

**Recombinant gelatin and collagen from  
methylophilic yeasts**

**Eric C. de Bruin**

CENTRALE LANDBOUWCATALOGUS



0000 0924 9166

- Promotor:** Prof. dr. N. C. M. Laane  
Voormalig hoogleraar in de Biochemie
- Copromotor:** Dr. F. A. de Wolf  
Agrotechnologisch Onderzoeks instituut (ATO B.V.)  
Departement Agrotechnologie en Voedingswetenschappen  
Wageningen UR
- Promotiecommissie:** Prof. dr. S. C. de Vries (Wageningen Universiteit)  
Prof. dr. M. Veenhuis (Rijks Universiteit van Groningen)  
Dr. J. Myllyharju (Universiteit van Oulu, Finland)  
Dr. ir. C. P. M. van Mierlo (Wageningen Universiteit)

NIJ 2101.31.3

# **Recombinant gelatin and collagen from methylotrophic yeasts**

**Eric C. de Bruin**

**Proefschrift**

**ter verkrijging van de graad van doctor  
op gezag van de rector magnificus  
van Wageningen Universiteit,  
Prof. dr. ir. L. Speelman,  
in het openbaar te verdedigen  
op vrijdag 15 maart 2002  
des namiddags te half twee in de Aula**

1 5431 87

E. C. de Bruin-Recombinant gelatin and collagen from methylotrophic yeasts-2002

PhD thesis Wageningen University, the Netherlands

Keywords: gelatin; collagen; methylotrophic yeast; *Hansenula polymorpha*; *Pichia pastoris*; fermentation; secretion; recombinant expression

ISBN: 90-5808-583-X

# Abstract

Based on its structural role and compatibility within the human body, collagen is a commonly used biomaterial in medical applications, such as cosmetic surgery, wound treatment and tissue engineering. Gelatin is in essence denatured and partly degraded collagen and is, as a result of its unique functional and chemical properties, also used in many medical and pharmaceutical products. Collagen and gelatin are traditionally extracted from animal tissues. The quality and the characteristics of the proteins are not very reproducible in today's batch-to-batch production processes and recently, potential contamination of collagen and gelatin with viruses and prions (causing the mad cow disease, BSE) became a matter of concern. BSE is thought to cause a new variety of the brain-wasting Creutzfeldt-Jacob disease in humans.

Recombinant DNA technology may provide safe collagen and gelatins from which the quality and characteristics can precisely be controlled and reproduced and, in addition, opens up possibilities for novel functional "tailor-made" proteins. For the heterologous production of animal proteins yeasts are frequently used. Since yeasts are eukaryotes, most translational modification, needed for functionality and stability of recombinant animal proteins, normally occur. However prolyl 4-hydroxylation, essential for gelling properties of recombinant gelatin and thermal stability of recombinant collagen, is generally considered to be absent in yeast systems.

In this study we explored the methylotrophic yeasts *Hansenula polymorpha* and *Pichia pastoris* for their use as recombinant production systems of natural and "tailor-made" gelatins and human collagen. We found that both yeasts are well able to cope with the repetitive gene sequences encoding animal gelatin and human collagen and showed that *P. pastoris* can produce synthetic gelatins with highly hydrophilic properties at high levels. Furthermore, it was discovered that *H. polymorpha* unexpectedly produced endogenous collagen-like proteins with 4-hydroxyproline amino acid residues. This finding indicated that the yeast *H. polymorpha*, in contrast to what was generally believed, must contain intrinsic prolyl 4-hydroxyase activity. Indeed, expression of murine gelatin in *H. polymorpha* yielded a hydroxylated product.

We also investigated if *H. polymorpha* could be used for the production of recombinant human collagen. Intact human collagen trimers were obtained but they

were not stable at temperatures higher than 15 °C, indicating that hydroxylation in the product was poor.

In the course of this study we found putative prolyl 4-hydroxylase genes in different eukaryotic microbial systems. In the future these genes may be used to further develop yeasts into cell factories for the production of animal gelatins and thermally stable human collagens.

# Stellingen

1. In contrast to what was generally believed, some microbial eukaryotic expression systems do contain prolyl 4-hydroxylase activity.

This thesis

2. Methylophilic yeasts are suitable hosts for recombinant gelatin production.

This thesis

3. The correct alignment and folding into triple helical conformation of recombinant collagen type I monomeric polypeptide chains lacking the C and N propeptides in *Saccharomyces cerevisiae* cells is caused by the absence of monomers of other collagen types.

Olsen *et al.*, 2001, "Production of human type I collagen in yeast reveals unexpected new insights into the molecular assembly of collagen trimers" *J. Biol. Chem.* 276: 24038-43

4. Het drinken van een borrel is een vorm van hersengymnastiek.

Ruitenbergh *et al.*, 2002, "Alcohol consumption and risk of dementia". *The Lancet* 359: 281-86

5. De inhoud van iemands koelkast zegt meer over zijn of haar persoonlijkheid dan de inhoud van diens boekenkast.

6. Alle vegetariërs houden van planten.

7. Stier Herman is dood beter af dan als levend museumstuk.

Bionieuws 1, 2002

8. Men leest geen boek, maar wordt door een boek gelezen.

9. Verdere privatisering van energie bedrijven leidt tot kortsluiting.

10. *In echte Fries is net snotterig.*

Stellingen behorend bij het proefschrift:

"Recombinant gelatin and collagen from methylophilic yeasts"

Eric C. de Bruin

Wageningen, 15 maart 2002

# Voorwoord

In dit voorwoord wil ik graag de mensen bedanken waarmee ik tijdens mijn promotie onderzoek samengewerkt heb. Ten eerste mijn promotor Colja Laane. Colja bedankt dat je mijn promotor wilde zijn en wilde blijven na je vertrek naar DSM. Ten tweede mijn copromotor Frits de Wolf. Frits bedankt dat je me de kans bood om dit onderzoek te verrichten en voor je goede begeleiding. Marc ook jij bedankt voor je goede ideeën, je betrokkenheid bij het project en dat je mijn paranimf wilde zijn. Jan, Gustaaf en Liesbeth bedankt voor jullie praktische inzet voor mijn onderzoek. Verder bedank ik de mensen van ATO-Bioconversie, Miriam, Pieternel, Marcia, Jacquélien, Henk, Kenneth, Gerrit, Jean, Gert, Gustaaf, Pieter, Twan, Hans, Giel, Andrea, Lolke, Jan, Martin, Nico, Bwee, Felix, Truus, Hetty, Ruud, Wouter en Emil, andere en ex-collega's van het ATO (te veel om op te noemen) voor hun belangstelling en gezelligheid. Ook alle mensen van Biochemie bedankt, met name Laura, Carlo, Willem en Harrold. Klaas Nico (uit de groep van Prof. Veenhuis uit Groningen) bedankt voor je "*Hansenula* hulp".

Verder wil ik mijn ouders, familie en vrienden bedanken voor hun interesse en steun en ja, ik ben nu eindelijk klaar met m'n studie.

Arianne, bedankt voor alles.

A handwritten signature in black ink, appearing to read "Emil". The signature is written in a cursive style and is underlined with a single horizontal stroke.



# Contents

## Abstract

## Voorwoord

<b>Chapter 1</b>	<b>General introduction</b>	<b>1</b>
	1.1 Collagens	
	1.2 Collagen biosynthesis	
	1.3 Applications of collagen	
	1.4 Applications of gelatin	
	1.5 Recombinant protein production in yeasts	
	1.6 Methylotrophic yeasts	
	1.7 Molecular biology	
	1.8 <i>Hansenula polymorpha</i> versus <i>Pichia pastoris</i>	
	1.9 Aim and outline thesis	
<b>Chapter 2</b>	Expression and secretion of human $\alpha 1(\text{I})$ procollagen fragment by <i>Hansenula polymorpha</i> as compared to <i>Pichia pastoris</i>	<b>19</b>
<b>Chapter 3</b>	Secreted production of a custom-designed, highly hydrophilic gelatin in <i>Pichia pastoris</i>	<b>29</b>
<b>Chapter 4</b>	Endogenous prolyl 4-hydroxylation in <i>Hansenula polymorpha</i> and its use for the production of hydroxylated recombinant gelatin	<b>49</b>
<b>Chapter 5</b>	Human collagen trimers from the methylotrophic yeast <i>Hansenula</i> <i>polymorpha</i>	<b>67</b>
<b>Chapter 6</b>	<b>General discussion</b>	<b>81</b>
	6.1 Endogenous collagen-like proteins of <i>Hansenula polymorpha</i>	
	6.2 Trigger of the prolyl 4-hydroxylase activity	
	6.3 Production of hydroxylated gelatin in <i>Pichia pastoris</i>	
	6.4 Prolyl 4-hydroxylase activity measurements	
	6.5 Fungal prolyl 4-hydroxylases	
	<b>Summary / Samenvatting voor niet-biotechnologen</b>	<b>91</b>
	<b>References</b>	<b>95</b>
	<b>List of publications</b>	<b>105</b>
	<i>Curriculum vitae</i>	<b>107</b>
	<i>Addendum</i>	<b>109</b>

# Chapter 1

## General introduction

### **Abstract**

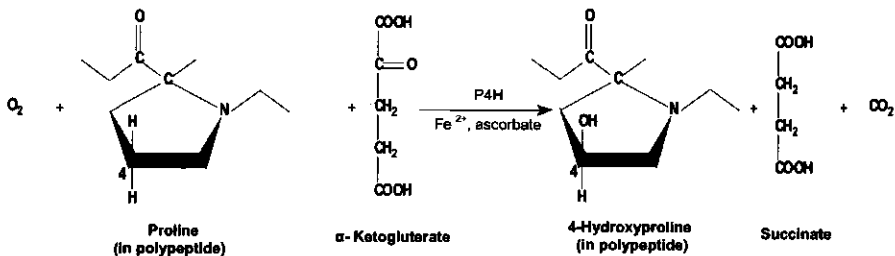
**This first chapter gives background information on the separate parts of this dissertation entitled “Recombinant gelatin and collagen from methylotrophic yeasts”. It starts with a description of what collagens and gelatins are, how these proteins are produced, for which applications these proteins are presently used and why recombinant production of gelatin and collagen is desirable.**

**Subsequent, an introduction is given in methylotrophic yeasts and how these organisms can be modified to use them as cell factories for the production of recombinant proteins. Finally, in the end of this chapter the objectives of this thesis are described.**

## 1.1 Collagens

The collagens, a family of extracellular matrix proteins, are the most abundant proteins in mammals accounting for up to 30 % of all proteins. To date 21 different types of collagens have been identified (Fitzgerald and Bateman, 2001). The most commonly occurring collagens are Type I, II and III, which form characteristic fibrils of up to several millimeters in length. These fibrils are the basis of connective tissues such as bone, tendon, skin, cartilage and blood vessels. The basic structure of fibrillar collagen is three individual polypeptide chains ( $\alpha$ -chains), which form a triple helical molecule. The characteristic  $\alpha$ -chain consists of G-X1-X2 amino acid repeats, where X1 and X2 can be any amino acid but are frequently the imino acids proline and 4-hydroxyproline, respectively (Kielty *et al.*, 1993, Prockop and Kivirikko, 1995).

The synthesis of the aminoacid hydroxyproline from proline is unique in that the reaction occurs after the proline has been incorporated into the protein. Prolyl hydroxylase, which catalyses this post-translational modification, is considered a dioxygenase because one oxygen atom is incorporated into proline to form hydroxyproline, while the second oxygen atom is used to convert  $\alpha$ -ketoglutarate to succinate. Ferrous iron and ascorbate (Vitamin C) are required as cofactors, but do not participate directly in the redox reaction (Kivirikko and Myllylä, 1982) (Fig. 1.1).



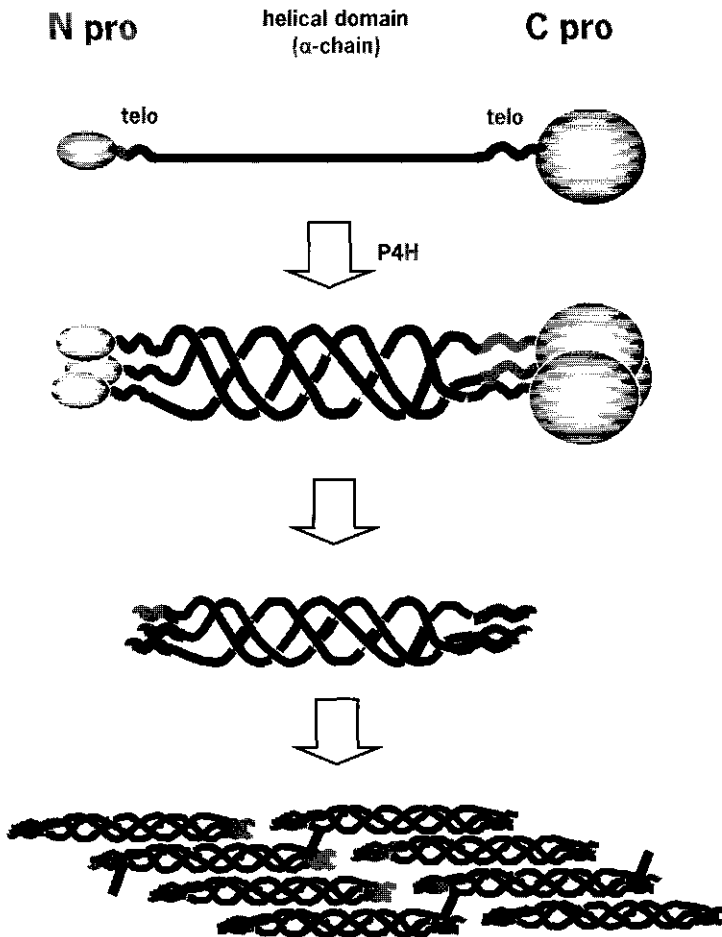
**Figure 1.1.** Prolyl 4-hydroxylase(P4H) uses molecular oxygen for the addition of one oxygen to a peptide-bound proline to form 4-hydroxyproline, while the second oxygen is used to convert  $\alpha$ -ketoglutarate to succinate. Ferrous iron and ascorbate are required as co-factors.

## 1.2 Collagen biosynthesis

The biosynthesis of collagens by animal fibroblasts is a complex process involving a number of different post-translational modifications, including glycosylation, disulphide-bond formation and lysyl- and prolyl-hydroxylation. Prolyl-hydroxylation of specifically the prolines in the X2 position to 4-hydroxyprolines is performed in the endoplasmic reticulum of the cells by prolyl 4-hydroxylase (EC 1.14.11.2). This reaction is essential for the formation of thermostable triple helices (Kivirikko *et al.*, 1992).

Collagen is synthesized as depicted schematically in Fig. 1.2. Collagen  $\alpha$ -chains are synthesized as precursor pro  $\alpha$ -chains (procollagens), containing globular N and C propeptides. As the procollagen is co-translationally translocated across the membrane of the endoplasmic reticulum, hydroxylation of proline residues occurs, specifically at the X2 position within the G-X1-X2 repeats. Once the polypeptide chain is fully translocated into the lumen of the endoplasmic reticulum the C propeptide folds (Kielty *et al.*, 1993).

The C propeptides play a crucial role in the inter-molecular recognition and alignment of the individual triple helix-constituting procollagen chains. The formation of the triple helix in newly synthesized procollagen molecules is known to proceed in a zipper-like manner from the C- to the N-terminus (Bruckner *et al.*, 1978). After assembly, the propeptides keep the triple helical procollagen soluble during its passage through the cell. After secretion, the procollagen is processed by specific extracellular proteases, to yield trimeric collagen, free of N and C propeptides (Fig. 1.2), which starts to assemble fibrils.



**Figure 1.2.** Collagen biosynthesis by fibroblasts; collagen is synthesized as procollagen subunits. After prolyl 4-hydroxylation (P4H) and pro-peptide mediated alignment, the helix-forming domains of the three subunits fold into a triple helix. After secretion, the propeptides are cleaved of by specific proteinases to yield trimeric collagen, which subsequent self-assembles into fibrils. After assembly covalent interchain crosslinks are formed in the course of time, which results in strength and stability of the matrix.

Fibril-forming collagens consist of one large central, triple helical domain flanked by short non-helical telo-peptides, which seem to ensure correct registration of the collagen molecules for self-assembly into well ordered fibrils (Leibovich and Weiss, 1970). However, a recent study indicates that this role is only seen in the kinetics of the process, as molecules lacking their entire telo-peptides form fibrils that are identical to those formed by full-length molecules (Nokelainen *et al.*, 2001). After self-assembly of collagen molecules into fibrils covalent interchain crosslinks are formed in the course of time, which results in strength and stability of the connective tissue matrices.

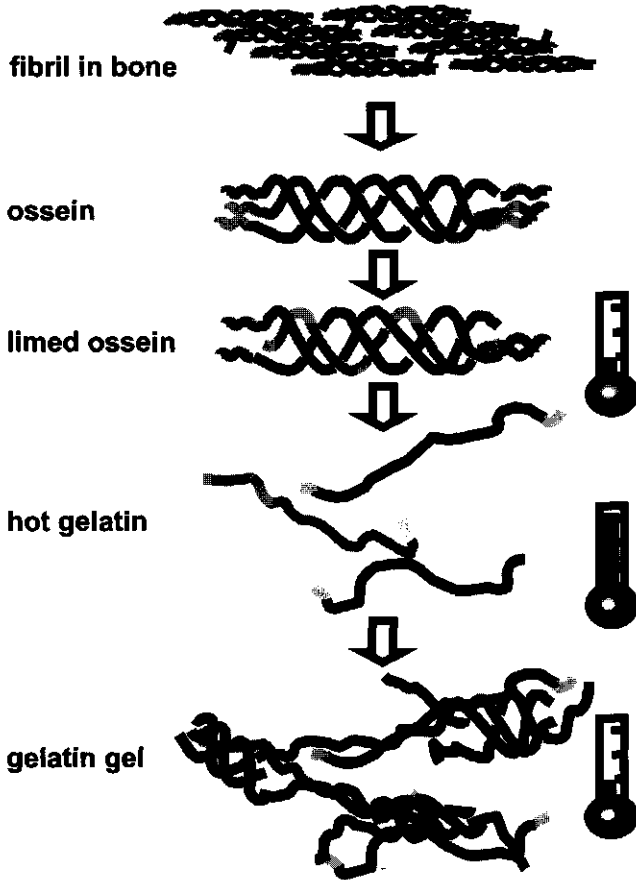
### **1.3 Applications of collagens**

Based on its structural role and compatibility within the body, collagen is a commonly used biomaterial in medical applications, such as cosmetic surgery, tissue engineering and wound treatment (Lee *et al.*, 2001). Collagen is regarded as one of the most useful biomaterials, due to its biological characteristics such as biodegradability, weak antigenicity and superior biocompatibility compared to other natural polymers. Human collagen, which is isolated from the human placenta, does not find wide application, because of the low availability of the source and the possible presence of disease-causing agents (e.g. human immunodeficiency virus). As a result of this, bovine collagen is used as a substitute for human collagen. The supply of bovine collagen, extracted from cowhides and tendons, is both abundant and inexpensive. However, bovine collagen can cause an inflammatory reaction in some people. Furthermore, concern has been raised about the potential contamination risk of bovine collagen with prions (bovine spongiform encephalopathy (BSE)), possibly causing the new variety of Creutzfeldt-Jakob disease in humans. As a result of this, efforts are increasing to develop an efficient and safe alternative (non-animal) production system of recombinant human collagen. In this thesis we describe routes to the production of human collagen by yeast systems.

## 1.4 Application of gelatin

Gelatin is denatured and partly degraded collagen, extracted from lime- or acid-treated bovine bones and skin (Fig. 1.3) (Asghar and Henrickson, 1982). Gelatin has unique physical and biological properties due to its repeating amino acid triplets of its collagen ancestor. These sequences are responsible for its ability to form helical regions in and between gelatin chains, thereby immobilising water. This results in the formation in a thermo-reversible gel (Fig. 1.3). Because of its gel-forming properties, its high water binding capacity and its retardation of ice- and sugar crystal formation, gelatin is widely used as a gelling agent in food substances (desserts, sweets, puddings, beverages, soups, etc.). Gelatin formulations are also used for the treatment of rheumatoid arthritis (Arborelius *et al.*, 1999). Gelatin is normally not allergenic and is used in various medical and pharmaceutical products such as intravenous infusions (Saddler and Horsey, 1987), matrix implants (Pollack, 1990), (injectable) drug delivery microspheres / capsules (Rao, 1995), vaccine / drug stabilisers and blood plasma expanders. Furthermore, gelatin finds its application in photographic industry as a film matrix material (Courts, 1980).

Despite its wide use, the suitability of traditional gelatin for some applications is limited. The above mentioned gelatin extraction procedure of traditional gelatins results in chemically modified peptide fragments, covering a wide range of molecular weights, which are not reproducible in today's batch-to-batch production processes. Recombinant gelatins may provide benefits in that the chemical composition and molecular weight can be precisely controlled and reproduced. Besides this, the mono-disperse gelatins, produced in a non-animal expression system does not bear the risk of associated infectious diseases such as BSE and viruses. Finally, for certain technical applications it would be beneficial to have gelatins with certain specific physico-chemical properties. When a suitable recombinant production system exists, these "tailor-made" gelatins, specially designed for a certain application, can be produced in large amounts.



*Figure 1.3. Schematic presentation of the gelatin preparation process from animal bones. After HCl-mediated decalcification, the obtained ossein is treated at extreme high pH to cleave the covalent cross-links between the separate collagen chains. When the material is subsequently heated, the denatured polypeptides lose their triple helical alignment and become soluble into hot water. Upon cooling, multiple helix variants are formed, which results in a thermo-reversible 3D-network of a gel.*



In this thesis we are exploring the suitability of two different yeast systems for the recombinant production of “natural” and “tailor-made” gelatins.

## 1.5 Recombinant protein production in yeasts

For the production of heterologous animal proteins yeasts offer certain advantages over bacteria. Since yeasts are eukaryotes most post-translational modifications, such as glycosylation, phosphorylation and formation of disulfide bonds, needed for functionality and/or stability of recombinant animal proteins, normally occur. However, prolyl 4-hydroxylation, essential for recombinant gelatin and collagen production, is generally considered to be absent in yeast systems (Kivirikko *et al.*, 1992, Bulleid *et al.*, 2000, Olsen *et al.*, 2001).

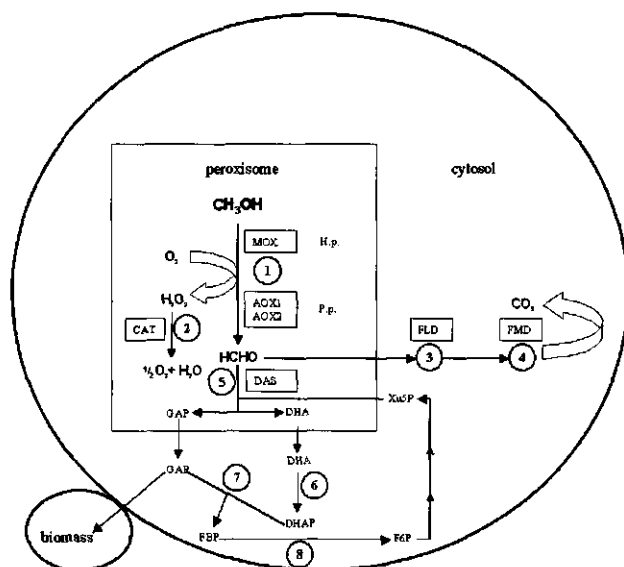
Yeasts can secrete proteins efficiently, which facilitates the isolation of the desired heterologous proteins and are well able to cope with highly repetitive coding sequences (Strausberg and Link, 1990), which is important in recombinant gelatin and collagen expression. In addition yeast, as an unicellular micro organism, retains the advantages of a bacterial system in the ease of manipulation and growth.

*Saccharomyces cerevisiae* (Baker’s yeast), from which the complete genome is known nowadays (Mewes *et al.*, 1997), was the first yeast being explored for the purpose of heterologous gene expression. Although there are some successful applications, the *S. cerevisiae* system shows some limitations (Romanos *et al.*, 1992, Sudbery *et al.*, 1996). These include, amongst others, instability of the production strains and undesired hyper-glycosylation of the secreted proteins. Furthermore, the yield of recombinant proteins is low in this system, because of the lack of strong promoters and relatively low cell densities during *S. cerevisiae* fermentations.

## 1.6 Methylotrophic yeasts

Methylotrophic yeasts are suitable alternative expression hosts, which exhibit some advantages over *S. cerevisiae* in above-mentioned aspects (Tschopp and Cregg, 1991, Sreekrishna and Kropp, 1996, Hollenberg and Gellissen, 1997, Gellissen, 2000, van Dijk *et al.*, 2000). Methylotrophic yeasts are able to utilise methanol as its sole carbon and energy source. All strains identified to date belong to only four genera:

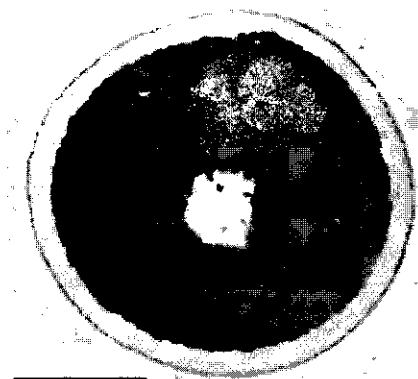
*Hansenula*, *Pichia*, *Candida* and *Torulopsis*, which can subsequently be divided into two groups: the asporogenous yeasts *Candida* and *Torulopsis* and the ascomycetous yeasts *Hansenula* and *Pichia*. The methylotrophic yeasts, which are isolated from soil, rotten fruits and the gut of insects, were initially considered to be good candidates for the production of single cell protein (SCP) from cheap carbon sources, like methanol. Advanced fermentation systems have been developed for these yeasts for the production of SCP from methanol. Studies on their physiology, biochemistry and ultra-structure revealed that these organisms share a specific methanol utilisation pathway (Egli *et al.*, 1980, van Dijken *et al.*, 1982, Veenhuis *et al.*, 1983). Methanol is oxidised to formaldehyde and hydrogen peroxide by alcohol oxidases (MOX and AOX1/AOX2 for *H. polymorpha* and *P.pastoris*, respectively), which are localised in specialised microbodies, the peroxisomes. In these organelles also catalase (CAT) is present, which decomposes the hydrogen peroxide into water and oxygen. The first step in the assimilatory (xylulose 5-phosphate) pathway to yield biomass is catalysed by dihydroxyacetone synthase (DAS), which is also present in the peroxisome. Formaldehyde generated by the alcohol oxidase reaction also enters the cytosolic dissimilatory pathway, via formaldehyde dehydrogenase (FLD) and formate dehydrogenase (FMD), to generate energy (Fig. 1.4).



**Figure 1. 4.** Methanol utilization pathway; 1 Methanol is oxidized by alcohol oxidases (MOX for *H. polymorpha* (H.p.), AOX1 and AOX2 for *P. pastoris* (P.p)) to generate formaldehyde and hydrogen peroxide. 2 The toxic hydrogen peroxide is converted to water and oxygen by catalase (CAT) 3 and 4 Formaldehyde generated by the alcohol oxidase reaction enters the cytosolic dissimilatory pathway, via formaldehyde dehydrogenase (FLD) and formate dehydrogenase (FMD), yielding energy and carbon dioxide. 5 For assimilation the formaldehyde reacts with xylulose-5 phosphate (Xu5P) by the action of dihydroxyacetone synthase (DAS) to generate the C3 compounds glyceraldehyde-3- phosphate (GAP) and dihydroxyacetone (DHA). 6 DHA is phosphorylated by dihydroxyacetone kinase to dihydroxyacetone phosphate (DHAP). 7 GAP and DHAP yield in an aldolase reaction fructose 1,6 biphosphate (FBP). 8 In further steps fructose 6- phosphate (F6P) and finally Xu5P is generated to enter the pathway again.

A striking feature of methanol-grown yeast is the excessive proliferation of the peroxisomes (Fig. 1.5). In methanol-limited continuous cultures the organelles can occupy up to 80 % of the total cell volume. In such cells alcohol oxidase and DAS may constitute over 60 % of the total cell protein, which illustrate that the genes encoding these proteins are controlled by very strong promoters. Also, the expression of the genes encoding FLD and FMD are under the control of strong inducible

promoters. All of these promoters represent suitable genetic components that can be applicable for the control of inducible heterologous gene expression in systems based on the respective methylotrophic hosts.



*Figure 1. 5. Micrograph of *H. polymorpha* wild-type cell, grown on methanol containing medium, showing two peroxisomes (P). Bar represents 0.5  $\mu\text{m}$ . (Kindly provided by Professor Marten Veenhuis, University of Groningen, the Netherlands).*

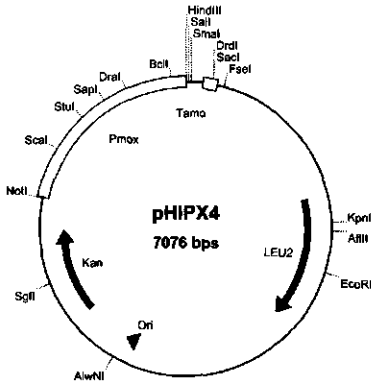
Other promoters (aside from those derived from methanol metabolism genes) can also be utilised for constitutive recombinant gene expression in the various methylotrophic yeast species. Typical examples of these promoters are *Pgap* (glyceraldehyde 3-phosphate dehydrogenase, Waterham *et al.*, 1997), *Ptef1* (transcription elongation factor, Kiel *et al.*, manuscript in preparation) and *Ppma1* ( $\text{H}^+$ -pumping ATPase, Cox *et al.*, 2000).

From all methylotrophic yeasts, *H. polymorpha* and *P. pastoris* have been studied most intensively. During the last decade molecular genetic methods have been developed for both organisms. The ability to modify *H. polymorpha* and *P. pastoris* at the molecular level has enabled basic, fundamental studies (unraveling the molecular mechanisms controlling the biogenesis of a specific cell organel, the peroxisome) as well as applied research (using the organisms as a host for the production of recombinant proteins).

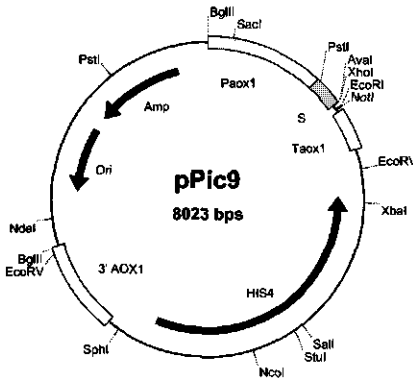
In this PhD study a comparison is drawn between *H. polymorpha* and *P. pastoris* with respect to recombinant gelatin and collagen expression. Therefore in the following part, the tools are described to genetically modify *H. polymorpha* and *P. pastoris* in order to use them as cell-factories, we note the differences between both organisms and focus on the recent developments in the field of recombinant gene expression in the two yeasts.

## 1.7 Molecular biology

Expression vectors are designed for recombinant protein expression in *H. polymorpha* (Fig. 1.6) and *P. pastoris* (Fig. 1.7), which have several common features. The expression cassette is one of those and is composed of a strong (methanol inducible) *H. polymorpha* or *P. pastoris* promoter (*P<sub>mox</sub>* and *P<sub>aox1</sub>* for vectors pHIPX4 and pPic9, respectively). The cassette is followed by one or more unique restriction sites for insertion of a heterologous gene, succeeded by the transcriptional termination sequence from a *H. polymorpha* or *P. pastoris* gene (*T<sub>mo</sub>* and *T<sub>aox1</sub>*, respectively), which directs efficient 3' processing and polyadenylation of the foreign mRNA. Vectors also include an auxotrophic marker gene (*LEU2 / HIS4*, respectively), or a dominant drug resistance marker gene for the selection of transformants. Because *Escherichia coli* is used plasmid amplification, the vectors also contain sequences required for plasmid replication and maintenance in *E. coli* (i.e. a replication origin and an antibiotic resistance gene).



**Figure 1. 6.** Plasmid map of *E. coli-H. polymorpha* shuttle vector pHIPX4 (Gietl et al., 1994) The expression vector contains the methanol oxidase (MOX) promoter (Pmox) / amino oxidase (AMO) terminator (Tamo) cassette and a LEU2 selectable marker. Furthermore the vector contains sequences required for plasmid replication and maintenance in *E. coli* (i.e. an origin of replication (Ori) and a kanamycin resistance gene (Kan)).



**Figure 1. 7.** Plasmid map of *E. coli-P.pastoris* shuttle vector pPic9 (Clare et al., 1991b). The expression vector contains the alcohol oxidase 1 (AOX 1) promoter (Paox1) / terminator (Taox1) cassette and a HIS4 selectable marker. For secretion of the recombinant product, the foreign gene is fused to the prepro leader region from the *S. cerevisiae*  $\alpha$ -mating factor (S). Furthermore, the vector contains sequences required for plasmid replication and maintenance in *E. coli* (i.e. an origin of replication (Ori) and an ampicillin resistance gene (Amp)).

Transformation procedures for introducing foreign DNA into *H. polymorpha* and *P. pastoris* are similar to what is described for other yeasts, including *S. cerevisiae*. They can be transformed using whole-cell methods with addition of PEG (Dohmen *et al.*, 1991), or by electroporation (Becker and Guarente, 1991, Faber *et al.*, 1994). Prior to transformation, the expression vectors are linearized by restriction enzymes in the *Pmox*, and in the *Paox1* or *HIS4* for pHIPX4 and pPIC9 respectively. This enhances the integration of the plasmid DNA at the homologous locus in the genome of *H. polymorpha* and *P. pastoris* (in the *MOX*-locus and the *AOX1*- or *HIS4*-locus, respectively), generating stable transformants. In most recombinant *H. polymorpha* strains foreign DNA is randomly integrated using HARS (*H. polymorpha* autonomously replicating sequence)-harboring circular plasmids, resulting in a variety of individual, mitotically stable strains containing single to multiple copies of the expression cassette (Faber *et al.*, 1992, Sohn *et al.*, 1996).

In recombination of *P. pastoris*, linear DNA can also generate stable transformants via a double crossover event between the *AOX1* promoter and the 3' *AOX1* region of the vector (Fig. 1.6) and the genome. This transplacement of the *AOX1* gene yields a phenotype that grows slowly on methanol (*Mut<sup>S</sup>*) caused by the loss of alcohol oxidase activity, because this strain is dependent on the transcriptionally weaker *AOX2* gene (Cregg *et al.*, 1989). This phenotype may give higher intracellular production of heterologous proteins than the *Mut<sup>+</sup>* phenotype, which has functional *AOX1* and *AOX2* genes and grows on methanol at wild-type rate. Selection of *P. pastoris* transformants is made by complementation of auxotrophic marker genes, such as *ADE1*, *URA3*, *ARG4* or *HIS4*, in the corresponding mutant strains, which are deficient in adenine, uracil, arginine or histidine biosynthesis, respectively (Higgins and Cregg, 1998). A collection of all possible combinations of mutant marker genes with corresponding mutant strains is presented by Cereghino *et al.*, 2001. This in order to design multi-gene systems for the production of foreign proteins that are heteromers, or that require a post-translational modification that is not or insufficient present in *P. pastoris*. For *H. polymorpha* the auxotrophic markers *URA3* and *LEU2* are present in combination with the respective uracil and leucine

deficient *H. polymorpha* strains. A screening method has been developed for selecting multicopy integrants in both organisms making use of zeocin or G418 resistance (Scorer *et al.*, 1994, Sohn *et al.*, 1999). The level of resistance of the strain to these antibiotics appeared to be proportional to the number of plasmids integrated in the genome.

Recombinant proteins produced by both expression systems can be accumulated in the cytosol, or secreted into the culture medium. For secretion of the proteins in *H. polymorpha* and *P.pastoris* the signal sequence of the *S. cerevisiae*  $\alpha$ -factor mating pheromone is frequently used. In some cases, secretion of heterologous proteins can also be driven by their own, heterologous signal sequence (see Chapter 2). Both *H. polymorpha* and *P. pastoris* grow on simple mineral media to very high cell-densities and do not secrete high amounts of endogenous proteins. Therefore the recombinant protein, when secreted, can be recovered from the culture medium in a relatively pure state (80-90 %). The tendency to hyper-glycosylate secreted proteins in methylotrophic yeasts is usually less in methylotrophic yeast than in *S. cerevisiae*. However, examples of hyper-glycosylated heterologous proteins have been described for both *P. pastoris* (Scorer *et al.*, 1993, Letourneur *et al.*, 2001) and *H. polymorpha* (Fellinger *et al.*, 1991). Glycosylation, or proteolytic degradation of foreign proteins may be prevented by targeting of the proteins into the peroxisomes of the yeasts. These microbodies can also be used as a storage compartment of cell-toxic proteins (Subramani, 1992, Faber *et al.*, 1996).



### 1.8 *H. polymorpha* versus *P. pastoris*

Despite their obvious relationship and physiological similarities, there are also important differences between the two species. There is a difference with respect to the carbon-source-dependent regulation of enzyme synthesis. While *P. pastoris* AOX1 promoter activation strongly depends on the presence of methanol, high levels of expression in *H. polymorpha* have been observed using the *MOX* promoter with glycerol or glucose in limiting amounts. The ability to express foreign genes, without the need of adding methanol, which is a highly toxic and explosive substance, is a unique feature of the *H. polymorpha* expression system. However, the tightly regulated *AOX1* promoter of *P. pastoris* may be profitable when cell-toxic proteins are expressed.

*H. polymorpha* has some specific advantages over *P. pastoris* being more thermo-tolerant. The relatively high optimal growth temperature for *H. polymorpha* ( $T_{opt}$ : 37-43 °C vs. 30 °C for *P. pastoris*) may be favourable for the production of mammalian (including human) proteins and, furthermore, has the advantage that it allows a better cooling management in large-scale fermentations. Finally, the growth rate of *H. polymorpha* is higher on methanol, which reduce both the risk of contamination and the fermentation time.

*H. polymorpha* and *P. pastoris* have become highly successful systems for the expression of heterologous genes. More than one hundred different proteins have been successfully produced in *P. pastoris* (Higgins and Cregg, 1998). *H. polymorpha* shows the same potential for high-level expression of foreign genes, although less examples have been described (van Dijk *et al.*, 2000). In Table I lists some examples of recombinant protein production in both expression systems.

**Table I.** Expression yields of foreign proteins in *Hansenula polymorpha* and *Pichia pastoris*.

<i>H. polymorpha</i>	<i>P. pastoris</i>	g/L (mode*)	Product (donor)	Reference
	x	22.0 (I)	Hydroxynitrile lyase (rubber tree)	Hasslacher <i>et al.</i> , 1997
	x	14.8 (S)	Non hydroxylated gelatin (mouse)	Werten <i>et al.</i> , 1999
x		13.5 (S)	Phytase ( <i>Aspergillus</i> )	Mayer <i>et al.</i> , 1999
	x	12.0 (I)	Tetanus toxin C ( <i>C. tetani</i> )	Clare <i>et al.</i> , 1991a
x		4.2 (I)	$\alpha$ -galactosidase (guar)	Giuseppin <i>et al.</i> , 1993
	x	0.4 (S)	$\alpha$ -galactosidase (coffee bean)	Zhu <i>et al.</i> , 1995
x		2.0 (S)	Hirudin (leech)	Weydemann <i>et al.</i> , 1995
	x	1.5 (S)	Hirudin (leech)	Rosefeld <i>et al.</i> , 1996
x		1.3 (S)	Serum albumin (human)	Kang <i>et al.</i> , 2001
	x	4.0 (S)	Serum albumin (human)	Sreekrishna <i>et al.</i> , 1996
	x	0.2 (S)	Endostatin (human)	Boehm <i>et al.</i> , 1999
x		0.1 (I)	Collagen (human)	Chapter 5
	x	0.5 (I)	Collagen (human)	Nokelainen <i>et al.</i> , 2001
x		0.35 (S)	Aprotinin (bovine)	Zurek <i>et al.</i> , 1996
	x	0.15 (S)	Aprotinin (bovine)	Vedvick <i>et al.</i> , 1991

\* I = Intracellular; S = Secreted

Table I shows that both expression systems generally produce heterologous, secreted proteins with yields of 0.2-15 g/L and that intracellular expression levels range from 0.1-22 g/L. This wide range in the expression yields may be due to the fermentation strategies used, the copy numbers of the heterologous gene and the properties of the target proteins in the individual expression machinery.

## 1.9 Aim and outline of this thesis

The objective of the study described in this thesis is to explore the methylotrophic yeasts *Hansenula polymorpha* and *Pichia pastoris* for their use as recombinant production systems of “natural” and “tailor-made” gelatins and of human collagen. This, in order to develop a safe microbial production system for natural gelatin and human collagen in the future and to develop a production system for non-natural, tailor-made gelatins with entirely novel physico-chemical properties, beneficial for certain technical applications. In Chapter 2 a comparison is made between both expression systems for the production of a human  $\alpha 1(I)$  collagen fragment. Secretion of the fragment was possible in *H. polymorpha*, using the native human collagen signal sequence. In Chapter 3, we showed that in *P. pastoris* production of non-hydroxylated, tailor-made gelatins, with new characteristics more beneficial for a certain technical application, could be produced at high levels.

Prolyl 4-hydroxylation, which is essential for the production of recombinant gelatin with gelling properties, was generally considered to be absent in yeast. However, in this study (Chapter 4) it was discovered that *H. polymorpha* possesses endogenous collagen-like sequences, containing 4-hydroxyproline. This indicated that this yeast must contain endogenous prolyl 4-hydroxylase activity. Expression of recombinant animal gelatin in this yeast yielded a secreted and hydroxylated product, without the need of co-expression of heterologous prolyl 4-hydroxylase genes. In Chapter 5, we investigated if *H. polymorpha* could be used for the production of human collagen trimers expressed. Finally, Chapter 6 presents a general discussion followed by a summary of the results obtained in this study.

## Chapter 2

### Expression and secretion of human $\alpha 1(\text{I})$ procollagen fragment by *Hansenula polymorpha* as compared to *Pichia pastoris*

Eric C. de Bruin, Frits A. de Wolf and Colja Laane

Published in: *Enzyme and Microbial Technology* 26: 640-644 (2000).

#### Abstract

Secretion of a fragment of the human collagen  $\alpha 1(\text{I})$  chain was achieved in *Hansenula polymorpha* using the native human  $\alpha 1(\text{I})$  procollagen secretory signal sequence. The N-terminal propeptide of the fragment was cleaved off during or after secretion, yielding the N-terminus of mature  $\alpha 1(\text{I})$  collagen. In *Pichia pastoris* transformants, the expression of the fragment could be detected on RNA-level, but the product could not be determined in the culture medium. After fusion of the fragment with a *myc*-HIS6 epitope, the intact product was found intracellularly. The difference in extracellular level of protein between the two expression hosts is most likely caused by difference in secretion efficiency.

## Introduction

Collagen, the most abundant protein found in the human body, provides structural support to bone, skin, tendons and blood vessels. In view of its function as a fibrillar, structural protein, there is a high demand of collagen and collagen-derived products for biomedical applications. Recently, the application of collagen from human and bovine sources has become subject to novel risks of contamination with disease-causing agents such as HIV and BSE, respectively. Thus, several groups are now studying the heterologous expression of collagen in safer hosts (Geddis and Prockop, 1993, Lamberg *et al.*, 1996). For the production of heterologous eukaryotic proteins yeasts are frequently used (Romanos *et al.*, 1992, Gellissen and Hollenberg, 1997). Apart from *Saccharomyces cerevisiae*, the most successful yeasts for heterologous protein expression are the methylotrophic yeast species *Hansenula polymorpha* and *Pichia pastoris* (Hollenberg and Gellissen, 1997). In a recent study of the production of triple helical human collagen type III in *P. pastoris* the extracellular production yield was very low. Although the native human type III collagen signal sequence was used for secretion, almost all product was found inside the cell (Vuorela *et al.*, 1997). To facilitate the isolation of recombinant collagen, secretion of the product is desired. In this study, we compared the secretion of a collagen type I ( $\alpha 1$ ) fragment in *P. pastoris* and *H. polymorpha* using the native human collagen type I ( $\alpha 1$ ) secretory signal sequence (pre-peptide). Our results show that *H. polymorpha* can use this sequence for secreting the fragment into the medium with a good yield.

## Materials and methods

### *Yeast strains and plasmids*

*H. polymorpha* strain NCYC495 [*leu1.1* (Gleeson *et al.*, 1986)] was used as a host for recombinant collagen production. The plasmid pHIPX4 (Gietl *et al.*, 1994), contains the *LEU2* selectable marker and an expression cassette of the methanol oxidase (*MOX*) promoter and the amino oxidase (*AMO*) terminator. *P. pastoris* strain GS115 [*his4*; (Cregg *et al.*, 1985)], ZEOCIN and the expression vector pPICZB were obtained

from Invitrogen (San Diego, CA, USA). pPICZB contains a ZEOCIN selectable resistance marker and the alcohol oxidase (*AOX1*) promoter/terminator cassette. For cloning of DNA restriction fragments, plasmid pMTL23 (Chambers *et al.*, 1988) was used. *Escherichia coli* strain XL1-blue (Stratagene, La Jolla, CA, USA) was used for vector construction. All molecular techniques were performed as described by Sambrook *et al.*, 1989 or according to manufacturers protocols.

### ***Media composition***

Minimal glucose plates, for selection of transformants, contained 1.34 % yeast nitrogen base without amino acids (Difco, Detroit, MI, USA), 1% glucose and 1.5 % agar. Fermentation basal salts medium contained per liter: 26.7 ml of phosphoric acid (85 %), 0.93 g calcium sulfate dihydrate, 18.2 g potassium sulfate, 14.9 g magnesium sulfate heptahydrate, 4.13 g potassium hydroxide and 4.3 ml of trace elements. Trace elements contained per litre; 4.5 g cupric chloride dihydrate, 0.09 g potassium iodide, 3.5 g manganese chloride tetrahydrate, 0.2 g sodium molybdate dihydrate, 0.02 g boric acid, 1.08 g cobalt sulfate heptahydrate, 42.3 g zinc sulfate heptahydrate, 65.0 g ferrous sulfate heptahydrate, 0.6 g thiamine, 0.2 g biotin and 5.0 ml sulfuric acid (96 %).

### ***DNA manipulation***

A 780 bp fragment of human  $\alpha 1(I)$  procollagen cDNA (Geddis and Prockop., 1993), encoding the pre-peptide, the N-terminal propeptide and a part of the helical domain, was cloned in pMTL23. Isolation of plasmid DNA from *E. coli* was performed using the Wizzard system (Promega, Madison, WI, USA). The fragment was digested from the resulting plasmid with *Xho* I and subsequently cloned in the *Sal* I site of pHIPX4 and in the *Xho* I site of pPICZB. DNA fragments from agarose gels were isolated using the Qiaquick columns from Qiagen (Chatsworth, CA, USA).

***Transformation of H. polymorpha***

Transformation of *H. polymorpha* was performed according to Faber *et al.*, 1994. After growth on minimal glucose plates at 37 °C for 3 days, several colonies were selected for PCR confirmation. Cells were used directly for PCR without any pretreatment using the MOX-primer: 5'-ACCCACTGCTGTCTGTGCCTGC-3' and the AMO-primer: 5'-TGTCCTTGGTCTCCTTGTGCACG-3'.

***Transformation of P. pastoris***

*P. pastoris* GS115 has a nonfunctional histidinol dehydrogenase gene (*HIS 4*). In order to enable endogenous production of histidinol dehydrogenase the gene was complemented as described by Munshi and Lee, 1997. *P. pastoris* was transformed according to Becker and Guarente, 1991 using a GenePulser (Bio-Rad, Richmond, CA, USA) set at 1500 V, 25  $\mu$ F and 200  $\Omega$  and using 0.2 cm cuvettes. After growth on minimal glucose plates with 100  $\mu$ g/ml ZEOCIN at 30 °C for 3 days, several colonies were selected for PCR confirmation. Cells were used directly for PCR without any pretreatment using the commercial vector primers 5'AOX1 and 3'AOX1 (Invitrogen)

***Small-scale expression in shake flasks***

Transformants verified by PCR were grown overnight in buffered minimal glucose medium for *H. polymorpha* and buffered minimal glycerol medium for *P. pastoris*, harvested and resuspended in buffered minimal methanol medium. Cells were grown in shaking flasks for 4 days at 37 °C and 30 °C for *H. polymorpha* and *P. pastoris*, respectively. Methanol was added to 0.5 % every day and cell samples were taken daily. Using the standard hot phenol method isolated RNA of cells. Northern blot analysis was performed using the Dig-system (Boehringer Mannheim, Indianapolis, IN, USA) with the exception that glyoxal gels were used (Mc Master and Carmichael, 1977). For western blot analysis cells were lysed by using the Y-per reagent of Pierce, Rockford, IL, USA.

***Productions in a bioreactor***

Fed-batch fermentation of transformants was performed in a 1-L bioreactor (Applikon, Schiedam, The Netherlands), essentially according to the *P. pastoris*

manual of Invitrogen. At the start of the fermentation, the fermenter contained 500 ml fermentation basal salts medium, to which 5g of casamino acids (Merck, Darmstadt, Germany) were added to reduce extracellular proteolysis (Clare *et al.*, 1991a). Glucose is reported to be a repressing substrate for the *MOX* promoter of *H. polymorpha*, whereas glycerol is a derepressing carbon source (Mayer *et al.*, 1999). Glycerol for *P. pastoris* instead, repress the *AOX1* promoter and is mainly used as carbon source for biomass production during fermentation. To compare both expression systems properly, in the *H. polymorpha* fermentation glucose and in the *H. polymorpha* fermentation glycerol was used as the carbon source. In the batch phase of the *H. polymorpha* fermentation 60 g/l glucose was used, which resulted in the same final wet cell weight/l as the initial 40 g/l glycerol used in the batch phase of the *P. pastoris* fermentation. The temperature was set at 37 °C for *H. polymorpha* and 30 °C for *P. pastoris*, agitation at 500 rev./min. and aeration rate at 1 vvm. The pH was adjusted to pH 5.0 with ammonium hydroxide (25 %). The fermenter was inoculated with 50 ml of a culture grown in minimal glucose medium. The batch culture was grown until the glucose or glycerol was completely consumed. After this, aeration and agitation were increased to 2 vvm and 1000 rpm, respectively and a glucose or glycerol fed-batch phase was initiated by feeding a 50 % solution, containing 12 ml/l trace salts at a rate of 10 ml/h. The pH was maintained at 5.0 by the addition of 25 % ammonium hydroxide. The glucose feed was stopped when the cell wet weight reached about 180 g/l. After complete consumption of carbon-source, 5 g of casamino acids were added and the culture was induced by feeding 100 % methanol containing 12 ml/l trace salts. The feed rate was initially set at 1 ml/h and was gradually increased to maximally 8 ml/h. Throughout the fermentation 2 ml culture samples were taken and spun at 20 000 X g. for 1 min., after which the wet cell pellets were weighed. The supernatants were filtered using disposable 0.22 µm filters.

#### ***SDS-PAGE and N-terminal protein sequencing***

SDS-PAGE (Laemmli, 1970) was performed in a Mini PROTEAN II system (Bio-Rad) under reducing denaturing conditions. For N-terminal protein sequencing, protein was blotted onto Immobilon PSQ (Millipore, Bedford, MA, USA) by applying



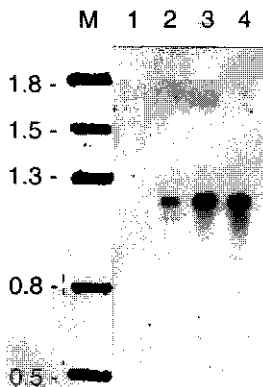
100 V for one hour in a Mini Trans-Blot Cell (Bio-Rad). Transfer buffer was 2.2 g CAPS per liter of 10 % methanol, pH 11. Blots were stained with Coomassie Brilliant Blue (CBB R-350, Pharmacia, Uppsala, Sweden) and selected bands were cut out. N-terminal sequencing was performed using Edman degradation.

For western blot analysis the filter was blocked with 5 % skim milk powder in TBST (0.1 M Tris -HCL, pH 7.5; 1.5 M NaCl; 0.1 % Tween-20) at room temperature for 1 h. The filter was incubated overnight with monoclonal anti-*myc* antibody (Invitrogen; 1 : 20 000 in 1 % skim milk in TBST). The filter was washed with TBST and then rinsed with Lumi-Light<sup>PLUS</sup> (Boehringer Mannheim) and exposed on a molecular light imager.

## Results and Discussion

### *Northern blot analysis*

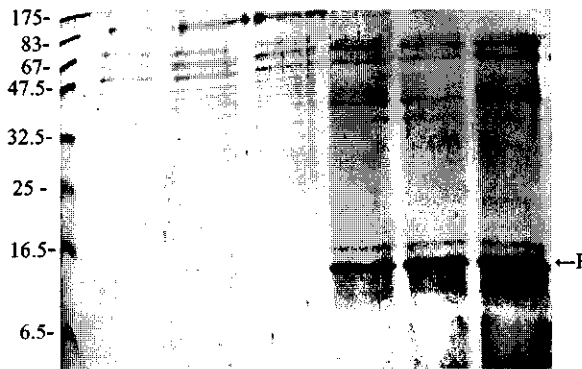
PCR positive transformants of *H. polymorpha* and *P. pastoris* were grown for four days in shaking flasks containing minimal glucose and minimal glycerol medium, respectively, with methanol being added to 0.5 % on the second and the third day. Every day a sample was taken and RNA was isolated. Northern blot analysis was performed to examine expression on RNA-level. Methanol induced expression of the collagen fragment on RNA-level was shown for transformants of both *P. pastoris* (Fig. 2.1) and *H. polymorpha* (data not shown). The sizes of the RNA bands in both hosts were about 1.2 kb, which corresponded to the size of the mRNA expected.



**Figure 2.1** Northern-blot analysis of the RNA isolated from the *P. pastoris* transformant containing the collagen type I cDNA fragment in the AOX1 promoter/terminator expression cassette. Lane M: RNA markers (Life Technologies, Rockville, MD, USA). Lane 1: growth on glycerol. Lane 2: after one day of methanol induction. Lane 3: after two days on methanol. Lane 4: after three days on methanol.

**Expression of human  $\alpha 1(I)$  procollagen chain fragment**

*H. polymorpha* and *P. pastoris* transformants expressing the collagen fragment on RNA-level were grown in a bioreactor as described in the material and method section. Protein production in the supernatant was followed in time by SDS-PAGE analysis (Fig. 2.2).



**Figure 2.2.** SDS-PAGE analysis of extracellular proteins produced in the fermentation broth of the *H. polymorpha* transformant containing the collagen type I cDNA fragment in the MOX promoter/ AMO terminator expression cassette. Lane M: pre-stained markers (Biolabs, Northbrook, IL, USA). Lane 1-3: 10  $\mu$ l of culture supernatant before methanol induction, harvested after 12, 18 and 24 hours of fermentation, respectively Lane 4-6: 10  $\mu$ l of culture supernatant after methanol induction, harvested after 36, 42 and 48 hours of fermentation, respectively. P: 14-kDa collagen product.

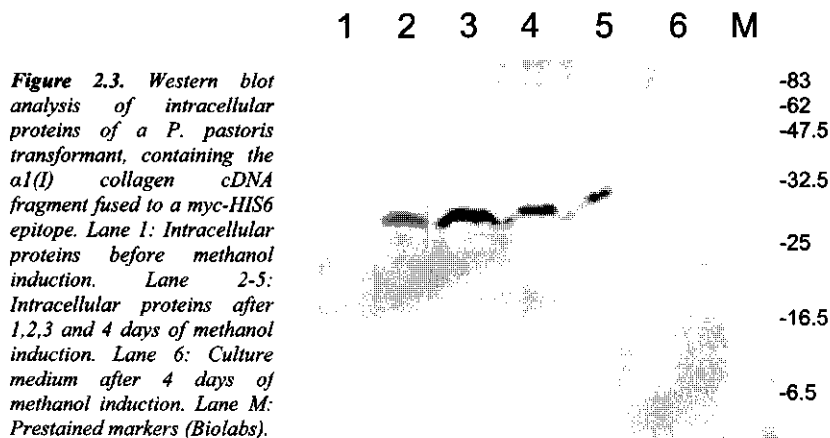
A protein with an apparent weight of 14 kDa on SDS-PAGE was secreted after methanol induction of the *H. polymorpha* transformant. The product was not purified, but the quantity was estimated by SDS-PAGE, using recombinant gelatin as a reference (Werten *et al.*, 1999) The estimated production level was 0.6 g/L in the clarified fermentation broth.

The size of the protein corresponds to the size of the helical part of the collagen fragment without the N- terminal propeptide attached. Lysis of cells in the bioreactor could be excluded, because the methanol oxidase, which is the most abundant protein inside the cell, was not found extracellularly in high amounts. This was in agreement with microscopic observations.

In *H. polymorpha* cells lysed with the Y-per reagent, the collagen product could not be detected by SDS-PAGE with CBB staining. Accumulation to high intracellular concentrations did not occur as far as could be deduced from the protein gel. Thus, it can be concluded that efficient secretion of a collagen chain fragment can be obtained with *H. polymorpha* by using the native human  $\alpha 1(I)$  collagen secretory signal.

During fermentation of a *P. pastoris* transformant, extracellular collagen production could not be determined by SDS-PAGE with CBB or silver staining. Extracellular degradation could be excluded based on earlier studies (Werten *et al.*, 1999). Probably, the product remains inside the cell.

To detect the product inside the cell the  $\alpha 1(I)$  collagen fragment was C-terminally tagged with a *myc*-HIS6 epitope. After small-scale expression and lysis of the *P. pastoris* cells, western blot analysis was performed (Figure 2.3).



After methanol induction the *myc*-antibody reacts with a 30 kDa protein intracellular protein (Lane 2-5). The size of this protein corresponds to the size of the intact  $\alpha 1(I)$  procollagen fragment. The quantity of the intracellular protein was not determined. Before methanol induction the intracellular product could not be detected (Lane 1). In the culture medium the *myc*-antibody did not react (Lane 6). Most likely, the native human  $\alpha 1(I)$  procollagen secretory signal sequence does not work for *P. pastoris*.

### ***N*-terminal sequencing**

*N*-terminal sequencing of the 14 kDa product in the fermentation medium of *H. polymorpha* yielded the sequence STGGISVPGPMG. This amino acid sequence corresponds to the *N* telopeptide, which is start of the helical domain of the collagen chain. Apparently, the *N*-terminal propeptide is cleaved off by an endogenous protease of *H. polymorpha* during, or after secretion. *H. polymorpha* can use the leader sequence (pre-peptide) of human  $\alpha 1(I)$  procollagen for secreting the recombinant  $\alpha 1(I)$  collagen fragment.

Surprisingly, we found only a difference of nine amino acid residues between the *N*-terminus of the produced  $\alpha 1(I)$  collagen fragment and the *N*-terminus of non-recombinant, mature human  $\alpha 1(I)$  collagen. From the latter the *N* propeptide is cleaved off by a specific extracellular procollagen *N*-proteinase (EC.3.4.24.14) (Colige *et al.*, 1995).

### **Conclusion**

The present data indicate that *H. polymorpha* can secrete a recombinant human  $\alpha 1(I)$  collagen fragment at a level of about 0.6 g/L by using the native human collagen pre-peptide and *N*-terminal propeptide. However, the human collagen pre-pro sequence could not be used to obtain secretion of recombinant collagen in *P. pastoris*. We have recently shown that secretion of animal type I gelatins can be obtained in *P. pastoris* by using the *S. cerevisiae*  $\alpha$ -mating factor prepro-signal, resulting in production levels of 2-3 g/L in the clarified broth (Werten *et al.*, 1999).

As expected amino acid analysis revealed that proline residues in the collagen product were not hydroxylated to hydroxyproline (data not shown). For production of triple helical collagens, hydroxylation of proline residues to hydroxyproline is essential. Because yeasts do not seem to contain prolyl hydroxylase activity, co-expression of a human prolyl 4-hydroxylase is needed to produce recombinant triple helical collagen.

Future studies are ongoing to evaluate the suitability of *H. polymorpha* for the secreted production of full-length, triple helical collagen type I.

### **Acknowledgements**

We thank D.J. Prockop, Thomas Jefferson University, Philadelphia, USA, for providing the COLIA1-CMV vector containing the cDNA of the  $\text{pro}\alpha 1$  chain of type I collagen and K.N. Faber, University of Groningen, Haren (Groningen), The Netherlands, for providing plasmid pHIPX4 and strain *H. polymorpha leu 1.1*. We thank M.W.T. Werten, ATO-DLO, Wageningen, The Netherlands for practical advice.

## Chapter 3

### Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*

Marc W.T. Werten, Wouter H. Wisselink, Tanja J. Jansen-van den Bosch,

Eric C. de Bruin and Frits A. de Wolf

Published in: *Protein Engineering* 14: 447-454 (2001).

#### Abstract

A custom-designed, highly hydrophilic gelatin was produced in *Pichia pastoris*. Secreted production levels in single-copy transformants were in the range of 3-6 g/l of clarified broth and purification to near homogeneity could be accomplished by differential ammonium sulfate precipitation. Despite the fact that gelatins are highly susceptible to proteolysis because of their unfolded structure, the recombinant protein was shown to be fully intact by SDS-PAGE, N-terminal sequencing, gel filtration chromatography and mass spectrometry. Owing to its highly hydrophilic nature, the migration of the synthetic gelatin in SDS-PAGE was severely delayed. Esterification of the carboxylic amino acid side chains resulted in normal migration. The high polarity of the synthetic gelatin also accounts for its negligible surface activity in water at concentrations up to 5% (w/v), as determined by tensiometry. Circular dichroism spectrometry showed that the non-hydroxylated gelatin did not form triple-helices at 4 °C. The spectrum was even more representative of the random coil conformation than the spectrum of natural non-hydroxylated gelatin.

## Introduction

Gelatin, in essence denatured and partially degraded collagen, is traditionally prepared by hot acid or alkaline extraction of animal bones and hides. Apart from its main use as a gelling agent in food (Asghar and Henrickson, 1982), gelatin is also used in medical and industrial applications such as intravenous infusions (Saddler and Horsey, 1987), matrix implants (Pollack, 1990), injectable drug delivery microspheres (Rao, 1995) and photographic film (Courts, 1980). Despite the diversity of current uses of natural gelatin, recombinant gelatins may provide benefits for specific applications, in that the chemical composition and molecular weight can be precisely controlled and reproduced (van Heerde *et al.*, 1999; de Wolf *et al.*, 2000). Furthermore, recombinant gelatins do not bear the risk of associated infectious diseases such as bovine spongiform encephalopathy (BSE).

Several reports have described the production of recombinant gelatin-like proteins in *Escherichia coli*. Analogously to the natural amino acid sequence of the collagen triple-helix forming domain, synthetic genes are constructed from repeating (Gly-Xaa-Yaa)<sub>n</sub>-encoding oligonucleotides, where Xaa and Yaa are often proline (Goldberg *et al.*, 1989; Obrecht *et al.*, 1991; Gardner *et al.*, 1993; Cappello and Ferrari, 1994). Gene instability problems are commonly observed with such highly repetitive genes (Capello and Ferrari, 1994). Also, expression levels usually obtained in *E. coli* are rather low and purification of the intracellularly produced protein can be difficult. Recently, Kajino *et al.*, 2000 reported the use of *Bacillus brevis* for the expression of gelatin-like proteins. They used sequence stretches selected from natural collagen genes and polymerized them to form semi-synthetic gelatin. Prior to their report, we reported the use of the methylotrophic yeast *Pichia pastoris* as a superior host for the secretion of recombinant gelatins having natural amino acid sequences, at up to 14.8 g/L of clarified broth (Werten *et al.*, 1999).

Having established the suitability of *P. pastoris* for the expression of natural recombinant gelatins, we set out to investigate the possibilities of producing entirely custom-designed gelatins having novel physico-chemical properties. A monomeric gene encoding a highly hydrophilic 9 kDa gelatin was designed such, as to allow

convenient polymerization into larger multimers. The monomeric gene is much longer than the single oligonucleotide monomers used in the expression of synthetic gelatins in *E. coli*, mentioned above. This offers more flexibility in the design of the amino acid sequence and a concomitant decrease in the overall repetitiveness of the gene. Here, we describe the high-level secretion of a fully synthetic, highly hydrophilic and non-degraded 36.8 kDa gelatin by *P.pastoris* and its characterization.

## Materials and Methods

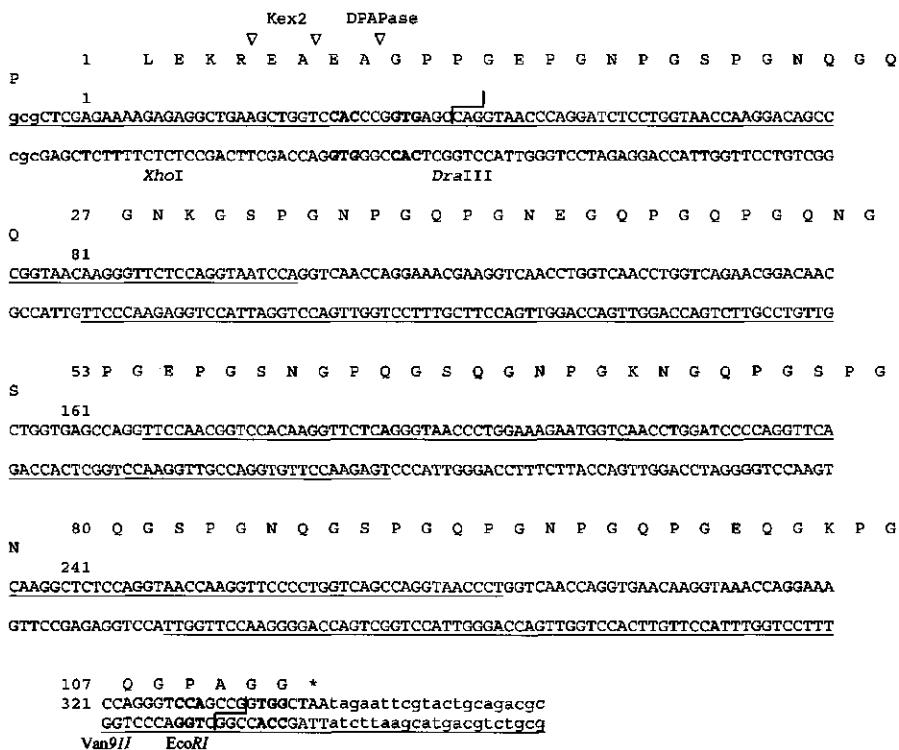
### *Vector construction*

The monomeric gelatin gene (referred to hereafter as "P" for "Polar") was constructed by overlap extension PCR (Ho *et al.*, 1989) of long oligonucleotides (underlined in Figure 3.1A). PCR was performed in a Perkin-Elmer GeneAmp 9700, using the proofreading enzyme *Pwo* DNA polymerase (Eurogentec). The 5' half of the gene was constructed by overlap extension of the first and second oligonucleotides and co-amplified by outer primers directed against nucleotides 1-26 (sense) and 197-174 (antisense). Likewise, the 3' half of the gene was constructed by overlap extension of the third and fourth oligonucleotides and co-amplified by primers directed against nucleotides 174-197 (sense) and 363-337 (antisense). The resulting PCR products were isolated from an agarose gel and were combined by another overlap extension PCR and co-amplified with the primers directed against nucleotides 1-26 (sense) and 363-337 (antisense). The resulting 0.3 kb PCR fragment was digested with *XhoI/EcoRI* and cloned in vector pMTL23 (Chambers *et al.*, 1988) to form vector pMTL23P. The sequence of the gene was verified by automated DNA sequencing of both strands.

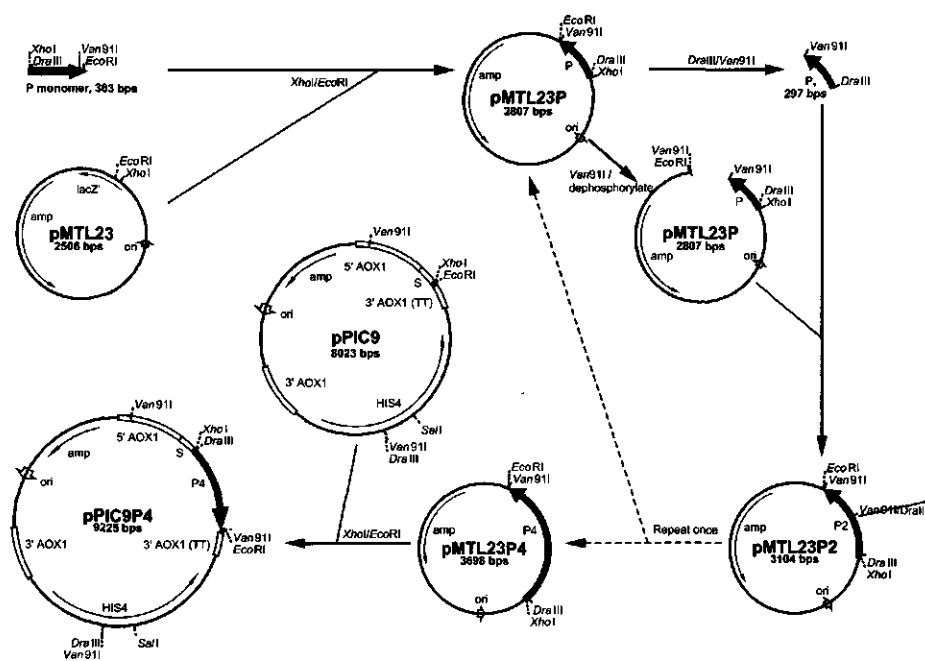
The monomeric gene was released by digesting pMTL23P with *DraIII/Van9II* (Fig. 3.1B). In a separate reaction the vector was linearized with *Van9II* and dephosphorylated. The *DraIII/Van9II* fragment was then inserted into this linearized vector to yield vector pMTL23P2. This process of insertional doubling can in principle be repeated to form multimers of any desired length, but was repeated only once here to form the vector pMTL23P4. The tetrameric gene (referred to hereafter as "P4") was then



cloned into the *XhoI/EcoRI* sites of vector pPIC9 (Invitrogen) that contains a HIS4 selectable marker, an alcohol oxidase 1 (AOX1) promoter/terminator cassette and a *Saccharomyces cerevisiae*  $\alpha$ -factor prepro secretory signal (Clare et al., 1991a). The Kex2 and dipeptidyl amino peptidase (DPAPase) cleavage sites of the  $\alpha$ -factor prepro sequence are lost from pPIC9 when using the *XhoI* site, but are restored upon ligation of the gelatin gene by the sequence between *XhoI* and *DraIII* (Figure 3.1A).



**Figure 3.1A.** Construction of the synthetic gelatin gene. Overlapping oligonucleotides and encoded P monomer. Letters in bold type indicate the non-ambiguous nucleotides recognized by *DraIII* and *Van9II*.



**Figure 3.1 B** Construction of the synthetic gelatin gene(B) Multimerization of P monomer to P4.

Note that Van911/DraIII hybrid sites are not recleavable.

### Transformation of *P.pastoris*

Plasmid pPIC9P4 was linearized with *SaI* in order to obtain preferentially  $Mut^+$  transformants [i.e. by integration at the *his4* locus rather than the *AOX1* locus and thus allowing normal growth on methanol (Clare *et al.*, 1991b)]. Transformation of *P.pastoris* strain GS115 (*his4*; Cregg *et al.*, 1985) by electroporation and selection of  $Mut^+$  transformants was as described previously (Werten *et al.*, 1999).

### Fermentative production of synthetic gelatin in *P.pastoris*

Fermentations were performed in 1-140 l fermenters (Applikon) in minimal basal salt medium (Invitrogen) supplemented with 0.2% (v/v) PTM<sub>1</sub> trace salts (Invitrogen). Methanol fed-batch fermentations were performed as described previously (Werten *et al.*, 1999), with the exception that no protease inhibiting supplements such as casamino acids were added and that the pH during methanol fedbatch was maintained at 3.0 for all fermentations.

***Small-scale purification of synthetic gelatin by differential acetone precipitation***

Differential acetone precipitation was as described previously (Werten *et al.*, 1999). Chilled acetone was added to fermentation supernatant at 40% (v/v), after which endogenous proteins were pelleted by centrifugation. The acetone concentration in the supernatant was then increased to 80% (v/v) and the pellet obtained after centrifugation was washed with 80% acetone and air-dried.

***Preparative purification of synthetic gelatin by differential ammonium sulfate precipitation***

Preparative purification of synthetic gelatin from fermentation supernatant consisted of twice-repeated ammonium sulfate precipitation at 40% saturation (4°C) and subsequent washing of the precipitate with 60% saturated ammonium sulfate. Depending on the scale of the purification, separation of the precipitate from the liquid was either by centrifugation or by depth-filtration using AKS-4 sheets (USF Seitz-Schenk). The protein was subsequently desalted by diafiltration and lyophilized.

***Bicinchoninic acid protein assay***

A commercially available bicinchoninic acid (BCA) protein assay was used according to the manufacturer's recommendations (Pierce). The reaction was performed at 60°C for 30 min. The calibration curve was prepared gravimetrically from lyophilized, desalted P4 gelatin, purified by differential ammonium sulfate precipitation (purity at least 98%).

***SDS-PAGE***

SDS-PAGE (Laemmli, 1970) was performed in a Mini-PROTEAN II system (Bio-Rad) under reducing denaturing conditions. Gels consisted of a 5% stacking and a 12.5% separating zone (2.7% crosslinking). Gels were stained using Coomassie PhastGel Blue R-350 (Amersham-Pharmacia Biotech) and were destained by heating in water using a microwave oven, similarly to Faguy *et al.*, 1996.

### ***Gel filtration chromatography***

Protein in 0.1 M sodium chloride was loaded on a 10 x 250 mm column packed with Superose 12 (Amersham-Pharmacia Biotech). Elution was carried out with 0.1 M sodium chloride at a flowrate of 0.2 ml/min, collecting 2 ml fractions and monitoring the absorbance at 214 nm.

### ***Mass spectrometry***

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed at the Department of Biochemistry, Wageningen University, The Netherlands, using a Voyager DE-RP delayed extraction mass spectrometer (PerSeptive Biosystems). Samples were prepared by the dried droplet method, using sinapinic acid dissolved in 30% (v/v) acetonitrile, 4% (v/v) trifluoroacetic acid, as matrix. Measurements were in positive, linear mode and accelerating voltage was 25,000 V. Cytochrome C and bovine serum albumin were used as external calibrants.

### ***Chemical modification of gelatins***

Esterification of carboxylic amino acid side chains was adapted from Wilcox (Wilcox, 1967). A 100 µg amount of protein was incubated in 500 µl of methanol, 0.1 M hydrochloric acid at 4°C for 72 h. The methanol was exchanged for 1 mM hydrochloric acid by diafiltration in a 3 kDa Microcon (Millipore). Removal of ester groups was performed by incubating the esterified protein in 100 mM Tris/HCl, pH 8.8 at 20°C for 72 h.

Hydrazination of carboxylic amino acid side chains was performed according to Matagne *et al.* (Matagne *et al.*, 1991). A 50 µg amount of protein was dissolved in 20 µl of 50 mM sodium phosphate buffer, pH 7. After addition of 170 µl of 8 M urea, 1 M hydrazine and 0.1 M 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, pH 4.5, the mixture was incubated at room temperature for two h.

### ***Surface tension measurements***

Surface tension at the liquid-air interface was measured according to the du Noüy ring method (Lecomte du Noüy, 1919) using a Krüss K6 tensiometer. The temperature of the sample vessel was maintained at 20°C. The actual measurements were performed five min after lowering of the ring onto the liquid surface, as suggested by Clarkson *et al.*, 1999. Bovine serum albumin was used as a reference protein and raw data were corrected for the hydrostatic volume effect according to Harkins and Jordan, 1930.

### ***Circular dichroism spectrometry***

Proteins were dissolved in Milli-Q water at 0.1 mg/ml. Measurements were performed at the Department of Biochemistry, Wageningen University, The Netherlands using a Jasco J-715 spectropolarimeter. The pathlength was 0.1 cm and the spectra were recorded from 190 to 260 nm at 4 °C, using a scanning speed of 20 nm/min at a resolution of 0.1 nm.

## **Results**

### ***Design and construction of the synthetic P4 gelatin gene***

The basic structure of natural gelatins consists of repeating Gly-Xaa-Yaa triplets, where Xaa and Yaa are often proline and hydroxyproline, respectively (the latter being post-translationally modified proline). This structure maintains the open, unfolded conformation characteristic of gelatin. Owing to this unfolded conformation, gelatin is fairly hydrophilic because its hydrogen bonds are highly exposed. Furthermore, only a small fraction of the protein is occupied by hydrophobic amino acids such as Trp, Tyr, Phe, Leu, Ile, Val and Met.

Our synthetic P4 gelatin design also provides the (Gly-Xaa-Yaa)<sub>n</sub> structure. To increase its hydrophilicity relative to that of natural gelatins, we designed a gelatin without any hydrophobic amino acids other than proline and with a high content of the hydrophilic amino acids asparagine and glutamine (Table I). To illustrate the high hydrophilicity of this synthetic gelatin compared with natural gelatins, the GRAVY values [grand average of hydropathy (Kyte and Doolittle, 1982)] of P4, natural

recombinant Col3a1 gelatin (Werten *et al.*, 1999) and cattle bone gelatin are indicated in Table I.

Amino acid	P4	Col3a1 <sup>b</sup>	Cattle bone gelatin <sup>c</sup>
Ala	1.0	7.9	12.0
Arg	0.0	4.8	5.0
Asn	12.0	3.5	-
Asp	0.0	3.1	-
Asx	-	-	4.3 <sup>d</sup>
Cys	0.0	0.0	0.0
Gln	16.0	3.5	-
Glu	4.0	3.5	-
Glx	-	-	6.7 <sup>d</sup>
Gly	33.7	33.8	33.4
His	0.0	1.8	0.5
Ile	0.0	1.3	1.2
Leu	0.0	1.8	2.4
Lys	3.0	3.5	2.6 <sup>e</sup>
Met	0.0	0.4	0.7
Phe	0.0	0.4	1.3
Pro	22.4	21.1	22.2 <sup>e</sup>
Ser	8.0	6.1	3.4
Thr	0.0	1.8	1.6
Trp	0.0	0.0	0.0
Tyr	0.0	0.4	0.2
Val	0.0	1.3	2.4
GRAVY	-1.77	-1.08	-0.75 to -1.09 <sup>e</sup>
Isoelectric point	4.9	9.7	4.7 to 5.4 <sup>f</sup>
Molecular weight (kDa)	36.8	20.6	65 to 300 <sup>f</sup>

<sup>a</sup>Theoretical amino acid composition (mol%), GRAVY values, isoelectric point and molecular weight were calculated using the ProtParam tool available at the ExPASy WWW server (Appel *et al.*, 1994). Experimental values were obtained as indicated below.

<sup>b</sup>Natural non-hydroxylated recombinant gelatin produced in *P.pastoris* (Werten *et al.*, 1999).

<sup>c</sup>The amino acid composition (mol%) of limed cattle bone gelatin is the mean of four measurements.

<sup>d</sup>Asn/Asp and Gln/Glu can not be distinguished by amino acid analysis and are given as Asx and Glx, respectively. This does not affect the GRAVY calculation.

<sup>e</sup>Values for Lys and Pro include hydroxylysine and hydroxyproline, respectively. The hydrophathy indices of these modified amino acids were assumed to lie between those of the respective unmodified amino acids and the lowest value of the Kyte and Doolittle hydrophathy scale. The range of GRAVY values indicated for cattle bone gelatin represents both extremes.

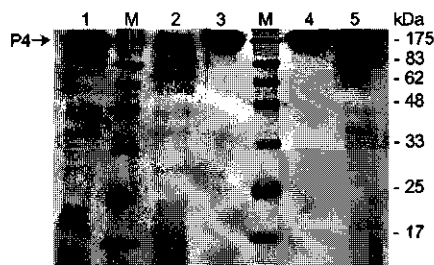
<sup>f</sup>Natural cattle bone gelatin is a heterogeneous mixture of molecules of different molecular weights and isoelectric points within the indicated ranges (Alleavitch *et al.*, 1988).

The content of acidic and basic residues was modulated to give an isoelectric point similar to that of common limed bone gelatin (Table I). Only Lys was used as a basic residue, because we anticipated proteolysis of mono-arginylic sites based on previous work (Werten *et al.*, 1999).

The P4 gene was constructed from four identical P monomers that were designed to have the codon usage of *P.pastoris* highly expressed genes (Sreekrishna and Kropp, 1996). The monomeric gene contains restriction sites for *Dra*III and *Van*91I. These enzymes allow the design of mutually complementary, non-palindromic overhangs that enable convenient elongation of the gene by insertional doubling in a fixed orientation (Figure 3.1). The process can be repeated until the desired polymer length has been achieved. This modular design offers flexibility in the construction of future gelatins by allowing the combination of different types and lengths of polymerized gelatins via the *Dra*III/*Van*91I sites. The *Xho*I and *Eco*RI sites provided by the sequence allow the direct insertion of the final synthetic gene into *P.pastoris* expression vector pPIC9, resulting in a fusion to the alpha-mating factor prepro secretion signal. Thus, the combined four P modules were cloned into pPIC9 to yield vector pPIC9P4.

#### ***Production of synthetic P4 gelatin***

Plasmid pPIC9P4 was used to transform *P.pastoris* GS115. Randomly chosen transformants were fermented and culture supernatants harvested throughout the fermentation were subjected to SDS-PAGE. The theoretical molecular weight of P4 is 36.8 kDa. Because collagenous proteins migrate in SDS-PAGE at an apparent molecular weight ~40% higher than the true molecular weight (Butkowski *et al.*, 1982; Werten *et al.*, 1999), one would expect a band of ~52 kDa. Instead, however, the Coomassie Blue stained SDS-PAGE gel showed a faint blurry band at the top of the separating gel that had a tendency to diffuse from the gel during methanol-acetic acid destaining. Migration of the gelatin into the gel was improved by running the gel at 4°C at twice the voltage recommended by the manufacturer of the electrophoresis system (*i.e.* 400 instead of 200 V). Diffusion of the protein from the gel during destaining was reduced by destaining the gel in water heated in a microwave oven (Faguy *et al.*, 1996), rather than performing the common lengthy incubations in methanol/acetic acid.



**Figure 3.2.** Purification of synthetic gelatin. Lane 1, fermentation supernatant; lane 2, precipitate obtained at 40% (v/v) acetone; lane 3, precipitate obtained by increasing the acetone concentration of the 40% (v/v) supernatant to 80% (v/v); lane 4, precipitate obtained at 40% of ammonium sulfate saturation; lane 5, supernatant obtained at 40% of ammonium sulfate saturation; Lane M, molecular weight marker.

Figure 3.2, lane 1 shows fermentation supernatant analyzed in this manner. N-terminal protein sequencing of this band (Sequencing Centre Utrecht, The Netherlands) revealed the expected amino acid sequence (GPPGEPGNPG). There was no indication of incomplete processing of the  $\alpha$ -factor derived Glu-Ala repeats by dipeptidyl amino peptidase, such as is occasionally observed when using this prepro sequence for secretion of heterologous proteins (Vedvick *et al.*, 1991; Briand *et al.*, 1999; Werten *et al.*, 1999; Goda *et al.*, 2000).

Recombinant gelatins with natural amino acid sequences can be purified from the fermentation broth by using differential acetone precipitation (Werten *et al.*, 1999). Endogenous extracellular proteins are precipitated at 40% (v/v) acetone, after which the gelatin is precipitated from the supernatant by addition of acetone to 80% (v/v). Figure 3.2, lanes 2 and 3 show that acetone precipitation was equally effective for the purification of synthetic gelatins. Amino acid analysis was performed (in triplicate) to estimate the purity of the precipitated gelatin. By linear least-squares fitting of the observed data in terms of the P4 amino acid composition, the average relative contributions (vector moduli) of the P4 component and the residuals component were found to be 96.2% and 3.8%, respectively ( $\pm 1.4\%$  SD). The purity of the preparation at



the protein level is thus estimated to be >96%, as stochastic fluctuations of measured values contribute to the residuals component. BCA protein analysis of acetone precipitates of different fermentations showed gelatin yields in the range of 3-6 g/L of clarified broth.

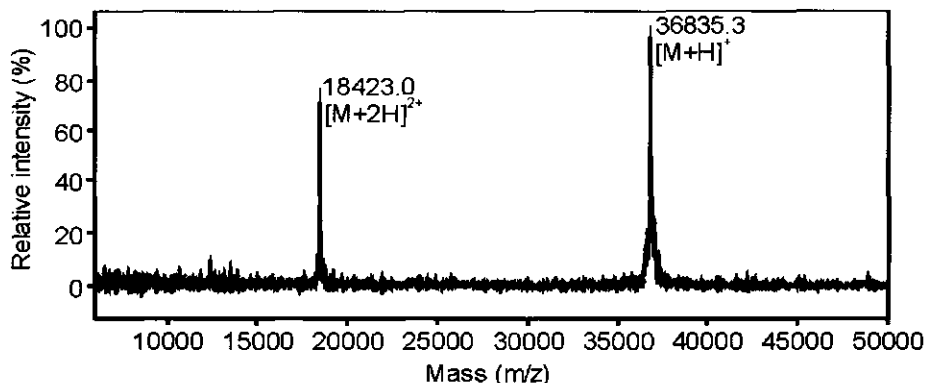
We previously found that recombinant gelatins with natural amino acid sequences can also be purified from *P.pastoris* fermentations using differential ammonium sulfate precipitation (unpublished data). Gelatinous proteins precipitate at 40% saturation, whereas endogenous extracellular *P.pastoris* proteins surprisingly do not precipitate at up to 80% saturation. In agreement with this, readily precipitable proteins such as  $\beta$ -lactoglobulin are rendered virtually unprecipitable upon mixing them with fermentation supernatant. We investigated if it was possible to purify P4 gelatin by differential ammonium sulfate precipitation. Indeed, Figure 3.2, lanes 4 and 5 show that synthetic gelatin is quantitatively precipitated at 40% ammonium sulfate saturation, whereas no endogenous proteins are visible. Based on amino acid analysis (in triplicate) and subsequent linear least-squares fitting of the observed data, the purity at the protein level was estimated to be >98.1% ( $\pm 0.7\%$  SD). Two-dimensional electrophoresis followed by silverstaining showed virtually no contaminants (not shown). Gelatin yields determined by BCA analysis of ammonium sulfate precipitates of several fermentations [possible only after thorough desalting because ammonium sulfate reduces BCA reactivity (Smith *et al.*, 1985)] were within 0.3 g/L of the values determined by acetone precipitation. This small difference (<10%) is largely due to interference of the BCA assay by the low amount of reducing exopolysaccharides that co-purifies in the acetone precipitation procedure, whereas exopolysaccharides are virtually eliminated when using ammonium sulfate precipitation (data not shown). Compared with differential acetone precipitation, the overall higher purity obtained and the higher amenability to scale-up render differential ammonium sulfate precipitation the method of choice for preparative purification of P4 gelatin.

### ***Establishing the molecular weight of synthetic P4 gelatin***

A possible explanation for the aberrant molecular weight observed in SDS-PAGE could be that P4 is glycosylated. N-linked glycosylation can be ruled out because no susceptible sites are present in the amino acid sequence. However, *P.pastoris* is also able to perform O-glycosylation and the structural determinants for such an event are unclear (Duman *et al.*, 1998). To rule out this possibility, periodic acid-Schiff's staining (Zacharius *et al.*, 1969) and Alcian Blue staining (Wardi and Michos, 1972) were performed on ammonium sulfate purified P4. No glycosylation was observed (not shown).

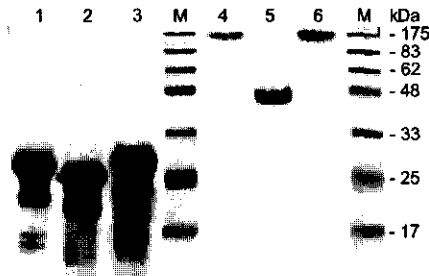
To determine whether the synthetic gelatins have in fact the correct molecular weight, but merely exhibit aberrant behavior in SDS-PAGE, analytical gel filtration chromatography was performed. The Superose 12 column was calibrated with a mixture of natural recombinant gelatin fragments (Werten *et al.*, 1999), giving a series of molecular weights of 53, 42, 28, 16, 12 and 8 kDa. Ammonium sulfate-purified P4 was subjected to gel filtration chromatography. Only one significant peak was observed. N-terminal protein sequencing of this fraction in solution (Sequencing Centre Utrecht) showed the correct N-terminus for P4. The molecular weight of P4 deduced from the chromatogram was 47 kDa. This is clearly much closer to the theoretical values of 36 kDa than the molecular weight apparent from SDS-PAGE, although the deviation is still significant.

Mass spectrometry was used to determine ultimately the molecular weight of P4. Materials purified by both ammonium sulfate precipitation and gel filtration chromatography were analyzed and the results were in good mutual agreement. Figure 3.3 shows the MALDI-TOF mass spectrum of P4 purified by ammonium sulfate precipitation. The observed molecular weight of 36,835 Da corresponds well with the theoretical value of 36,818 Da. This result shows that the apparent high molecular weight observed in SDS-PAGE is indeed the result of aberrant migration behavior. Furthermore, the SDS-PAGE and gel filtration chromatography results are confirmed, in that there is no presence of proteolytically degraded fragments.



**Figure 3.3.** MALDI-TOF mass spectrum of synthetic gelatin purified by differential ammonium sulfate precipitation. The  $[M+H]^+$  and  $[M+2H]^{2+}$  peaks correspond to singly and doubly charged molecular ions, respectively.

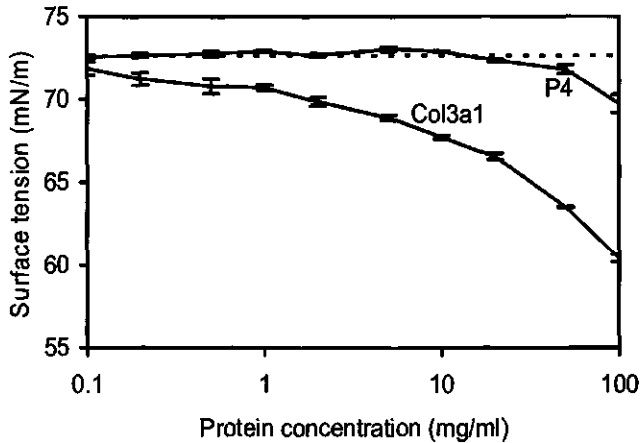
We speculated that the aberrant migration rate of synthetic gelatin compared with normal gelatin was due to low binding of SDS in view of its high hydrophilicity, as the interaction of SDS with proteins is mainly of a hydrophobic nature (Reynolds and Tanford, 1970). Esterification of the carboxylic amino acid side chains with methanol/hydrochloric acid would increase the protein's hydrophobicity and thus its SDS binding capacity and migration rate. The migration of natural gelatins treated in this way was only slightly affected (*i.e.* a decrease in apparent molecular weight of  $\sim 2$  kDa), whereas the esterified synthetic gelatin migrated much faster than the unmodified protein (Figure 3.4). The molecular weight observed for esterified P4 was  $\sim 50$  kDa. As natural gelatins migrate  $\sim 40\%$  more slowly than common proteins (Butkowski *et al.*, 1982; Werten *et al.*, 1999), this value is in good agreement with the value expected for natural gelatins of the same molecular weight as P4 (*i.e.* a 36.8 kDa natural gelatin would migrate at  $\sim 52$  kDa). Removal of the ester groups at high pH restored the aberrant migration of P4 (Figure 3.4). Hydrazination of the carboxylic amino acid side chains of P4 did not result in an altered migration rate (not shown), indicating that the effect of esterification was not due to the mere removal of negative charge.



**Figure 3.4.** Esterification of natural recombinant Col3a1 and synthetic P4 gelatin. Lane 1, unmodified Col3a1 natural recombinant gelatin; lane 2, esterified Col3a1; lane 3, esterified Col3a1 after hydrolysis of the ester groups; lane 4, unmodified P4; lane 5, esterified P4; lane 6, esterified P4 after hydrolysis of the ester groups; lane M, molecular weight marker.

#### **Characterization of synthetic P4 gelatin**

In marked contrast with recombinant gelatins having natural sequences, we noticed that highly concentrated solutions of P4 showed essentially no foaming. A direct relationship exists between protein (surface) hydrophobicity, surface tension and foam stability (Horiuchi *et al.*, 1978). Therefore, the surface activity of P4 relative to that of Col3a1 natural recombinant gelatin was determined, using the du Noüy ring method (Lecomte du Noüy, 1919). Figure 3.5 shows that P4 does not show any significant lowering of the surface tension of water at concentrations up to 5% (w/v), whereas Col3a1 already has an effect at 0.01%. Within the range of up to 10% of protein, it was not possible to determine the apparent critical micelle concentration (CMC; *i.e.* the concentration whereby the surface tension curve reaches a plateau phase) for either of the gelatin types. For comparison, bovine serum albumin has an apparent CMC of ~0.003% (Clarkson *et al.*, 1999).



**Figure 3.5.** Surface tension of natural recombinant *Col3a1* and synthetic *P4* gelatin as a function of concentration. Error bars indicate the standard deviation of five measurements. The dotted line indicates the surface tension of water

We previously showed that natural, non-hydroxylated recombinant gelatins produced in *P.pastoris*, do not form collagen triple-helices even at 4 °C (Werten *et al.*, 1999). Synthetic *P4* gelatin was also non-hydroxylated, as shown by the amino acid analyses and N-terminal sequencing described above. Circular dichroism spectrometry was performed to see whether the greater hydrophilicity of *P4* somehow influences triple-helix formation. Figure 3.6 shows the spectrum of *P4* and natural recombinant *Col3a1* gelatin at 4 °C. Both proteins show a clear absence of the positive peak at ~221 nm characteristic of the collagen triple-helix (de Wolf and Keller, 1996; Rossi *et al.*, 1996). The mean residue ellipticity of *P4* at that position of the spectrum is somewhat lower than that of *Col3a1*.

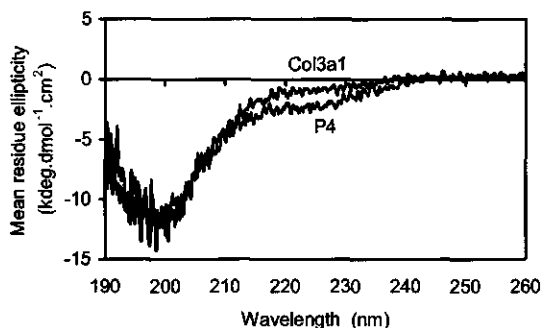


Figure 3.6. Circular dichroism spectra of natural recombinant Col3a1 and synthetic P4 gelatin at 4°C.

## Discussion

This paper describes the extracellular production of a recombinant, custom-designed and highly hydrophilic gelatin in *P.pastoris*. Yields in single-copy transformants were 3 to 6 g/L of clarified broth, which is comparable to the yields obtained in single-copy transformants for natural gelatins (Werten *et al.*, 1999). Expression of the same hydrophilic gelatins in *Hansenula polymorpha* did not succeed, probably due to proteolytic degradation of the synthetic proteins in this host.

Despite the obvious physico-chemical differences, both synthetic P4 and natural recombinant gelatin could be purified from the fermentation broth by both differential acetone precipitation and differential ammonium sulfate precipitation. The universality of both purification techniques is probably due to the hydrophilicity and unfolded structure of gelatins in general. Especially differential ammonium sulfate precipitation allowed convenient large-scale purification of P4 gelatin to near homogeneity.

Secreted synthetic gelatin was fully intact, as evidenced by SDS-PAGE, gel filtration chromatography, N-terminal sequencing and mass spectrometry. This is in contrast with natural recombinant gelatins produced in *P. pastoris*, which were partly

degraded (Werten *et al.*, 1999). Apart from the occurrence of some minor background degradation, collagen type I-derived natural recombinant gelatins were cleaved into several major bands by a Kex2-like protease. Cleavage occurred C-terminal of two occurrences of the mono-arginylic sequence Met-Gly-Pro-Arg. We speculated that the amino acids occupying the -2 and -4 positions in this motif (relative to the site of cleavage) were a major factor determining the cleavage efficiency (Werten *et al.*, 1999), which is in accordance with recent data on the substrate specificity of *S.cerevisiae* Kex2 (Bevan *et al.*, 1998; Suzuki *et al.*, 2000). It may well be that the above-mentioned minor background degradation represented a limited extent of cleavage at Arg residues having “suboptimal” residues at the -2 and -4 positions. Therefore, in the design of the synthetic gelatin described here, only Lys was used as a basic residue to control the isoelectric point. The finding that secreted P4 was completely intact does indeed suggest a general susceptibility of Arg residues in recombinant gelatins to proteolysis, and may thus have implications for the rational design of (partially) unfolded proteins to be expressed extracellularly in *P. pastoris*.

While the molecular weight of P4 as determined by mass spectrometry was in good agreement with the value deduced from the amino acid sequence, the molecular weight apparent from gelatin-calibrated gel filtration chromatography was ~10 kDa higher. Ionic interactions with Superose 12 are negligible in the presence of salt and only small hydrophobic peptides appear to show significant hydrophobic interactions with this matrix (Andersson *et al.*, 1985). It is therefore not very likely that a lower degree of such interactions of P4 relative to natural gelatins causes the seemingly aberrant molecular weight. Possibly, the effect is due to an increased hydrodynamic size of P4 relative to natural gelatins, as a result of increased interaction of this highly hydrophilic protein with water.

Natural gelatins migrate ~40% more slowly in SDS-PAGE than expected. Several possible explanations for the aberrant migration behavior of gelatins have been suggested (Furthmayr and Timpl, 1971; Freytag *et al.*, 1979; Hayashi and Nagai, 1980; Noelken *et al.*, 1981; Butkowski *et al.*, 1982). It is most likely not the result of anomalously low SDS binding, but is at least in part due to the low average residue molecular weight of gelatin, resulting in a relatively high number of residues (*i.e.* molecular length) per unit of molecular weight (Freytag *et al.*, 1979; Noelken *et al.*,

1981; Butkowski *et al.*, 1982). Most other reports on aberrant protein migration rates in SDS-PAGE involve highly acidic proteins that show reduced binding of SDS due to electrostatic repulsion by the protein's high negative net charge (Ohara and Teraoka, 1987; Matagne *et al.*, 1991; Casarégola *et al.*, 1992; McGrath *et al.*, 1992). SDS-PAGE showed that P4 migrates at a highly reduced rate, even much more slowly than natural gelatins. We showed here that esterification of the carboxylic side chains of P4 restores its migration rate roughly to that expected for normal gelatins (*i.e.* ~40% more slowly than common proteins). In contrast, hydrazination did not affect the migration rate of P4. Hydrazination eliminates the negative charge of the same carboxylic residues as does esterification, but reduces the protein's hydrophobicity whereas esterification increases it. As the binding of SDS to proteins is primarily hydrophobic in nature (Reynolds and Tanford, 1970), the extremely slow migration of the highly polar P4 gelatin in SDS-PAGE is therefore most likely the result of insufficient SDS binding and a concomitant low negative net charge. The finding that the resolution of the SDS-PAGE was improved by increasing the field strength to twice that recommended by the manufacturer of the electrophoresis system indicates that the higher field strength aids protein migration in overcoming diffusive forces.

Surface activity is a major determinant in a protein's function as a protective colloid (*e.g.* in photographic emulsions). Solutions of P4 showed essentially no foaming and tensiometric analysis of solutions of P4 in water showed only negligible surface activity. P4 thus represents a novel hydrocolloid combining some of the characteristics unique to gelatins and a low surface activity commonly expected only for the most hydrophilic polysaccharide hydrocolloids.

We previously showed that non-hydroxylated, natural recombinant gelatins do not show triple-helical structure in circular dichroism spectrometry (Werten *et al.*, 1999). It is a well-established fact that hydroxyproline residues play a crucial role in the stabilization of the collagen triple-helix. This role is easily recognized when examining the amino acid compositions and thermal stabilities of natural collagens from different species (Privalov, 1982). X-ray crystallography showed that the triple-helix is surrounded by a cylinder of hydration (Bella *et al.*, 1994). Although recently questioned (Holmgren *et al.*, 1999; Nagarajan *et al.*, 1999), the role of hydroxyproline in the stabilization of the triple-helix is generally attributed to its hydrogen bonding



with this water network (Brodsky and Shah, 1995). In view of the high polarity of P4, we considered it prudent to investigate its conformation using circular dichroism spectrometry. No triple-helical structure was observed at 4°C and P4 gelatin is thus an essentially non-gelling gelatin. Non-gelling gelatins permit novel applications such as low-temperature silver halide crystallization in the preparation of photographic emulsions (de Wolf *et al.*, 2000). Comparison of the circular dichroism spectrum of P4 with that of Col3a1 natural recombinant gelatin showed that the mean residue ellipticity of P4 at the discriminating wavelength of 221 nm was actually lower than that of Col3a1. Although the latter is essentially in a random coil conformation, the higher hydrophilicity of P4 probably reduces minor intramolecular and intermolecular interactions, thereby resulting in a slightly lower ellipticity.

Current research is directed towards the production of other synthetic gelatins with distinct functionalities and combining them to form chimeric tailor-made biopolymers.

### **Acknowledgements**

We thank Ir Friso B.J. van Assema, Ir Harald G. de Jonge, Ing Pieter H. Lodder, Ir Marcia P.M. Dielissen and Ing Marjo Koets for assistance, Ir Harrold A. van den Burg (Department of Biochemistry, Wageningen University, The Netherlands) for mass spectrometry, Dr ir Carlo P.M. van Mierlo (Department of Biochemistry, Wageningen University, The Netherlands) for help with circular dichroism measurements and Dr ir Johan M.H. Stoop for critical reading of the manuscript.

## Chapter 4

### Endogenous prolyl 4-hydroxylation in *Hansenula polymorpha* and its use for the production of hydroxylated recombinant gelatin

Eric C. de Bruin, Marc W.T Werten, Colja Laane and Frits A. de Wolf

Published in: FEMS Yeast Research (*in press*).

#### Abstract

Several yeast systems have recently been developed for the recombinant production of gelatin and collagen. Amino acid sequence-specific prolyl 4-hydroxylation is essential for the gel-forming capacity of gelatin and for the proper folding of (pro)collagen. This post-translational modification is generally considered to be absent in microbial eukaryotic systems and therefore co-expression of heterologous (human or animal) prolyl 4-hydroxylase would be required. However, we found that the well-known protein expression host *Hansenula polymorpha* unexpectedly does have the endogenous capacity for prolyl 4-hydroxylation. Without co-expression of a heterologous prolyl 4-hydroxylase, both an endogenous collagen-like protein and a heterologously expressed collagen fragment were found to be sequence-specifically hydroxylated.

were subsequently treated for 20 min at 70 °C in a water bath, placed on ice for 1 min and centrifuged at 20,000 g in a micro-centrifuge. Microscopic analysis of cells showed that the heat treatment did not cause detectable cell-lysis.

We showed previously that differential acetone precipitation can be used to separate non-collagenous extracellular *P. pastoris* proteins, precipitating at 40 % (v/v) of acetone, from recombinant collagen-like proteins, precipitating at 80 % (v/v) of acetone (Werten et al., 1999). This procedure was applied here for the separation of putative collagen-like proteins and non-collagenous proteins, released by the heat-treatment described above. Acetone, previously chilled to 0 °C, was added dropwise to chilled supernatant of heat-treated *H. polymorpha* cells to 40 % (v/v). Proteins were then precipitated by centrifugation for 15 min. at 20,000 g in a micro-centrifuge. After removal of the 40 % (v/v) acetone precipitate and increasing the acetone concentration in the supernatant to 80 % (v/v), putative collagen-like proteins were precipitated.

#### **SDS-PAGE and N-terminal protein sequencing**

SDS-PAGE (sodiumdodecylsulfate polyacrylamide gel electrophoresis) was performed in a Mini PROTEAN II system (Bio-Rad, Richmond, CA, USA) under reducing, denaturing conditions. Gels were stained with Coomassie Brilliant Blue (CBB R-350). For N-terminal protein sequencing, protein was blotted onto Immobilon PSQ (Millipore, Bedford, MA, USA) by applying 100V for one hour in a Mini Trans-Blot Cell (Bio-Rad). Transfer buffer was 2.2 g CAPS per liter of 10 % methanol, pH 11. Blots were stained with Coomassie Brilliant Blue (CBB R-350) and selected bands were cut out. N-terminal sequencing using Edman-degradation was performed by Sequencing Centre University of Utrecht, Utrecht, The Netherlands or by Protein Research Facility, ECSI, Amsterdam, The Netherlands.

***Determination of the degree of hydroxylation from N-terminal sequencing data***

The degree of hydroxylation of proline residues in the X2 position of [G-X1-X2] triplets was estimated from the data of N-terminal amino acid sequencing as follows. The development and the decay of glycine, proline and hydroxyproline peaks in successive amino acid sequencing steps were analyzed by comparing the relative signal intensities of each amino acid. First, the decay of signals of previous steps was analyzed in following steps that by itself could not give rise to a new signal of the same amino acid, in view of the sequence. In sequencing steps that gave rise to new proline or hydroxyproline signals, the proline or hydroxyproline signals remaining from previous steps could then be calculated by interpolation. This remaining signal was subtracted from the new signal in order to calculate the corrected signal, newly obtained in each step. The signals were also linearly corrected for the slow overall decay of sensitivity observed for successive triplets. As the corrected signals in successive steps correspond to approximately equimolar amounts of material, the sum of the corrected proline and hydroxyproline signals in the X2 positions can be compared with the corrected proline signals in the preceding X1 positions:

$$P_{X1} = P_{X2} + C \cdot O_{X2}$$

Here, P and O are the corrected proline and hydroxyproline peak areas, respectively and C is an arbitrary conversion factor, relating the relative intensities of the proline and hydroxyproline signals. As C can be calculated from this equation, the degree of hydroxylation of proline residue X2 can be calculated as:

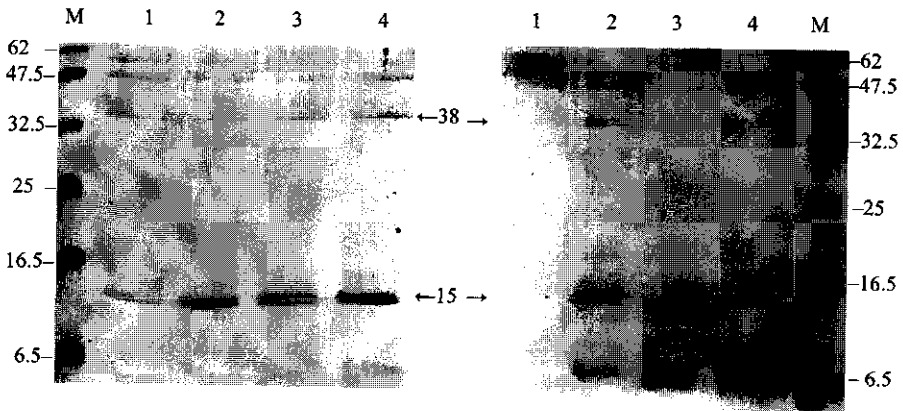
$$\% \text{ hydroxylation} = 100 \cdot C \cdot O_{X2} / (P_{X2} + C \cdot O_{X2}).$$

## Results

### *Expression of collagenous proteins by H. polymorpha in bioreactors*

A 28 kDa fragment of the helical domain of mouse  $\alpha 1(I)$  collagen, Coll1a1, which we have previously expressed in *P. pastoris* (Werten *et al.*, 1999), was cloned into the *H. polymorpha* expression vectors pHIPX7 and pHIPX4 for constitutive and methanol-induced Coll1a1 expression, respectively. Fermentations in mineral medium were performed with selected transformants.

SDS-PAGE showed the constitutive as well as the methanol-induced production of gelatin in the culture medium using the pHIPX7-1A1 and pHIPX4-1A1 transformants (see Fig. 4.1A and 4.1B, respectively).



**Figure 4.1. (A)** SDS-PAGE of the extracellular proteins during glucose fed-batch fermentation of *H. polymorpha* pHIPX7-1A1 (constitutively producing recombinant gelatin) grown in mineral medium. 10  $\mu$ L of culture supernatant was loaded in each lane. Lane 1, 2, 3, 4, sample after 12, 24, 40 and 70 hours of fermentation, respectively. Lane M: prestained marker (New England Biolabs, Northbrook, IL, USA). The arrows indicate the 38 kDa endogenous protein and the recombinant 15 kDa Coll1a1 product.

**Figure 4.1. (B)** SDS-PAGE of the extracellular proteins during methanol-induced expression of recombinant gelatin by *H. polymorpha* pHIPX4-1A1 in mineral medium. 10  $\mu$ L of culture supernatant was loaded in each lane. Lane 1, 2, 3, 4, sample after 12, 24, 40 and 70 hours of fermentation, respectively. Methanol-induction was started after 12 hours of fermentation. Lane M, prestained marker (New England Biolabs). The arrows indicate the 38 kDa endogenous protein and the recombinant 15 kDa Coll1a1 product.

Collagenous proteins are known to migrate in SDS-PAGE at an apparent weight approximately 1.4 times higher than the true molecular weight, and thus, the intact 28 kDa Colla1 protein is known to migrate at an apparent molecular weight of 38 kDa (Werten *et al.*, 1999). In addition to a band with the expected apparent molecular weight of about 38 kDa, a more abundant band of about 15 kDa was produced by both the constitutive transformant (Fig. 4.1A) and the methanol-induced transformant (Fig. 4.1B). N-terminal amino acid sequencing of the 38 and 15 kDa protein bands from both fermentations was performed. Instead of the expected Colla1 sequence, the 38 kDa protein yielded another sequence, not present in Colla1: GPPGPPGPPG (Table I). Surprisingly, this sequence is of a collagenous nature. *H. polymorpha* apparently expresses an endogenous protein with a collagen-like N-terminal amino acid sequence. The N-terminal sequence of the 15 kDa protein corresponded to an internal sequence of the expressed recombinant *COL1A1* gene product: GFQGPPGEPG (Table I), indicating that the product was partially degraded by endogenous proteases of *H. polymorpha*. Gelatin is extremely susceptible to proteolytic degradation, due to its unfolded conformation. In *P. pastoris*, proteolytic degradation of the recombinant Colla1 product was also observed, but to a lesser extent (Werten *et al.*, 1999).

It has been reported that supplementation of the fermentation medium with casamino acids can reduce extracellular proteolysis in *P. pastoris* fermentations (Werten *et al.*, 1999, Clare *et al.*, 1991). However, addition of 1 % (w/v) of casamino acids to the medium during gelatin expression in *H. polymorpha* did not decrease proteolysis (not shown), and the same 38 and 15 kDa products were found, having the same N-terminal amino acid sequences (Table I). In a further attempt to reduce proteolysis, the fermentations were repeated in medium supplemented with 1 % (w/v) peptone, but the same degradation product of 15 kDa was obtained (not shown). The identity of the band was confirmed by N-terminal sequencing of the isolated band. However, hydroxyprolines instead of prolines were now present in the X2 positions of the triplets (Table I). In view of this unexpected result, and in view of the collagen-like sequence of the 38 kDa band, we also checked the potential presence of 4-hydroxyproline in the 38 kDa protein. Indeed, this protein also contained hydroxyproline specifically at the X2 position of the triplets (Table I).

**Table I.** Medium-dependant prolyl 4-hydroxylation of recombinant Coll1a1 and of an endogenous 38 kDa protein by *H. polymorpha*.

Mode of expression	Supplement to the medium	Occurrence of 4-hydroxyproline in the N-terminal amino acid sequence	
Constitutive	-	-	-
	Casamino acids (1 % w/v)	-	-
	Peptone (1 % w/v)	+	+
Methanol-induced	-	-	-
	Casamino acids (1 % w/v)	-	-
	Peptone (1 % w/v)	+	+
	< 10 kDa peptone fraction <sup>e</sup>	+	+
	4-hydroxyproline (0.12 % w/v)	-	-

<sup>a</sup>15 kDa recombinant gelatin band shown in Fig. 4.1A and 4.1B

<sup>b</sup>The N-terminal sequence was : GFQGPPGEPG (referred to as -) and GFQGPOGEOG (referred to as +), where O is 4-hydroxyproline.

<sup>c</sup>Endogenous 38 kDa protein band shown in Fig. 4.1A and 4.1B.

<sup>d</sup>The N-terminal sequence was : GPPGPPGPPG (referred to as -) and GPOGPOGPOG (referred to as +), where O is 4-hydroxyproline.

<sup>e</sup>These protein bands were analyzed in the > 10 kDa ultra-filtration fraction of the culture medium that contained the < 10 kDa peptone fraction (see Fig. 4.2).

From the positions of the hydroxyproline peaks in the sequencing chromatograms, it appeared that the hydroxyproline residues in both proteins were hydroxylated at the C4 atom, being 4-hydroxyprolines. The degree of hydroxylation in position X2 of the endogenous 38 kDa protein was calculated to be in the range of 50-65 mol %, based on the relative intensities of the proline and hydroxyproline signals obtained from N-terminal amino acid sequencing (see Materials and Methods section). Due to the nature of the sequence of the recombinant gelatin, as compared to the uniform poly-[GPP] stretch of the endogenous *H. polymorpha* protein described above, the degree of hydroxylation in the X2 position of the recombinant gelatin was difficult to calculate. The estimates of the degree of hydroxylation in the X2 position of various determinations varied from 25 to 50 mol %, with an average value of about

35 mol %. At a cell wet-weight of 200-300 g/L, about 0.5 g/L of hydroxylated recombinant 15 kDa gelatin was produced, as visually estimated from a Coomassie Blue-stained SDS-PAGE gel, using recombinant gelatin (Werten *et al.*, 1999) as a reference.

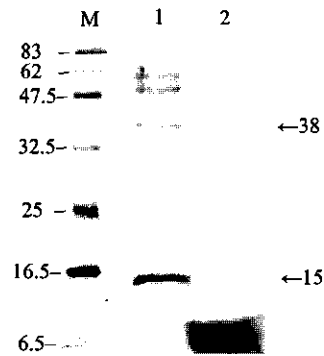
It is apparent from the above, that *H. polymorpha* not only has an endogenous collagen-like protein, but is also able to perform prolyl 4-hydroxylation on both the endogenous protein and the recombinant Colla1 protein. This hydroxylation occurred only at the X2 position of [G-X1-X2] triplets, which represents a sequence-specificity identical to that of animal prolyl 4-hydroxylases. It is therefore most likely that *H. polymorpha* has an endogenous prolyl 4-hydroxylase. As a highly unlikely alternative explanation for endogenous hydroxylation of protein-incorporated proline residues, one could consider sequence-specific, ribosome-mediated incorporation of 4-hydroxyproline during translation. To exclude this hypothetical possibility, the expression experiment was repeated, using mineral medium supplemented with free 4-hydroxyproline, instead of peptone. The 4-hydroxyproline was added at 0.12 % (w/v), corresponding to the 4-hydroxyproline content of the peptone supplement (as determined by amino acid analysis). Subsequent N-terminal sequencing showed that both the endogenous and the recombinant protein were not hydroxylated (Table I).

One could argue that the hydroxylated 38 and 15 kDa proteins could be derived from the peptone (a pancreatic digest of animal tissue) added to the medium. This is highly unlikely as the identical, but non-hydroxylated sequences were found in experiments using mineral medium. Furthermore, the signal-strength and quality of the N-terminal sequencing reactions seem too high to have originated from random collagen fragments present in peptone. Also, a 4-hydroxyproline contribution by a supposed peptone background to the sequencing signal of the 38 and 15 kDa bands is improbable, as in that case one would not expect this amino acid to occur exclusively at the X2 position in the [G-X1-X2] triplets. Nevertheless, to further exclude this possibility, a low molecular weight fraction of peptone was used in a new fermentation. A 10 % (w/v) peptone solution was filtrated using an ultra-filtration device with a 10 kDa cut-off (Millipore Centricon). Only the low molecular weight components and peptides of the peptone that passed the membrane (Fig. 4.2, lane 2), were added to the fermentation medium of an expression experiment with the



pHIPX4-1A1 strain. The proteins produced during fermentation were subsequently ultra-filtrated with another 10 kDa filter and washed by diafiltration with distilled water to remove residual < 10 kDa peptone (Fig. 4.2, lane 1). Subsequent N-terminal sequencing of the 38 and 15 kDa proteins, revealed again the presence of hydroxylated prolines (Table I).

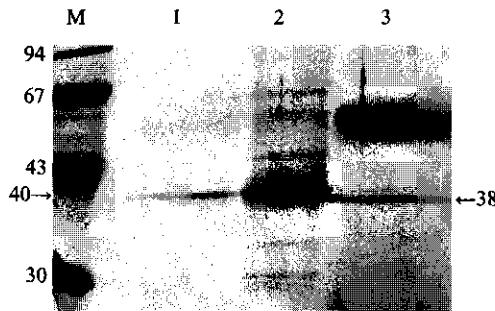
**Figure 4.2.** SDS-PAGE of the > 10 kDa ultra-filtrated fraction of the culture medium of the *H. polymorpha* pHIPX4-1A1 transformant (lane 1). The medium was supplemented with a < 10 kDa ultra-filtrated peptone fraction (lane 2, ten times concentrated relative to what was present in the medium), which was subsequent removed from the medium using a 10 kDa filter (lane 1). Lane M, prestained marker (New England Biolabs, Northbrook, IL, USA).



### **Isolation of the endogenous collagen-like protein by heat treatment of cells grown in shake flasks**

Although never reported for yeasts prior to this report, endogenous collagen-like proteins have been described in the smut fungus *Microbotryum violaceum* (Celerin *et al.*, 1996). These proteins could be isolated from previously washed cells by mechanical shearing (Celerin *et al.*, 1996), as well as by heat treatment (Celerin, M., personal communication). We performed the heat treatment procedure on thoroughly washed *H. polymorpha* cells and the 38 kDa protein was again observed. The isolated protein behaved very similar to collagen-like proteins in differential acetone precipitation as described previously (Werten *et al.*, 1999). The 38 kDa protein, precipitating at 80 % (v/v) of acetone, could be separated from non-collagenous proteins, precipitating at 40 % (v/v) (Fig. 4.3). N-terminal sequencing of a 40 kDa protein precipitated at 40 % (v/v)

(Fig. 4.3, lane 2) yielded a non-collagenous sequence, being EASLGFDLGVQATDGS PKTA. N-terminal amino acid sequencing of the 38 kDa protein precipitated at 80 % (v/v) (Fig. 4.3, lane 3) showed that the N-terminus consisted of at least seven successive [GPX] amino acid repeats, of which at least six triplets are [GPO], the sequence being GPOGPOGPOGPOGPOGPOGPX. Such a long stretch of contiguous [GPP] triplets is not present in any animal collagen, discarding the notion that it might have originated from the peptone added to the medium. Possibly, there are even more contiguous [GPO] triplets, as sequencing was terminated after 21 amino acids.



**Figure 4.3.** SDS-PAGE of proteins isolated from *H. polymorpha* NCYC495 grown in YPD-medium. Lane 1, supernatant after heat treatment at 70 °C and removal of cells; lane 2, protein precipitated at 40 % (v/v) acetone (a protein of 40 kDa is present); lane 3, protein precipitated after bringing the 40 % (v/v) acetone supernatant to 80 % (v/v) acetone (a protein of 38 kDa is present). Protein-precipitates in lanes 2 and 3 are ten times concentrated relative to the supernatant in lane 1. Lane M, molecular weight marker (Amersham Pharmacia Biotech, Uppsala, Sweden).

### ***Prolyl 4-hydroxylation of recombinant proteins by P. pastoris***

As *H. polymorpha* and *P. pastoris* are related species, we wondered if the addition of peptone to the fermentation medium would also induce the hydroxylation of recombinant gelatin in *P. pastoris*. Production of Coll1a1 in *P. pastoris* strains was performed as described previously (Werten *et al.*, 1999), with the exception that 1 % (w/v) of peptone was now present in the fermentation medium. N-terminal amino acid sequencing revealed that the secreted, recombinant Coll1a1 was not hydroxylated.

## **Discussion**

Prolyl 4-hydroxylation was hitherto considered to be absent in yeast systems Kivirikko *et al.*, 1992, Vuorela *et al.*, 1997, Vaughan *et al.*, 1998, Toman *et al.*, 2000, Olsen *et al.*, 2001, Nokelainen *et al.*, 2001, Bulleid *et al.*, 2000), and thus the production of hydroxylated gelatin or collagen in *S. cerevisiae* and *P. pastoris* required co-expression of an animal prolyl 4-hydroxylase (Vuorela *et al.*, 1997, Vaughan *et al.*, 1998, Toman *et al.*, 2000, Olsen *et al.*, 2001, Nokelainen *et al.*, 2001). We have shown in this study however, that the methylotrophic yeast *H. polymorpha* is able to hydroxylate both an endogenous collagen-like protein and a heterologously expressed gelatin, without the need for co-expression of a heterologous prolyl 4-hydroxylase.

Hydroxylation of the endogenous collagen-like protein and the heterologous Coll1a1 collagen fragment, as determined by N-terminal sequencing, occurred in both constitutive glucose-fed fermentations as well as methanol-induced fermentations when peptone was included in the medium. Both proteins showed the same, but non-hydroxylated sequences when produced in fermentations on mineral medium only. Apparently, the hydroxylation activity in *H. polymorpha* is triggered by a component contained in the peptone. *P. pastoris* fermentations supplemented with peptone did not result in the production of hydroxylated recombinant gelatin, at least suggesting that the endogenous hydroxylation activity and/or its induction by peptone may be unique for *H. polymorpha*.

Both the endogenous and heterologous proteins were hydroxylated at the C4 atom of prolines in the X2 position of [G-X1-X2] triplets. Given the apparent identical sequence and atom-specificities of the hydroxylation events found in *H. polymorpha* as compared to those of animal prolyl 4-hydroxylases, the obvious conclusion is that this yeast must have a prolyl 4-hydroxylase-like enzyme. Attempts to confirm this using an *in vitro* activity assay developed for animal prolyl 4-hydroxylases (based on the hydroxylation-coupled decarboxylation of [1-14C] alpha-ketoglutarate (Kivirikko and Myllylä, 1982)) failed. This was probably due to the high background caused by decarboxylation of the Krebs cycle-intermediate alpha-ketoglutarate by other enzymes in the cell extract. It is of course also uncertain whether the yeast hydroxylase would be active at all under the conditions developed for the animal hydroxylases. Nevertheless, in order to exclude the highly unlikely notion that *H. polymorpha* might be able to take up 4-hydroxyproline from the medium and sequence-specifically incorporate it during protein translation, we performed a fermentation supplemented with free 4-hydroxyproline instead of peptone. No hydroxyproline was found and thus *H. polymorpha* must indeed have a prolyl 4-hydroxylase.

The level of hydroxylation found in the recombinant protein was 25-50 mol %. This is lower than the 50-65 mol % found in the endogenous collagen-like protein. This could be due to a lower affinity of the endogenous hydroxylase for the recombinant sequence as compared to the endogenous collagen-like sequence. Alternatively, the high flux of recombinantly over-expressed gelatin may exceed the capacity of the endogenous hydroxylation machinery.

The intriguing question remains why *H. polymorpha* produces hydroxylated collagen-like proteins. The N-terminal sequence of the collagen-like protein consisted of at least seven successive [G-X1-X2] triplets, of which at least six were [GPO] triplets, where O is 4-hydroxyproline. Such long stretches of uniformly repeating [GPO] or [GPP] triplets have not yet been described, except for the extended [GPP] stretches found in a tumor virus (Geck *et al.*, 1990). Although this virus replicates in animal cells, the occurrence of 4-hydroxyprolines has not yet been demonstrated in this protein. Hydroxylated, collagenous proteins have never been described for yeasts, but have been described for the smut fungus *Microbotryum violaceum* (Celerin *et al.*,

1996). These cell surface-attached proteins seem to have a more irregular sequence, often deviating from the consistent [G-X1-X2] triplet structure. They can be released by mechanical shear (Celerin *et al.*, 1996) or heat treatment (Celerin, M., personal communication). The *H. polymorpha* collagen-like protein behaves very similar in that it is released into the medium of high cell density fermentations (possibly due to the shearing forces caused by the high-speed agitation), and that it can also be released from thoroughly washed, intact cells by heat-treatment. This suggests that the *H. polymorpha* protein may also be cell surface-attached. As for its function, it may well be involved in cell-cell recognition and mating as has been suggested for the *M. violaceum* protein (Celerin *et al.*, 1994). On the other hand, the triple-helical motif is found in such diverse proteins as collagens, host-defense proteins, membrane receptors, surfactant proteins, mannose binding protein etc (Brodsky and Shah, 1995) and thus the *H. polymorpha* collagen-like protein may serve one of many possible purposes.

Although functional yeast systems for the production of recombinant, hydroxylated gelatins by co-expression of animal prolyl 4-hydroxylases have been described (Vuorela *et al.*, 1997, Vaughan *et al.*, 1998, Toman *et al.*, 2000, Olsen *et al.*, 2001, Nokelainen *et al.*, 2001), certain problems may be associated with this strategy. Over-expression in *P. pastoris* of the human prolyl 4-hydroxylase reportedly caused inhibition of growth (Keizer-Gunnink *et al.*, 2000), possibly due to aggregation of the catalytic  $\alpha$ -subunit of the heterologous hydroxylase (Vuori *et al.*, 1992) and/or cellular interference with the  $\beta$ -subunit (protein disulfide isomerase, EC 5.3.4.1) (Philajaniemi *et al.*, 1987). We expect that the endogenous prolyl 4-hydroxylation described here will offer new possibilities for the development of *H. polymorpha* and related yeasts into suitable non-animal systems for the production of hydroxylated gelatin and collagen.

### **Acknowledgements**

We gratefully acknowledge Martina Celerin, Indiana University, Bloomington, USA for communicating the heat treatment method for the isolation of cell-attached yeast proteins and Klaas Nico Faber and Jan Kiel, University of Groningen, Haren, The Netherlands, for the gift of the vectors pHIPX4 and pHIPX7 and the strain *H. polymorpha leu 1.1*. We thank Gustaaf de Haas, ATO B.V., Wageningen, The Netherlands for his technical assistance.

## Chapter 5

### Human collagen trimers from the methylotrophic yeast *Hansenula polymorpha*

Eric C. de Bruin, Gustaaf G. de Haas, Klaas Nico Faber, Marc W.T. Werten,

Colja Laane, and Frits A. de Wolf

publication in preparation

#### Abstract

Our group discovered that *Hansenula polymorpha* possesses intrinsic prolyl 4-hydroxylase activity (Chapter 4, de Bruin *et al.*, 2002). Because amino acid sequence-specific prolyl 4-hydroxylation is essential for proper folding of (pro)collagen into trimers that are stable at 37 °C, we investigated the expression of a human procollagen sequence in this yeast. Electron microscopy on immunogold labelled *H. polymorpha* cells producing human procollagen showed that recombinant procollagen molecules were present in the endoplasmic reticulum. Pepsin treatment of cell extracts, yielded trimeric collagen molecules, with a low melting temperature. This low melting temperature could point to a low level of hydroxylation.

## Introduction

Based on its structural role and compatibility within the body, collagen is a commonly used biomaterial in medical applications, such as cosmetic surgery, tissue engineering and wound treatment (Lee *et al.*, 2001). Collagen is regarded as one of the most useful biomaterials, due to its biological characteristics such as biodegradability, weak antigenicity and superior bio-compatibility compared to other natural polymers. Human collagen, which is isolated from the human placenta, does not find wide application, because of the low availability of its source and the possible presence of disease-causing agents (e.g. HIV). As a result of this, bovine collagen is used as a substitute of human collagen. The supply of bovine collagen, extracted from cowhides and tendons, is both abundant and inexpensive. However, bovine collagen is degraded faster and even has the potential to cause an inflammatory reaction in some people. Furthermore, concern has been raised about the potential contamination risk of bovine collagen with prions (bovine spongiform encephalopathy), possibly causing the new variety of Creutzfeldt-Jakob disease in humans.

In human fibroblasts, collagen is synthesized as the precursor procollagen containing globular N and C propeptides. The propeptides keep the procollagen soluble during its passage through the cell. As the procollagen is co-translationally translocated across the membrane of the endoplasmic reticulum, hydroxylation of proline residues occur, specifically at the X2 position within the G-X1-X2 amino acid repeats. Once the polypeptide chain is fully translocated into the lumen of the endoplasmic reticulum the C propeptide folds. The latter is essential for alignment of three individual helix-constituting procollagen chains to form a trimer. The formation of this triple helix in newly synthesized procollagen molecules is known to proceed in a zipper-like manner from the C- to the N-terminus (Bruckner *et al.*, 1978).

At the moment, efforts are increasing to develop an efficient and safe production system for recombinant human collagens (Bulleid *et al.*, 2000, Ruggiero *et al.*, 2000). Several groups have used yeast cells for recombinant collagen production (Vuorela *et al.*, 1997, Vaughan *et al.*, 1998, Toman *et al.*, 2000, Olsen *et al.*, 2001).



Especially, methylotrophic yeasts have a reputation of high yield production of heterologous proteins (Cregg *et al.*, 2000, Gellissen, 2000 van Dijk *et al.*, 2000). As yeasts are considered to be incapable of prolyl 4-hydroxylation (Vuorela *et al.*, 1997, Bulleid *et al.*, 2000, Toman *et al.*, 2000), production of recombinant collagens in yeasts was assumed to require co-expression of an animal prolyl 4-hydroxylase (EC 1.14.11.2). However, we recently discovered that the methylotrophic yeast *Hansenula polymorpha* expressed endogenous collagen-like sequences with 4-hydroxyproline residues, exclusively in the X2 position of G-X1-X2 repeats, exactly like in animal collagens (Chapter 4, de Bruin *et al.*, 2002). Most importantly, it was shown that the endogenous activity of *H. polymorpha* could be used to obtain sequence-specific hydroxylation of secreted, recombinant gelatin (Chapter 4, de Bruin *et al.*, 2002). In this study we investigated the suitability of *H. polymorpha* for the production of recombinant human collagen trimers.

## Materials en methods

### *Yeast strains and plasmids*

The *H. polymorpha* strain NCYC 495 (*leu2*; (Gleeson *et al.*, 1986), which is deficient in beta-isopropylmalate dehydrogenase, was used in this study. For methanol induced procollagen expression in *H. polymorpha* we used the vector pHIPX4 (Gietl *et al.*, 1994), which contains a LEU2 selectable marker and an expression cassette of the *H. polymorpha* methanol oxidase (MOX) promoter and the amino oxidase (AMO) terminator. The *P. pastoris* strain GS115 (*his4*, Cregg *et al.*, 1985) and the vector pPICZ $\alpha$ A (Invitrogen, San Diego, CA, USA) were used for methanol-induced procollagen expression. pPICZ $\alpha$ A contains a ZEOCIN selectable resistance marker and the *P. pastoris* AOX1 promoter/terminator expression cassette.

In order to investigate the secreted production of human collagen, we used a modified  $\alpha$  1 (I) procollagen construct. The *S. cerevisiae*  $\alpha$ -mating factor pre-pro sequence was fused to a truncated human  $\alpha$  1 (I) procollagen gene missing the natural procollagen pre-, N pro- and N telopeptide sequence plus approximately 15 % of the

helical domain. For detection of the recombinant product a C-terminal *myc*-tag was fused to the C-propeptide. The modified procollagen encoding cDNA fragment was amplified by the polymerase chain reaction (PCR) using the primers HCOL-FW 5'-CTGGTACCTCTGGCTTCCCTGGTGCT GT-3' and HCOL-RV 5'-TACCGCGGAAGTTCCAGGAAGCAGACAG-3' and COL1A1-CMV (Geddis and Prockop, 1993) as a template. The resulting fragment was *KpnI* / *SacII* digested and inserted into the *KpnI* / *SacII* sites of *P. pastoris* vector pPICZαA. The resulting vector was subsequently digested with *HindIII* / *SalI*, yielding a 3.8 kb fragment containing the *S. cerevisiae* α-mating factor secretion sequence, the modified form of the human α 1 (I) procollagen and the C-terminal *myc*-tag. This fragment was finally ligated into the *HindIII* / *SalI* sites of *H. polymorpha* vector pHIPX4.

### Media composition

Minimal glucose plates, for selection of transformants, contained 1.34 % yeast nitrogen base without amino acids (Difco, Detroit, MI, USA), 1% glucose and 1.5 % agar. For small scale (1L shake flask, containing 100 ml of culture medium) expression of procollagen in *H. polymorpha*, transformants were precultured in mineral medium (van Dijken *et al.*, 1976) containing 0.5 % (w/v) of glucose. Production of procollagen was induced in mineral medium containing 0.5 % (v/v) of methanol and 1 % (w/v) of peptone. The preculture was diluted in this medium to an optical density of 0.100 at 600 nm, as measured in a Corning colorimeter 254, using disposable cuvettes with 1 cm pathlength. Culture samples were taken at 12 h intervals for western blot analysis and immuno cytochemistry.

### Transformation of *H. polymorpha*

Transformation of *H. polymorpha* by electroporation was performed according to Faber *et al.*, 1994, using a GenePulser (Bio-Rad, Richmond, CA, USA). Prior to transformation, the pHIPX4 procollagen construct was linearized in the *Pmox* by *Sca I* digestion. After growth for 3 days on minimal glucose plates at 37 °C, obtained transformants were grown for about 20 generations on non-selective medium (YPD) to stabilize them. After subsequent growth on minimal glucose plates, mitotically stable clones were subjected to PCR to confirm the transformation procedure. Cells

were used directly for PCR without any pretreatment using a primer based on the human procollagen gene and on the *H. polymorpha* AMO-terminator.

### **Transformation of *P. pastoris***

*P. pastoris* GS115 (*his4*, Cregg et al., 1985) has a nonfunctional histidinol dehydrogenase gene. In order to enable endogenous production of histidinol dehydrogenase, the gene was restored as described by Munshi and Lee, 1997. *P. pastoris* was transformed by electroporation according to Werten *et al.*, 1999, using a GenePulser (Bio-Rad). Prior to transformation the pPicZ $\alpha$  procollagen construct was linearized in the *Paox1* by *Pme* I digestion. After growth for 3 days on minimal glucose plates, containing 100  $\mu$ g / ml ZEOCIN at 30 °C, several colonies were selected for PCR confirmation. Cells were used directly for PCR without any pretreatment, using primers based on the human procollagen gene and on the *P. pastoris* AOX1-terminator.

### **Procollagen production by *H. polymorpha* and *P. pastoris* in bioreactors**

For the production of procollagen in *H. polymorpha* and *P. pastoris* in bioreactors, fermentations were performed as described previously (Chapter 2, de Bruin *et al.*, 2000), with the exception that the medium of *H. polymorpha* contained 1 % (w/v) of peptone (Duchefa, Haarlem, the Netherlands) to induce the endogenous prolyl 4-hydroxylase activity. In the culture medium of *P. pastoris*, 1 % (w/v) of casamino acids (Merck, Darmstadt, Germany) was added to reduce proteolysis. We showed previously that the addition of peptone to the medium did not induce hydroxylation activity in *P. pastoris* (Chapter 4, de Bruin *et al.*, 2002). After inducing the cells for 24 hours with methanol, cells were separated from the culture medium by centrifugation. The cells were subsequently washed with PBS (150 mM NaCl, 20 mM NaPi, pH=7.5) and crude cell extracts were prepared using acid-washed glass beads as described by Waterham *et al.*, 1994. The triple-helical conformation of the recombinant  $\alpha$  1 (I) procollagen was examined by pepsin digestion as described by Myllyharju, 2000, using porcine pepsin obtained from Sigma, St. Louis, MO, USA. Pepsin was added to cell extracts to a final concentration of 200  $\mu$ g / ml, and samples were digested over

night at 4 °C. After digestion, pepsin was inactivated by adjusting the pH to 7.5 and samples were analyzed by SDS-PAGE.

#### **SDS-PAGE, N-terminal protein sequencing and immunoblotting**

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed in a Mini PROTEAN II system (Bio-Rad) under reducing, denaturing conditions. Gels were stained with Coomassie Brilliant Blue (CBB R-350) For N-terminal protein sequencing, protein was blotted onto Immobilon P<sup>SQ</sup> (Millipore) by applying 100V for one hour in a Mini Trans-Blot Cell (Bio-Rad). Transfer buffer was 2.2 g CAPS per liter of 10 % methanol, pH 11. Blots were stained with Coomassie Brilliant Blue (CBB R-350) and selected bands were cut out. N-terminal sequencing using Edman-degradation was performed by Midwest Analytical, St. Louis (MO), USA.

For immunoblotting, protein was electrophoretically transferred onto a PVDF filter (Millipore). The filter was blocked with 5 % skim milk powder in TBST (0.1 M Tris-HCl, pH 7.5; 1.5 M NaCl; 0.1 % Tween-20) at room temperature for 1h. The filter was incubated overnight with monoclonal anti-*myc* antibody (Roche; 1 : 20.000 in 1% skim milk in TBST), washed with TBST, and incubated for 1h with a secondary antibody-conjugated to alkaline phosphatase (AP) (goat anti-mouse, Sigma; 1 : 10.000 in 1% skim milk in TBST). The filter was washed with TBST and then rinsed with AP buffer (0.1 M Tris-HCl, pH 9.5; 0.5 M MgCl<sub>2</sub>; and 0.1 M NaCl). Antibody-binding was detected by incubating the filter in 10 ml AP buffer containing 33 µl of 5-bromo 4-choro 3-indoyl phosphate (50 mg/ml) and 66 µl of nitro-blue tetrazolium (50mg/ml) (Amersham Pharmacia Biotech, Uppsala, Sweden).

#### ***Electron microscopy***

Procollagen producing *H. polymorpha* cells were fixed and prepared for electron microscopy and immuno gold labelling as described by Keizer-Gunnink *et al.*, 2000 and viewed on a Philips EM 400 electron microscope.

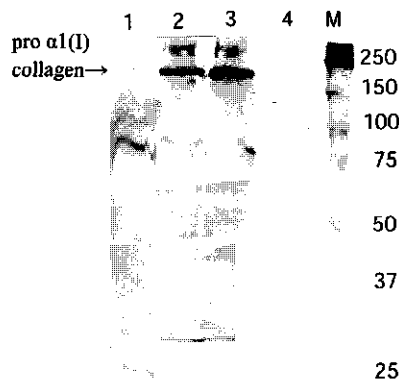
## Results

### *Production of homotrimeric $\alpha 1(I)$ collagen by *H. polymorpha**

Expression of human procollagen in *H. polymorpha* transformants was induced by growing the cells in medium containing 0.5 % (v/v) of methanol. 1 % (w/v) of peptone was added to induced the endogenous prolyl 4-hydroxylase activity in *H. polymorpha* (see Chapter 4, de Bruin *et al.*, 2002). The recombinant procollagen, as identified by the attached *myc*-epitope, was only found intracellular as shown by western blot analysis (Fig. 5.1) and SDS-PAGE analysis (not shown). Also without the addition of 1 % (w/v) peptone to the culture medium, procollagen trimers were only present intracellularly (not shown). According to its molecular weight, the procollagen product was apparently intact inside the cells. Despite the presence of the  $\alpha$ -MF leader sequence of *S. cerevisiae*, the recombinant procollagen was not secreted by *H. polymorpha*. This is in agreement with results obtained earlier with recombinant *P. pastoris* and *S. cerevisiae* strains producing procollagen trimers (Vuorela *et al.*, 1997, Vaughan *et al.*, 1998, Toman *et al.*, 2000).

**Figure 5.1.** Western blot analysis of crude extracts

(30  $\mu$ g of total cell protein) during methanol induced expression of  $\alpha 1(I)$  procollagen in *H. polymorpha*. Lane 1, intracellular proteins before methanol induction. Lane 2 and 3, intracellular proteins after 12 and 24 hours of methanol induction, respectively. Lane 4, culture medium after 24 hours of methanol induction. The arrow indicates the recombinant  $\alpha 1(I)$  procollagen. Lane M: precision protein standards (Bio-Rad, Richmond, CA, USA).





**Figure 5.2 (A).** Electron micrograph of an ultrathin section of the *H. polymorpha* strain producing  $\alpha 1(I)$  procollagen. In the cells a strong proliferation of membranes, as indicated by arrows, and swelling of the cell wall (\*) was observed. The arrows indicate the proliferated endoplasmic reticulum. The bar represents 0.5  $\mu\text{m}$ . **(B).** Immuno gold labelling, using anti-myc antibodies, showed specific labelling in the endoplasmic reticulum. Abbreviations: N= nucleus, P= peroxisome, M= mitochondrion.

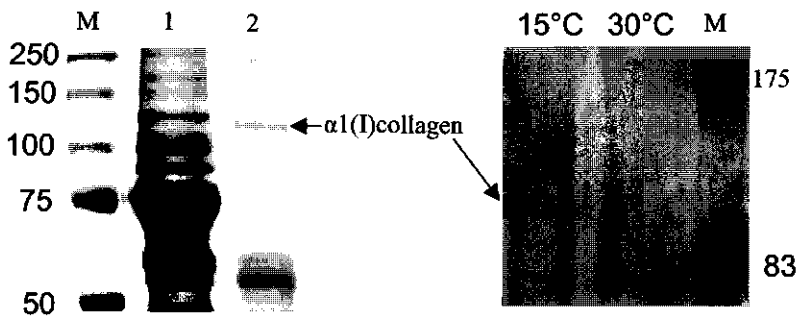
Ultra-structural analysis of the *H. polymorpha* strain producing procollagen showed a strong proliferation of membranes (Fig. 5.2A). Since these membranous structures were continuous with the nuclear membrane, they presumably attribute to the endoplasmic reticulum. In addition, in many cells a swelling of the cell wall was observed (Fig 5.2A). The structures mentioned were invariably absent in wild-type cells, indicating that the specific structures developed in response to the synthesis of recombinant procollagen. Strong ER proliferation and cell wall swellings have also been observed during *Xenopus laevis* magainin II expression in *H. polymorpha* (Faber *et al.*, 1996) and are probably related to general ER stress (K.N. Faber, personal communication).

The intracellular distribution of the recombinant procollagen was studied by immunocytochemistry on ultrathin sections (Fig. 5.2B). The recombinant  $\alpha 1$  (I) procollagen was located within the lumen of the proliferated compartments attributed to the endoplasmic reticulum. Probably, the procollagen molecules are accumulating in the endoplasmic reticulum and are not further proceeding into the secretory pathway of the yeast.

In bioreactors, the expression level of intracellular procollagen was 0.2 % of the total cellular protein as quantified by western blot analysis using scanning densitometry (Bio Rad GS-800 calibrated densitometer). This corresponds to about 100 mg / L, at 300 g / L wet cell weight. The Positope<sup>TM</sup> protein (Invitrogen) was used as a reference for quantification.

To examine if procollagen trimers were produced by *H. polymorpha*, cell-extracts were prepared and subjected to pepsin digestion. The triple helical domain of collagen can be discerned from unfolded collagen and other proteins because it is resistant to a variety of proteolytic enzymes, including pepsin. One distinct band was seen on SDS-PAGE after pepsin treatment of cell extracts of methanol-induced *H. polymorpha* cells (Fig. 5.3A). The pepsin resistant protein has a lower apparent molecular weight (120 kDa) than the original procollagen (175 kDa) and migrated in SDS-PAGE to the position expected of the intact helical domain of the  $\alpha 1$  (I) collagen, devoid of the *S. cerevisiae*  $\alpha$ MF-propeptide and procollagen C propeptide. The N-terminal amino acid sequence of the pepsin resistant protein (PGAVGAKGE)

indeed corresponded to the start of the helical domain of the recombinant collagen molecule. Due to a lack of material, it was not possible to obtain reliable data on the presence or absence of 4-hydroxyproline, either by amino acid sequencing or by analysis of the amino acid composition. The pepsin resistance showed, however, that the  $\alpha 1(I)$  procollagen molecules produced are properly folded to form homotrimeric  $\alpha 1(I)$  collagen molecules.



**Figure 5.3 (A).** SDS-PAGE analysis of pepsin treated cell extract of  $\alpha 1(I)$  procollagen producing *H. polymorpha*. Lane 1, cell extract, Lane 2, pepsin treated cell extract. The arrow indicates the pepsin resistant  $\alpha 1(I)$  collagen, as identified by N-terminal amino acid sequencing. Lane M, prestained markers (Bio-Rad, Richmond, CA, USA). **(B).** The thermal stability of the recombinant collagen homotrimers, during pepsin digestion at different temperatures. Lane M, prestained markers (New England Biolabs, Northbrook, IL, USA).

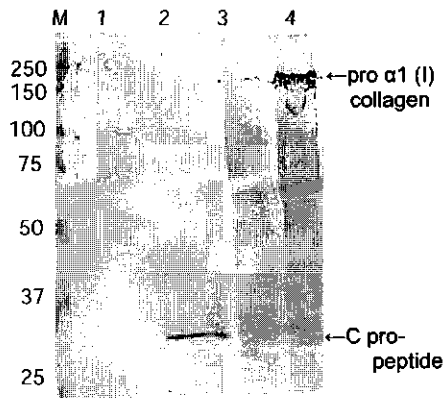
The thermal stability of the triple helical molecules produced, was monitored by their pepsin resistance at different temperatures (Fig. 5.3 B). The results show that the melting temperature ( $T_m$ ) of the recombinant collagen trimers was between 15-30 °C. This low  $T_m$  could point to a low level of hydroxylation. It has been previously reported that the thermal stability of collagens is directly related to the hydroxyproline content. Intact, unhydroxylated triple helical  $\alpha 1(I)$  collagen homotrimers have a  $T_m$  of 24 °C, while fully hydroxylated  $\alpha 1(I)$  collagen homotrimers have a  $T_m$  of about 40 °C (Geddis and Prockop, 1993, Toman *et al.*, 2000). On the other hand, in our construct 15 % of the N-terminus of the helical domain was missing, which also lowers the  $T_m$  of the produced recombinant collagen.



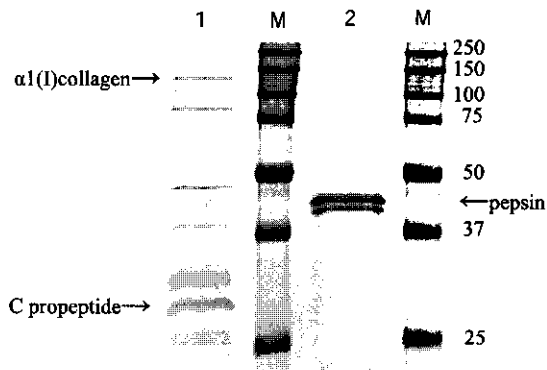
**Expression of human procollagen in *P. pastoris***

In this thesis a comparison is drawn between *H. polymorpha* and *P. pastoris* with regard to recombinant gelatin and collagen production. Therefore, in addition to *H. polymorpha*, we also expressed the same procollagen construct in *P. pastoris*. In contrast to *H. polymorpha* intracellular accumulation of procollagen trimers did not occur in *P. pastoris* as determined by western blot analysis (Fig. 5.4A). Only a low amount of recombinant procollagen was found inside *P. pastoris* cells. The blot also showed that, after methanol induction, a *myc*-labelled protein of about 30 kDa was present in the culture medium (Fig. 5.4A). N-terminal amino acid sequencing of the protein yielded SLQQIEN, which is present in N-terminal part of the C propeptide. The C propeptide was apparently removed from the helical domain by an endogenous protease of *P. pastoris*. SDS-PAGE analysis of supernatants showed that during fermentation a 120 kDa protein was produced into the medium (Fig. 5.4B). N-terminal sequencing showed that this protein appeared to be the non-hydroxylated, intact helical domain of  $\alpha 1$  (I) collagen, devoid of the C propeptide. Pepsin digestion showed that the intact helical domains of  $\alpha 1$  (I) collagen in the medium were not present as collagen trimers (Fig. 5.4B).  $\alpha 1$  (I) collagen homotrimers were also not detected after pepsin digestion of cell extracts (not shown).

**Figure 5.4 (A).** Western blot analysis of a *P. pastoris* strain producing human procollagen. Lane 1, culture medium before methanol induction. Lane 2, culture medium after 24 hours of methanol induction. Lane 3, intracellular proteins before methanol induction. Lane 4, intracellular proteins after 24 hours of methanol induction. Lane M, precision protein standards (Bio-Rad, Richmond, CA, USA).



**Figure 5.4 (B).** SDS-PAGE analysis of the *P. pastoris* strain producing human procollagen. Lane 1, culture medium after 24 hours of methanol induction. The arrows indicate the intact  $\alpha 1(I)$  collagen and the C propeptide as identified by N-terminal amino acid sequencing. Lane 2, culture medium after pepsine digestion. Lane M, precision protein standards (Bio-Rad, Richmond, CA, USA).



## Discussion

In this study we showed that protease resistant  $\alpha 1(I)$  collagen homotrimers were intracellularly produced by *H. polymorpha*. The lack of secretion and the intracellular localisation of the product in the endoplasmic reticulum of the yeast showed that, despite the presence of the  $\alpha$  MF secretion signal, the procollagen molecules did not further proceed in the secretory pathway. The recombinant collagen trimers were not stable at temperatures higher than 20 °C, indicating that the product was poorly hydroxylated.

N propeptides are believed to play no role in the assembly of procollagen molecules into trimers. Even deletion of both the N pro and the C propeptide peptide domains, only leaving the N telo and C telo peptides, resulted in correctly aligned triple helical collagen (Olsen et al., 2001). In this study, we expressed a modified form of  $\alpha 1(I)$  procollagen without N pro but also without the N telo peptide and showed that collagen trimers were formed. Our results indicate that the N telo peptide is also not required for procollagen trimer formation and that the C pro and C telo domains are much more important for collagen trimer formation than the N pro and N telo domains.

Expression of the modified procollagen in *P. pastoris*, which did not contain recombinant prolyl 4-hydroxylase, resulted in secretion of the intact, non-hydroxylated helical domain (120 kDa), without the C propeptide attached. The propeptide was removed from the helical domain and was found in the culture medium. This suggests that an endogenous protease of *P. pastoris* spontaneously processed the C propeptide at a single site, 18 amino acids upstream of the native

cleavage site. In transgenic plants producing procollagen spontaneous processing of the propeptides was also found (Ruggiero *et al.*, 2000). Probably, cleavage of the propeptides occurs in a region, which is more sensitive to proteases. In animal cells, cleavage of the C propeptide is performed by a specific extracellular C proteinase (EC 3.4.24.19) (Hojima *et al.*, 1985).

It has been described that recombinant prolyl 4-hydroxylase containing *P. pastoris* and *S. cerevisiae* strains have difficulties in secreting procollagen trimers (Vuorela *et al.*, 1997, Vaughan *et al.*, 1998, Keizer-Gunnink, *et al.*, 2000, Toman *et al.*, 2000, Nokelainen *et al.*, 2001). The lack of secretion may be related to the large size of the procollagen trimers. We showed here that, as opposed to *H. polymorpha*, procollagen trimers were not formed inside *P. pastoris* cells, which did not contain recombinant prolyl 4-hydroxylase. This lack of trimer formation probably led to secretion of the procollagen molecules in *P. pastoris*.

### **Acknowledgements**

The authors thank Dr. D.J. Prockop, Thomas Jefferson University, Philadelphia, USA, for providing the COLIA1-CMV vector containing the cDNA of the pro $\alpha$ 1 chain of type I collagen.

## Chapter 6

### General discussion

#### Abstract

**In this thesis the methylotrophic yeasts *Hansenula polymorpha* and *Pichia pastoris* were investigated in their use for the recombinant production of gelatin and collagen. Prolyl 4-hydroxylation (essential for the recombinant production of gelatins with gelling properties and thermally stable collagens) was generally considered to be absent in yeasts. However, in this thesis we describe the discovery of endogenous collagen-like protein sequences with 4-hydroxyproline residues, indicating that *H. polymorpha* contains intrinsic prolyl 4-hydroxylase activity. In this last chapter, the results reported in the preceding chapters of this dissertation, together with some additional experimental results, are discussed.**

### 6.1 Endogenous collagen-like proteins of *H. polymorpha*

We discovered that *H. polymorpha* produces at least two different collagen-like proteins (or protein fragments), which are present in the extracellular medium. One with an apparent weight of 38 kDa with and with an N-terminal sequence of GPOGPOGPOGPO GPOGPOGPOG (Chapter 4) and one with an apparent molecular weight of about 120 kDa with an N-terminal sequence of GGOGGGOGGOG, where O represents 4-hydroxyproline. Possibly, both sequences are triple helical parts of proteins. At least they were found to be protease resistant (data not shown).

Hydroxylated, collagenous proteins have never been found in yeasts, but have been described for the smut fungus *Microbotryum violaceum* (Celerin *et al.*, 1996). These cell surface-attached proteins of *M. violaceum* seem to have a more irregular sequence, often deviating from the consistent [G-X1-X2] triplet structure. They can be released by mechanical shear (Celerin *et al.*, 1996), or by heat treatment (Celerin, M., personal communication). The collagen-like proteins of *H. polymorpha* behave very similar. Firstly, because they are released into the medium of high cell density fermentations (possibly due to the shearing forces caused by the high-speed agitation). Secondly, these collagen-like proteins can be released from thoroughly washed, intact cells by heat-treatment (Chapter 4). This suggests that the *H. polymorpha* proteins may also be cell surface-attached. As for its function, they may be structural cell wall proteins involved in cell-cell interactions and mating as has been suggested for the *M. violaceum* protein (Celerin *et al.*, 1994). Alternatively, the triple-helical motif is found in such diverse mammalian proteins as host-defense proteins, membrane receptors, surfactant proteins, mannose binding protein etc (Brodsky and Shah, 1995).

Likewise, the *H. polymorpha* collagen-like proteins may serve one of many possible purposes. Research is currently ongoing to further characterise the collagenous proteins of *H. polymorpha*.

## 6.2 Trigger of the prolyl 4-hydroxylase activity

The finding of endogenous collagen-like proteins containing 4-hydroxyproline residues, indicated that *H. polymorpha* must contain intrinsic prolyl 4-hydroxylase activity. Expression of recombinant animal gelatin in *H. polymorpha* yielded a secreted and hydroxylated product, without the need of co-expressing the recombinant, animal prolyl 4-hydroxylase genes (Chapter 4). The addition of peptone to the cultivation medium of *H. polymorpha* was found to be necessary for the hydroxylation of the endogenous and recombinant collagen-like peptides (Chapter 4). Incorporation of 4-hydroxyproline into native and recombinant proteins could be excluded (Chapter 4). Therefore, components in the peptone probably induce the prolyl 4-hydroxylase expression/activity in *H. polymorpha*. In order to find out which components are responsible for the induction of the endogenous prolyl hydroxylase activity in *H. polymorpha*, some additional experiments were performed. First, we checked if a mixture of free amino acids, mimicking the amino acid balance of peptone, could trigger hydroxylation of proline residues in *H. polymorpha*. Supplementation of the fermentation medium with such a mixture of free amino acids, however, did not induce hydroxylation. (Table I). When a trypsinized and ultra-filtrated (< 10 kDa) gelatin solution was added to the culture medium, the recombinant product was hydroxylated (Table I). Addition of the recombinant, non-hydroxylated COL3A1 gelatin (Werten *et al.*, 1999) did not have an effect on the hydroxylation (Table I). This indicated that oligo-peptides present in peptone but not in COL3A1 induced prolyl 4-hydroxylation in *H. polymorpha*. Possibly, prolyl 4-hydroxylation in the oligo-peptides is important, as COL3A1 is not hydroxylated. With two synthetic oligo-peptides (one hydroxylated and the other not hydroxylated) we could, however, not induce prolyl 4-hydroxylation in the recombinant protein (Table I).

Maybe the *H. polymorpha* cells contain collagen receptors, which recognize a certain specific hydroxylated oligo-peptide sequence as described for platelets. The receptors located on the platelet surface recognize the collagen sequence GFOGER, in which O is 4-hydroxyproline, initiating haemostasis (Monnet *et al.*, 2000). By a

similar mechanism, a cascade may be elicited in *H. polymorpha*, which finally leads to prolyl 4-hydroxylase expression or activation in this organism.

**Table I.** Medium dependant prolyl 4-hydroxylation of recombinant Colla1 by *Hansemula polymorpha*.

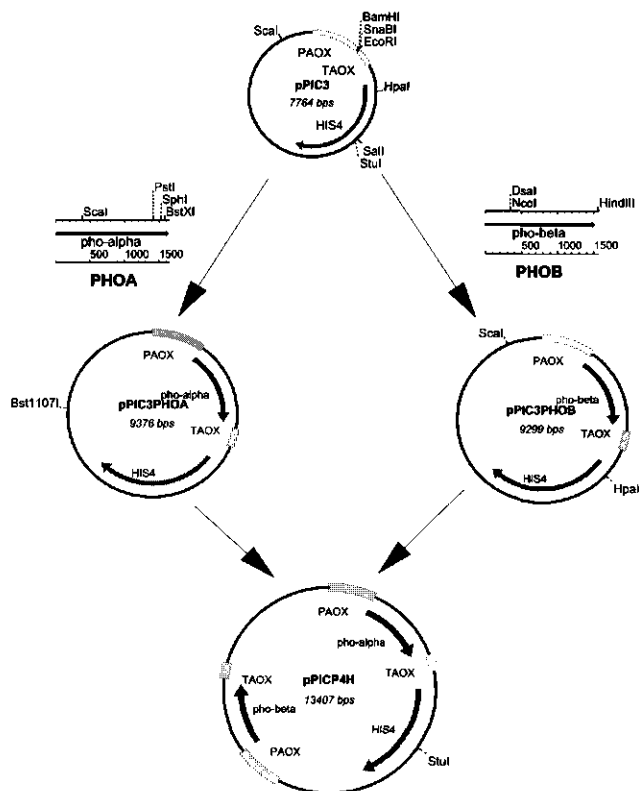
Mode of expression	Supplement to cultivation medium	4-hydroxyproline in Colla1 sequence <sup>a</sup>
Methanol-induced	Peptone-mimicking mixture of free amino acids	–
	Co-factors <sup>b</sup>	–
	< 10 kDa trypsinized gelatin <sup>c</sup>	+
	Non-hydroxylated COL3A1	–
	Synthetic collagenous oligo-peptide (GAPGPPGSPGPPGAPGYP)	–
	Synthetic collagenous oligo-peptide (GAOGPOGSOGPOGAOGYO)	–
	poly-L-proline	–
	poly-L-4-hydroxyproline	–

<sup>a</sup> The N-terminal sequence was : GFQGPPGEPG (referred to as –) and GFQGPOGEOG (referred to as +), where O is 4-hydroxyproline. <sup>b</sup> The protein band was analyzed in the > 10 kDa ultra-filtration fraction of the culture medium that contained the < 10 kDa trypsinized gelatin fraction. <sup>c</sup> 0,3 % w/v of FeSO<sub>4</sub> (10mM), 0,3 % w/v of α-ketoglutarate(20 mM) and 0,175 % w/v of vitamin C (10 mM).

### 6.3 Production of hydroxylated gelatin in *P. pastoris*

*P. pastoris* is very good producer of non-hydroxylated gelatins (Werten *et al.*, 1999), but it appeared to contain no endogenous prolyl 4-hydroxylase activity (Chapter 4). Therefore, the production of hydroxylated gelatins in *P. pastoris* requires co-expressing of a heterologous prolyl 4-hydroxylase. In an attempt to produce hydroxylated gelatins in *P. pastoris*, we isolated the two subunits of the human prolyl 4-hydroxylase by RT-PCR using human diploid lung fibroblasts (HFL-1, Breul *et al.*, 1980). The signal sequence derived from the *P. pastoris* acid phosphatase gene (*PHO1*) was used for targeting of the subunits to the endoplasmic reticulum of the yeast (see Fig. 6.1. for the cloning-strategy).

**Figure 6.1.** Construction of the plasmid pPIC4H required for recombinant prolyl 4-hydroxylase expression in *P. pastoris*. The sequences encoding the native signal peptides of the human  $\alpha$  (I)-subunit and of the human  $\beta$ -subunit (PDI) were replaced by the *P. pastoris* acid phosphatase (*PHO1*) leader sequence. The C-terminal ER-retention signal -KDEL of  $\beta$ -subunit was modified to an -HDEL signal by PCR. The subunits were ligated into the *P. pastoris* vector pPIC3. The resulting expression cassettes were subsequently combined to yield plasmid pPIC4H.





Intracellular expression of the human prolyl 4-hydroxylase subunits separately, and in combination was obtained in *P. pastoris*, as determined by western blot analysis using mono-clonal antibodies against the subunits (ICN Biomedical, Costa Mesa, CA, USA). Subsequent introduction of the *COL1A1* encoding sequence in a strain expressing both hydroxylase subunits however, did not yield hydroxylated gelatin. Probably, in contrary to other studies (Vuorela *et al.*, 1997, Nokelainen *et al.*, 2001), no active recombinant  $\alpha_2 / \beta_2$  tetramer was formed in *P. pastoris*.

Gelatin is extremely susceptible to proteolytic degradation due to its open, unfolded structure. When the *H. polymorpha* and *P. pastoris* systems are compared, the extent of proteolytic degradation of secreted, recombinant gelatin is much greater in *H. polymorpha* than in *P. pastoris*. On the other hand, *H. polymorpha* contains the endogenous capacity to hydroxylate and secrete collagenous proteins. To combine both favourable features, we tried to isolate the gene(s) responsible for prolyl 4-hydroxylation in *H. polymorpha*, which subsequently could be introduced into *P. pastoris*. When isolated, the homologous hydroxylase could, in principle, also be targeted into the endoplasmic reticulum of *H. polymorpha* to increase hydroxylation of recombinant animal gelatin and human procollagen produced by this organism.

We used PCR and RT-PCR with degenerate primers based on conserved regions of the  $\alpha$ -subunits of animal prolyl 4-hydroxylases and used low stringency oligo hybridization of gDNA with heterologous hydroxylase probes, but unfortunately without success. Also a subtracted cDNA library from two pools of mRNA of *H. polymorpha* grown in different media (with and without peptone) did not yield prolyl 4-hydroxylase-like genes.

#### 6.4 Prolyl 4-hydroxylase activity measurements

Attempts to measure prolyl 4-hydroxylase activity in cell extracts of *H. polymorpha* using an *in vitro* assay based on the hydroxylation-coupled decarboxylation of [1-<sup>14</sup>C]  $\alpha$ -ketoglutarate (Kivirikko and Myllylä, 1982) failed, due to high backgrounds caused by decarboxylation of the Krebs cycle-intermediate alpha-ketoglutarate by other enzymes in the cell extract. An attempt has been made to purify the prolyl 4-

hydroxylase from *H. polymorpha* by affinity chromatography on non-hydroxylated Col3a1-gelatin columns. By this method a cis-trans-peptidylprolyl isomerase (PPIase, EC 5.2.1.8) was purified from *H. polymorpha*. This enzyme is important for the folding of proteins (Walsh *et al.*, 1992), especially for the folding of collagen-like proteins into triple helical molecules (Steinman *et al.*, 1991). Prolyl 4-hydroxylase activity measurements of column fractions showed activity, but N-terminal sequencing of proteins (other than PPIase) present in the fractions failed, thus far, because the amount of protein in these fractions was too low. Research is ongoing to identify the prolyl 4-hydroxylase of *H. polymorpha*.



**Figure 6.2.** Alignment of the amino acid residues of the putative prolyl 4-hydroxylase of *Aspergillus niger* (A) and *Neurospora crassa* (N), with the C-terminal and catalytically active part of the human  $\alpha$  (I)-subunit (H1), the human  $\alpha$  (II)-subunit (H2) and the *Drosophila melanogaster*  $\alpha$ -subunit (D) of the respective prolyl 4-hydroxylase. The amino acid residues in H1, H2 and D, which precede the alignment region are not shown. The three Fe<sup>2+</sup>-binding residues, two histidines and an asparate are indicated in bold with gray backgrounds. The lysine, which ionically binds the C-5 carboxyl group of  $\alpha$ -ketoglutarate is indicated in dark background. The Clustal W program (Thompson *et al.*, 1994) was used for the alignment of the multiple sequences.

### 6.5 Fungal prolyl 4-hydroxylase genes

In the course this study, we discovered fungal proteins, which showed homology with the catalytic part of different  $\alpha$ -subunits of animal prolyl 4-hydroxylases (Fig. 6.2).

The proteins were found in the genomic databases of *Neurospora crassa* and *Aspergillus niger*, from which part and the complete genomic sequence is elucidated, respectively. Additionally, an overall identity of about 30 % was found between the fungal hydroxylases with a viral prolyl 4-hydroxylase, which acts as a monomer (Eriksson *et al.*, 1999). Also in this case the iron-binding triad and the lysine residue involved in binding of  $\alpha$ -ketoglutarate were conserved (Fig. 6.3).

After isolation of the hydroxylase-like genes from *Neurospora crassa* and *Aspergillus niger* by RT-PCR, the respective genes were over-expressed in *E. coli*. We could successfully produce the putative hydroxylase of *N. crassa*, but expression of the putative *A. niger* hydroxylase was not obtained in *E. coli*. Purification of the recombinant hydroxylase-like peptide of *N. crassa*, containing a 6HIS affinity tag, was obtained under denaturing conditions by immobilized-Ni affinity chromatography. After refolding of the protein an activity assay was carried out with different substrates. Prolyl 4-hydroxylase activity was, however, not found, using poly-GPP or poly-PPP as a substrate. Maybe, the putative hydroxylase was not properly refolded, or co-expression of the fungal PDI is required for activity. Yet, the finding of these fungal hydroxylase-like genes could be a new lead for the isolation of the *H. polymorpha* gene encoding the prolyl 4-hydroxylase as soon as the genomic information on *H. polymorpha* will be available. In this study we showed that the latter enzyme is suitable for the hydroxylation of recombinant animal gelatin. In the future, knowledge about the endogenous processing of collagen in *H. polymorpha* can be applied to further develop the methylotrophic yeasts *H. polymorpha* and *P. pastoris* into high yield production systems of hydroxylated gelatins and collagens.

```

N.crassa      1 MSGLGLTKAAGLVVANTVIVALSAPYLSSETFSQISSLTFIISGLPVPDHPPLAKFLKLFRN
A.niger      1 -----MPKLSLSLTLTSLFYLLPLYLFVIAPILLRQLFPVTQDIFDPDTLATTESDISLD
PBCV-1      1 -----MTNKFISYNKMETREYLLTILFVIACFMVLNLERREGFETSDR

N.crassa      TATCTEKPRQFFFDVSGSYKIELVSLDPLIYYIHNLITPSEIISLLVTAEPMPKPEVVT
A.niger      PEILS-LPFDGIIFTCPEPTYRVHLLSKDPLIYYIEDFLSTAEADHLVKISEPNYSPEIYY
PBCV-1      PGVCDGKYYEKILGFLIEDIECDVLIN-----AAIKKGLIKSEVGGATENDPIKLEPKSRN
      .           :           .           *;.           *.. :   !* .           *.

N.crassa      KHGRQQQTQDRTESSAGLPRDDPAVKCVLDRANGFMGTMLRDGWDGMPPEQ-LVRVRAGQ
A.niger      NGQTIKVDPSKRLSDRALLPRDNTVRCLESRAAAFQGWKPHLYIERMWAQ----RYNWSG
PBCV-1      SEQ-----TWFMPEHEVVDKIQKKTREFLNSKKHCIDKYNFEDVQVARYKPGQ
      .           :           . :   :   :   * .           :           .           **.

N.crassa      RENVEDMFDRPLLANDGSNRVWNRVASFFAILQDNCIKGE--TFPHAKEFILLSSPWE
A.niger      HYKHEDMAE-----SVARSGDRLSTFMVYLGDECEGGG--TFPPLK-----
PBCV-1      YYYEDGGD-----CDDACPKDQRLATLMVYLKAPEGGGGGETIFPTLK-----
      :   *.* .           .           *:::!: . *   *   *.* *

N.crassa      FMWRGGLEBQNLKNVTDLKPLWBEHEDGGLAFREVAAGN-AIFWVNLHANGTGDERTMAAG
A.niger      -MPRGDQWQDFLD-----CEMAQENNVSGITFKFIRGN-AIFWENLRSDGTGYRETMAAA
PBCV-1      -----TRIKERKGTSIFFWADVTRKLYKETMAAG
      :!*   * .   !*           .   *   **

N.crassa      LFPGGGLKLTAMNIWPRRYWEY----- 318
A.niger      FPVTGGKLVGLNIWNYQPEKGRRYIDSY 293
PBCV-1      LPVKSGENIJANQWIRAVK----- 242
      :*   * *   * *
    
```

**Figure 6.3.** Alignment of the amino acid residues of the putative prolyl 4-hydroxylase of *Neurospora crassa* and *Aspergillus niger* with the residues of the monomeric prolyl 4-hydroxylase of *Paramecium bursaria* *Chlorella virus* (PBCV-1) (Eriksson et al., 1999). The signal peptides sequences, as predicted using the computation parameters of Nielsen et al., 1997, are indicated in bold. The three Fe<sup>2+</sup>-binding residues, two histidines and an aspartate are indicated in bold with gray backgrounds. The lysine binding the C-5 carboxyl group of the  $\alpha$ -ketoglutarate is indicated with gray background. The Clustal W program (Thompson et al., 1994) was used for the alignment of the multiple sequences.

### **Acknowledgements**

Marco van den Berg, DSM, Delft, The Netherlands is acknowledged for providing the DNA sequence of the putative prolyl 4-hydroxylase gene of *Aspergillus niger*, Jan Springer and Liesbeth Verhagen for subtractive cDNA screening and cloning of hydroxylase genes of *Neurospora crassa* and *Aspergillus niger*, and Marc Werten for performing the affinity column procedure in a first attempt to isolate the prolyl 4-hydroxylase from *H. polymorpha*.

## Summary

Gelatin and collagens are used in the cosmetics and food industries as well as for the production of medical and surgical supplies. They are extracted from animal tissues and are therefore at risk from contamination by infectious agents (Chapter 1). In this study a comparison is drawn between two different methylotrophic yeasts in respect to the production of recombinant gelatin and collagen.

Biotechnological interest in methylotrophic yeasts arose when it became clear that methanol grown cells synthesise the proteins involved in methanol metabolism to very high levels. The genes encoding these proteins possess strong, tightly regulated promoter elements. These promoter elements are now used in expression vectors for the control of inducible gene expression in the organisms (Chapter 1).

We found that the methylotrophic yeast species *Hansenula polymorpha* and *Pichia pastoris* are well able to cope with the repetitive gene-sequences encoding gelatin and collagen. Secretion of a human  $\alpha 1(I)$  collagen fragment was possible in *H. polymorpha*, using the native collagen leader sequence (Chapter 2). The N propeptide was cleaved from the helical part of the expressed fragment. This suggested that an endogenous protease of *H. polymorpha* spontaneously processed the N propeptide. We found a difference of only nine amino acid residues between the N-terminus of mature human  $\alpha 1(I)$  collagen and the N-terminus of the recombinant product. In *P. pastoris*, expressing the same construct, the collagen fragment was produced intracellularly and not secreted.

Chapter 3 showed that in single copy transformants of *P. pastoris*, non-hydroxylated, synthetic gelatins with highly hydrophilic properties can be produced at high levels (3-6 g / L). This showed the great potential of *P. pastoris* for the production of tailor-made gelatins (and probably also for other synthetic proteins) with specific physio-chemical properties beneficial for certain technical applications. Expression of the same synthetic gelatins in *H. polymorpha* did not succeed, probably due to proteolytic degradation of the recombinant proteins in this host.

Prolyl 4-hydroxylation, which is essential for the production of recombinant gelatin with gelling properties and for the production of recombinant, thermally stable collagen trimers, was generally considered to be absent in yeasts. However, in this thesis we described the discovery of endogenous collagen-like protein sequences, containing 4-hydroxyproline residues (Chapter 4). This finding indicated that *H. polymorpha* must contain intrinsic prolyl 4-hydroxylase activity to catalyse this post-translational modification. Expression of recombinant animal gelatin in *H. polymorpha* yielded a secreted and hydroxylated product, without the need of co-expressing of heterologous, animal prolyl 4-hydroxylase genes (Chapter 4).

We also investigated if *H. polymorpha* could be used for the production of recombinant human collagen (Chapter 5). Intracellular production of intact collagen trimers was obtained (at levels of 0.2 % of the total cell protein), but the recombinant trimers were not stable at temperatures higher than 15 °C, indicating that the hydroxylation in the product was poor.

In Chapter 6, we describe putative prolyl hydroxylase genes isolated from the filamentous fungi *Neurospora crassa* and *Aspergillus niger*. When the genomic information of *H. polymorpha* will available, these sequences may lead to the isolation of the genes encoding the prolyl 4-hydroxylase of *H. polymorpha*. This information can be applied to further develop the methylotrophic yeasts *H. polymorpha* and *P. pastoris* into high yield production systems of fully hydroxylated gelatins and collagens.

## Samenvatting

Collageen is het meest voorkomende eiwit in het menselijk lichaam en zorgt voor stevigheid van botten en weefsels. Collageen wordt daarom veel gebruikt als biomateriaal in de plastische en cosmetische industrie, zoals voor het behandelen van brandwonden (kunst huid) en het opvullen van rimpels. Gelatine is gedeeltelijk afgebroken collageen en wordt ook op grote schaal gebruikt voor medische en farmaceutische toepassingen. Zo zitten medicijnen bijvoorbeeld in een capsule van gelatine die, na inname, langzaam oplost, zodat het medicijn vrijkomt en z'n werk kan doen. Vanwege z'n unieke eigenschappen wordt gelatine ook grootschalig gebruikt in levensmiddelen zoals in toetjes, ijs, snoepgoed, pudding, soep enz.

Collageen en gelatine worden verkregen uit dierlijke bronnen. Menselijk collageen wordt bijvoorbeeld geïsoleerd uit de placenta (nageboorte). Deze bron is slechts zeer beperkt beschikbaar en is bovendien onderhevig aan het risico van besmetting met HIV. Gelatine wordt geïsoleerd uit de botten en huiden van varkens en koeien. Recent werd echter de eventuele mogelijkheid van besmetting met virussen (de veroorzakers van varkenspest en mond- en klauwzeer) en prionen (de veroorzaker van de gekkekoeienziekte, BSE) een bron van zorg. Van BSE wordt gedacht dat het de hersen-ziekte van Creutzfeldt-Jacob veroorzaakt bij mensen.

Recombinant DNA technologie kan een oplossing bieden van bovengenoemde problemen. Het DNA dat codeert voor menselijk collageen en dierlijk gelatine kan in een micro-organisme gezet worden, zodat deze collageen en gelatine uitscheidt in vloeistof van het kweekmedium. Een gist is een voorbeeld van zo'n micro-organisme en is vrij van allerlei ziektes. Bovendien kunnen gisten allerlei enzymatische reacties uitvoeren die belangrijk zijn voor stabiliteit van recombinant geproduceerde dierlijke eiwitten. Ook kunnen met recombinant DNA technologie verschillende nieuwe eiwitten gemaakt worden die in de natuur niet voorkomen, maar die voor bepaalde toepassingen erg goed gebruikt kunnen worden.



In dit onderzoek hebben we twee verschillende gisten, genaamd *Hansenula polymorpha* en *Pichia pastoris*, met elkaar vergeleken met betrekking tot de productie van recombinant gelatine en collageen. We hebben ontdekt dat *Pichia pastoris* heel goed in staat is om gelatine te maken dat niet voorkomt in de natuur, maar dat gunstige eigenschappen heeft voor een bepaalde toepassing. Ook hebben we ontdekt dat *Hansenula polymorpha* van nature zelf collageen-achtige eiwitten produceert. Omdat deze gist zelf deze eiwitten maakt, kan *Hansenula polymorpha* ook goed recombinant gelatine en collageen produceren. Er vindt zelfs een enzymatische reactie plaats waarvan gedacht werd dat gisten deze zelf niet uit kon voeren, maar die wel erg belangrijk is voor de functionaliteit en stabiliteit van collageen en gelatine. Uit deze studie is gebleken dat de gisten *Hansenula polymorpha* en *Pichia pastoris* uitermate geschikte kandidaten zijn voor de productie van recombinant collageen en gelatine. Wanneer positieve eigenschappen van de twee afzonderlijke gisten kunnen worden gecombineerd in een verder onderzoek zal in de toekomst recombinant gelatine en collageen geproduceerd door gist een goed en veilig alternatief zijn voor traditioneel verkregen dierlijk gelatine en menselijk collageen.

## References

- Alleavitch, J., Turner, W.A., Finch, C.A. (1988), "Gelatin" In: Ullmann's Encyclopedia of Industrial Chemistry, vol. A12. Wiley-VCH Verlag, Weinheim, pp. 307-317.
- Andersson, T., Carlsson, M., Hagel, L., Pernemalm, P. and Janson, J. (1985), "Agarose-based media for high-resolution gel filtration of biopolymers". *J. Chromatogr.* 326: 33-44.
- Appel, R.D., Bairoch, A. and Hochstrasser, D.F. (1994), "A new generation of information retrieval tools for biologists: the example of the ExPASy WWW server". *Trends Biochem. Sci.* 19: 258-260.
- Asghar, A. and Henrickson, R.L. (1982), "Chemical, biochemical, functional, and nutritional characteristics of collagen in food systems". *Adv. Food Res.* 28: 231-372.
- Arborelius, M., Jr., Konttinen, Y.T., Nordstrom, D.C. and Solovieva, S.A. (1999), "Gly-X-Y repeat sequences in the treatment of active rheumatoid arthritis". *Rheumatol. Int.* 18: 129-135.
- Becker, D.M. and Guarente, L. (1991), "High efficiency transformation of yeast by electroporation". In: *Guide to yeast genetics and molecular biology* (Gutrie, C. and Fink, G.R. eds.). Meth. Enzym. 194, Academic Press, London, pp. 182-187.
- Bella, J., Eaton, M., Brodsky, B. and Berman, H.M. (1994), "Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution". *Science* 266: 75-81.
- Bevan, A., Brenner, C. and Fuller, R.S. (1998), "Quantitative assessment of enzyme specificity *in vivo*: P2 recognition by Kex2 protease defined in a genetic system". *Proc. Natl. Acad. Sci. USA* 95: 10384-10389.
- Boehm, T., Pirie-Shepherd, S., Trinh, L.B., Shiloach, J. and Folkman, J. (1999), "Disruption of the KEX1 gene in *Pichia pastoris* allows expression of full-length murine and human endostatin". *Yeast* 15: 563-572.
- Breul, S.D., Bradley, K.H., Hance, A.J., Schafer, M.P., Berg, R.A. and Crystal, R.G. (1980), "Control of collagen production by human diploid lung fibroblasts". *J. Biol. Chem.* 255: 5250-5260.
- Briand, L., Perez, V., Huet, J.C., Danty, E., Masson, C. and Pernollet, J.C. (1999), "Optimization of the production of a honeybee odorant-binding protein by *Pichia pastoris*". *Protein Expr. Purif.* 15: 362-369.
- Brodsky, B. and Shah, N.K. (1995), "Protein motifs 8. The triple-helix motif in proteins". *FASEB J.* 9: 1537-1546.
- Bruckner, P., Bachinger, H.P., Timpl, R. and Engel, J. (1978), "Three conformationally distinct domains in the amino-terminal segment of type III procollagen and its rapid triple helix leads to and comes from coil transition". *Eur. J. Biochem.* 90: 595-603.

- Bulleid, N.J., John, D.C. and Kadler, K.E. (2000), Recombinant expression systems for the production of collagen. *Biochem. Soc. Trans.* 28: 350-353.
- Butkowski, R.J., Noelken, M.E. and Hudson, B.G. (1982), "Estimation of the size of collagenous proteins by electrophoresis and gel chromatography". *Methods Enzymol.* 82: 410-423.
- Capello, J. and Ferrari, F. (1994), "Plastics from Microbes". In: Mobley, D.P. (ed.), Carl Hanser, Munich, pp. 35-92.
- Casarégola, S., Jacq, A., Laoudj, D., McGurk, G., Margaron, S., Tempête, M., Norris, V. and Holland, I.B. (1992), "Cloning and analysis of the entire *Escherichia coli* *ams* gene. *ams* is identical to *hmp1* and encodes a 114 kDa protein that migrates as a 180 kDa protein". *J. Mol. Biol.* 228: 30-40.
- Celerin, M., Laudenbach, D.E., Bancroft, J.B. and Day, A.W. (1994). "Evidence that fimbriae of the smut fungus *Microbotryum violaceum* contain RNA". *Microbiol.* 140: 2699-2704.
- Celerin, M., Ray, J.M., Schisler, N.J., Day, A.W., Stetler-Stevenson, W.G. and Laudenbach, D.E. (1996). "Fungal fimbriae are composed of collagen". *EMBO J.* 15: 4445-4453.
- Cereghino, G.P., Cereghino, J.L., Say Sunga, A., Johnson, M.A., Lim, M., Gleeson, M.A. and Cregg, J.M. (2001), "New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris*". *Gene* 263:159-169.
- Chambers, S.P., Prior, S.E., Barstow, D.A. and Minton, N.P. (1988), "The pMTL nic- cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing". *Gene* 68: 139-149.
- Clare, J.J., Romanos, M.A., Rayment, F.B., Rowedder, J.E., Smith, M.A., Payne, M.M., Sreekrishna, K. and Henwood, C.A. (1991a), "Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strains containing multiple gene copies". *Gene* 105: 205-212.
- Clare, J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K. and Romanos, M.A. (1991b), "High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene". *Biotechnol.* 9: 455-460.
- Clarkson, J.R., Cui, Z.F. and Darton, R.C. (1999), "Protein denaturation in foam". *J. Colloid. Interface Sci.* 215: 333-338.
- Collige, A., Beschin, A., Samyn, B., Goebels, Y., Van Beeumen, J., Nusgens, B.V. and Lapiere, C.M. (1995), "Characterization and partial amino acid sequencing of a 107-kDa procollagen I N-proteinase purified by affinity chromatography on immobilized type XIV collagen". *J. Biol. Chem.* 270: 16724-16730.
- Courts, A. (1980), In: Grant B.A. (ed.), *Applied Protein Chemistry*. Applied Science Publishers, London, pp. 1-29.
- Cox, H., Mead, D., Sudbery, P., Eland, R.M., Mannazzu, I. and Evans, L. (2000), "Constitutive expression of recombinant proteins in the methylotrophic yeast *Hansenula polymorpha* using the PMA1 promoter". *Yeast* 16: 1191-1203.
- Cregg, J.M., Barringer, K.J., Hessler, A.Y. and Madden, K.R. (1985), "*Pichia pastoris* as a host system for transformations". *Mol. Cell. Biol.* 5: 3376-3385.

- Cregg, J.M., Madden, K.R., Barringer, K.J., Thill, G.P. and Stillman, C.A. (1989), "Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*". *Mol. Cell. Biol.* 9: 1316-1323.
- Cregg J.M., Cereghino J.L., Shi J. and Higgins D.R. (2000), "Recombinant protein expression in *Pichia pastoris*". *Mol. Biotechnol.* 16: 23-52.
- de Bruin, E.C., de Wolf, F.A. and Laane, N.C.M. (2000), "Expression and secretion of human alpha1(I) procollagen fragment by *Hansenula polymorpha* as compared to *Pichia pastoris*". *Enzyme Microb. Technol.* 26: 640-644.
- de Bruin, E.C., Werten, M.W.T, Laane N.C.M. and de Wolf, F.A. (2002), "Endogenous prolyl 4-hydroxylation in *Hansenula polymorpha* and its use for the production of hydroxylated recombinant gelatin". *FEMS Yeast Research*, *in press*.
- de Wolf, F.A. and Keller, R.C.A. (1996), "Characterization of helical structures in gelatin networks and model polypeptides by circular dichroism". *Progr. Colloid Polym. Sci.* 102: 9-14.
- de Wolf, F.A., Werten, M.W.T, Wisselink, H.W., Jansen-van den Bosch, T.J., Toda, Y., van Heerde, G.V., Bouwstra, J.B. (2000), *Eur. Pat. Appl.* 1014176
- Dohmen, R.J., Strasser, A.W., Honer, C.B. and Hollenberg, C.P. (1991), "An efficient transformation procedure enabling long-term storage of competent cells of various yeast genera". *Yeast* 7: 691-692.
- Duman, J.G., Miele, R.G., Liang, H., Grella, D.K., Sim, K.L., Castellino, F.J. and Bretthauer, R.K. (1998), "O-Mannosylation of *Pichia pastoris* cellular and recombinant proteins". *Biotechnol. Appl. Biochem.* 28: 39-45.
- Egli, T, van Dijken, J.P., Veenhuis, M., Harder, W. and Fiechter A. (1980) "Methanol metabolism in yeasts: regulation of the synthesis of catabolic enzymes". *Arch. Microbiol.* 124: 115-124.
- Eriksson, M., Myllyharju, J., Tu, H., Hellman, M. and Kivirikko, K.I. (1999), "Evidence for 4-hydroxyproline in viral proteins. Characterization of a viral prolyl 4-hydroxylase and its peptide substrates". *J. Biol. Chem.* 274: 22131-22134.
- Faber, K.N., Swaving, G.J., Faber, F., Ab, G., Harder, W., Veenhuis, M. and Haima, P. (1992), "Chromosomal targeting of replicating plasmids in the yeast *Hansenula polymorpha*". *J. Gen. Microbiol.* 138: 2405-2416.
- Faber, K.N., Haima, P., Harder, W., Veenhuis, M. and Ab, G. (1994), "Highly-efficient electrotransformation of the yeast *Hansenula polymorpha*". *Curr. Genet.* 25: 305-310.
- Faber, K.N., Westra, S., Waterham, H.R., Keizer-Gunnink, I., Harder, W. and Veenhuis, G.A. (1996), "Foreign gene expression in *Hansenula polymorpha*. A system for the synthesis of small functional peptides". *Appl. Microbiol. Biotechnol.* 45: 72-79.
- Faguy, D.M., Bayley, D.P., Kostyukova, A.S., Thomas, N.A. and Jarrell, K.F. (1996), "Isolation and characterization of flagella and flagellin proteins from the Thermoacidophilic archaea *Thermoplasma volcanium* and *Sulfolobus shibatae*". *J. Bacteriol.* 178: 902-905.

- Fellinger, A.J., Verbakel, J.M., Veale, R.A., Sudbery, P.E., Bom, I.J., Overbeeke, N. and Verrips, C.T. (1991), "Expression of the alpha-galactosidase from *Cyamopsis tetragonoloba* (guar) by *Hansenula polymorpha*". *Yeast* 7: 463-473
- Fitzgerald, J. and Bateman, J.F. (2001), "A new FACIT of the collagen family: COL21A1". *FEBS Lett.* 505: 275-280.
- Freytag, J.W., Noelken, M.E. and Hudson, B.G. (1979), "Physical properties of collagen-sodium dodecyl sulfate complexes". *Biochemistry* 18: 4761-4768.
- Furthmayr, H. and Timpl, R. (1971), "Characterization of collagen peptides by sodium dodecylsulfate-polyacrylamide electrophoresis". *Anal. Biochem.* 41: 510-516.
- Gardner, K., Lock, R.L., O'Brien, J.P. and Salemme, F.R. (1993), PCT Pat. Appl. 931023
- Geck, P., Whitaker, S.A., Medveczky, M.M. and Medveczky, P.G. (1990). "Expression of collagen-like sequences by a tumor virus, herpesvirus saimiri". *J. Virol.* 64: 3509-3515.
- Geddis, A.E and Prockop D.J.(1993), "Expression of human COL1A1 gene in stably transfected HT1080 cells: The production of a thermostable homotrimer of type I collagen in a recombinant system". *Matrix* 13: 399-405.
- Gellissen G. and Hollenberg, C.P. (1997), " Application of yeasts in gene expression studies: a comparison of *Saccharomyces cerevisiae*, *Hansenula polymorpha* and *Kluyveromyces lactis* - a review". *Gene* 190: 87-97.
- Gellissen, G. (2000), "Heterologous protein production in methylotrophic yeasts". *Appl. Microbiol. Biotechnol.* 54: 741-750.
- Giefl, C., Faber, K.N., van der Klei, I.J. and Veenhuis, M. (1994), "Mutational analysis of the N-terminal topogenic signal of watermelon glyoxysomal malate dehydrogenase using the heterologous host *Hansenula polymorpha*". *Proc. Natl. Acad. Sci. USA* 91: 3151-3155.
- Giuseppin, M.L., Almkerk, J.W., Heistek, J.C. and Verrips, C.T. (1993), "Comparative study on the production of guar alpha-galactosidase by *Saccharomyces cerevisiae* SU50B and *Hansenula polymorpha* 8/2 in continuous cultures". *Appl. Environ. Microbiol.* 59: 52-59.
- Gleeson, M.A., Otori, G.S. and Sudbery, P.E. (1986), "Transformation of the methylotrophic yeast *Hansenula polymorpha*". *J. Gen. Microbiol.* 132: 3459-3465.
- Goda, S., Takano, K., Yamagata, Y., Katakura, Y. and Yutani, K. (2000), "Effect of extra N-terminal residues on the stability and folding of human lysozyme expressed in *Pichia pastoris*". *Protein Eng.* 13: 299-307.
- Goldberg, I., Salerno, A.J., Patterson, T. and Williams, J.I. (1989), "Cloning and expression of a collagen-analog-encoding synthetic gene in *Escherichia coli*". *Gene* 80: 305-314.
- Harkins, W.D. and Jordan, H.F. (1930), "A method for the determination of surface and interfacial tension from the maximum pull on a ring", *J. Am. Chem. Soc.* 52: 1751-1771.

- Hasslacher, M., Schall, M., Hayn, M., Bona, R., Rumbold, K., Luckl, J., Griengl, H., Kohlwein, S.D. and Schwab, H. (1997), "High-level intracellular expression of hydroxynitrile lyase from the tropical rubber tree *Hevea brasiliensis* in microbial hosts". *Protein Expr. Purif.* 11: 61-71.
- Hayashi, T. and Nagai, Y. (1980), "The anomalous behavior of collagen peptides on sodium dodecyl sulfate-polyacrylamide gel electrophoresis is due to the low content of hydrophobic amino acid residues". *J. Biochem. (Tokyo)* 87: 803-808.
- Higgins, D.R. and Cregg, J.M. (1998), "Introduction to *Pichia pastoris*". *Methods Mol. Biol.* 103: 1-15.
- Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989), "Site-directed mutagenesis by overlap extension using the polymerase chain reaction". *Gene* 77: 51-59.
- Hojima, Y., van der Rest, M. and Prockop, D.J. (1985), "Type I procollagen carboxyl-terminal proteinase from chick embryo tendons. Purification and characterization". *J. Biol. Chem.* 260: 15996-16003.
- Hollenberg, C.P. and Gellissen G. (1997), "Production of recombinant proteins by methylotrophic yeasts". *Curr. Opin. Biotech.* 8: 554-560.
- Holmgren, S.K., Bretscher, L.E., Taylor, K.M. and Raines, R.T. (1999), "A hyperstable collagen mimic". *Chem. Biol.* 6: 63-70.
- Horiuchi, T., Fukushima, D., Sugimoto, H. and Hattori, T. (1978), "Studies on enzyme-modified proteins as foaming agents: effect of structure on foam stability". *Food Chemistry* 3: 35-42.
- Kajino, T., Takahashi, H., Hirai, M. and Yamada, Y. (2000), "Efficient production of artificially designed gelatins with a *Bacillus brevis* system". *Appl. Environ. Microbiol.* 66: 304-309.
- Kang, H.A., Kang, W., Hong, W.K., Kim, M.W., Kim, J.Y., Sohn, J.H., Choi, E.S., Choe, K.B. and Rhee, S.K. (2001), "Development of expression systems for the production of recombinant human serum albumin using the MOX promoter in *Hansenula polymorpha* DL-1". *Biotechnol. Bioeng.* 76: 175-185.
- Keizer-Gunnink, I., Vuorela, A., Myllyharju, J., Pihlajaniemi, T., Kivirikko, K.I. and Veenhuis, M. (2000), "Accumulation of properly folded human type III procollagen molecules in specific intracellular membranous compartments in the yeast *Pichia pastoris*". *Matrix Biol.* 19: 29-36.
- Kielty, C.M., Hopkinson, I. and Grant, M.E. (1993), "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. and Steinmann, B., Eds.), Wiley-Liss, New York, pp. 103-147.
- Kivirikko, K.I. and Myllylä, R. (1982), "Post-translational enzymes in the biosynthesis of collagen: Intracellular enzymes". *Methods Enzymol.* 82: 245-304.
- Kivirikko, K.I., Myllylä, R. and Pihlajaniemi, T. (1992), "Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins". In: *Post-translational modifications of proteins.* (Harding, J.J. and Crabbe, M.J.C., Eds.) CRC Press, Boca Raton, pp. 1-51.

- Koch, M., Foley, J.E., Hahn, R., Zhou, P., Burgeson, R.E., Gerecke, D.R. and Gordon, M.K. (2001), "alpha 1(XX) collagen, a new member of the collagen subfamily, fibril-associated collagens with interrupted triple helices". *J. Biol. Chem.* 276: 23120-23126.
- Kyte, J. and Doolittle, R.F. (1982), "A simple method for displaying the hydropathic character of a protein". *J. Mol. Biol.* 157: 105-132.
- Laemmli, U.K. (1970), "Cleavage of structural proteins during the assembly of the head of bacteriophage T4". *Nature* 227: 680-685.
- Lamberg, A., Helaakoski, T., Myllyharju, J., Peltonen, S., Notbohm, H., Pihlajaniemi, T. and Kivirikko K. I. (1996), "Characterization of human type III collagen expressed in a baculovirus system- Production of a protein with a stable triple helix requires coexpression with the two types of recombinant prolyl 4-hydroxylase subunit". *J. Biol. Chem.* 271: 11988-11995.
- Lecomte du Noüy, P. (1919), "A new apparatus for the measuring surface tension". *J. Gen. Physiol.* 1: 521-524.
- Lee, C.H., Singla, A. and Lee, Y. (2001), "Biomedical applications of collagen". *Int. J. Pharm.* 221: 1-22.
- Leibovich, S.J. and Weiss, J.B. (1970), "Electron microscope studies of the effects of endo- and exopeptidase digestion on tropocollagen. A novel concept of the role of terminal regions in fibrillogenesis". *Biochim. Biophys. Acta* 214: 445-454.
- Letourneur, O., Gervasi, G., Gaia, S., Pages, J., Watelet, B. and Jolivet, M. (2001), "Characterization of *Toxoplasma gondii* surface antigen 1 (SAG1) secreted from *Pichia pastoris*: evidence of hyper O-glycosylation". *Biotechnol. Appl. Biochem.* 33: 35-45.
- Matagne, A., Joris, B. and Frère, J.M. (1991), "Anomalous behaviour of a protein during SDS/PAGE corrected by chemical modification of carboxylic groups". *Biochem. J.* 280: 553-556.
- Mayer, A.F., Hellmuth, K., Schlieker, H., Lopez-Ulibarri, R., Oertel, S., Dahlems, U., Strasser, A.W. and van Loon, A.P. (1999), "An expression system matures: a highly efficient and cost-effective process for phytase production by recombinant strains of *Hansenula polymorpha*". *Biotechnol. Bioeng.* 63: 373-381.
- McGrath, K.P., Fournier, M.J., Mason, T.L. and Tirrell, D.A. (1992), "Genetically directed syntheses of new polymeric materials. Expression of artificial genes encoding proteins with repeating -(AlaGly)<sub>3</sub>ProGluGly- elements". *J. Am. Chem. Soc.* 114: 727-733.
- Mc Master, G.K and Carmichael, G.G. (1977), Analysis of single and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. *Proc. Natl. Acad. Sci. USA* 74: 4835-4838.
- Mewes, H.W., Albermann, K., Bahr, M., Frishman, D., Gleissner, A., Hani, J., Heumann, K., Kleine, K., Maierl, A., Oliver, S.G., Pfeiffer, F. and Zollner, A. (1997), "Overview of the yeast genome". *Nature* 387: 7-65.
- Monnet, E. and Fauvel-Lafeve, F. (2000), "A new platelet receptor specific to type III collagen. Type III collagen-binding protein". *J. Biol. Chem.* 275: 10912-10917.

- Munshi, C. and Lee, H.C.(1997), "High-level expression of recombinant *Aplysia* ADP- ribosyl cyclase in *Pichia pastoris* by fermentation". *Prot. Expr. Pur.* 11:104-110.
- Myllyharju, J. (2000), "Recombinant collagen trimers from insect cells and yeast". *Methods Mol. Biol.* 139: 39-48.
- Nagarajan, V., Kamitori, S. and Okuyama, K. (1999), "Structure analysis of a collagen-model peptide with a (Pro-Hyp-Gly) sequence repeat". *J. Biochem. (Tokyo)* 125: 310-318.
- Nielsen, H., Engelbrecht, J., Brunak S. and von Heijne, G. (1997), "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites". *Protein Eng.* 10: 1-6.
- Noelken, M.E., Wisdom, B.J.J. and Hudson, B.G. (1981), "Estimation of the size of collagenous polypeptides by sodium dodecyl sulfate-polyacrylamide gel electrophoresis". *Anal. Biochem.* 110: 131-136.
- Noelkainen, M., Tu, H., Vuorela, A., Notbohm, H., Kivirikko, K.I. and Myllyharju, J. (2001), "High-level production of human type I collagen in the yeast *Pichia pastoris*". *Yeast* 18: 797-806.
- Obrecht, G., Lefèvre, J.F. and Meyruis, P. (1991) French Pat. Appl. 91 16215.
- Ohara, O. and Teraoka, H. (1987), "Anomalous behavior of human leukocyte interferon subtypes on polyacrylamide gel electrophoresis in the presence of dodecyl sulfate". *FEBS Lett.* 211: 78-82.
- Olsen, D.R., Leigh, S.D., Chang, R., McMullin, H., Ong, W., Tai, E., Chisholm, G., Birk, D.E., Berg, R.A., Hitzeman, R.A. and Toman, P.D. (2001), "Production of human type I collagen in yeast reveals unexpected new insights into the molecular assembly of collagen trimers". *J. Biol. Chem.* 276: 24038-24043.
- Philajaniemi, T., Helaakoski, T., Tasanen, K., Myllylä, R., Huhtala, M.L., Koivu, J. and Kivirikko, K.I. (1987), "Molecular cloning of the beta-subunit of human prolyl 4-hydroxylase. This subunit and protein disulphide isomerase are products of the same gene". *EMBO. J.* 6: 643-649.
- Pollack, S.V. (1990), "Silicone, fibrel, and collagen implantation for facial lines and wrinkles". *J Dermatol. Surg. Oncol.* 16: 957-961.
- Prockop, D.J. and Kivirikko, K.I. (1995), "Collagens: molecular biology, diseases, and potentials for therapy". *Annu. Rev. Biochem.* 64: 403-434.
- Privalov, P.L. (1982), "Stability of proteins. Proteins which do not present a single cooperative system". *Adv. Protein Chem.* 35: 1-104.
- Rao, K.P. (1995), "Recent developments of collagen-based materials for medical applications and drug delivery systems". *J. Biomater. Sci. Polym. Ed.* 7: 623-645.
- Reynolds, J.A. and Tanford, C. (1970), "Binding of dodecyl sulfate to proteins at high binding ratios. Possible implications for the state of proteins in biological membranes". *Proc. Natl. Acad. Sci. USA* 66: 1002-1007.
- Romanos, M.A., Scorer, C.A. and Clare, J.J. (1992). "Foreign gene expression in yeast: a review". *Yeast* 8: 423-428.



- Rosenfeld, S.A., Nadeau, D., Tirado, J., Hollis, G.F., Knabb, R.M. and Jia, S. (1996), "Production and purification of recombinant hirudin expressed in the methylotrophic yeast *Pichia pastoris*". *Protein Expr. Purif.* 8: 476-482.
- Rossi, A., Zuccarello, L.V., Zanaboni, G., Monzani, E., Dyne, K.M., Cetta, G. and Tenni, R. (1996), "Type I collagen CnBr peptides- Species and behaviour in solution". *Biochemistry* 35: 6048-6057.
- Ruggiero, F., Exposito, J.Y., Bournat, P., Gruber, V., Perret, S., Comte, J., OLAGNIER, B., Garrone, R. and Theisen, M. (2000), "Triple helix assembly and processing of human collagen produced in transgenic tobacco plants". *FEBS Lett.* 469: 132-136.
- Saddler, J.M. and Horsey, P.J. (1987), "The new generation gelatins. A review of their history, manufacture and properties". *Anaesthesia* 42: 998-1004.
- Sambrook J., Fritsch E.F. and Maniatis, T. (1989), "Molecular cloning: a laboratory manual" Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.
- Scorer, C.A., Buckholz, R.G., Clare, J.J. and Romanos, M.A. (1993), "The intracellular production and secretion of HIV-1 envelope protein in the methylotrophic yeast *Pichia pastoris*". *Gene* 136:111-119.
- Scorer, C.A., Clare, J.J., McCombie, W.R., Romanos, M.A. and Sreekrishna, K. (1994), "Rapid selection using G418 of high copy number transformants of *Pichia pastoris* for high-level foreign gene expression". *Biotechnol.* 12: 181-184.
- Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J. and Klenk, D.C. (1985), "Measurement of protein using bicinchoninic acid". *Anal. Biochem.* 150: 76-85.
- Sohn, J.H., Choi, E.S., Kim, C.H., Agaphonov, M.O., Ter-Avanesyan, M.D., Rhee, J.S. and Rhee, S.K. (1996), "A novel autonomously replicating sequence (ARS) for multiple integration in the yeast *Hansenula polymorpha* DL-1". *J. Bacteriol.* 178: 4420-4428.
- Sohn, J.H., Choi, E.S., Kang, H.A., Rhee, J.S., Agaphonov, M.O., Ter-Avanesyan, M.D. and Rhee, S.K. (1999), "A dominant selection system designed for copy-number-controlled gene integration in *Hansenula polymorpha* DL-1". *Appl. Microbiol. Biotechnol.* 51: 800-807.
- Sreekrishna, K. and Kropp, K.E. (1996), "Nonconventional Yeasts in Biotechnology. A Handbook" In: Wolf, K. (Ed.), Springer-Verlag, Berlin, pp. 203-253.
- Steinmann, B., Bruckner, P. and Superti-Furga, A. (1991), "Cyclosporin A slows collagen triple-helix formation in vivo: indirect evidence for a physiologic role of peptidyl-prolyl cis-trans-isomerase". *J. Biol. Chem.* 266: 1299-1303.
- Strausberg, R.L. and Link, R.P. (1990), "Protein-based medical adhesives". *Trends Biotechnol.* 8: 53-57.
- Subramani, S. (1992), "Targeting of proteins into the peroxisomal matrix". *J. Membr. Biol.* 125:99-106.
- Sudbery, P.E. (1996), "The expression of recombinant proteins in yeasts". *Curr. Opin. Biotechnol.* 7: 517-524.

- Suzuki, Y., Ikeda, N., Kataoka, E. and Ohsuye, K. (2000), "Effect of amino acid substitution at the P3 and P4 subsites of fusion proteins on Kex2 protease activity." *Biotechnol. Appl. Biochem.* 32: 53-60.
- 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. and Hitzeman, R.A. (2000), "Production of recombinant human type I procollagen trimers using a four-gene expression system in the yeast *Saccharomyces cerevisiae*". *J. Biol. Chem.* 275: 23303-23309.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994), "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice". *Nucleic Acids Research* 22: 4673-4680.
- Tschopp, J.F. and Cregg, J.M. (1991), "Heterologous gene expression in methylotrophic yeast". *Biotechnol.* 18: 305-322.
- van Heerde, G.V., van Rijn, A.C., Bouwstra, J.B., de Wolf, F.A., Mooibroek, H., Werten, M.W.T., Wind, R.D. and van den Bosch, T.J. (1999), *Eur. Pat. Appl.* 0926543.
- van Dijk, R., Faber, K.N., Kiel, J.A., Veenhuis, M. and van der Klei, I. (2000), "The methylotrophic yeast *Hansenula polymorpha*: a versatile cell factory". *Enzyme Microb. Technol.* 26: 793-800.
- van Dijken, J.P., Otto, R. and Harder, W. (1976), "Growth of *Hansenula polymorpha* in a methanol-limited chemostat. Physiological responses due to the involvement of methanol oxidase as a key enzyme in methanol metabolism". *Arch. Microbiol.* 111: 137-144.
- van Dijken, J.P., Veenhuis, M. and Harder, W. (1982), "Peroxisomes of methanol-grown yeasts". *Ann. NY Acad. Sci.* 386: 200-216.
- Vaughn, P.R., Galanis, M., Richards, K.M., Tebb, T.A., Ramshaw, J.A. and Werkmeister, J.A. (1998), "Production of recombinant hydroxylated human type III collagen fragment in *Saccharomyces cerevisiae*". *DNA Cell. Biol.* 17: 511-518.
- Vedvick, T., Buckholz, R.G., Engel, M., Urcan, M., Kinney, J., Provow, S., Siegel, R.S. and Thill, G.P. (1991), "High-level secretion of biologically active aprotinin from the yeast *Pichia pastoris*". *J. Ind. Microbiol.* 7:197-201.
- Veenhuis, M., van Dijken, J.P. and Harder, W. (1983), "The significance of peroxisomes in the metabolism of one-carbon compounds in yeast". *Adv. Microb. Physiol.* 24: 1-76.
- Vuorela, A., Myllyharju, J., Nissi, R., Pihlajaniemi, T. and Kivirikko, K.I. (1997), "Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast *Pichia pastoris*: formation of a stable enzyme tetramer requires coexpression with collagen an assembly of stable collagen requires coexpression with prolyl 4-hydroxylase". *EMBO J.* 16: 6702-6712.
- Vuori, K., Pihlajaniemi, T., Myllylä, R. and Kivirikko, K.I. (1992), "Site-directed mutagenesis of human protein disulphide isomerase: effect on the assembly, activity and endoplasmic reticulum retention of human prolyl 4-hydroxylase in *Spodoptera frugiperda* insect cells". *EMBO J.* 11: 4213-4217.
- Ward, A.H. and Michos, G.A. (1972), "Alcian blue staining of glycoproteins in acrylamide disc electrophoresis". *Anal. Biochem.* 49: 607-609.

- Walsh, C.T., Zydowsky, L.D. and McKeon, F.D. (1992), "Cyclosporin A, the cyclophilin class of peptidylprolyl isomerases, and blockade of T cell signal transduction". *J. Biol. Chem.* 267: 13115-13118.
- Waterham, H.R., Titorenko, V.I., Haima, P., Cregg, J.M., Harder, W. and Veenhuis, M. (1994), "The *Hansenula polymorpha* PER1 gene is essential for peroxisome biogenesis and encodes a peroxisomal matrix protein with both carboxy- and amino-terminal targeting signals". *J. Cell Biol.* 127: 737-749.
- Waterham, H.R., Digan, M.E., Koutz, P.J., Lair, S.V. and Cregg, J.M. (1997), "Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter". *Gene* 186: 37-44.
- Wilcox, P.E. (1967), "Esterification". *Methods Enzymol.* 11: 605-617.
- Werten, M.W.T., van den Bosch, T.J., Wind, R.D., Mooibroek, H. and de Wolf, F.A. (1999), "High-yield secretion of recombinant gelatins by *Pichia pastoris*". *Yeast* 15: 1087-1096.
- Werten, M.W.T., Wisselink, W.H., Jansen-van den Bosch, T.J., de Bruin, E.C. and de Wolf, F.A. (2001), "Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*". *Protein Eng.* 14: 447-454.
- Weydemann, U., Keup, P., Piontek, M., Strasser, A.W., Schweden, J., Gellissen, G. and Janowicz, Z.A. (1995), "High-level secretion of hirudin by *Hansenula polymorpha* - authentic processing of three different preprohirudins". *Appl. Microbiol. Biotechnol.* 44: 377-385.
- Zacharius, R.M., Zell, T.E., Morrison, J.H. and Woodlock, J.J. (1969), "Glycoprotein staining following electrophoresis on acrylamide gels". *Anal. Biochem.* 30: 148-152.
- Zhu, A., Monahan, C., Zhang, Z., Hurst, R., Leng, L. and Goldstein, J. (1995), "High-level expression and purification of coffee bean alpha-galactosidase produced in the yeast *Pichia pastoris*". *Arch. Biochem. Biophys.* 324: 65-70.
- Zurek, C., Kubis, E., Keup, P., Hörlein, D., Beunink, J., Thömmes, J., Kula, R.M., Hollenberg, C.P. and Gellissen, G. (1996), "Production of two aprotinin variants in *Hansenula polymorpha*". *Process Biochem.* 31: 679-689.

## List of publications

Zhu, Y., Rinzema, A., Tramper, J., **de Bruin, E.C.** and Bol, J. (1998), "Fed-batch fermentation dealing with nitrogen limitation in microbial transglutaminase production by *Streptoverticillium mobaraense*". *Applied Microbiology and Biotechnology* 49: 251-257.

Bosma, T., Kruizinga, E., **de Bruin, E.C.**, Poelarends, G.J. and Janssen, D.B. (1999), "Utilization of trihalogenated propanes by *Agrobacterium radiobacter* AD1 through heterologous expression of the haloalkane dehalogenase from *Rhodococcus* sp. strain M15-3". *Applied and Environmental Microbiology* 65: 4575-4581.

**de Bruin, E.C.**, de Wolf, F.A. and Laane, N.C.M. (2000), "Expression and secretion of human  $\alpha 1(I)$  procollagen fragment by *Hansenula polymorpha* as compared to *Pichia pastoris*". *Enzyme Microbial Technology* 26: 640-644.

Werten, M.W.T., Wisselink, W.H., Jansen-van den Bosch, T.J., **de Bruin E.C.** and de Wolf, F.A. (2001), "Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*". *Protein Engineering* 14: 447-454.

**de Bruin, E.C.**, Werten, M.W.T. de Wolf, F.A. (2001), "Method for the production of hydroxylated collagen-like compounds". Patent Application WO: 0204615.

**de Bruin, E.C.**, Werten, M.W.T., Laane N.C.M. and de Wolf, F.A. (2002), "Endogenous prolyl 4-hydroxylation in *Hansenula polymorpha* and its use for the production of hydroxylated recombinant gelatin". *FEMS Yeast Research*, *in press*.

**de Bruin, E.C.**, de Haas, G.G., Faber, K.N., Werten, M.W.T., Laane, N.C.M. and de Wolf, F.A., "Human collagen trimers from the methylotrophic yeast *Hansenula polymorpha*". *in prep*.

## *Curriculum vitae*

Eric Christian de Bruin werd op 7 april 1974 geboren te Harlingen. Na het behalen van het VWO-diploma aan het Lauwers College te Buitenpost begon hij in 1992 met de studie biologie aan de Rijks Universiteit van Groningen. Al snel koos hij voor de hoofdrichting biotechnologie en startte in 1995 een doctoraal onderzoek naar de afbraak van halopropanen door verschillende haloalkaan dehydrogenases bij de milieu-biotechnologie groep van Prof. dr. D.B. Janssen. Na deze periode vertrok Eric uit Groningen voor een stage bij de divisie industriële microbiologie van TNO-Voeding te Zeist. Onder leiding van Dr. Zhu Yang hield hij zich een half jaar bezig met de microbiële productie van het enzym transglutaminase door *Streptoverticillium mobaraense*. Na afgestudeerd te zijn als biotechnoloog begon hij in 1997 met zijn promotie onderzoek en werd als onderzoeker in opleiding van Wageningen Universiteit gedetacheerd bij de afdeling bioconversie van het agrotechnologisch onderzoeks-instituut ATO B.V. te Wageningen. Hier heeft hij, onder begeleiding van Prof. dr. N.C.M. Laane en Dr. F.A. de Wolf, gedurende vier en een half jaar gewerkt aan het onderzoek beschreven in dit proefschrift. Vanaf september 2001 is hij werkzaam bij Sanquin Research te Amsterdam als onderzoeker voor de biotech start-up Vitaleech Bioscience N.V.

## *Addendum*

Dit werk werd uitgevoerd bij het agrotechnologisch onderzoeks instituut, ATO B.V. , Postbus 17, 6700 AA Wageningen, in samenwerking met het laboratorium voor Biochemie van Wageningen Universiteit. Beiden maken onderdeel uit van het Departement Agrotechnologie en Voedingwetenschappen van Wageningen UR.

Het onderzoek beschreven in dit proefschrift werd grotendeels gefinancierd door een beurs (cSEO97) van stichting Dienst Landbouwkundig Onderzoek (DLO) en maakte deel uit van het onderzoeksprogramma van de onderzoekschool VLAG (Voeding, Levensmiddelentechnologie, Agrobiotechnologie en Gezondheid).

Het drukken van dit proefschrift werd mede mogelijk gemaakt door :

