JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

ARTICLE

pubs.acs.org/JAFC

New Strategy for Expression of Recombinant Hydroxylated Human-Derived Gelatin in *Pichia pastoris* KM71

Huiming Duan, Sirajo Umar, Runsong Xiong, and Jinchun Chen*

College of Life Science and Technology, Beijing University of Chemical Technology, Beijing 100029, People's Republic of China

ABSTRACT: Gelatin is a well-known biopolymer, and it has a long history of use mainly as a gelling agent in the food industry. This paper reports a new method for producing recombinant hydroxylated human-derived gelatin in *Pichia pastoris* KM71. Three independent expression cassettes encoding for specific length of gelatin, prolyl 4-hydroxylase (P4H, EC 1.14.11.2), α-subunit (αP4H), and protein-disulfide isomerase (PDI) were individually cloned in one expression vector, pPIC9K. The modified gelatin gene and two subunit genes of P4H were under the control of two different inducible promoters, namely, alcohol oxidase 1 promoter (PAOX1) and formaldehyde dehydrogenase 1 promoter (PFLD1), respectively. The results of sodium dodecylsulfate-polyacrylamide gel electrophoresis show that a recombinant gelatin was successfully expressed in *P. pastoris* KM71 by methanol induction. Liquid chromatography coupled with tandem mass spectrometry analysis indicates that the expressed gelatin was hydroxylated with approximately 66.7% of proline residues in the Y positions of Gly-X-Y triplets. The results of nuclear magnetic resonance spectroscopy of recombinant gelatin test show that the ^{1}H and ^{13}C spectra have many corresponding characteristic displacement peaks, and amino acids composition analysis shows that it contains hydroxyproline and its UV absorption is consistent with the characteristics of gelatin.

KEYWORDS: recombinant gelatin, *Pichia pastoris* KM71, coexpression, prolyl 4-hydroxylase, hydroxylation

INTRODUCTION

Gelatin is denatured collagen that is widely used as a gelling agent in pharmaceuticals, photography, and cosmetics, and particularly in the food industry, because of its unique properties. The current methods of making gelatin involve the extraction of collagen from animal tissue (skin and bone of bovine or porcine) and its conversion to gelatin, which is later isolated by acid, base, or enzymatic extraction methods.[1] However, the products obtained through these processes are heterogeneous mixtures of polypeptides with different sizes and charges, conditions that affect the products' gel-forming capacity. In addition to the risk of associated infectious diseases, such as bovine spongiform encephalopathy, animal-derived gelatin may cause immune hypersensitivity when consumed by humans. On the contrary, production of human gelatin in a microbial expression system may eliminate the aforementioned problems because the size and charge of human gelatin can be easily controlled. Moreover, scaling up production of the desired product by DNA manipulation and bioengineering-related techniques in microbial expression system is feasible. These advantages of recombinant gelatin have attracted increasing interest from research and industrial circles.

All gelatins share the most common structural feature of their ancestral source (type I collagen) of a triple-helical domain, which consists of three intertwined polypeptide chains.[2,3] Collagen is made up of peptide triplets, Gly-X-Y, where X and Y can be any of the amino acids in theory. However, proline has a preference for the X position, whereas hydroxyproline has a preference for the Y position. Biosynthesis of triple-helical proteins involves a complex pathway, including many post-translational processing events.[4] A key modification is the hydroxylation of proline residues in the Y position of Gly-X-Y triplets by the enzyme P4H to produce 4-hydroxyproline,[3] which is essential for gel formation, gel-forming capacity of gelatin, and proper folding of procollagen.[5] Hydroxyproline strongly affects the solution behavior of recombinant gelatins. It prevents the physical cross-linking of recombinant gelatin molecules in aqueous solution and, consequently, the temperature-dependent sol−gel transition.[6]

Werten et al.[7] discussed the utility of *P. pastoris* as an expression system for recombinant gelatin. Fragments of rat type III and mouse type I collagen ranging from 21 to 74 kDa in size were expressed and secreted into the extracellular media by transfected *P. pastoris* strains.[7] *P. pastoris* does not contain an endogenous active P4H; thus, the obtained gelatins were non-hydroxylated segments of collagen α-chains, which were essentially in a random coil conformation, and the obtained gelatins were nongelling and susceptible to proteolytic degradation.[7] This result indicates that coexpression of P4H with gelatin genes is necessary for the production of functional recombinant gelatin with physical and chemical properties similar to those of the animal-derived type.

To our knowledge, no study on the production of recombinant hydroxylated gelatin by coexpression of gelatin and P4H genes has been conducted. Nevertheless, several researchers have reported that proline hydroxylation can be achieved by coexpression of recombinant collagen and heterologous P4H in the yeast, *P. pastoris* or *S. cerevisiae*.[8,9] This observation may imply that hydroxylation of proline in the Y position of Gly-X-Y triplets of recombinant gelatin is also feasible in *P. pastoris*.

Received: February 24, 2011
Accepted: May 23, 2011
Revised: April 30, 2011
Published: May 23, 2011

Table 1. List of Primers Used for PCR

primer	sequence (5′−3′)	characteristics
3′AOX1 (TT)	CGGGGTACCCCTTAGACATGACTGTTCCTCAG	forward
	CCGCTCGAGAAGCTTGCACAAACGAACGTCT	reverse
PFLD1	GCGCTCGAGGCATGCAGGAATCTCTGGCACG	forward
	GCGGGATCCTGTGAATATCAAGAATTGTAT	reverse
α	CGCGGATCCGCCACCATGATCTGGTATATA	forward
	CGGGGTCCTCATTATTCCAATTCTGACAACGTACAAG	reverse
β	CCGGAATTCGACGCCCCCGAGGAGGAGGACCAC	forward
	CCTAGGTCATTACAGTTCATCGTGCACAGCTTTCTG	reverse
αMF	CGCGGATCCGCCACCATGAGATTTCCTTCAATTTTTACT	forward
	CCGGAATTCAGCTTCAGCCTCTCTTTTCTCGAG	reverse

We apply a new strategy for coexpressing the segment of genes coding for gelatin, and the α and *β* domains of P4H. These three genes were designed and cloned on an expression vector (pPIC9K) under the control of two different strong inducible promoters. The gelatin gene was put under the control of PAOX1, and each of the P4H genes under the control of PFLD1. The cloned vector was transfected into *P. pastoris* KM71 strain to produce recombinant hydroxylated gelatin.

MATERIALS AND METHODS

Strains, Plasmids, Reagents, and Media. *Escherichia coli* strain DH5α was used for plasmid construction and amplification. Plasmid pGEM-T (Promega) was used to construct plasmid vector. *P. pastoris* KM71 (Invitrogen) and pPIC9K (Invitrogen) were used as the host and expression vector, respectively. *P. pastoris* vector pPIC9K, which contains the *S. cerevisiae* alpha factor leader and PAOX1, was used as a backbone for the expression of the gelatin gene. Taq DNA polymerase, T4 DNA ligase, and restriction enzymes were purchased from Takara. Yeast nitrogen base (YNB) containing ammonium sulfate, but with no added amino acids, was bought from Difco. Yeast extract and tryptone were purchased from Oxoid. All other chemicals were of the highest quality. *Escherichia coli* strain DH5α was grown at 37 °C in LB medium (1% NaCl, 1% tryptone, 0.5% yeast extract, pH 7.0) containing ampicillin (100 μg/mL). *P. pastoris* strains were cultivated in YPD medium (1% yeast extract, 2% soya peptone, 2% glucose, pH 7.0). *P. pastoris* transformants were selected on minimal dextrose medium (MD) plates (1.34% YNB with ammonium sulfate without amino acids, 4×10^{-5}% biotin, 2% dextrose and 1.5% agar, pH 7.0). The growth and induction media for *P. pastoris* were BMGY (1.34% YNB, 4×10^{-5}% biotin, 1% yeast extract, 2% soya peptone, 100 mM potassium phosphate, 1% glycerol, pH 6.0) and BMMY (same as BMGY, except that glycerol was replaced by 0.5% methanol), respectively.

Construction of Expression Vector. The various DNA components used in this experiment were acquired by polymerase chain reaction (PCR) and assembled on pGEM-T vector to make a fusion of expression cassettes. All the procedures with pGEM-T vectors were conducted in *E. coli* DH5α stain. The fusion was excised from pGEM-T vector and cloned into pPIC9K expression vector before it was transfected into *P. pastoris*. All the primers used in this section are shown in Table 1.

Transformation of *P. pastoris*. A plasmid with the expected DNA, pPIC9K/gelatin-TT-PFLD1-α-TT-PFLD1-αMF-*β*, was recovered, linearized with *Sal*I, and subsequently retransformed into *P. pastoris* KM71 cells by electroporation using a GenePulser (Bio-Rad) set at 1500 V, 25 μF, and 400 Ω, as well as 0.2 cm cuvettes. Selection of His^+ transformants and G418-resistant clones was performed in accordance with the instruction manual (Invitrogen). All of the His^+ transformants in KM71 are Mut_S (methanol utilization slow) because of the disruption of the alcohol oxidase 1 (AOX1) gene. Transformants were selected on MD plates containing 4 mg/mL of G418 to enrich strains with multiple copies of the integrated DNA.

Cultivation Conditions. To induce protein expression in *P. pastoris*, a single colony was selected on the MD plates containing 4 mg/mL of G418, and cultivated overnight in 5 mL of YPD at 30 °C. Moreover, 1 mL of the culture medium was transferred into 50 mL of BMGY and incubated for approximately 24 h until the culture reached OD600 = 2−6. The cells were harvested by centrifugation and resuspended in 15 mL of BMMY, which contained 1534 mg/L of sodium glutamate, and incubated at 250 rpm for 3 d at 30 °C. Methanol was added into the medium every 24 h to a final concentration of 0.5% to maintain induction. Meanwhile, 80 μg/mL of L-ascorbic acid (sodium salt) was added every 12 h during the induction of gelatin expression. Samples were taken every 12 h and centrifugated. The culture supernatants were checked by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

SDS-PAGE Analysis of the Expressed Products. Denaturing SDS-PAGE was conducted in accordance with the method of Laemmli.[10] The supernatant samples (20 μL) were resolved on 10% SDS-PAGE under reducing conditions and visualized by Coomassie Blue R-250 staining.

Protein Digestion and LC-MS/MS Analysis. Coomassie-stained gel bands were excised and destained with 50% acetonitrile containing 50 mM $(NH_4)_2CO_3$. After washing with Milli-Q water until the gels became clear, the gel pieces were dehydrated with 100% acetonitrile for 5 min and dried by Speed-Vac for 20−30 min, followed by rehydration with 15 μL of cold trypsin solution (12.5 ng/μL) and incubation at 37 °C for 16−24 h.

Nano-LC-MS/MS was performed on an HPLC system composed of two LC-20AD nanoflow LC pumps, an SIL-20 AC autosampler, and an LC-20AB microflow LC pump (Shimadzu), which are connected to an LTQ-Orbitrap mass spectrometer (ThermoFisher). The sample was loaded on a CAPTRAP column (0.5 mm ×2 mm, MICHROM Bioresources) for 6 min at a flow rate of 25 μL/min. The sample was subsequently separated by a C18 reverse-phase column (0.10 mm ×150 mm, packed with 3 μm Magic C18-AQ particles, MICHROM Bioresources) at a flow rate of 500 nL/min.

NMR Spectrum Examination. The weighted recombinant hydroxylated gelatin was soaked in heavy water for 2 h, water bathed at 40 °C, and heated and stirred for 10 min. When the recombinant hydroxylated gelatin was completely dissolved, it was moved to a 5 mm NMR sample tube for the ^1H NMR test.

The recombinant hydroxylated gelatin was soaked in heavy water for 2 h, water bathed at 40 °C, and heated and stirred for 20 min until completely dissolved. Subsequently, it was moved to a 10 mm NMR sample tube for the ^{13}C NMR test.

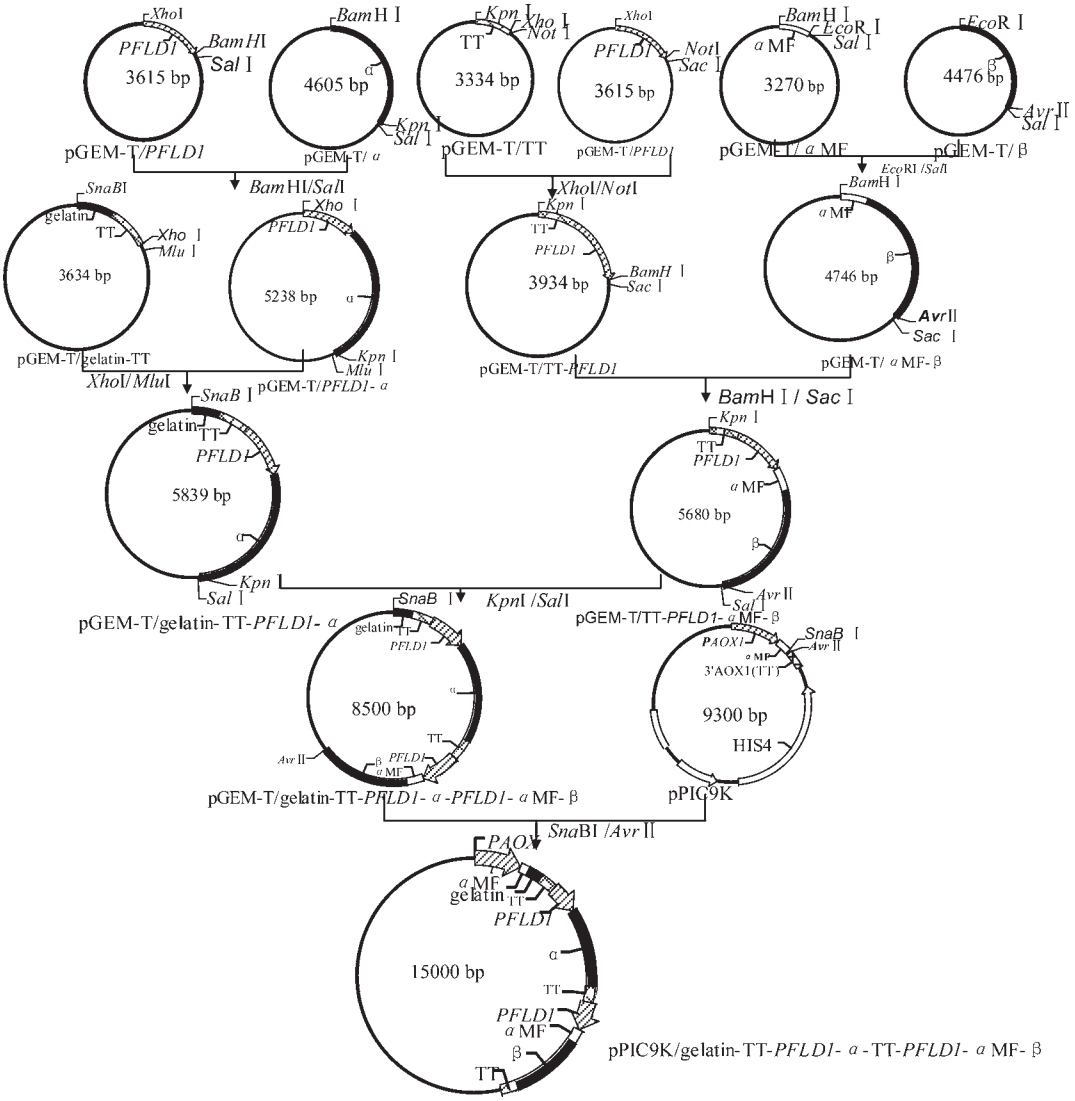


Figure 1. Physical maps of the vectors and the construction strategy for the *P. pastoris* expression plasmid pPIC9K/gelatin-TT-PFLD1-α-TT-PFLD1-αMF-β used in the current study.

UV Absorption Spectra Examination. UV absorption spectra examination was conducted under the following conditions: measurement type, wavelength scan; data mode, Abs; starting wavelength, 400 nm; ending wavelength, 190 nm; scan speed, 800 nm/min; sampling interval, 0.5 nm; slit width, 2 nm; lamp change, 340 nm; baseline correction, user; response, fast; path length, 10 mm; peak integration; method, rectangular; sensitivity, 1; threshold, 0.0100.

Hydroxyproline Examination. An amino acid automatic analyzer L8900 was used for hydroxyproline examination. The sample in the hydrolysis tube was weighed and added with 6 M HCl solution to obtain a final volume of 10 mL. The sample was frozen in ice water for 3−5 min, vacuumed, and poured with nitrogen; this procedure was repeated thrice. The tube under the protection of nitrogen was sealed, and placed in a 110 °C incubator for hydrolysis for 20 h. Afterward, the sample was allowed to cool, and then the hydrolyzate was transferred in a 50 mL volumetric flask, filtered, vacuum pumped, and diluted with 0.02 M HCl. The examination was started after filtration. The operating conditions were as follows: analytical column, 4.6 mm ×60 mm resin 2619; ammonia removal column, 4 mm ×50 mm resin 2650; flow rate, A pump 0.4 mL/min, B pump 0.35 mL/min; column temperature, 57 °C; reaction column temperature, 135 °C.

RESULTS

Construction of Vector. In this study, we cloned and expressed gene fragments of gelatin and the enzyme P4H in *P. pastoris* KM71. To obtain the modified gelatin gene, a cDNA encoding 100 amino acids [nos. 531−630 of the human α 1 (I) procollagen] was constructed by overlap PCR with long oligonucleotides. The gelatin gene was linked to pPIC9K. The fragment gelatin-transcription terminator (TT) was then acquired by PCR with the vector pPIC9K/gelatin as a template, which was inserted into pGEM-T to form pGEM-T/gelatin-TT. The cloned gelatin-TT fragment was confirmed by sequencing.

For the two P4H subunits controlled by PFLD1, *P. pastoris* GS115 genomic DNA was utilized as a template to generate PFLD1 by PCR.

The sequence coding the signal peptide of human PDI was replaced with the *S. cerevisiae* α mating factor (αMF) prepro sequence. A cDNA for human PDI extending from the codon for the first amino acid after the signal peptide cleavage site to the stop codon was synthesized using PCR. The C-terminal endoplasmic

dx.doi.org/10.1021/jf200778r |*J. Agric. Food Chem.* 2011, 59, 7127–7134

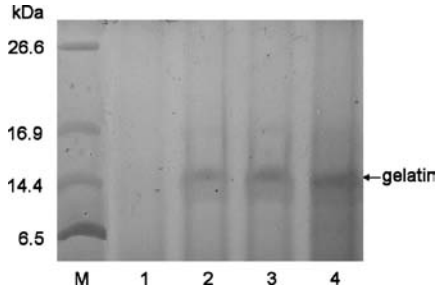


Figure 2. SDS-PAGE electrophoresis of the recombinant gelatin without coexpression P4H. Shown are 1 supernatant before induction and 2−4 supernatant at 24, 48, and 72 h, respectively, after induction. *M* is the molecular weight of standard protein.

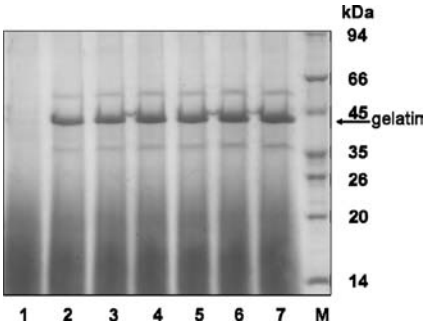


Figure 3. SDS-PAGE analysis of the expression level of recombinant gelatin with coexpression P4H at different induction times. Lanes 1−7: the cultured solutions incubated for 0, 12, 24, 36, 48, 60, and 72 h, respectively, after the induction of expression by methanol addition. *M* indicates protein standard.

reticulum retention signal KDEL of PDI was modified to an HDEL signal by PCR. The αP4H gene cDNA was obtained from a human source with its native secretion signal.

In all reactions, the PCR products, which were extracted and identified by DNA sequencing, were inserted into the vector pGEM-T and transfected into the host strain, *E. coli* DH5α. The recombinant plasmid was constructed as follows: In addition to previously formed pGEM-T/gelatin-TT, every other two adjacent vectors were linked to form pGEM-T/PFLD1-α, pGEM-T/TT-PFLD1, and pGEM-T/αMF-β with corresponding restriction enzymes. Subsequently, every two adjacent vectors mentioned above were connected to form pGEM-T/gelatin-TT-PFLD1-α and pGEM-T/TT-PFLD1-αMF-β. The two newborn vectors were then connected to form pGEM-T/gelatin-TT-PFLD1-α-TT-PFLD1-αMF-β. The gelatin-TT-PFLD1-α-TT-PFLD1-αMF-β fusion from the pGEM-T/gelatin-TT-PFLD1-α-TT-PFLD1-αMF-β vector was cut with *Sna*BI and *Avr*II, and inserted into the same restriction sites of pPIC9K to yield pPIC9K/gelatin-TT-PFLD1-α-TT-PFLD1-αMF-β (shown in Figure 1). The 5′-terminus of gelatin was fused to the α-factor signal sequence of the parent expression vector, which made the three expression cassettes (located downstream of the α-factor signal sequence) under the control of PAOX1, PFLD1, and PFLD1, with each gene having its own independent promoter and transcription terminator. A Kozak sequence (GCCACC) was added to the 5′ end of the start codon ATG of the α gene and the second αMF signal peptide to enhance their expression. To prevent reading through the neighboring DNA sequence, two consecutive stop codons TAATGA were added between the 3′ end of each gene and its TT sequence.

SDS-PAGE Analysis and LC-MS/MS Examination. Our experimental results are consistent with the report that collagenous proteins migrate in SDS-PAGE at a molecular weight more than 40% higher than the true molecular weight because of the high content of proline.[7] When only the gelatin gene fragment was cloned and expressed in *P. pastoris*, nonhydroxylated gelatin was obtained. In theory, the molecular weight of nonhydroxylated gelatin is supposed to be 9 kDa on the basis of the cDNA sequence of the cloned fragment; however, SDS-PAGE shows bands approximately 14.5 kDa in size (shown in Figure 2), which is more than 40% higher than expected.

Similarly, when gelatin and enzyme genes were coexpressed in *P. pastoris*, hydroxylated gelatin was obtained with a molecular weight of 42 kDa instead of the theoretical molecular weight of 27 kDa (shown in Figure 3). The expression also shows typical induction pattern as gelatin concentration increases with time (Figure 3). After trypsin digestion, the samples submitted for LC-MS/MS analysis were found to be gelatin proteins, which matched with the proteome database from UniProt. Among the amino acids detected in the peptide sequence were hydroxyprolines, which have been identified as P[113] (shown in Table 2). Approximately 66.7% of the proline residues (in the Y position of Gly-X-Y) in the analyzed peptides were hydroxylated.

^{1}H Spectra of Recombinant Hydroxylated Gelatin. The ^{1}H spectra of the recombinant gelatin (shown in Figure 4) are as follows: 3.958 spectral peak is the α-CH_2 peak of the gelatin characteristic amino acid glycine; 3.858 spectral peak is the ν-CH_2 peak of the gelatin characteristic amino acid hydroxyproline; and 3.635 spectral peak is δ-CH_2 peak of the gelatin characteristic amino acid proline. All these displacement values of the peaks are close to those of natural gelatin.[11]

^{13}C Spectra of Recombinant Hydroxylated Gelatin. The ^{13}C spectra of the recombinant gelatin (shown in Figure 5) are as follows: 72.606 spectral peak is the γ-C peak of the gelatin characteristic amino acid hydroxyproline; and 63.268 spectral peak is the α-C peak of the gelatin characteristic amino acid proline. All these displacement values of the peaks are close to those of natural gelatin.[11]

UV Absorption Spectra of Recombinant Hydroxylated Gelatin. In the 190−400 nm UV region, the recombinant gelatin has a maximum absorption peak at 225 nm, which is close to the absorption peak of type I collagen.[12] This maximum absorption peak was obtained because the amino acids that constitute this recombinant gelatin contain only one phenylalanine without tyrosine and tryptophan, have very low absorption, and are almost invisible at 280 nm (see Figure 6). Furthermore, the absorption peak of the natural gelatins at 280 nm was slightly higher than the recombinant gelatin because they also generally contain tyrosine.

Hydroxyproline Examination. The recombinant gelatin samples were hydrolyzed with hydrochloric acid and examined by an amino acids automatic analyzer. The 4.953 peak is the peak of hydroxyproline (shown in Figure 7).

DISCUSSION

The *E. coli* expression system is not suitable for handling human-derived gelatin genes with relatively high repetitive sequence because of its congenital defects on the post-translational modification of some eukaryotic proteins. For this reason, we chose the *P. pastoris* expression system for the

Table 2. LC-MS/MS Analysis of the Target Protein

protein	protein probability	percent share of spectrum id's	description	peptide sequence
sp|P02452|CO1A1_HUMAN, sp|P02453|CO1A1_BOVIN, sp|Q9XSJ7|CO1A1_CANFA	1	64.2	collagen α-1(I) chain OS = *Homo sapiens* GN = COL1A1 PE = 1 SV=5, collagen α-1(I) chain OS = *Bos taurus* GN = COL1A1 PE = 1 SV = 3, collagen α-1(I) chain OS = Canis familiariz GN = COL1A1 PE = 1 SV = 1	GLTGSP[113]GSP[113]GPDGK
sp|P02452|CO1A1_HUMAN, sp|P02453|CO1A1_BOVIN, sp|Q9XSJ7|CO1A1_CANFA	1	64.2	collagen α-1(I) chain OS = *Homo sapiens* GN = COL1A1 PE = 1 SV = 5, collagen α-1(I) chain OS=*Bos taurus* GN = COL1A1 PE = 1 SV = 3, collagen α-1(I) chain OS = Canis familiariz GN = COL1A1 PE = 1 SV = 1	GAAGEP[113]GK
sp|P02452|CO1A1_HUMAN, sp|P02453|CO1A1_BOVIN, sp|Q9XSJ7|CO1A1_CANFA	1	64.2	collagen α-1(I) chain OS = *Homo sapiens* GN = COL1A1 PE = 1 SV = 5, collagen α-1(I) chain OS = *Bos taurus* GN = COL1A1 PE = 1 SV = 3, collagen α-1(I) chain OS = Canis familiariz GN = COL1A1 PE = 1 SV = 1	TGPP[113]GPAGQDGR
sp|P02452|CO1A1_HUMAN, sp|P02453|CO1A1_BOVIN, sp|Q9XSJ7|CO1A1_CANFA	1	64.2	collagen α-1(I) chain OS = *Homo sapiens* GN = COL1A1 PE = 1 SV = 5, collagen α-1(I) chain OS = *Bos taurus* GN = COL1A1 PE = 1 SV = 3, collagen α-1(I) chain OS = Canis familiariz GN = COL1A1 PE = 1 SV = 1	GVPGPPGAVGPAGK
sp|P02452|CO1A1_HUMAN, sp|P02453|CO1A1_BOVIN, sp|Q9XSJ7|CO1A1_CANFA	1	64.2	collagen α-1(I) chain OS = *Homo sapiens* GN = COL1A1 PE = 1 SV = 5, collagen α-1(I) chain OS = *Bos taurus* GN = COL1A1 PE = 1 SV = 3, collagen α-1(I) chain OS = Canis familiariz GN = COL1A1 PE = 1 SV = 1	PGPPmac_opsb;113]GPPGAR
sp|P02452|CO1A1_HUMAN, sp|P02453|CO1A1_BOVIN, sp|Q9XSJ7|CO1A1_CANFA	1	64.2	collagen α-1(I) chain OS = *Homo sapiens* GN = COL1A1 PE = 1 SV = 5, collagen α-1(I) chain OS = *Bos taurus* GN = COL1A1 PE = 1 SV = 3, collagen α-1(I) chain OS = Canis familiariz GN = COL1A1 PE = 1 SV = 1	GQAGVMGFP[113]GPK

expression of recombinant gelatin. Although mammalian expression systems have unique advantages in terms of post-translational modification of the expressed product, these systems are very expensive, and in many cases, the target product yield is low. In this regard, *P. pastoris* may be a suitable candidate because of its post-translational modification capacity. Our results also demonstrated the gelatin gene expression and its hydroxylation when P4H genes were introduced. In addition, a *P. pastoris expression* vector, pPIC9K, is an ideal choice for carrying macromolecule DNA because of its high molecular weight.

To express a functional human gelatin successfully, we cloned three genes coding for gelatin, αP4H, and PDI on one expression vector, pPIC9K, instead of cloning each gene on a separate vector. Each of the cloned genes was controlled by a promoter and had its own terminator sequence. The purpose of this procedure is to facilitate integration and coexpression of the cloned expression cassettes as a whole in the yeast genome. This procedure also saves time and resources.

The cloned gelatin gene was placed under the control of PAOX1. This promoter is tightly repressed during the growth of the yeast on most common carbon sources, such as glucose, glycerol, or ethanol, but it is highly induced during growth on methanol. The αP4H and PDI genes were controlled by PFLD1 separately. PFLD1 is another strong promoter of *P. pastoris*, which is strongly and independently induced when methanol is used as the only carbon source (with ammonium sulfate as a nitrogen source) or methylamine as the sole nitrogen source (with glucose or sorbitol as a carbon source).[13] Therefore, PFLD1 and PAOX1 can be induced simultaneously by a common inducer (methanol) when they exist on one vector. The successful expression and hydroxylation of the recombinant gelatin in the current study proved that the two promoters can function effectively without interference even when they exist close to each other.

P4H, an $\alpha\beta$ tetrameric enzyme composed of the αP4H and the β subunit in higher eukaryotes, is the enzyme that modifies proline residues to hydroxyproline. This enzyme is retained and functions within the endoplasmic reticulum through a C-terminal endoplasmic reticulum retention signal sequence, which is predominantly KDEL in mammalian systems and HDEL in yeast.[14] In our construct, we replaced the signal sequence of the human PDI with αMF prepro sequence of *S. cerevisiae*. We performed this step because the authentic signal sequence of the human PDI polypeptide is highly ineffective in transporting the recombinant polypeptide into the lumen of the endoplasmic reticulum in *P. pastoris* cells, whereas the efficiency of this transport increases considerably when the human signal sequence is replaced by the *S. cerevisiae* αMF prepro sequence.[8]

During gelatin biosynthesis, the gelatin gene is cotranslationally transported into the lumen of the endoplasmic reticulum where it is hydroxylated by P4H, which is in charge of the conversion of proline residues into 4-hydroxyprolines in the Y position of Gly-X-Y triplets that make up the triple helical regions of collagen and gelatins.[8] P4H requires Fe^{2+}, 2-oxoglutarate, O_2, and ascorbate.[15] Glutamate, possibly acting as a precursor for the synthesis of α-ketoglutarate (an essential cofactor for proline hydroxylation), is required to generate high levels of triple helical procollagen molecules.[16] Only when the Y-position prolyl residues are properly hydroxylated to 4-hydroxyproline by P4H can the newly synthesized chains fold into a triple-helical conformation at 37 °C. Moreover, if hydroxylation does not occur, the polypeptides remain nonhelical, are poorly secreted by cells, and cannot self-assemble into collagen fibrils. Coexpression of the

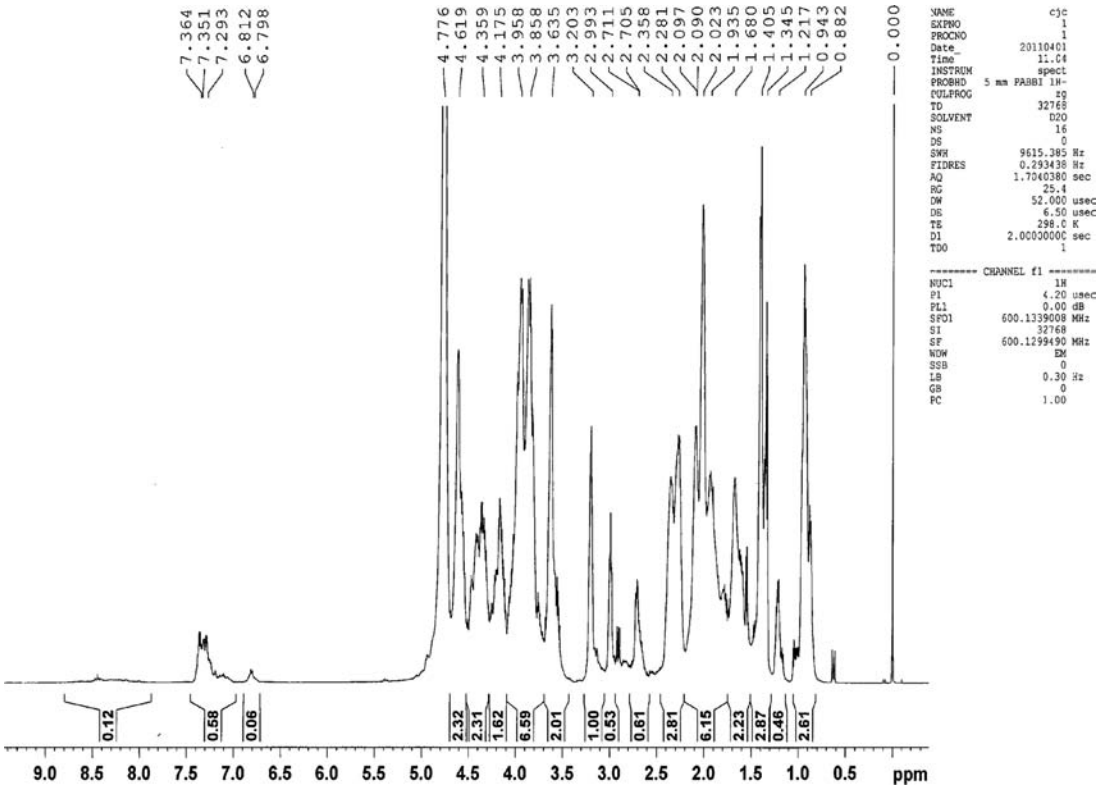


Figure 4. [1]H spectra of recombinant gelatin.

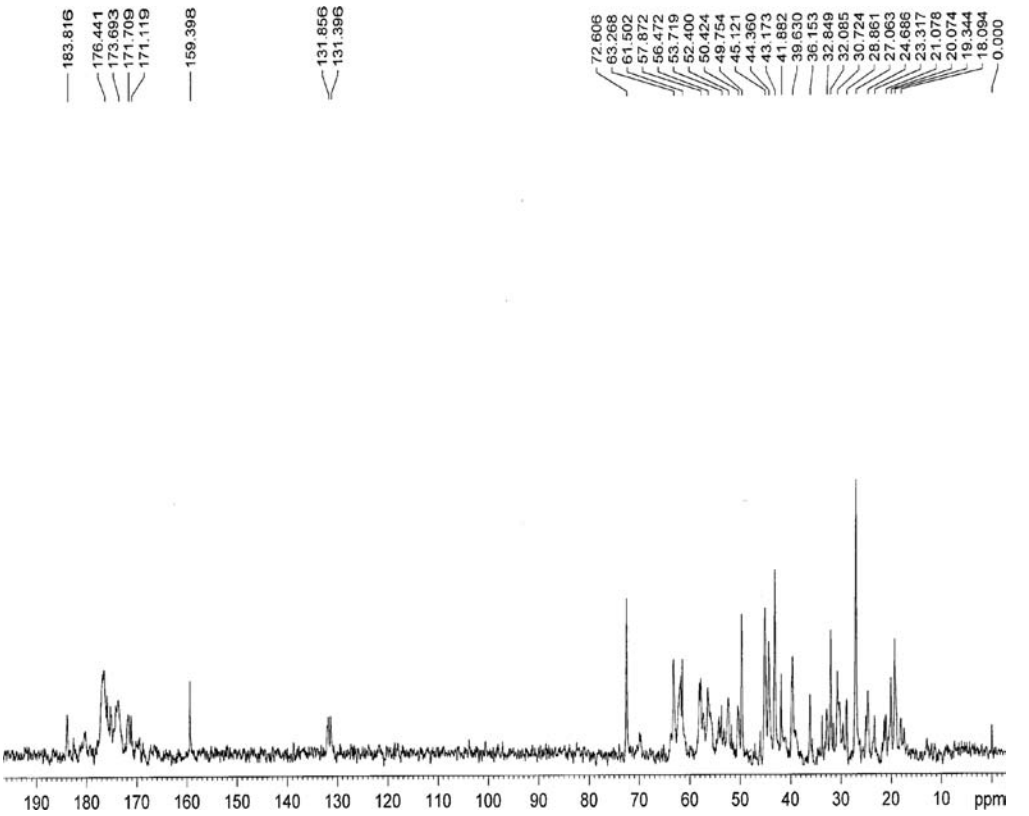


Figure 5. [13]C spectra of recombinant gelatin.

dx.doi.org/10.1021/jf200778r |*J. Agric. Food Chem.* 2011, 59, 7127–7134

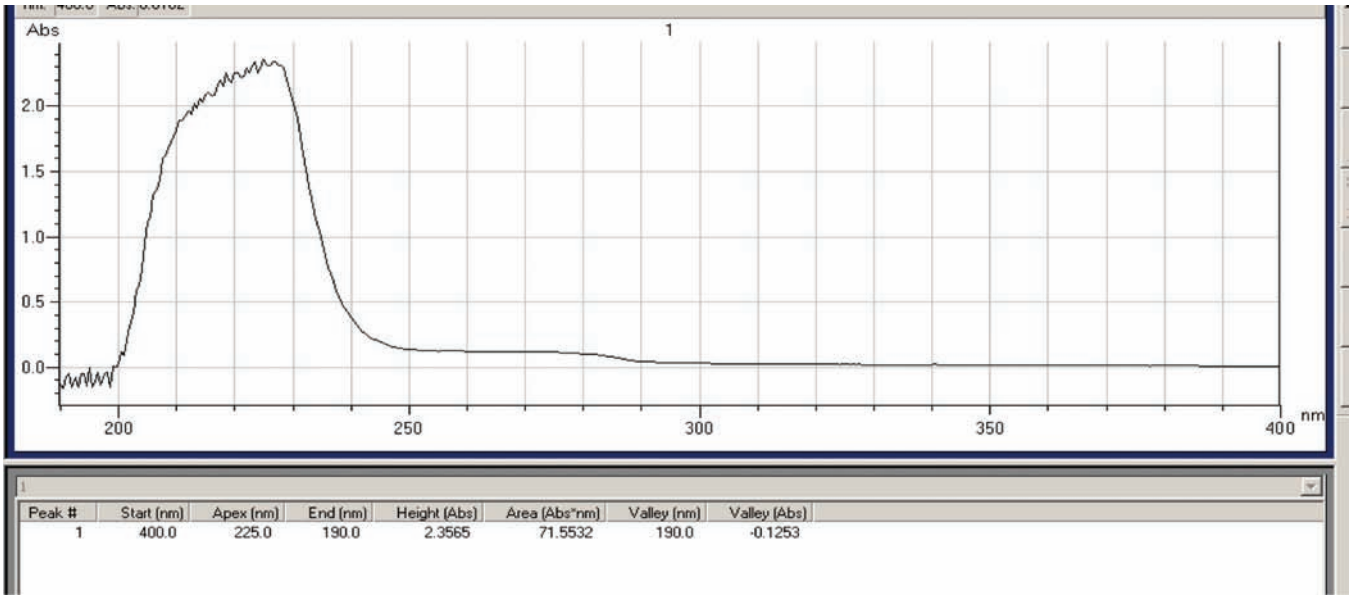
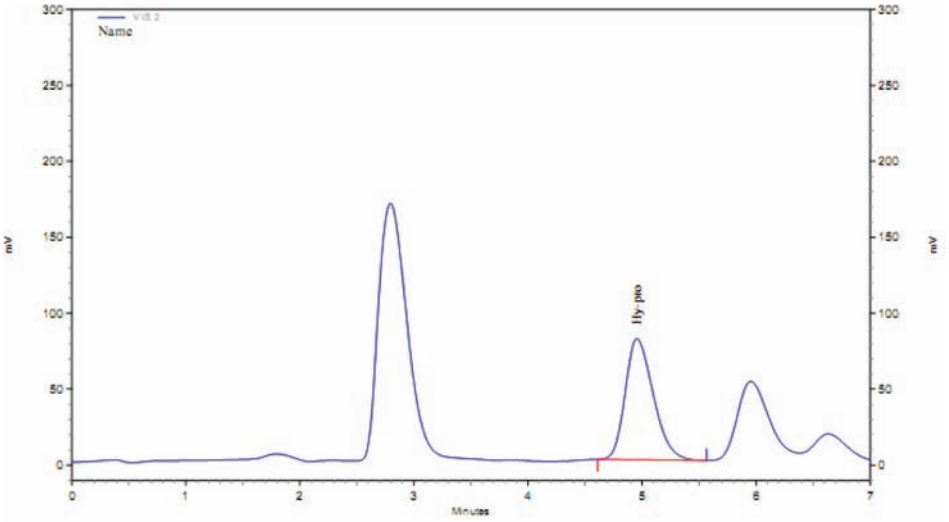


Figure 6. UV spectra of recombinant gelatin.



VIS 2 Results

Pk #	RT	Name	Height	Area
1	4.953	Hy-pro	267600	4571768
Totals			267600	4571768

Figure 7. Hydroxyproline analysis of recombinant gelatin.

three cloned genes holds a control position in the production of functional gelatin. Our experimental results confirm that, without expressing P4H genes, gelatin is expressed in the form of non-hydroxylated monomer with a molecular weight of 14.5 kDa. Hydroxylated gelatin of 42 kDa, approximately three times the weight of monomer (14.5 kDa), was produced only when the genes encoding for gelatin and P4H were coexpressed. LC-MS/MS analysis further verified the presence of hydroxyproline residues. The small amount of hydroxylation that did not occur indicates that the P4H activity was not very high.

We used *P. pastoris* GS115 as the expression strain. However, the plasmid could not be linearized with *Bgl*II because the αP4H gene has the *Bgl*II restriction site. Thus, Mut^S could not be obtained. To address this limitation, we linearized the plasmid with *Sal*I. Afterward, Mut^+ (methanol utilization plus) was produced. After induction, no expression of gelatin occurred. We then switched to *P. pastoris* KM71 as the expression strain. The plasmid was linearized with *Sal*I, the resulting transformant of which was Mut^S because of the disruption of the AOX1 gene. The successful expression of gelatin in *P. pastoris* KM71 indicates

that Mut^S is possibly the only suitable genotype for the expression.

The catalytic substrate of P4H is Gly-X-Y triplet; hydroxylation occurs only when proline is in the Y position and not when it is in the X position. The hydroxylation rate was 66.7%; six peptides were detected, but nine prolines in the Y position in Gly-X-Y should be hydroxylated. Given that only six prolines were hydroxylated, the rate of hydroxylation was therefore six divided by nine.

Nonhydroxylated recombinant gelatin is a monomer that is susceptible to enzymatic digestion. After hydroxylation, the recombinant gelatin is more stable and resistant to enzymatic degradation. No study on the differences between nonhydroxylated and hydroxylated recombinant gelatin in terms of physical and chemical properties, such as viscosity, tensile strength, etc., has been conducted.

AUTHOR INFORMATION

Corresponding Author

*Phone: +86-10-64439673. Fax: +86-10-64439673. E-mail: jingchunchen@hotmail.com.

ABBREVIATIONS USED

P4H, prolyl 4-hydroxylase; rG, recombinant gelatin; PCR, polymerase chain reaction; AOX1, *P. pastoris* gene encoding alcohol oxidase 1; PAOX1, the alcohol oxidase 1 promoter; PFLD1, formaldehyde dehydrogenase promoter; YNB, yeast nitrogen base; YPD, yeast extract peptone dextrose medium; BMGY, buffered minimal glycerol medium; BMMY, buffered minimal methanol medium; MD, minimal dextrose medium; PDI, the protein-disulfide isomerase; ER, endoplasmic reticulum; Mut^+, methanol utilization plus; Mut^S, methanol utilization slow; αMF, *S. cerevisiae* α mating factor; rG, recombinant gelatin; TT, transcription terminator

REFERENCES

(1) Asghar, A.; Henrickson, R. L. Chemical, biochemical, functional, and nutritional characteristics of collagen in food systems. *Adv. Food. Res.* **1982**, *28*, 231–372.

(2) Kielty, C. M.; Hopkinson, I.; Grant, M. E. The collagen family: structure, assembly and organization in the extracellular matrix. In *Connective Tissue and Its Heritable Disorders. Molecular, Genetic and Medical Aspects*; Royce, P. M., Steinmann, B., Eds.; Wiley-Liss: New York. 1993; pp 103−147.

(3) Prockop, D. J.; Kivirikko, K. I. Collagens: molecular biology, diseases, and potentials for therapy. *Annu. Rev. Biochem.* **1995**, *64*, 403–434.

(4) Bateman, J. F.; Lamande, S.; Ramshaw, J. A. M. *Collagen Superfamily in Extracellular Matrix Molecular Components and Interactions*; Comper, W.D., Ed.; Harwood Academic Publishers: Amsterdam, 1995; Vol. 2, pp 22−67.

(5) Kivirikko, K. I.; Myllyla, R.; Pihlajaniemi, T. Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. In *Post-Translational Modifications of Proteins*; Harding, J. J., Crabbe, M. J. C., Eds.; CRC Press: Boca Raton, FL, 1992; pp 1−51.

(6) Olsen, D.; Jiang, J.; Chang, R.; Duffy, R.; Sakaguchi, M.; Leigh, S.; Lundgard, R.; Ju, J.; Buschman, F.; Truong-Le, V.; Pham, B.; Polarek, J. W. Expression and characterization of a low molecular weight recombinant human gelatin: development of a substitute for animal-derived gelatin with superior features. *Protein Expression Purif.* **2005**, *40*, 346–357.

(7) Werten, M. W. T.; van den Bosch, T. J.; Wind, R. D.; Mooibroek, H.; de Wolf, F. A. High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast* **1999**, *15*, 1087–1096.

(8) Vuorela, A.; Myllyharju, J.; Nissi, R.; Pihlajaniemi, T.; Kivirikko, K. I. Assembly of human prolyl 4-hydroxylase and type III collagen in the *Pichia pastoris*: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase. *EMBO J.* **1997**, *16*, 6702–6712.

(9) Toman, P. D.; Chisholm, G.; McMullin, H.; Giere, L. M.; Olsen, D. R.; Kovach, R. J.; Leigh, S. D.; Fong, B. E.; Chang, R.; Daniels, G. A.; Berg, R. A.; Hitzeman, R. A. Production of recombinant human type I procollagen trimers using a four-gene expression system in the yeast Saccharomyces cerevisiae. *J. Biol. Chem.* **2000**, *275*, 23303–23309.

(10) Laemmli, U. K. Cleavage of structural proteins during assembly of the head of the bacteriophage T4. *Nature* **1970**, *227*, 680–685.

(11) He, Y.; Zhang, M. The standard ^1H and ^{13}C NMR spectra of gelatin. *Leather Sci. Eng.* **1990**, *2*, 5−10 (in Chinese).

(12) Lin, Y. K.; Liu, D. C. Comparison of physical-chemical properties of type I collagen from different species. *Food Chem.* **2006**, *99*, 244–251.

(13) Shen, S.; Sulter, G.; Jeffries, T. W.; Cregg, J. M. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* **1998**, *216*, 93–102.

(14) Kivirikko, K. I.; Helaakoski, T.; Tasanen, K.; Vuori, K.; Myllyla, R.; Parkkonen, T.; Pihlajaniemi, T. Molecular biology of prolyl 4-hydroxylase. *Ann. N.Y. Acad. Sci.* **1990**, *580*, 132–142.

(15) Kivirikko, K. I.; Myllyla, R.; Pihlajaniemi, T. Protein hydroxylation: prolyl 4-hydroxylase, an enzyme with four cosubstrates and a multifunctional subunit. *FASEB J.* **1989**, *3*, 1609–1617.

(16) Toman, P. D.; Chisholm, G.; McMullin, H.; Giere, L. M.; Olsen, D. R.; Kovach, R. J.; Leigh, S. D.; Fong, B. E.; Chang, R.; Daniels, G. A.; Berg, R. A.; Hitzeman, R. A. Production of recombinant human type I procollagen trimers using a four-gene expression system in the yeast Saccharomyces cerevisiae. *J. Biol. Chem.* **2000**, *275*, 23303–23309.