

CORRELATION BETWEEN PROLINE-PROLINE/HYDROXY-PROLINE IN COLLAGEN PROTEINS AND THE MOISTURE CONTENT OF THE STRATUM CORNEUM AFTER ORAL INGESTION

Yuri Mukai¹, Keitaro Machino², Kenji Etchuya³, Hiroataka Tanaka⁴, Takanori Sasaki⁵

^{1,2,5}Department of Electronics and Bioinformatics, School of Science and Technology,
Meiji University, Kawasaki, Japan

^{1,3,4,5}Faculty of Electrical Engineering, Graduate School of Science and Technology,
Meiji University, Kawasaki, Japan
Email: ¹yuri@isc.meiji.ac.jp

ABSTRACT

While collagen was previously thought to be related to the moisture of skin, the correlation between orally ingested collagen and the moisture of skin had not been clarified. Based on recent studies regarding collagen peptide's ability to assist hyaluronan synthesis and to change the moisture content of the stratum corneum after oral ingestion of fish/porcine collagen, the relationship between the di-peptide in fish/porcine collagen, Proline-Proline/hydroxy-Proline (Pro-Pro/HyP), and the moisture content of the stratum corneum was evaluated in this study. Moreover, collagen proteins in Eukaryotic genomes which are thought to be able to effectively increase the moisture content of the stratum corneum were retrieved from the UniProtKB/Swiss-Prot protein sequence database. Fish/porcine collagens which have high Pro-Pro/HyP frequency were components of filaments in spines, cartilage, ligament and bone. In addition, human and mouse organs with frequent elastic actions, including the kidney, lung, heart, muscles and skin, Pro-Pro/HyP-rich collagen proteins were found by comprehensive database scanning.

Key words: collagen, hyaluronan, hydroxy-Proline (HyP), peptide, Proline (Pro), sequence analysis, Welch's *t*-test

I. INTRODUCTION

Collagen is known as a structural protein in the extracellular matrix which is found in corium, bone, dens and tendons of multicellular organisms. Collagen proteins have several Glycine (Gly)-X-Y sequences called (GXY)_n domain in their primary sequences and whose form takes the shape of triple-strand helices. While X and Y are arbitrary amino acid residues, Proline (Pro) frequency is higher than other residues in both X and Y. Especially, two hydroxylated residues including hydroxy-Proline (HyP) and hydroxy-Lysine (HyK) are commonly found in Y which are rarely found in other proteins except collagen protein families (1, 2). Moreover, these hydroxyl groups become modified by oligosaccharides including galactose and glucose. Based on high hydrophilicity of (GXY)_n domains as described above, it was previously believed that a correlation between hydroxyl groups, oligosaccharides of collagen and the moisture content of the stratum corneum existed.

Since collagens are degraded to monomeric amino acids and small peptides by pepsin when collagen is ingested orally, the collagens are not

thought to be absorbed into the stratum corneum directly. Collagen peptides made from solubilized collagen proteins are available on the market as beneficial supplements for the skin. However, the correlation between orally ingested collagen and the moisture of skin had not been clarified.

Recent studies reported that di- and tri-peptides derived from collagens were detected in human blood two hours after the oral ingestion of fish scale and porcine skin collagen hydrolysates and in particular, the small peptides including HyP, peculiar to collagen, remained in the blood (3). Moreover, it was indicated that Pro-HyP enhances both cell mitotic activity and hyaluronic mediated by the activation of hyaluronan synthase II transcriptions (4). Hyaluronan is a component of the stratum corneum and a mucoid which has extremely high hydrophilicity. A possibility was noted that the peptide obtained from collagen raises the amount of hyaluronan and affects the moisture content of the stratum corneum. On the other hand, oral ingestion of fish collagen increased the moisture content of the stratum corneum more than that of porcine collagen (5). Therefore, a hypothesis was

Table 1. Change of the moisture content of the stratum corneum after oral ingestion of fish/porcine collagen peptide (5)

Test substance	Subject (number)	Moisture content before ingestion	Moisture content four weeks after ingestion	Amount of change in the moisture content
Placebo 10 g	Total Subject (37)	249.3 ± 83.6	314.5 ± 21.9	65.2 ± 107.0
	Subject over thirty years old (22)	256.1 ± 87.5	291.3 ± 19.3	35.2 ± 95.6
Fish scales collagen 10 g	Total Subject (39)	237.5 ± 79.9	336.6 ± 111.3	99.1 ± 95.4
	Subject over thirty years old (28)	239.7 ± 87.5	338.6 ± 102.1	98.9 ± 76.9
Porcine skin collagen 10 g	Total Subject (37)	255.8 ± 89.4	340.4 ± 127.3	84.6 ± 10.3
	Subject over thirty years old (27)	267.5 ± 86.8	344.9 ± 124.7	77.4 ± 102.5

* Average ± Standard deviation

formed using the aforementioned facts as follows; The enhancement ability of hyaluronan synthesis is due to higher Pro-Pro/HyP frequency in fish collagen than in porcine collagen, resulting in an increase in the moisture content of the stratum corneum. In this research, the correlation Yuri Mukai. *et al.*: Correlation between Pro-Pro/HyP in collagen proteins and the moisture content between Pro-Pro/HyP frequencies in fish/porcine collagens and the moisture content of the stratum corneum was evaluated. Moreover, collagen proteins which are thought to be able to effectively increase the moisture content of the stratum corneum were retrieved from Eukaryotic genomes.

II. METHODS

A. Dataset extraction and evaluation by Welch's *t*-test

Datasets of fish and porcine collagens were extracted from UniProtKB/Swiss-Prot protein sequence database release 2011_07 by conducting a search with keywords "Actinopterygi" in OC lines or "Pig" in OS lines, and "Collagen" or "collagen" in KW lines. Pro-Pro/HyP pair frequencies of each entry in the fish and the porcine datasets were counted. The frequency of Pro-Pro-Pro was defined as two times. Difference of Pro-Pro/HyP frequencies between fish and porcine

collagen proteins was evaluated by Welch's *t*-test. P-value less than 0.05 were considered as a statistically significant difference.

B. Comprehensive database scanning of ProPro/HyP-rich proteins

Pro-Pro/HyP-rich collagen proteins in Eukaryotic genomes were retrieved from the UniProtKB/Swiss-Prot release 2011_07, and Pro-Pro/HyP pair frequencies of selected entry were counted. The frequency of Pro-Pro-Pro was defined as two times.

III. RESULTS AND DISCUSSION

A. Correlation between the amount of Pro-Pro/HyP and the moisture content of the stratum corneum

Fish and porcine collagen data (fish: 15 entries and porcine: 20 entries) was obtained from UniProtKB/Swiss-Prot protein sequence database release 2011_07. Annotations about hydroxylation of Pro residues were still not sufficient in UniProtKB/Swiss-Prot, therefore all "PP" characters in SQ lines in UniProtKB/Swiss-Prot were considered as Pro-Pro/HyP pairs.

Table 2. Pro-Pro/HyP frequencies of fish/porcine collagen proteins

Fish	Pro-Pro/HyP	Porcine	Pro-Pro/HyP
ACES_DANRE	2	A4_PIG	1
CCBE1_DANRE	7	BGH3_PIG	2
CO1A2_ONCMY	15	C1QA_PIG	2
COL11_DANRE	1	DPP4_PIG	2
COL12_DANRE	9	FCN1_PIG	0
CRA1A_DANRE	25	FCN2_PIG	2
CRA1B_DANRE	18	IL1A_PIG	0
GT251_DANRE	2	IL1B_PIG	0
IL1B_ONCMY	1	ITB1_PIG	0
OTO1A_DANRE	10	MBL1_PIG	2
OTOL1_ONCKE	8	MBL2_PIG	0
SCAR5_DANRE	1	MMP1_PIG	0
TSP3A_DANRE	4	MMP14_PIG	2
TSP3B_DANRE	2	PGS1_PIG	1
TSP4B_DANRE	1	PGS2_PIG	3
		PLMN_PIG	2
		SFTPA_PIG	4
		SFTPD_PIG	6
		SPRC_PIG	1
		TIMP1_PIG	1

Table 3. Difference of Pro-Pro/HyP frequencies between fish/porcine collagen proteins

	Fish (20)	Porcine (20)	P value
Pro-Pro/HyP (Number of pair)	7.07±7.08	1.55±1.50	0.012

* Average ± Standard deviation Yuri Mukai. *et al.*: Correlation between Pro-Pro/HyP in collagen proteins and the moisture content

Table 2 shows the entries of fish and porcine collagen proteins retrieved from UniProtKB/Swiss-Prot and Pro-Pro/HyP frequencies in their sequences. While an entry was found in the porcine collagen dataset which has a maximum number of six pairs of Pro-Pro/HyP, almost all entries have less than two pairs. Six entries have no Pro-Pro/HyP pair. On the other hand, fish collagen proteins contained at least one or more Pro-Pro/HyP pairs. An entry was found in the fish collagen dataset which has a maximum number of 25 pairs of Pro-Pro/HyP.

Almost all Pro-Pro/HyP-rich entries in the fish dataset were derived from *Danio rerio* (Zebra fish). CRA1A_DANRE (25 Pro-Pro/HyP pairs) and CRA1B_DANRE (18 pairs) were collagen α -1 proteins in spine and cartilage. CO1A2_ONCMY (15 pairs) was a collagen α -2 protein in ligament and bone derived from *Oncorhynchus mykiss* (Rainbow trout), while SFTPD_PIG (6 pairs) was a collagen protein found in porcine lung.

Table 3 shows the difference of Pro-Pro/HyP frequencies between fish and porcine collagen proteins evaluated by Welch's *t*-test. As seen in Table 3, a statistically significant difference in Pro-Pro/HyP frequencies was found among fish and porcine groups ($P < 0.05$). Yuri Mukai. *et al.*: Correlation between Pro-Pro/HyP in collagen proteins and the moisture content

As Table 1 and Table 3 show, the amount of Pro-Pro/HyP in collagen proteins was positively correlated with the moisture content of the stratum corneum. Therefore, the difference of the moisture contents among fish and porcine groups in reference (5) was able to be explained as follows; Higher amounts of di-or tri-peptide including Pro-Pro/HyP were released into the blood after oral ingestion of fish collagen proteins than that of porcine collagens. Fish collagen ingestion increased the expression amount of hyaluronan synthesis protein than did porcine collagens. As a result of the enhancement of hyaluronan synthesis with higher moisturizing capability, the moisture content of the stratum corneum rose more significantly.

Table 4. Pro-Pro/HyP-rich collagen proteins in UniProtKB/Swiss-Prot

Species	Entry number	UniProt/Swiss-Prot ID (Pro-Pro/HyP frequencies)
<i>Homo sapiens</i> (Human)	31	CO4A5_HUMAN (69), CO4A4_HUMAN (63), CO1A1_HUMAN (62), CO5A1_HUMAN (58), CO5A3_HUMAN (52), CO7A1_HUMAN (52), CO1A1_HUMAN (49), COMA1_HUMAN (49), USH2A_HUMAN (49), COBA1_HUMAN (48), COBA2_HUMAN (48), COGA1_HUMAN (46), CO4A1_HUMAN (45), CO4A3_HUMAN (45), TENX_HUMAN (45), CO3A1_HUMAN (42), CORA1_HUMAN (41), CO2A1_HUMAN (40), COFA1_HUMAN (40), WDR33_HUMAN (39), COCA1_HUMAN (36), CO4A2_HUMAN (35), CO1A2_HUMAN (33), CO5A2_HUMAN (33), PGBM_HUMAN (29), CO8A1_HUMAN (27), CO9A1_HUMAN (27), CO9A3_HUMAN (27), COJA1_HUMAN (27), CO4A6_HUMAN (25), CO8A2_HUMAN (25)
<i>Mus musculus</i> (Mouse)	25	CO4A4_MOUSE (65), CO1A1_MOUSE (57), CO5A1_MOUSE (53), CO7A1_MOUSE (51), USH2A_MOUSE (50), CO4A3_MOUSE (49), CO1A1_MOUSE (47), CO4A1_MOUSE (43), CO3A1_MOUSE (42), COBA1_MOUSE (42), COGA1_MOUSE (40), COBA2_MOUSE (38), COFA1_MOUSE (37), CORA1_MOUSE (36), CO2A1_MOUSE (35), COCA1_MOUSE (34), CO5A2_MOUSE (32), MIA3_MOUSE (31), COJA1_MOUSE (30), CO8A1_MOUSE (29), CO9A1_MOUSE (29), CO4A2_MOUSE (28), CO1A2_MOUSE (25), CODA1_MOUSE (25), PGBM_MOUSE (25)
<i>Rattus norvegicus</i> (Rat)	6	CO5A1_RAT (52), CO1A1_RAT (45), CO3A1_RAT CO2A1_RAT (35), CORA1_RAT (34), CO1A2_RAT (29)
<i>Bos taurus</i> (Bovine)	4	CO1A1_BOVIN (49), COBA2_BOVIN, (49), CO2A1_BOVIN (40), CO2A2_BOVIN (33)
<i>Canis familiaris</i> (Dog) (<i>Canis lupus familiaris</i>)	3	CO4A5_CANFA (64), CO1A1_CANFA (50), CO1A2_CANFA (33)
<i>Gallus gallus</i> (Chicken)	3	COCA1_CHICK (39), CO8A1_CHICK (28), CO9A1_CHICK (25)
<i>Perna viridis</i> (Tropical green mussel)	2	FP1V1_PERVI (29), FP1V2_PERVI (25)
<i>Caenorhabditis elegans</i>	1	BLI1_CAEEL (34)
<i>Cynops pyrrhogaster</i> (Japanese common newt)	1	CO1A1_CYNPY (35)
<i>Xenopus laevis</i> (African clawed frog)	1	CO2A1_XENLA (31)
<i>Xenopus tropicalis</i> (Western clawed frog)	1	CO2A1_XENTR (30)
<i>Ascaris suum</i> (Pig roundworm)	1	CO4A2_ASCSU (38)
<i>Cricetulus longicaudatus</i> (Chinese hamster)	1	CO5A1_CRILO (53)
<i>Oryctolagus cuniculus</i> (Rabbit)	1	CO8A1_RABIT (29)
<i>Rhabditis briggsae</i>	1	COL99_CAEBR (26)
<i>Drosophila melanogaster</i> (Fruit fly)	1	COLL1_DROME (27)
<i>Danio rerio</i> (Zebra fish) (<i>Brachydanio rerio</i>)	1	CRA1A_DANRE (26)
<i>Mytilus coruscus</i> (Sea mussel)	1	FP1_MYTGA (48)
<i>Mytilus galloprovincialis</i> (Mediterranean mussel)	1	FP1_MYTCO (56)

B. Comprehensive database scanning of ProPro/HyP-rich proteins

Eighty-six entries with more than 25 of Pro-Pro/HyP pairs were retrieved from all proteins in UniProtKB/Swiss-Prot release 2011_07 (Table 4). In mammalian, there were large amount of Pro-Pro/HyP-rich collagen proteins. In particular in *Homo sapiens* (human), there were 31 entries which have more than 25 pairs. CO4A5_HUMAN has the maximum number of 69 pairs of Pro-Pro/HyP, was human collagen α -5 protein found in kidney. Sixty-five and 64 pairs of Pro-Pro/HyP were found in mouse collagen α -4 protein (CO4A4_MOUSE) and collagen α -5 protein of *Canis familiaris* (Dog) (CO4A5_CANFA) in kidney, lung, heart, muscle and skin. Pro/HyP-rich region were also found in Amphibian collagen proteins in lung of *Xenopus laevis* (African clawed frog), *Xenopus tropicalis* (Western clawed frog) and *Cynops pyrrhogaster* (Japanese common newt), mussel byssus collagens of *Mytilus coruscus* (sea mussel) and *Mytilus galloprovincialis* (Mediterranean mussel), larva and pupa collagens of *Drosophila melanogaster* (Fruit fly), and a collagen in *Ascaris suum* (Pig roundworm). Mentioned above, Pro-Pro/HyP-rich collagen proteins were existed in the organization with both rich moisturizing and elasticity. In particular in mammalia, collagen proteins in kidney, lung and heart were found significantly. Since these organs have the functions as an acceptor and a pump of large amount of fluid, it is surmised that strong hydrophilic environment and the elasticity of internal organs are required. Pro-Pro/HyP-rich collagens are thought to be strongly related in these features. It is considered that the tension tolerance of collagen filaments is maintained due to rich cross-linkages between Pro-Pro/HyP domains among three collagen molecules.

Data registration of collagen proteins to the UniProt/Swiss-Prot and UniProt/TrEMBL databases is

behind, and a large number of collagen proteins which are not identified yet are thought to be still remained in various genomes. Therefore, proteins with more abundant Pro-Pro/HyP pairs may be found in unknown collagens.

In the future, the content of Pro-Pro/HyP pairs should be investigate comprehensively from predicted collagen proteins based on (GXY)_n motifs. Alternatively, Pro-Pro/HyP-rich ORFs may be able to predict as collagen proteins. It would be expected to predict collagen proteins suitable for supplements with the effect which can increase effectively the moisture content of skin.

ACKNOWLEDGEMENT

The authors are grateful to Dr. T. Terasaki and Dr. A.S. Magdalena for providing useful comments and proofreading.

REFERENCES

- [1] Ramshaw, J.A.M., Shah, N.K., Brodsky, B., 1998, J. Struct. Biol., 122, pp. 86-91.
- [2] Bos K.J., Rucklidge, G.J., Dunbar, B., Robins, S.P., 1999, Matrix Biol., 18, pp. 149-153.
- [3] Ohara, H., Matsumoto, H., Ito, K., Iwai, K., Sato, K., 2007, J. Agric. Food Chem., 55, pp. 15321535.
- [4] Ohara, H., Ichikawa, S. Matsumoto, H., Akiyama, M., Fujimoto, N., Kobayashi, T. Tajima, S., 2010, J. Dermatol., 37, pp. 330-338.
- [5] Ohara, H., Ito, K., Iida, H., Matsumoto, H., 2009, Nihonshokuhinkagakukougakkaishi (Japanese), 56 (3), pp. 137-145.



Yuri Mukai is a lecturer of Department of Electronics and Bioinformatics, School of Science and Technology, Meiji University, Japan.