Pakistan J. Zool., vol. 52(1), pp 321-330, 2020. **DOI**: https://dx.doi.org/10.17582/journal.pjz/2020.52.1.321.330

New Strategy for Expression of Recombinant Human Prolyl-4-Hydroxylase in *Pichia pastoris*

Yating Cheng[1], Wenlong Shi[1], Xue Xiao[2], Qirong Zhang[2], Qihao Zhang[1], Zhijian Su[1], Qi Xiang[1,2]* and Yadong Huang[1,2]

[1]*Institute of Biomedicine and Guangdong Provincial Key Laboratory of Bioengineering Medicine, Jinan University, Guangzhou 510632, PR China*
[2]*Biopharmaceutical RD Center of Jinan University, Guangzhou 510630, PR China*

Yating Cheng and Wenlong Shi contributed equally to this study.

Article Information
Received 25 July 2019
Revised 30 August 2019
Accepted 03 October 2019
Available online 05 November 2019

Authors' Contribution
QX and YH designed the study. YC and WS performed experimental work, also helped in the experiment of rhP4H expression and functional identification of enzyme. XX and QZ helped in genetic engineering work. ZS, QZ and YC analyzed the data. YC and QX wrote the article.

Key words
Recombinant human prolyl 4-hydroxylase (rhP4H), Collagen, Hydroxylation, Activity

ABSTRACT

Collagen is one of the most abundant proteins in humans and plays an essential role in cell maintenance and organization. Collagen has a unique triple-helix structure composed of three polypeptides, and hydroxylation of proline residues is vital for the stability of this triple-helix. Prolyl-4-hydroxylase (P4H) is an enzyme with an α2β2 tetramer arrangement that functions to post-translationally hydroxylate proline residues in the collagen chain. The α subunit (P4Hα) of the P4H α2β2 tetramer contains the catalytic site, whereas the β subunit (PDI) maintains the solubility and activity of the α subunit. In this study, two plasmids were used to produce an active human P4H tetramer (rhP4H) in *Pichia pastoris*. The P4Hα was cloned into pPICZαA, and PDI was cloned into PHIL. Real-time PCR showed that the P4H gene was expressed at the mRNA level. SDS-PAGE and western blotting analysis showed that the secreted expression of rhP4H was achieved successfully in *P. pastoris* GS115 cells. The catalytic activity of rhP4H was tested in a system that contained O_2, Fe^{2+}, α-ketoglutaric acid and ascorbic acid. Liquid chromatography coupled with tandem mass spectrometry was used to monitor the hydroxylation of proline. The results showed that rhP4H specifically hydroxylated proline residues only in the Gly-X-Y structure but did not react with free proline. We hypothesize that rhP4H can be engineered to control structural properties of recombinant collagen by selecting a certain degree of hydroxylation of proline residues, which should contribute to creating engineered collagens with specific end uses.

INTRODUCTION

Collagen plays an essential role in cell maintenance and organization, and is used widely in pharmaceutical, chemical, health products, food, cosmetics and feed industries. Native collagen adopts a triple-helix conformation, which consists of three α chains that are characterized by the repeating sequence motif Gly-X-Y (Exposito *et al*., 2010; Jo *et al*., 2019; Zhu *et al*., 2018). Daily intake of collagen is beneficial for blood pressure and blood lipids (Lugo *et al*., 2016; Sato, 2017) and improves osteoarthritis-related symptoms and skin conditions (Lugo *et al*., 2016; Proksch *et al*., 2014; Liu *et al*., 2019). Dysfunctional collagen promotes pathological fibrosis, such as osteogenesis imperfecta, dysplasia and osteoporosis, aneurysm and osteoarthrosis. In the secretory pathway, collagen can act as a chaperone protein to promote the proper folding of synthetic polypeptides (Shigemura *et al*., 2018).

In the Golgi and endoplasmic reticulum, many intracellular collagens undergo post-translational modifications. In the endoplasmic reticulum, post-translational modification, folding and transport all take place and are important for collagen integrity. Newly translated procollagen chain specific proline and lysine residues are prolyl hydroxylase and lysyl modified (Heard *et al*., 2016). Hydroxylation of collagen and crosslinking can result in loss of connective tissue defect in vivo, including low bone mass and skin fragility (Moskowitz, 2000).

Collagen biosynthesis involves several post-transcriptional modification enzymes. An important member of these enzymes is prolyl-4-hydroxylase (P4H). P4H is a member of the dioxygenase family. These enzymes require α-ketoglutaric acid, Fe^{2+} and O_2 for catalytic activity (Gorres and Raines, 2010; Myllyharju, 2003). During hydroxylation, α-ketoglutaric acid is decarboxylated by oxygen to form succinic acid, with one oxygen atom incorporated into proline to form hydroxyproline, and the other oxygen atom used to synthesize succinic acid. Ascorbic acid functions to maintain the activity of the enzyme. P4H modifies the

* Corresponding author: txiangqi@jnu.edu.cn; yangf@aset.ac.cn
0030-9923/2020/0001-0321 $ 9.00/0

proline residue at the Y position of the Gly-X-Y repeat in collagen to give 4-hydroxyproline (4Hyp). This frequent modification imparts thermal stability to the folded triple-helix structure (Gorres and Raines, 2010; Pozzolini *et al*., 2015a).

In vertebrates, prolyl hydroxylase is an α2β2 heterotetramer (Buechter *et al*., 2003). The α subunit (P4Hα) of the heterotetramer exists as three types, P4Hα1, P4Hα2 and P4Hα3, and each of these α subunits contains a hydroxylated catalytically active site of proline (Stein *et al*., 2009; Myllyharju, 2003). The β subunit (PDI) is a disulfide isomerase containing two thioredoxin domains and has multiple roles in cells. The main role of PDI in the tetramer of rhP4H is to maintain the solubility and activity of the P4Hα and to retain the enzyme in the endoplasmic reticulum (Kersteen and Raines, 2003). Stein *et al*. (2009) have expressed recombinant human P4H in tobacco, whereas Peng *et al*. (2018) expressed type I human collagen and the P4H α-protein in silkworm. Expression of the human P4H tetramer (rhP4H) in Escherichia coli has also been achieved (Pinkas *et al*., 2011). However, using such expression systems in these host organisms for heterologous protein production has a number of issues such as incorrect folding and refolding into the biologically active state, and this can lead to dramatic reductions in recombinant protein yields. These issues are not easy to overcome.

In this report, a *Pichia pastoris* expression system was used to produce rhP4H. The activity of rhP4H was confirmed by UPLC-MS. rhP4He specifically catalyzed the hydroxylation of proline in the Gly-X-Y motif both in the fermentation broth of yeast cells or after isolation and purification. The Gly-X-Y unit is an exclusive structure of human collagen. The results demonstrate that rhP4H hydroxylates a human collagen-derived peptide and should promote functional analysis of P4H and further analysis of how this enzyme maintains the stability of collagen.

MATERIALS AND METHODS

Strains, plasmids, reagents and instrumentation

E. coli strain Top 10 (Takara, Guangzhou, China), vectors pPICZαA and PHIL and *P. pastoris* strain *GS115* (Invitrogen, Guangzhou, China) were used for cloning and heterologous expression. *Eco*RI and *Xba*I enzymes and DNA polymerase were from Dalian Takara (Dalian, China), and PCR purification kits, gel extraction kits and micro preparation kits were purchased from Tiangen (Beijing, China). Primers were composed by Invitrogen. The monoclonal mouse antihuman P4Hα2 antibody was from Abcam (UK) and the monoclonal mouse antihuman PDI antibody was purchased from Genetex (USA). 4 peptide-1 (dansyl-Gly-Phe-Pro-Gly-OEt) and dansyl-Gly-Phe-Hyp-Gly-OEt were from GL Biochem Ltd (Shanghai, China). Palmitoyl-4-peptide-7 (Pal-Gly-Gln-Pro-Arg) was from JiTai (Guangzhou, China). All other chemicals were of the highest quality and used without further purification.

Ultra-Performance Liquid Chromatography (UPLC) and hybrid quadrupole orthogonal time-of-flight (Q-TOF) mass spectrometry were performed with a system from Waters Corporation, Manchester (UK). The CFX96 real-time quantitative PCR detection system was from Bio-Rad (USA).

Construction of human P4Hα2 and PDI expression vectors

Our plan for constructing active P4H tetramers in *P. pastoris* was to clone cDNAs encoding α and β subunits into different plasmids. Both cDNAs were transcribed from the same two kinds of AOX1 promoter, and the two subunits had their own translation initiation ribosome binding site.

The cDNA encoding human PDI was obtained by RT-PCR using primers The PCR fragment was put among the *Eco*RI and *Xba*I sites of the PHIL to get a recombinant plasmid PHIL-PDI.

The cDNA encoding P4Hα2 was obtained by RT-PCR using primers. The PCR fragment was put among the *Eco*RI and *Xba*I sites of the pPICZαA to get a recombinant plasmid pPICZαA-P4Hα2. Finally, the two recombinant plasmids were used to generate full-length P4H tetramers.

Transformation of P. pastoris *GS115*

Plasmid PHIL-PDI used for the transformation was recovered and linearized. The linearized expression vector was transformed into *P. pastoris* GS115 cells by electro-transformation using a Gene Pulser Xcell (Bio-Rad, USA). Positive clones were collected on YPD plates containing 0.1 mg / mL Zeocin.

Plasmid pPICZαA-P4Hα2 was linearized using the same approach as employed for pHIL-PDI. The pPICZαA-P4Hα2 expression vector was transformed into *P. pastoris* P1 by electro-transformation. Positive clones were collected on YPD plates including 0.1, 0.5, 1.5, 2.5 and 3.5 mg/mL of Zeocin. The strains were named 1, 2, 3, 4 and 5.

1, 2, 3, 4 and 5 strains were cultured in 5 mL of YPD medium at 30 °C overnight and 1% of the culture was transmitted to 25 mL of YPG. In a shaking incubator, the cells were cultured at 30 °C till the OD600 reached 1.5. The cell pellets were then induced in YPM. At the same time, methanol (0.5, 1.0, 1.5, 2.0%) was added to the culture and 1-mL samples were collected every 24 h post induction. All collected samples were centrifuged at the maximum speed in a microcentrifuge at room temperature. The supernatant was transferred to a separate tube and

used to analyze the expression levels and determine the optimal time post induction to harvest cells.

SDS-PAGE and western blotting

P4H expression and purification was analyzed by SDS-PAGE and gel staining was carried out with Coomassie Brilliant Blue R-250. Western blot analysis was performed using monoclonal antibodies against human PDI and human P4Hα2.

Real-time quantitative PCR

All the samples of total RNA was collected with the Yeast RNA Kit (Omega) from 1, 2, 3, 4 and 5 bacteria after inducing at 72 h.

A total of 500 ng RNA was used in reverse transcription reactions by using the PrimeScript™ RT Master Mix (TaKaRa). The components were added in the advisable order were as follow: 4 μL of 5* PrimeScript RT Master Mix, 500 ng RNA and RNase-free ddH2O to a total volume of 20 μL. The sample was mixed gently, centrifuged briefly and then the reverse-transcription reaction was performed at 37 °C for 15 min followed by 85 °C for 5 s to inactivate the reverse transcriptase. For determining the RT-qPCR, primers were listed in Table I.

To analyze the relative mRNA expression levels of P4Hα2 and PDI, RT-qPCR reactions were performed in a 96-well plate that contained 10 μL of Maxima SYBR Green qPCR Master Mix, 0.5 μL of primer, 2 μL of cDNA (diluted) and 7 μL of deionized water. The thermal cycling program using the CFX96 real-time quantitative PCR detection system was: 95 °C for 3 min, 40 cycles of 95 °C for 5 s and 60 °C for 30 s. From the fluorescent values obtained through the RT-qPCR analysis, the threshold cycles (CT value) were obtained.

All sample data were analyzed by the $2^{-\Delta\Delta CT}$ method. For this method, the His3 gene of *P. pastoris* was used as the internal reference gene.

Table I. Primers for determining the RT-qPCR.

Name	Primer sequence
PDIF	5'-AGGCTGATGACATCGTGAACT-3'
PDIR	5' GGTATTTGGAGAACACGTCACTG-3'
P4Hα2F	5'-CAAACTGGTGAAGCGGCTAAA-3'
P4Hα2R	5'-GCACAGAGAGGTTGGCGATA-3'
HISF	5'-TCAGTAGTGCTGTTAGCGGCGACAT-3'
HISR	5'-ACTTAAGGTCACAGTGCGTTCCGGTAT-3'

Purification of rhP4H

Ammonium sulfate fractionation purified the P4H crude enzyme solution. Ammonium sulfate between 20% and 70% was added to the crude enzyme solution. The precipitant was collected by centrifugation at 12,000 rpm for 30 min and resuspended in 100 mM glycine, 10 mM Tris and 100 mM NaCl, pH 7.8. SDS-PAGE analyzed the 20%–70% ammonium sulfate fractions.

Amino acid analysis of P4Hα2 and PDI

The relative molecular mass of the recombinant P4Hα2 and PDI monomers were identified by mass spectrometry. HPLC-based peptide mapping confirmed the identity. All samples were analyzed by Bio-Tech Pack Technology Company Ltd (Beijing, China).

Assays of rhP4H enzymatic activity

Based assay was developed to monitor UPLC-MS rhP4H hydroxylation of proline in vitro. Experiments were performed at 30 °C in 1 mL of 50 mM Tris-HCl buffer (pH = 7.8) containing 1 mg/mL bovine serum album (BSA), 100 U/mL catalase, 0.1 mM dithiothreitol (DTT), 0.1 mM $FeSO_4 \cdot 7H_2O$, 0.5 mM ketopentane acid and 2 mM ascorbic acid. The tetra-peptide substrate (4 peptide-1 or palmitoyl 4 peptide-7, 1 mM stock in methanol) or monomeric L-proline (1 mM) was used to initiate the reaction. The reaction was incubated at 30 °C overnight and then boiled in a water bath for 1 min.

Based assay was developed to monitor UPLC-MS rhP4H hydroxylation of proline in vivo. The *P. pastoris GS115* strain containing the recombinant plasmids P4Hα2 and PDI was induced to express in YPM medium, which included 4 peptide-1 (Dansyl-Gly-Phe-Pro-Gly-OEt), palmitoyl 4 peptide-7 (Pal-Gly-Gln-Pro-Arg) or L-proline. Ascorbic acid and Fe2+ were added to the YPM medium every 24 h. The supernatant was collected 72 h after induction.

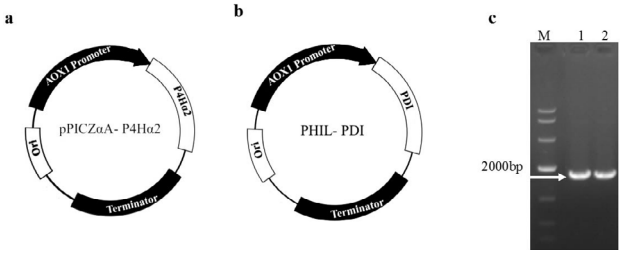

Fig. 1. Schematic diagram of the pPICZαA-P4Hα2 and PHIL-PDI constructs. (a) Construction of the expression plasmid pPICZαA-P4Hα2. (b) Construction of the expression plasmid PHIL-PDI. (c) Nucleic acid electrophoresis of recombinant plasmid pPICZαA-P4Hα2 and PHIL-PDI. M: DNA Ladder 10000; lane 1: the monoclones of recombinant plasmid pPICZαA-P4Hα2; lane 2: the monoclones of recombinant plasmid PHIL-PDI.

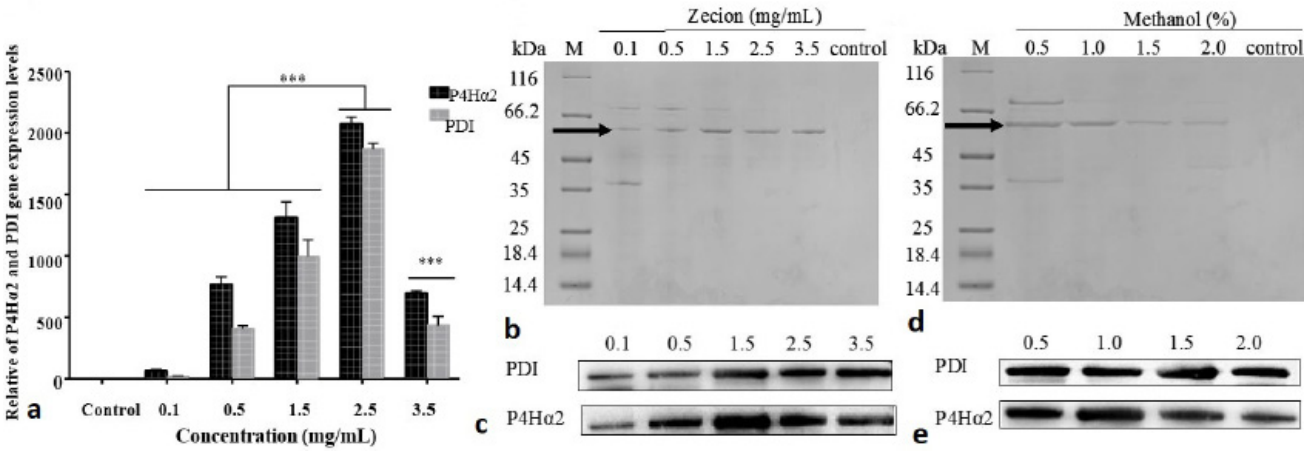

Fig. 2. Analysis of the expression of rhP4H in *P. pastoris* GS115. The apparent molecular mass of rhP4H is ~60 kDa. (a) Detection of P4Hα2 and PDI at the mRNA level at different concentrations of Zecion. (b) SDS-PAGE analysis of the expression of rhP4H selected by 0.1, 0.5, 1.5, 2.5 and 3.5 mg/mL Zecion, respectively. M: middle molecular weight protein markers. (c) Western blotting analysis of PDI and P4Hα2 expression at different concentrations of Zecion. (d) SDS-PAGE analysis of the expression of rhP4H induced by 0.5, 1.0, 1.5, 2.0% methanol, respectively. M: middle molecular weight protein markers. (e) Western blotting analysis of the expression of PDI and P4Hα2 at different concentrations of methanol.

All of the above sample analyses were performed on an ACQUITY™ UPLC I-Class System (Waters Corporation) equipped with a binary solvent system, an automated sample manager and a photodiode array (PDA) detector. The separation was carried out on an ACQUITY UPLC™ BEH C18 column (3.0 × 150 mm, 1.7 µm, Waters Corporation) at a temperature of 35 °C. The UPLC system was then connected to Q-TOF mass spectrometer equipped with electrospray ionization. All experimental data were analyzed by Masslynx™ 4.1 software.

Statistical analysis

Data are reported as mean ± SD (n = 3). Groups were compared in Graph Pad Prism 7 (GraphPad Software Inc., La Jolla, CA, USA) by one-way analysis of variance followed by Tukey's test. P values < 0.05 were considered statistically significant.

RESULTS

Vector construction

The P4Hα2 and PDI genes were obtained from human placenta by RT-PCR. Figure 1A shows the P4Hα2 gene insert into pPICZαA vector, which was digested with restriction enzymes EcoRI and XbaI, and the PDI gene was inserted into the PHIL vector in the same manner as the P4Hα2 gene (Fig. 1B). The results from nucleic acid electrophoresis showed that the two genes were ligated into the expression vectors successfully (Fig. 1C) and DNA sequencing confirmed that the correct target genes were present in the vectors.

SDS-PAGE and western blot analysis

The strains were screened with different concentrations of Zeocin and induced for 72 h to collect the yeast supernatant. Figure 2 shows the SDS-PAGE results of the yeast secreted supernatant. Because the molecular weight of P4Hα2 and PDI are similar, the expression of the two proteins were confirmed by western-blot (Fig. 2b). All strains were found to express P4Hα2 and PDI. The relative expression of the proteins by the strains was highest when the concentration of Zeocin-resistance was 2.5 mg/mL. Subsequently, four strains were induced with different methanol concentrations. The protein expression levels were analyzed by SDS-PAGE and western blotting. We found that the expression levels of P4HA2 were higher than PDI after 72 h induction with 1% methanol.

RT-qPCR

cDNA samples from the control and sample groups were subjected to RT-qPCR. The $2^{-\Delta\Delta CT}$ method was used to calculate the relative expression, and the His 3 gene of *P. pastoris* GS115 without the expression vectors pPICZαA-P4Hα2 and PHIL-PDI was used as a control gene to calculate the relative P4Hα2 and PDI gene. The results of RT-qPCR (Fig. 2a) and western blotting showed that the expression of the target genes at different concentrations of Zeocin-resistance varied (Fig. 2c). The relative expression levels of mRNA were highest when the concentration of Zeocin-resistance was 2.5 mg/mL, and the expression level of P4Hα2 was higher than that of PDI.

Table II. Mass Spectrum identification information statistics of PDI and P4Hα2.

Enzymatic	Total number of spectra	Identification number of spectra	Identification of peptides	Protein coverage	Protein scoring
PDI	14399	551	70	92.5%	323.31
P4Hα2	14120	509	28	62.9%	323.31

Table III. Activity spectrum of rhP4H, P4Hα2 or PDI with different substrates both in supernatant of *pastoris* extract and fermentation broth of yeast cells.

Samples	Substrates			Product		
	Dansyl-Gly-Phe-Pro-Gly-OEt	Pal-Gly-Gln-Pro-Arg	L-Pro	Dansyl-Gly-Phe-Hyp-Gly-OEt	Pal-Gly-Gln-Hyp-Arg	L-Hyp
rhP4H	✓	✓	✓	✓	✓	-
P4Hα2	✓	✓	✓	-	-	-
PDI	✓	✓	✓	-	-	-

"✓" indicated that products or substrates were detected; "-" indicated that products were undetected.

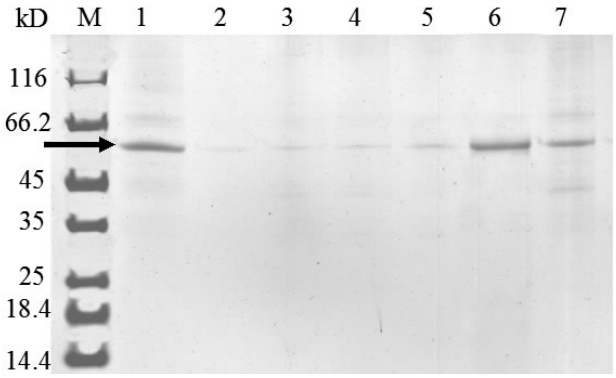

Fig. 3. SDS-PAGE analysis of the expression of rhP4H. rhP4H is present in the various ammonium sulfate fractions. M: middle molecular weight protein markers; lane 1: supernatant sample of rhP4H after induction; lanes 2–7: reconstituted sample of rhP4H after precipitation with 20%, 30%, 40%, 50%, 60% and 70% ammonium sulfate, respectively.

Purification and amino acid analysis of rhP4H

The highest yield of partially purified P4H was achieved when extracted with 60% saturated ammonium sulfate. The purity of the protein is relatively high. Amino acid analysis was carried out and the results of this analysis were confirmed by MALDI-TOF MS peptide mapping (Table II).

Catalysis of rhP4H

Functional verification of rhP4H produced in *P. pastoris in vitro* was performed by measuring the hydroxylation process of L-proline, 4 peptide-1 (dansyl-Gly-Phe-Hyp-Gly-OEt) or palmitoyl 4 peptide-7 (Pal-Gly-Gln-Pro-Arg). The two peptides interacted with P4H at 30 °C, and the mass spectrometry results were consistent with the previously published results of recombinant P4H produced in E. coli. Functional verification of the full-length P4H enzyme secreted by *P. pastoris in vivo* was performed by using the same approach described above. The two peptides were added to the medium during the process of inducing secretion of P4H, and the results were in agreement with the fermentation broth of yeast cells in the former results.

UPLC-MS-based assay for enzymatic activity

UPLC-MS-based assays were developed to measure the activity of the P4H enzyme in vitro and in vivo. The method is based on the previously disclosed substrate, and the substrate (dansyl-Gly-Phe-Pro-Gly-OEt) was separated from its hydroxylated product (dansyl-Gly-Phe-Hyp-Gly-OEt). UPLC-MS confirmed the status of each peak.

In the process of inducing expression of rhP4H, 4 peptide-1 (Dansyl-Gly-Phe-Pro-Gly-OEt) was added to the medium. All samples were then collected and analyzed by UPLC after 72 h (Fig. 4). rhP4H produced by *P. pastoris* was found to play a role in hydroxylation. At the same time, the fragment [M+H] peak of 4 peptide-1 was detected in the reaction system by the same detection method used for the in vitro reaction.

Similarly, we also used the same method when palmitoyl 4 peptide-7 (Pal-Gly-Gln-Pro-Arg) was added (Fig. 5) and fragment peaks of the product were detected by UPLC. However, when L-proline was used as the substrate, no fragment peaks of the expected product were detected by UPLC (Table III). This result is consistent with a previous report that rhP4H does not hydroxylate L-proline.

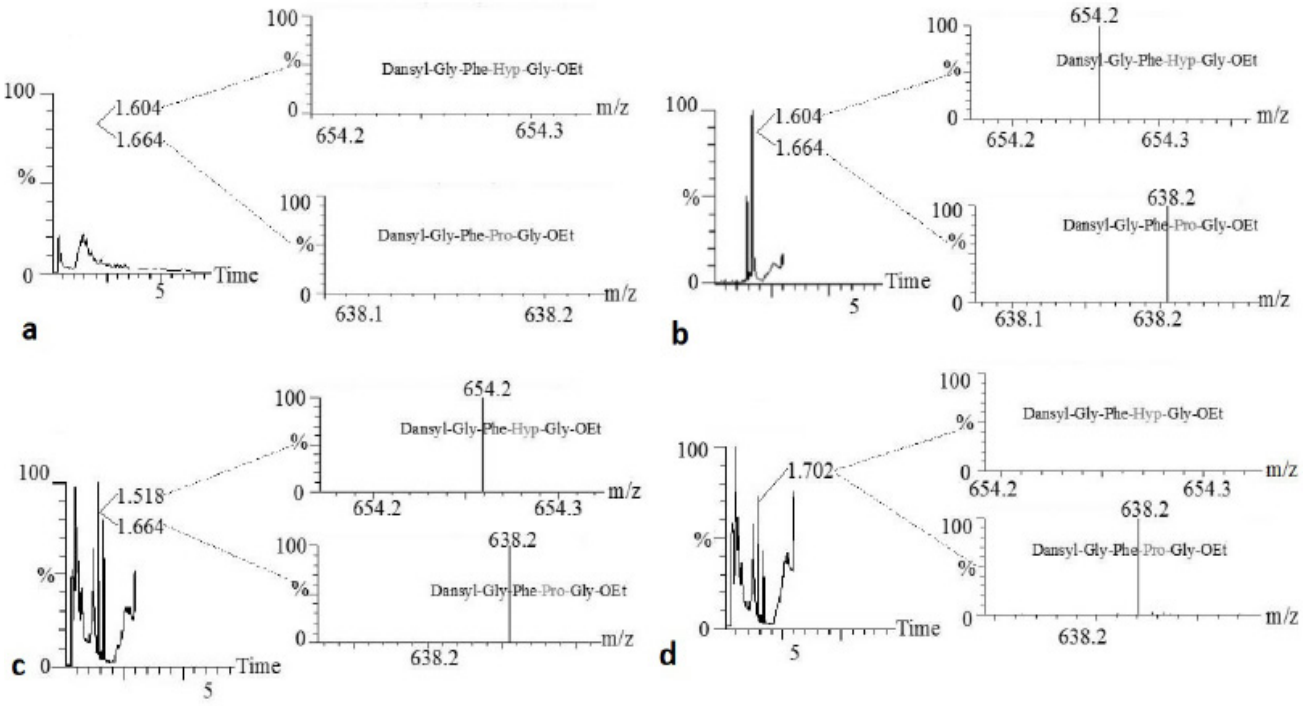

Fig. 4. Activity spectrum of rhP4H with Dansyl-Gly-Phe-Pro-Gly-OEt as the substrate in supernatant of pastoris extract or in fermentation broth of yeast cells. (a) Blank control. (b) Substrate and product standard. (c)The activity identification mass spectrum of rhP4H. (d) The activity identification mass spectrum of PDI.

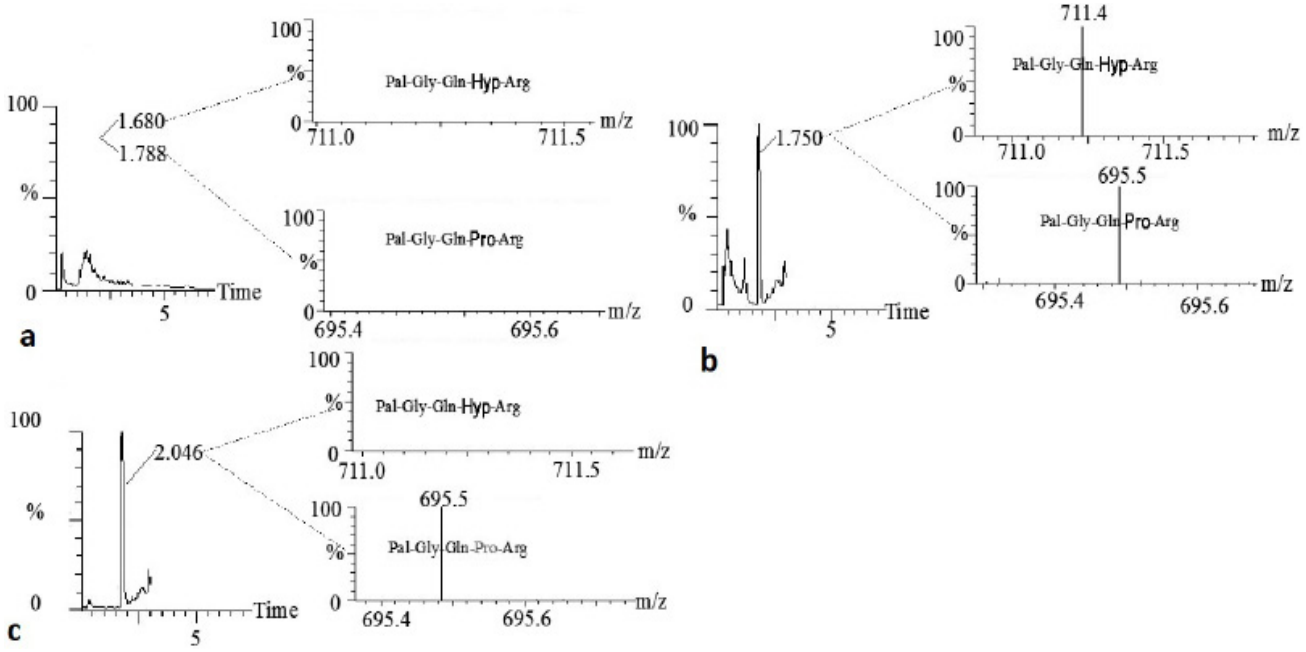

Fig. 5. Activity spectrum of rhP4H with Pal-Gly-Gln-Pro-Arg as the substrate in supernatant of pastoris extract or in fermentation broth of yeast cells. (a) Blank control. (b) The activity identification mass spectrum of rhP4H. (c) The activity identification mass spectrum of PDI.

DISCUSSION

Collagen is a common biological material used in a variety of applications; however, there are many limitations with its use in industry, including weak processing ability, poor hydrolysis properties, low homology and strong immunogenicity (Saito and Marumo, 2015; Shoulders and Raines, 2009). Currently, commercial collagens are taken from animal tissues and skin extracts. Collagen taken from animals and used in applications can cause zoonotic diseases. In addition, the method to extract collagen from animal tissue has disadvantages, including poor sustainability and the process is not environmentally friendly (Sewing *et al*., 2017). Therefore, identifying safe, effective and bioactive collagen material has received increasing attention (Saito and Marumo, 2015). In recent years, the use of genetic engineering to produce collagen can effectively overcome the abovementioned drawbacks (Lee *et al*., 2001; He *et al*., 2015). However, because of the complex triple-helix structure of collagen, production of recombinant collagen is associated with a series of challenges. In particular, maintaining the stability of the collagen triple-helix is a hot topic.

The formation and stability of the triple-helix conformation of collagen are related to hydroxyproline. Hydroxyproline in mature collagen is produced by the action of prolyl-4-hydroxylase and is a post-translational modification process (Sipilä *et al*., 2018; Song *et al*., 2017; Walker *et al*., 2017). Human collagen produced by genetic engineering lacks hydroxyproline because the post-transcriptional modification process does not occur, which affects the formation of the normal triple-helix conformation (Kim *et al*., 2014). Moreover, the natural physical and chemical properties and biological activity of recombinant collagen are also significantly affected

rhP4H is an α2β2 isomeric tetramer in which the α subunit has catalytic activity and the β subunit retains the α subunit in the catalytically active, non-aggregated form. The complex assembly mechanisms of the α and β subunits may affect the amount produced and activity of P4H. Current studies have identified that human tissue contains three α subunits, P4Hα1, P4Hα2 and P4Hα3. Most cell types exist P4Hα1, whereas P4Hα2 mainly exists in chondrocytes, osteoblasts and capillary endothelial cells and P4Hα3 is present in relatively low amounts in adult and fetal tissues (John and Bulleid, 1996; Neubauer *et al*., 2005). Studies had shown that P4Hα2 is associated with the expression of type I, III and IV of collagen. Kersteen *et al.* and Neubauer *et al.* have shown that the monomeric form of P4H can hydroxylate collagen and collagen peptides to form hydroxyproline to ensure the stability of the triple-helix structure (Kersteen *et al*., 2004; Kersteen and Raines, 2003). In this report, P4Hα2 and PDI expression systems were constructed in separate plasmids to produce an active human P4H tetramer (rhP4H) in *P. pastoris*. rhP4H with only the α2 and β subunits can hydroxylate proline in either the 4 peptide-1 (dansyl-Gly-Phe-Hyp-Gly-OEt) or palmitoyl 4 peptide-7 (Pal-Gly-Gln-Pro-Arg) substrates.

Compared with the *E. coli* expression system (Kersteen *et al*., 2004; Neubauer *et al*., 2007), the recombinant *P. pastoris* eukaryotic expression system has several advantages: rapid growth rate, bright genetic background, post-translational modification capabilities and non-pathogenic traits. Such features ensure heterologous proteins are correctly folded and assemble into native conformations (Li *et al*., 2018). In this report, rhP4H was expressed in vitro using a *P. pastoris* system. In a past report, P4Hα2 and PDI were simultaneously constructed on the same expression vector to induce expression of the two subunits using separate promoters (Pozzolini *et al*., 2015b). In this study, we constructed the two subunits in separate expression vectors that contained the same promoter. We then separately transformed the two recombinant plasmids into *P. pastori*s for co-expression of P4Hα2 and PDI, which reduced the complexity of protein expression and rhP4H was obtained at higher purity. Expression conditions were optimized to reduce challenges associated with subsequent purification. Structural properties of rhP4H produced by genetically engineered yeast strains can be controlled by modulating the degree of hydroxylation. This enables particular properties of recombinant collagen to be engineered for specific end uses, for example, for incorporation into a variety of non-residual and green bio-manufactured medical materials (Bazrafshan and Stylios, 2019; Koski *et al*., 2017).

The traditional P4H catalytic activity assay monitors the release of [^{14}C] CO_2 from [1-^{14}C] α-ketoglutarate. Kersteen *et al.* developed a detection method based on HPLC, monitored for hydroxylation of proline P4H (Kersteen *et al*., 2004). The study used a traditional P4H radioisotopes to confirm the new HPLC-based assay's effectiveness. According to this report with minor changes, we applied a UPLC-QTOF-based hydroxylation assay to monitor P4H catalytic activity. We found that rhP4H was not only active in the supernatant of the *P. pastoris* extract, but also active in fermentation broth of yeast cells, indicating both the α subunit and PDI of rhP4H were co-expressed in *P. pastoris* to produce an active recombinant enzyme. Clearly, *P. pastoris* not only contains the appropriate chaperone protein and conditions required for assembly of active rhP4H, but also a system for transporting all of the co-substrates required for the enzyme into the lumen of the endoplasmic reticulum. However, when either the α subunit or PDI was expressed,

there was no prolyl-4-hydroxylase activity observed, which supports previous studies that showed that the α subunit in the absence of PDI forms inactive aggregates. Moreover, lack of P4Hα, the PDI did not recognize collagen and collagen peptides (Hieta *et al*., 2003; Koski *et al*., 2017; Vasta and Raines, 2018).

Interestingly, we found that the enzymatic reaction could not be carried out in the absence of ascorbic acid, proving that ascorbic acid is related to the catalysis of P4H. As previously reported, P4H can catalyze the decarboxylation of α-ketoglutaric acid, resulting in uncoupling of the co-substrate conversion to the inactivate enzyme. Ascorbic acid can rescue inactivated P4H by reducing Fe^{3+} to the active Fe^{2+} form (Kivirikko1 *et al*., 1997; Tuderman *et al*., 1977; Pinkas *et al*., 2011).

Besides P4Hs being found in animals, they are also found in plants and microbes. Compared with the rhP4H, the plant P4H substrate specificity is quite different. Plant P4Hs mainly use poly-L-proline or a similar sequence as the substrate, rather than human collagen with the Gly-X-Y motif. Though in conotoxins, the hydroxylated prolines are not part of any consensus sequence. Plant P4Hs hydroxylate a sequence rich in proline residues and have a variety of repeat motifs, and bacteria are unique in that they hydroxylate proline to free amino acids. The activity of rhP4H was also tested against L-proline both in *P. pastoris* extracts and in the yeast cell. According to the UPLC-MS results, rhP4H did not catalyze the conversion of L-proline to hydroxyproline, proving that rhP4H only hydroxylates peptides, which contain Gly-X-Y repeat units.

With the rapid development of tissue engineering, researchers are focusing increasingly more on biomedical materials. Collagen has moderate antigenic activity, low irritation and low cytotoxicity, and is therefore an ideal biomedical material and is used widely in dressings, artificial organs and regenerative medicine (Bazrafshan and Stylios, 2019; Myllyharju, 2004). The expression of recombinant human collagen has made significant progress whose material is used in specific applications, such as the production of bioengineered corneal substitutes. The production of collagen having a stable triple helix structure requires hydroxyproline depending on the activity of prolyl-4-hydroxylase. Realizing to express and manipulate signals biologically in recombinant collagens to meet the demands of biomaterials and biomedical needs is very important, which constitutes a creative foundation for collagen applications.

ACKNOWLEDGEMENT

This work was supported by grants from the Guangzhou Science and Technology Program key project [grant number 201803010044], the Science and Technology Program of Tianhe District, Guangzhou City [grant number 201704YG066], and the Major Scientific and Technological Special Project of the Administration of Ocean and Fisheries of Guangdong Province [Yuecainong, 2017, no. 17, grant number GDME-2018C013].

Statement of conflict of interest

There are no conflicts of interest to declare.

REFERENCES

Bazrafshan, Z. and Stylios, G.K., 2019. Spinnability of collagen as a biomimetic material: A review. *Int. J. biol. Macromol.,* **129:** 693-705. https://doi.org/10.1016/j.ijbiomac.2019.02.024

Buechter, D.D., Paolella, D.N., Leslie, B.S., Brown, M.S., Mehos, K.A. and Gruskin, E.A., 2003. Co-translational incorporation of trans-4-hydroxyproline into recombinant proteins in bacteria. *J. biol. Chem.*, **278**: 645-650. https://doi.org/10.1074/jbc.M209364200

John, D.C. and Bulleid, N.J., 1996. Intracellular dissociation and reassembly of prolyl 4-hydroxylase: The α-subunits associate with the immunoglobulin-heavy-chain binding protein (BiP) allowing reassembly with the β-subunit. *Biochem. J.*, **317**: 659-665. https://doi.org/10.1042/bj3170659

Lee, C.H., Singla, A. and Lee, Y.Y., 2001. Biomedical applications of collagen. *Int. J. Pharmaceut.,* **221**: 1-22. https://doi.org/10.1016/S0378-5173(01)00691-3

Exposito, J.Y., Valcourt, U., Cluzel, C. and Lethias, C., 2010. The fibrillar collagen family. *Int. J. mol. Sci.*, **11**: 407-426. https://doi.org/10.3390/ijms11020407

Gorres, K.L. and Raines, R.T., 2010. Prolyl 4-hydroxylase. *Crit. Rev. Biochem. mol. Biol.*, **45**: 106-124. https://doi.org/10.3109/10409231003627991

He, J., Ma, X.X., Zhang, F.L., Li, L.B., Deng, J.J., Xue, W.J., Zhu, C.H. and Fan, D.D., 2015. New strategy for expression of recombinant hydroxylated human collagen α1(III) chains in *Pichia pastoris* GS115. *Biotechnol. appl. Biochem.*, **62**: 293-299. https://doi.org/10.1002/bab.1264

Heard, M.E., Besio, R., Weis, M., Rai, J., Hudson, D.M., Dimori, M., Zimmerman, S.M., Kamykowski, J.A., Hogue, W.R., Swain, F.L., Burdine, M.S., Mackintosh, S.G., Tackett, A.J., Suva, L.J., Eyre, D.R. and Morello, R., 2016. Sc65-Null mice provide evidence for a novel endoplasmic reticulum complex regulating collagen lysyl hydroxylation. *PLoS Genet.*, **12**: e1006002. https://doi.org/10.1371/

journal.pgen.1006002

Hieta, R., Kukkola, L., Permi, P., Pirilä, P., Kivirikko, K.I., Kilpeläinen, I. and Myllyharju, J., 2003. The peptide-substrate-binding domain of human collagen Prolyl 4-Hydroxylases. *J. biol. Chem.*, **278**: 34966-34974. https://doi.org/10.1074/jbc.M303624200

Kivirikko, K.I. and Myllyharju, J., 1997. Characterization of the iron- and 2- oxoglutarate-binding sites of human prolyl 4-hydroxylase. *The EMBO J.,* **16**: 1173-1180. https://doi.org/10.1093/emboj/16.6.1173

Jo, Y.J., Karbstein, H.P. and van der Schaaf, U.S., 2019. Collagen peptide-loaded W1/O single emulsions and W1/O/W2 double emulsions: influence of collagen peptide and salt concentration, dispersed phase fraction and type of hydrophilic emulsifier on droplet stability and encapsulation efficiency. *Fd. Funct*, **10**: 3312-3323. https://doi.org/10.1039/C8FO02467G

Kersteen, E.A., Higgin, J.J. and Raines, R.T., 2004. Production of human prolyl 4-hydroxylase in *Escherichia coli*. *Protein Expr. Purif.,* **38**: 279-291. https://doi.org/10.1016/j.pep.2004.09.008

Kersteen, E.A. and Raines, R.T., 2003. Catalysis of protein folding by protein disulfide isomerase and small-molecule mimics. *Antioxid. Redox Signal*, **5**: 413-424. https://doi.org/10.1089/152308603768295159

Kim, H.K., Kim, M.G. and Leem, K.H., 2014. Collagen hydrolysates increased osteogenic gene expressions via a MAPK signaling pathway in MG-63 human osteoblasts. *Fd. Funct.,* **5**: 573-578. https://doi.org/10.1039/c3fo60509d

Koski, M.K., Anantharajan, J., Kursula, P., Dhavala, P., Murthy, A.V., Bergmann, U., Myllyharju, J. and Wierenga, R.K., 2017. Assembly of the elongated collagen prolyl 4-hydroxylase α2β2 heterotetramer around a central α2 dimer. *Biochem. J.,* **474**: 751-769. https://doi.org/10.1042/BCJ20161000

Tuderman, L., Myllylä, R. and Kivirkko, K.I., 1977. Mechanism of the prolyl hydroxylase reaction. *Eur. J. Biochem,* **80**: 341-348. https://doi.org/10.1111/j.1432-1033.1977.tb11888.x

Li, X., Yang, Y.K., Zhan, C.J., Zhang, Z.Y., Liu, X.X, Liu, H.B. and Bai, Z.H., 2018. Transcriptional analysis of impacts of glycerol transporter 1 on methanol and glycerol metabolism in *Pichia pastoris. FEMS Yeast Res.,* **18**: 1-14. https://doi.org/10.1093/femsyr/fox081

Liu, Y., Liu, B.M., Li, C.C. and An, M.M., 2019. Regulation of collagen metabolism in keratinocyte-fibroblast organotypic co-culture by pressure. *Pakistan J. Zool.,* **51**: 21-29. http://dx.doi.org/10.17582/journal.pjz/2019.51.1.21.29

Lugo, J.P., Saiyed, Z.M. and Lane, N.E., 2016. Efficacy and tolerability of an undenatured type II collagen supplement in modulating knee osteoarthritis symptoms: a multicenter randomized, double-blind, placebo-controlled study. *Nutr. J.*, **15**: 1-15. https://doi.org/10.1186/s12937-016-0130-8

Moskowitz, R.W., 2000. Role of collagen hydrolysate in bone and joint disease. *Semin. Arthritis Rheum.*, **30**: 87-99. https://doi.org/10.1053/sarh.2000.9622

Myllyharju, J., 2003. Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biol.*, **22**: 15-24. https://doi.org/10.1016/S0945-053X(03)00006-4

Myllyharju, J., 2004. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet.,* **20**: 33-43. https://doi.org/10.1016/j.tig.2003.11.004

Neubauer, A., Neubauer, P. and Myllyharju, J., 2005. High-level production of human collagen prolyl 4-hydroxylase in *Escherichia coli. Matrix Biol.,* **24**: 59-68. https://doi.org/10.1016/j.matbio.2004.11.004

Neubauer, A., Soini, J., Bollok, M., Zenker, M., Sandqvist, J., Myllyharju, J. and Neubauer, P., 2007. Fermentation process for tetrameric human collagen prolyl 4-hydroxylase in *Escherichia coli*: Improvement by gene optimisation of the PDI/β subunit and repeated addition of the inducer anhydrotetracycline. *J. Biotechnol.,* **128**: 308-321. https://doi.org/10.1016/j.jbiotec.2006.10.017

Peng, Y.Y., Nebl, T., Glattauer, V., Ramshaw and J.A.M., 2018. Incorporation of hydroxyproline in bacterial collagen from *Streptococcus pyogenes. Acta Biomat.,* **80**: 169-175. https://doi.org/10.1016/j.actbio.2018.09.012

Pinkas, D.M., Ding, S., Raines, R.T. and Barron, A.E., 2011. Tunable, post-translational hydroxylation of collagen domains in *Escherichia coli*. *ACS Chem. Biol.*, **6**: 320-324. https://doi.org/10.1021/cb100298r

Pozzolini, M., Scarfì, S., Mussino, F., Ferrando, S., Gallus, L. and Giovine, M., 2015a. Molecular cloning, characterization, and expression analysis of a prolyl 4-hydroxylase from the marine sponge *Chondrosia reniformis*. *Mar. Biotechnol.*, **17**: 393-407. https://doi.org/10.1007/s10126-015-9630-3

Pozzolini, M., Scarfì, S., Mussino, F., Salis, A., Damonte, G., Benatti, U. and Giovine, M., 2015b. *Pichia pastoris* production of a prolyl 4-hydroxylase derived from *Chondrosia reniformis* sponge: A new biotechnological tool for the recombinant production of marine collagen. *J. Biotechnol.,* **208**: 28-36. https://doi.org/10.1016/j.jbiotec.2015.05.007

Proksch, E., Segger, D., Degwert, J., Schunck, M., Zague,

V. and Oesser, S., 2014. Oral supplementation of specific collagen peptides has beneficial effects on human skin physiology: a double-blind, placebo-controlled study. *Skin Pharmacol. Physiol.,* **27**: 47-55. https://doi.org/10.1159/000351376

Saito, M. and Marumo, K., 2015. Effects of collagen crosslinking on bone material properties in health and disease. *Calcif. Tissue Int.,* **97**: 242-261. https://doi.org/10.1007/s00223-015-9985-5

Sato, K., 2017. The presence of food-derived collagen peptides in human body-structure and biological activity. *Fd. Function*, **8**: 4325-4330. https://doi.org/10.1039/C7FO01275F

Sewing, J., Klinger, M. and Notbohm, H., 2017. Jellyfish collagen matrices conserve the chondrogenic phenotype in two- and three-dimensional collagen matrices. *J. Tissue Engineer. Regener. Med.*, **11**: 916-925. https://doi.org/10.1002/term.1993

Shigemura, Y., Suzuki, A., Kurokawa, M., Sato, Y. and Sato, K., 2018. Changes in composition and content of food-derived peptide in human blood after daily ingestion of collagen hydrolysate for 4 weeks. *J. Sci. Fd. Agric.*, **98**: 1944-1950. https://doi.org/10.1002/jsfa.8677

Shoulders, M.D. and Raines, R.T., 2009. Collagen structure and stability. *Annu. Rev. Biochem.*, **78**: 929-958. https://doi.org/10.1146/annurev.biochem.77.032207.120833

Sipilä, K.H., Drushinin, K., Rappu, P., Jokinen, J., Salminen, T.A., Salo, A.M., Käpylä, J., Myllyharju, J. and Heino, J., 2018. Proline hydroxylation in collagen supports integrin binding by two distinct mechanisms. *J. biol. Chem.*, **293**: 7645-7658. https://doi.org/10.1074/jbc.RA118.002200

Song, X., Zhu, C.H., Fan, D.D., Mi, Y., Li, X., Fu, R.Z., Duan, Z.G., Wang, Y. and Feng, R.R., 2017. A novel human-like collagen hydrogel scaffold with porous structure and sponge-like properties. *Polymers,* **9**: 2-16. https://doi.org/10.3390/polym9120638

Stein, H., Wilensky, M., Tsafrir, T., Rosenthal, M., Amir, R., Avraham, T., Ofir, K., Dgany, O., Yayon, A. and Shoseyov, O., 2009. Production of bioactive, post-translationally modified, heterotrimeric, human recombinant Type-I collagen in transgenic tobacco. *Biomacromolecules*, **10**: 2640-2645. https://doi.org/10.1021/bm900571b

Vasta, J.D., Raines and R.T., 2018. Collagen prolyl 4-hydroxylase as a therapeutic target. *J. med. Chem.*, **61**: 10403-10411. https://doi.org/10.1021/acs.jmedchem.8b00822

Walker, K.T., Nan, R., Wright, D.W., Gor, J., Bishop, A.C., Makhatadze, G.I., Brodsky, B. and Perkins, S.J., 2017. Non-linearity of the collagen triple helix in solution and implications for collagen function. *Biochem. J.,* **474**: 2203-2217. https://doi.org/10.1042/BCJ20170217

Zhu, C.H., Lei, H., Wang, S.S., Duan, Z.G., Fu, R.Z., Deng, J.J., Fan, D.D., Lv, X.Q., 2018. The effect of human-like collagen calcium complex on osteoporosis mice. *Mater. Sci. Eng. C. Mater. Biol. Appl.*, **93**: 630-639. https://doi.org/10.1016/j.msec.2018.08.011