Production of Recombinant Human Type I Procollagen Trimers Using a Four-gene Expression System in the Yeast *Saccharomyces cerevisiae*

P. David Toman‡§¶, George Chisholm‡, Hugh McMullin‡, Lynne M. Giere‡, David R. Olsen‡, Robert J. Kovach‡, Scott D. Leigh‡, Bryant E. Fong‡, Robert Chang‡, Gregory A. Daniels‡, Richard A. Berg‡, and Ronald A. Hitzeman§

From ‡Cohesion Technologies, Palo Alto, California 94303 and ¶Genotronics Incorporated, South San Francisco, California 94080

The expression of stable recombinant human collagen requires an expression system capable of post-translational modifications and assembly of the procollagen polypeptides. Two genes were expressed in the yeast *Saccharomyces cerevisiae* to produce both propeptide chains that constitute human type I procollagen. Two additional genes were expressed coding for the subunits of prolyl hydroxylase, an enzyme that post-translationally modifies procollagen and that confers heat (thermal) stability to the triple helical conformation of the collagen molecule. Type I procollagen was produced as a stable heterotrimeric helix similar to type I procollagen produced in tissue culture. A key requirement for glutamate was identified as a medium supplement to obtain high expression levels of type I procollagen as heat-stable heterotrimers in *Saccharomyces*. Expression of these four genes was sufficient for correct assembly and processing of type I procollagen in a eucaryotic system that does not produce collagen.

Collagen is the single most abundant protein found in animals. It is found in all animals, including sponges. It is not expressed in yeast. In mammals, it is expressed in most tissues and plays both a structural as well as a signaling role in the development, maintenance, and repair of tissues and organs. More than 30 gene products compose the collagen family of molecules (1). Procollagens have several features and require numerous steps for production of functional molecules, including post-translational modifications (2). Key features in the collagen family are the formation of a triple helix composed of three polypeptide chains and the post-translational modification of proline residues to hydroxyproline, which provides stability of the triple helix against thermal denaturation and unfolding (T_m) at the animal’s body temperature (3). The content of proline and hydroxyproline is correlated with the temperature of an animal’s environment (4). The triple helical domain of procollagen consists of -(GXY)_n- repeats, where X and/or Y is frequently proline or hydroxyproline in the mature molecule. Prolyl 4-hydroxylase, an α2β2 tetrameric enzyme composed of the prolyl hydroxylase α-subunit (αPH) and the protein-disulfide isomerase (PDI) subunit in higher eucaryotes, is the enzyme that modifies proline residues to hydroxyproline. Additional steps for procollagen production include carbohydrate attachment, folding into a triple helix, secretion into the extracellular matrix, and cleavage by specific proteases to remove the propeptide domains to form mature collagen helices. A C-terminal non-helical propeptide facilitates the assembly of trimeric collagen molecules, leading to helix formation (5); the N-terminal propeptide may limit fiber diameter (6). The association and folding steps of three polypeptide chains that compose the triple helix potentially require chaperone functions in the endoplasmic reticulum, with PDI (7) and Hsp47 (8) as two proteins that have been implicated in the assembly of a procollagen trimer.

A fundamental question regarding collagen biosynthesis is which genes are essential for the expression of collagen in cells and which are nonessential. Expression of recombinant collagen has been performed using mammalian, baculoviral, and transgenic systems. Single procollagen genes were expressed in mammalian cells to produce homotrimeric type I procollagen (9), type II procollagen (10), and homotrimeric type V collagen (11). In baculovirus, prolyl hydroxylase was transfected and shown to be a functional enzyme (12). Subsequently, type I and III procollagens were transfected and expressed and were shown to be capable of modification by prolyl hydroxylase (13–15). Recently, homotrimeric type I procollagen and an engineered form of α2(I) procollagen have been expressed in the milk of transgenic mice (16, 17). In contrast, no report of procollagen expression and assembly has been published using a bacterial expression system.

The yeast *Pichia pastoris* was first engineered to express prolyl hydroxylase and subsequently shown to produce functional type III procollagen if the gene for type III procollagen was introduced (18, 19). Like *Saccharomyces*, *Pichia* contains endogenous PDI, but not αPH, and it does not synthesize procollagen. It was therefore a useful system to test the requirements for genes to produce type III procollagen. In this system, the type III procollagen gene and the two genes for prolyl hydroxylase were sufficient to produce stable type III procollagen molecules. However, *Saccharomyces* is evolutionarily diverse from *Pichia*. Furthermore, type I procollagen is composed of polypeptides generated from two distinct genes to form an (α1)3 structure, whereas type II and III procollagens require only one gene product to form an (α1β1)3 structure.

To our knowledge, this is the first report to describe a multigene system in *Saccharomyces* that results in both the assem-
bly and non-native post-translational modification of a multimeric protein to produce a functional heterologous molecule. A total of four gene products were required in Saccharomyces to generate a thermally stable triple helical type I procollagen: two genes that code for the polypeptide chains of type I procollagen and two additional genes that code for the subunits of prolyl hydroxylase. No other added genes were required to produce a functional procollagen. We further optimized our expression system at the molecular level, but also optimized the addition of medium components to significantly increase the level of expression of type I procollagen in Saccharomyces.

**EXPERIMENTAL PROCEDURES**

**Plasmid Constructions**—The precursor plasmid pG7E100 and plasmid pG7E150, which contains the GAL1/GAL10 dual promoter and is the base plasmid for other constructions, were made as follows. YEp9T, containing yeast (2α origin, FLPI gene terminator in 2α DNA, and the yeast TRPI gene) and bacterial (pBR322 functions) sequences (20), was modified between an Ndel site in the 2α DNA and a second Ndel site near the origin of pBR322 with the polynucleotide sequence (Ndel-PvuII-Apal-BglII-ClaI-Nhel-Xhol-EcoRI-BamHI-AFII-NotI-Ndel) to create pGET100 (p2α site closest to 2α DNA). Genomic DNA from yeast strain ST1989 (MATα trpI his4 ade2 gal2) was used as the template for PCR. A DNA sequence (21) contained the polylinker element (Ndel-BglII-AflII) which was placed into the polylinker of pG7E100 to create pGET333. Plasmids pG7E100 and pGET333 were constructed by cloning synthetic oligonucleotides into the polylinker region using the restriction enzymes PmlI and EcoRI. The synthesis of both plasmids was confirmed by restriction analysis.

The full-length cDNA for human type I procollagen has been deduced from the cDNA sequence to the position of a potential translation stop codon (32). The cDNA sequence was used as a template for PCR using primers (based on sequence information (21)) containing synthetic oligonucleotides (22). A fragment containing the promoter region (10 bp upstream of ATG), the prepro-HSA signal (23) secretion signal-AfII-GAL10 promoter/GAL1 promoter-AfII-chicken α-3-phosphoglycerate kinase gene terminator (the natural 330-bp HindIII fragment containing the URA3 yeast gene (31) was converted to a Sall fragment and placed into the above Xhol site. A NotI fragment containing the ADH1 terminator-chicken PDI yeast invertase secretion signal-AfII-GAL10 promoter/GAL10 promoter-AfII-chicken α-3-phosphoglycerate kinase gene terminator was added to the NotI site in the polylinker. This new plasmid (pGET829) containing the URA3 gene and dual expression units for chicken PDI and αPH within the disrupted TRPI gene was cut with PvuII (61 bp from EcoRI in the promoter region) and ApaI (10 bp in from the PvuI site on the other end of the TRPI gene) and therefore has homology to both ends of the TRPI gene. Integration of this linear fragment was performed by a double crossover during yeast transformation. Western assays using antibodies to chicken αPH and PDI (32) identified the highest producing transformants. Subsequent analysis of the resultant single (chicken PDI) and double (chicken PDI and PDI) transformants indicated that it contains multiple integrations of the chicken PDI and αPH expression unit at the TRPI locus.

**Type I Procollagen Expression in Yeast**

**Quantitative Assay for Collagen—**Collagen yield was determined by a luminometric immunoassay utilizing a goat anti-type I collagen antibody from BIODESIGN International (Kennebunkport, ME) derivatized with either biotin or ruthenium chelate. Samples were analyzed by lysis cells as described above and centrifuging to remove cell debris. The clarified supernatant samples from cell lysis were diluted in matrix buffer (100 m PIPES (pH 6.5) and 1% (w/v) bovine serum albumin) in duplicate. A 25-μl aliquot was mixed with 50 μl of an antibody solution containing 1 μg/ml ruthenium chelate-conjugated antibody and 1.5 μg/ml biotin-conjugated antibody in diluent (matrix buffer plus 1.5% Tween 20). Samples were incubated for 2 h at ~20 °C with shaking. A 25-μl aliquot of origin assay buffer (IGEN Inc., Gaithersburg, MD) was added to each sample and placed in an ORIGEN analyzer (IGEN Inc.). Total protein was determined using the BCA assay (Pierce) using a microtiter plate format.

**Gels and Western Blots**—The equivalent of 20 ml of cells at A595 = 1.0 were collected, resuspended in 200 μl of buffer, and lysed as described above. SDS sample buffer was added; the samples were incubated at 100 °C for 5 min; and the debris was collected by centrifugation. Clarified supernatants were loaded onto 5 or 10% SDS-polyacrylamide gels,
electrophoresed, and stained using either GELCODE blue (Pierce) or silver stain.

Protein was transferred from gels onto polyvinylidene difluoride membranes (Millipore Corp., Bedford, MA) using Western blotting and probed using an antibody against the N-propeptide region (LP-39, 1:15,000 dilution) or C-propeptide region (LP-41, 1:10,000 dilution) of human α(I) procollagen (34) and a horseradish peroxidase-linked secondary detection antibody (Pierce), followed by development using the ECL detection kit (Amersham Pharmacia Biotech).

Carbohydrate Analysis—Plasmid pGET327 in strain GY5344, an early strain containing the integrated αPH and PDI expression cassette, was grown, and the cells were lysed as described above. Procollagen was precipitated from the clarified supernatant with 4.5 M NaCl and resuspended in 0.1 M Tris-HCl (pH 7.4). Recombinant procollagen C-proteinase/BMP-1 (35) was used to cleave at the C-propeptide junction (36). The digest was treated with endoglycosidase H (New England Biolabs Inc., Beverly, MA) to remove N-linked carbohydrates as described by the manufacturer. The digests were analyzed by Western blotting using the α(I) procollagen C-propeptide-specific antibody LF-41 (34).

Determination of Thermal Stability—Pepsin digestions were performed on yeast extracts at pH 2.5 using 640 units/ml pepsin with incubation for 15 min. The samples were neutralized with 1 M Tris base. SDS sample buffer was added, and the samples were boiled and then loaded onto a 5% SDS-polyacrylamide gel. Type I procollagen purified from yeast cells conditioned medium of human skin fibroblasts was treated with a mixture of trypsin (100 μg/ml) and chymotrypsin (250 μg/ml) (37). The samples were preheated to the desired temperature for 15 min in a thermal cycler (Perkin-Elmer Model 480), followed by addition of proteases and further incubation for 2 min. The digestion was stopped by addition of SDS sample buffer, followed by immediate boiling of the samples.

Amino Acid Analysis—Aliquots of the purified protein samples were dried and then subjected to vapor-phase hydrolysis overnight at 116 °C under N2 in vacuo. The hydrolyzed amino acids were derivatized with the AccQ-Tag chemistry kit from Waters and analyzed on an AccQ-Tag column using a Hewlett-Packard Model 1100 high pressure liquid chromatography apparatus.

Circular Dichroism Analysis—Purified samples were diluted to ~100 μg/ml using 200 mM sodium phosphate (pH 7.0). Aliquots of 200 μl were analyzed using a Jasco (Easton, MD) Model J-715 CD spectropolarimeter with a Peltier controlled sample holder. The samples were equilibrated for 5 min at each temperature and then scanned from 250 to 185 nm. The results were plotted as the molar ellipticity at a given wavelength as a function of temperature. A first derivative plot of the data was used to determine Tm.

RESULTS

Expression of α(I) Procollagen Gene and Analysis of Collagen Protein—The initial approach to producing human type I procollagen was to express the α(I) procollagen polypeptide. The native human α(I) procollagen signal sequence was tested as well as the prepro-HSA (23) and prepro-α-factor (38) sequences that are commonly used to express heterologous proteins in yeast (Fig. 1). Yeast cells were transfected with plasmids that contain the α(I) procollagen gene constructs varying only in their signal sequence (Fig. 2A). The native procollagen signal sequence resulted in the highest level of α(I) procollagen produced based on the intensity of the bands on the Western blot; the prepro-α-factor and prepro-HSA regions also directed the synthesis of human procollagen, but to a lesser degree. As expression of α(I) procollagen in the absence of α2(I) procollagen results in homotrimer formation in mammalian cells (9), yeast extracts were treated with pepsin to digest susceptible proteins. A light band was detected at the expected size for collagen on SDS-polyacrylamide gel (Fig. 2B), indicating the presence of a homotrimeric collagen triple helix.

Since earlier reports described hyperglycosylation in several yeast strains (39), we compared the sizes of N-linked oligosaccharide at the single acceptor site located in the C-propeptide of α(I) procollagen (40). The C-propeptide trimer was removed from procollagen with C-proteinase/BMP-1 (41). Following deglycosylation of human skin fibroblast- and yeast-derived human procollagen C-propeptides with endoglycosidase H, identical decreases in molecular mass were seen by Western blotting (Fig. 2C). This result suggests that similar levels of carbohydrate were added to human procollagen expressed in Saccharomyces and in fibroblast cultures.

The same signal sequence substitutions tested for α(I) procollagen plus the yeast invertase signal were used for expression of the individual gene products α2(I) procollagen, αPH, and PDI. The native signal sequences for α2(I) procollagen and αPH gave the highest expression detected by Western blotting, whereas PDI was more efficiently expressed and secreted into the endoplasmic reticulum with the preinvertase signal sequence than with its native signal sequence. In addition, no difference was measured in the expression levels or retention of PDI in the endoplasmic reticulum by Western blotting when the endoplasmic reticulum retention signal KDEL, present in PDI of higher eucaryotes, was replaced with HDEL, present in yeast PDI (data not shown).

The formation of a triple helix of homotrimeric type I procollagen stable at >25 °C would indicate the functionality of the prolyl hydroxylase enzyme. The Tm for non-hydroxylated procollagen is ~25 °C (3). We compared pepsin-digested extracts from Saccharomyces expressing human α(I) procollagen at 30 °C with and without the genes that code for prolyl hydroxylase. One prominent band was seen on SDS-polyacrylamide gel that comigrated with the expected size of α(I) procollagen in extracts from cells containing the αPH and PDI genes, but this band was absent from cells without prolyl hydroxylase (Fig. 2D). In the absence of the hydroxylation system, the pro-α1 chains were not able to fold into a triple helix that was stable at the 30 °C growth temperature. These data suggest the yeast cells are expressing and assembling an active prolyl hydroxylase tetramer, reinitiating in the formation of hydroxyproline, which stabilizes the triple helix at elevated temperatures.

Expression of Heterotrimeric Type I Procollagen—The next step was to express four genes in Saccharomyces to test whether they are sufficient for production of type I procollagen. To accurately measure procollagen expression levels, an immunoassay was developed to quantify human heterotrimeric
Western blot of yeast extracts from strains expressing the N-terminal sequence of type I procollagen is pGET327; that containing the prepro-HSA signal is pGET335. B, pepsin digestions at different temperatures of yeast extracts expressing α1(I) procollagen (pGET327). Pepsin-digested yeast extracts were electrophoresed on SDS-polyacrylamide gel, and proteins were visualized by silver staining.

Characterization of type I procollagen homotrimer. A, Western blot of yeast extracts from strains expressing α1(I) procollagen with different signal sequences. Transfected Saccharomyces cerevisiae cells were grown at 20 °C, and expression of the α1(I) procollagen gene was induced by galactose. Yeast extracts were subjected to electrophoresis and Western blotting using antibody LF-41 directed against human/human or human/chicken prolyl hydroxylase subunits. Transformed yeast cells were reduced and electrophoresed on SDS-polyacrylamide gel, followed by Western blotting. The C-propeptide was identified using the LF-41 endo H-insensitive antibody. C, carbohydrate analysis of the homotrimeric type I procollagen C-propeptide region expressed from plasmid pGET327. Endoglycosidase H (EndoH) was used to cleave N-linked oligosaccharide from the C-propeptide of both the human skin fibroblast (HSF)- and yeast-expressed human type I procollagen homotrimer. The digests were reduced and electrophoresed on SDS-polyacrylamide gel, followed by Western blotting. The C-propeptide was identified using the LF-41 antibody. D, pepsin digestions of the type I procollagen homotrimer from pGET327 expressed in yeast with and without prolyl hydroxylase/protein-disulfide isomerase genes integrated into the yeast. Pepsin-digested yeast extracts were electrophoresed on SDS-polyacrylamide gel, and proteins were visualized by silver staining.

Expression constructs that contained chicken/chicken αPH and PDI subunits expressed type I procollagen at higher levels than human/human or human/chicken prolyl hydroxylase subunits (Table I). The integration of the chicken prolyl hydroxylase genes into the yeast genome further increased type I procollagen expression. Other plasmid constructs tested but not reported here included different combinations of yeast promoters driving the four genes. In addition, a new strain was created by integrating the α1(I) and α2(I) procollagen genes into the yeast chromosome. The results of these experiments were lower or undetectable levels of procollagen expression (data not shown).

**Characterization of Recombinant Type I Collagen**—The folding of type I procollagen into a heterotrimeric helix, the $T_m$ of the resulting helix, and the level of hydroxyproline in the helical region were determined. The thermal stability of recombinant type I collagen was evaluated by treatment with a mixture of trypsin and chymotrypsin at various temperatures (37). The yeast-derived recombinant collagen heterotrimer (Fig. 3A) was resistant to the proteases at temperatures as high as 40 °C. The melting curves for the fibroblast-derived collagen (Fig. 3B) suggested that this collagen had a slightly higher $T_m$ relative to the recombinant collagen. Circular dichroism measurement of purified type I collagen showed a $T_m$ of 35 °C (Fig. 4), which is slightly below the $T_m$ of tissue-derived collagen. To directly demonstrate the presence of hydroxyproline in the recombinant collagen, amino acid analysis was performed on a purified sample. This analysis showed hydroxyproline levels that were 82 ± 2% ($n = 7$) of values for collagen from tissue-derived sources, which is in agreement with the CD measurements. Direct detection of hydroxyproline residues by amino acid analysis unequivocally demonstrated the functionality of the prolyl hydroxylase enzyme in our strains.

**Medium Optimization**—During development of the expression system, experiments had shown that 0–2% casamino acid supplementation of the medium influenced the level of detectable human procollagen, with 0.5% casamino acids as the optimal concentration. Further analysis of the medium using three additional supplementations at 0–1% was undertaken to optimize procollagen production (Table II). Casamino acid supplementation was compared with the medium supplements Bacto-Tryptone, Bacto-peptone, and yeast extract for their influence on procollagen expression. A level of 0.5% casamino acid supplementation supported the highest levels of procollagen production of the different supplements tested.

Since casamino acids were the most stimulatory for procollagen production, smaller amino acid mixtures, based on the concentrations that would be found in medium containing 0.5% casamino acids, were tested to identify the stimulatory component(s). Several combinations of amino acid mixtures were tried. Increased levels of proline and glycine had no effect on procollagen production levels. Ultimately, yeast nitrogen base supplemented with arginine, glutamate, and lysine or with glutamate alone supplied the needed component necessary for high level procollagen production (Table II). Procollagen expression levels were 3–4 μg/mg of total protein, or 0.3–0.4%. This requirement of hydroxylated procollagen for precursors of α-ketoglutarate in the media suggests that not enough α-ketoglutarate is made in vivo for the hydroxylation reaction. Ascorbic acid (another cofactor of prolyl hydroxy-

---

**Table I**

| Plasmid       | Type I procollagen genes | αPH/PDI | Yeast strain | Expression level (μg/mg protein) |
|--------------|--------------------------|---------|--------------|---------------------------------|
| pGET150      | α1/α2                    | h/c     | GY5196       | 0                               |
| pGET337      | α1/α2                    | c/c     | GY5196       | 0.11                            |
| pGET391      | α1/α2                    | h/c     | GY5196       | 0.27                            |
| pGET903      | α1/α2                    | h/c     | GY5196       | 0.05                            |
| pGET150      | α1/α2                    | c/c     | GY5382       | 0                               |
| pGET737      | α1/α2                    | c/c     | GY5382       | 0.71                            |

* a h denotes the human gene; c denotes the chicken gene.

* Multiple αPH/PDI expression units were integrated into the chromosome.
The results of this study demonstrate that a complex, multisubunit procollagen molecule can be synthesized and assembled into a triple helix in *Saccharomyces cerevisiae* containing a functional multisubunit prolyl hydroxylase enzyme. Three procollagen chains, coded by two genes, are synthesized and assembled into a triple helix. Two genes coding for prolyl hydroxylase form an active enzyme, presumably a tetramer. Prolyl hydroxylase post-translationally modified procollagen, resulting in a thermally stable molecule. Therefore, the post-translational modification must occur in the same location within the endoplasmic reticulum as the assembly of the three procollagen chains. *Saccharomyces* apparently lacks only the prolyl hydroxylase required for production of thermally stable human type I procollagen. The prolyl hydroxylase stabilizes the triple helix through the generation of hydroxyproline residues, but also is associated with the folding of the triple helix. PDI supports the folding and disulfide formation of procollagen C-propeptides (7). The low levels of triple helix detected by expressing the *α*1(I) procollagen gene in the absence of transfected prolyl hydroxylase genes suggest the existence of a less efficient or rate-limited mechanism in *Saccharomyces* for the generation of triple helical procollagen. It can be postulated that the endogenous yeast PDI may assemble the three procollagen polypeptide chains, but the winding of the triple helix is more efficient in the presence of prolyl hydroxylase.

The procollagen molecule produced in this expression system had several features of tissue-derived procollagen synthesized by mammalian cells. The procollagen molecule was triple helical and thermally stable with a *T*<sub>m</sub> of 35 °C based on CD analysis. Non-hydroxylated collagen has a *T*<sub>m</sub> of 23–25 °C, whereas collagen isolated from mammalian tissues has a *T*<sub>m</sub> of 39–40 °C (3). The *T*<sub>m</sub> was consistent with the stability of the triple helix to proteolytic digestion and was in agreement with the level of hydroxyproline determined by amino acid analysis equivalent to 82% of levels found in tissue-derived type I collagen. N-Linked carbohydrate was also detected on recombinant type I procollagen. Since earlier reports describe hyperglycosylation of heterologous proteins in several yeast strains, the size of N-linked oligosaccharide at the single acceptor site located in the C-propeptide of *α*1(I) procollagen was determined. The C-propeptide contained carbohydrate with a mass equivalent to that found in procollagen isolated from skin fibroblasts. No evidence of hyperglycosylation was observed for the *α*1(I) procollagen polypeptide.

Both homotrimeric and heterotrimeric type I procollagens were synthesized by *Saccharomyces* using plasmids that contained one or both procollagen genes, respectively. The procollagen genes and the genes for prolyl hydroxylase were sufficient to assemble procollagen polypeptides and to generate thermally stable procollagen. To obtain and to optimize expression of procollagen, gene location, use of different promoters,
signal sequence modifications, and species origin of the post-translational machinery were varied. All of these parameters played key roles in the production of thermally stable human type I procollagen, with some combinations producing little or no detectable procollagen material. The configuration that gave the highest type I procollagen yield was placement of the human α1(I) and α2(I) procollagen genes on a 2μ vector with integration of the two chicken prolyl hydroxylase genes into the yeast genome. We initially tested chicken αPH and PDI, but expected human αPH and PDI to offer equal or better post-translational modification efficiency and to potentially increase procollagen levels indirectly through increased folding and procollagen thermal stability. Instead, higher expression of human procollagen was measured with chicken prolyl hydroxylase genes (αPH and PDI) compared with their human counterparts. Two possibilities to explain the higher expression of human procollagen using chicken prolyl hydroxylase are potentially higher enzymatic activity and enhanced interaction of the gene with its promoter to increase transcription. Sequence homologies between chicken and human αPH protein and cDNA sequences are 88 and 81%, respectively (28). Chicken and human PDI protein sequences are >90% homologous, whereas cDNA sequence comparison shows only 76% homology (26).

Developmental work on this yeast expression system showed the requirement of casamino acids for higher level production of recombinant procollagen. Addition of individual components of casamino acids to the medium showed that glutamate alone was sufficient to provide procollagen expression levels equivalent to those observed using the entire mixture of amino acids found in casamino acids. One possible explanation for the role of glutamate is its ability to undergo intracellular oxidation to create α-ketoglutarate, an intermediate in the tricarboxylic acid cycle and an essential cofactor for proline hydroxylation. Glycine and proline supplementation did not affect procollagen expression levels even though procollagen consists of high amounts of these two amino acids. In addition, this medium does not contain ascorbic acid, a cofactor for prolyl hydroxylase and necessary for the expression of thermally stable procollagens in other recombinant expression systems (14, 18). "Saccharomyces" must generate sufficient levels of this cofactor to supply the prolyl hydroxylase enzyme. These results are important for the production of proteins that could be used in medical applications, as use of a medium free of animal-derived components provides an extra margin of safety and avoids a potential regulatory hurdle defining the source of a raw material.

Fibroblasts have been used to delineate the biosynthetic pathway of collagen for 3 decades. Since fibroblasts express procollagen ubiquitously, many suggestions have been made about the function of various proteins in the biosynthesis of procollagen at the level of transcription, translation, and assembly. Hsp47 is a serpin-like collagen-specific chaperone localized in the endoplasmic reticulum that transiently binds to type I–V collagens and that is involved in the assembly and/or packaging of collagens (6). Hsp47 is associated with polysome-bound α1(I) procollagen chains (42), and it prevents overmodification of type III procollagen in transfected 293 kidney cells (43). BiP/Gpr78 and Grp94 form a complex with Hsp47 during the maturation of newly synthesized type IV collagen (44). PDI interacts with the C-propeptide of collagen chains prior to trimer formation, and prolyl hydroxylase remains associated with the triple helical domain if triple helical formation is prevented (7). Yeast was chosen as a model system for the synthesis of procollagen because it is a well characterized eucaryotic organism that does not normally synthesize procollagen. To synthesize type I procollagen, only four genes were shown to be required. This study shows that other fibroblast-specific genes are not needed for the basic mechanism of recognition and assembly of the procollagen polypeptides to form a stable triple helical molecule. Recombinant Hsp47 was not required for assembly of triple helical type I procollagen. PDI expressed with αPH increased homotrimeric type I procollagen synthesis, suggesting an enhancement of the assembly of procollagen with increased PDI and/or αPH protein.

In summary, we have demonstrated the minimum requirements for type I procollagen expression in "Saccharomyces". A total of four genes are required. Two genes code for the two polypeptide chains of human type I procollagen. Two additional genes code for prolyl hydroxylase, a modification enzyme that post-translationally hydroxylates proline residues within the triple helical domain of the procollagen polypeptides. Prolyl hydroxylase or its individual subunits also enhance the level of procollagen synthesized. Glutamate, possibly acting as a precursor for the synthesis of α-ketoglutarate, is required for generating higher levels of triple helical procollagen molecules. Other proteins associated with procollagen synthesis in mammalian cells are not required in "Saccharomyces". It remains to be determined if specific chaperone proteins play a subtle role in the assembly of procollagens in eucaryotic cells.

Acknowledgments—We thank Arcie Alea, Meghan Bowzer, Hein Bui, Betty Elder, Lucas Hanscom, Nathan C. Hitzeman, Doug Hodges, Caroline Lanigan, Chin Y. Loh, Paul Nguyen, Winson Ong, Naomi Sakai, and Ernie Tai for excellent technical work on this project. We also thank Rudolf Jaenisch and Daniel Greenspan for chicken PH and PDI cDNA clones, and Larry Fisher for LF-39 and LF-41 antibodies.

REFERENCES

1. Haralson, M. A., and Hassell, J. R. (1995) in Extracellular Matrix: A Practical Approach (Haralson, M. A., and Hassell, J. R., eds) pp. 1–30, IRL Press, New York
2. Prokop, D. J., and Kivirikko, K. I. (1984) N. Engl. J. Med. 311, 376–386
3. Berg, R. A., and Prockop, D. J. (1973) Biochim. Biophys. Res. Commun. 52, 115–120
4. Pivivalov, P. L. (1982) Adv. Protein Chem. 35, 53–104
5. McLaughlin, S. H., and Bulleid, N. J. (1998) Matrix Biol. 16, 369–377
6. Fleischmajer, R., Olsen, B. R., Timpl, R., Perlis, J. S., and Lovelance, O. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 3354–3358
7. Wilson, R., Lees, J. F., and Bulleid, N. J. (1996) J. Biol. Chem. 271, 9637–9643
8. Nagata, K. (1996) Trends Biochem. Sci. 21, 22–26
9. Geddins, A. E., and Prokop D. J. (1993) Matrix 13, 399–405
10. Fertala A., Sieron, A. L., Ganguly, A., Li, S. W., Alas-Kokko, L., Anumula, K. R., and Prokop, D. J. (1994) Biochem. J. 298, 31–37
11. Fichard, A., Tillet, E., Delacoux, F., Garrone, R., and Ruggiero, F. (1997) J. Biol. Chem. 272, 30083–30087
12. Vueri, K., Pihlajaniemi, T., Marttila, M., and Kivirikko, K. I. (1992) Proc. Natl.
Type I Procollagen Expression in Yeast

23309

13. Tomita, M., Kitajima, T., and Yoshizato, K. (1997) J. Biochem. (Tokyo) 121, 1061–1069
14. Myllyharju, H., Lamberg, A., Notbohm, H., Fietzek, P. P., Pihlajaniemi, T., and Kivirikko, K. I. (1997) J. Biol. Chem. 272, 21824–21830
15. Lamberg, A., Helaakoski, T., Myllyharju, J., Peltonen, S., Notbohm, H., Pihlajaniemi, T., and Kivirikko, K. I. (1996) J. Biol. Chem. 271, 11988–11995
16. Toman, P. D., Pieper, F., Sakai, N., Karatzas, C., Platenburg, E., de Wit, I., Samuel, C., Dekker, A., Daniels, G. A., Berg, R. A., and Platenburg, G. (1999) Transgenic Res. 8, 415–427
17. John, D. C. A., Watson, R., Kind, A. J., Scott, A. R., Kadler, K. E., and Bulleid, N. J. (1999) Nature Biotechnol. 17, 385–389
18. Vuorela, A., Myllyharju, J., Nissi, R., Pihlajaniemi, T., and Kivirikko, K. I. (1997) EMBO J. 16, 6702–6712
19. Vaughan, P. R., Galanis, M., Richards, K. M., Tebb, T. A., Ramshaw, J. A. M., and Werkmeister, J. F. (1998) Connect. Tissue Res. 21, 43–50
20. Wozney, J. M., Rosen, V., Celeste, A. J., Mitsiok, L. M., Whitters, M. J., Dzir, R. W., Hewick, R. M., and Wang, E. A. (1988) Science 292, 1528–1534
21. Chukin, Y., van der Rest, M., and Prockop, D. J. (1988) J. Biol. Chem. 263, 15996–16003
22. Hill, S. T., Smith, B. D., and Greenspan, D. S. (1988) J. Biol. Chem. 263, 5858–5865
23. De Baetselier-Van Broeckhoven, A. (1994) Bioprocess Technol. 19, 431–447
24. Clark, C. C. (1979) J. Biol. Chem. 254, 10798–10802
25. Kessler, E., Adar, R., Goldberg, B., and Neice, R. (1986) Collagen Relat. Res. 6, 249–266
26. Sauk, J. J., Smith, T., Norris, K., and Ferreira, L. (1994) J. Biol. Chem. 269, 3941–3946
27. Hosokawa, N., Hohenadl, C., Satoh, M., Kuhn, K., and Nagata, K. (1998) J. Biol. Chem. (Tokyo) 124, 654–662
28. Ferreira, L. R., Norris, K., Smith, T., Hetert, C., and Sauk, J. J. (1996) Connect. Tissue Res. 33, 265–273