fractions with bioactive peptides, having as background the results obtained here.

Acknowledgments

The work was funded by the Grant project Fondecyt 1210056.

References

1. Silvipriya, K.S., et al. *J App Pharm Sci*. **5**, 123-127 (2015), https://doi.org/10.7324/JAPS.2015.50322
2. Al-Nimry, S., et al. *Mar Drugs* **19**,145, (2021), https://doi.org/10.3390/md19030145
3. Geahchan, S., Baharlouei, P., Rahman, A. *Mar Drugs* **20**(1), 61 (2022), https://doi.org/10.3390/md20010061
4. Jana, S., et al. *J Mater Chem B.* **10**(4), 489-505 (2022), https://doi.org/10.1039/d1tb02125g
5. Yamada, S., et al. *Dent Mater J.* **40**(6), 1295-1302(2021), https://doi.org/10.4012/dmj.2020-446
6. Wu, R., et al. *Food Chem.* **248**, 346-352 (2018), https://doi.org/10.1016/j.foodchem.2017.12.035
7. Derkach, S.R., et al. *Polymers (Basel)* **12**(12), 3051 (2020), https://doi.org/10.3390/polym12123051
8. Nuñez, S.M., et al. *Electron J Biotech.* **48**, 101-108 (2020), https://doi.org/10.1016/j.ejbt.2020.09.009

9. Osorio, D., Rondon-Villarreal, P., Torres, R. *The R Journal* **7**(1), 4-14, ISSN 2073-4859 (2015), https://journal.r-project.org/archive/2015/RJ-2015-001.pdf
10. RStudio Team (2015) *RStudio: Integrated Development for R. RStudio*, Inc., Boston, MA URL http://www.rstudio.com/
11. Coppola, D., et al. *Mar Drugs* **19**(2), 116 (2021), https://doi.org/10.3390/md19020116
12. Caruso, G., et al. *Mar Drugs* **18**(12), 622 (2020), https://doi.org/10.3390/md18120622
13. Minkiewicz, P., Iwaniak, A., and Darewicz, M. (2019) BIOPEP-UWM Database of Bioactive Peptides: Current Opportunities. *International journal of molecular sciences* **20**(23), 5978 (2020), https://doi.org/10.3390/ijms20235978. https://biochemia.uwm.edu.pl/biopep-uwm/
14. Kolde, R., *Pheatmap: pretty heatmaps. V1.0.12. R package*. (2012). https://cran.r-project.org/web/packages/pheatmap/index.html