Assembly of Stable Human Type I and III Collagen Molecules from Hydroxylated Recombinant Chains in the Yeast *Pichia pastoris*

**EFFECT OF AN ENGINEERED C-TERMINAL OLIGOMERIZATION DOMAIN FOLDON**

Pichia pastoris

**Hydroxylated Recombinant Chains in the Yeast Assembly of Stable Human Type I and III Collagen Molecules from**

The Journal of Biological Chemistry Vol. 278, No. 34, Issue of August 22, pp. 32478–32483, 2003

Received for publication, April 28, 2003, and in revised form, June 3, 2003

Published, JBC Papers in Press, June 12, 2003, DOI 10.1074/jbc.M304405200

Outi Pakkanen‡, Eija-Riitta Hämäläinen§, Kari I. Kivirikko¶, and Johanna Myllyharju††

From the ‡Collagen Research Unit, Biocenter Oulu and Department of Medical Biochemistry and Molecular Biology, University of Oulu, FIN-90014 Oulu and §Fibrogen Europe, Medipolis, FIN-90230 Oulu, Finland

The C-propeptides of the proα chains of type I and type III procollagens are believed to be essential for correct chain recognition and chain assembly in these molecules. We studied here whether the 30-kDa C-propeptides of the human pC(I), pCα2(I), and pCα1(III) chains, i.e. proα chains lacking their N-propeptides, can be replaced by foldon, a 29-amino acid sequence normally located at the C terminus of the polypeptide chains in the bacteriophage T4 fibritin. The αfoldon chains were expressed in *Pichia pastoris* cells that also expressed the two types of subunit of human prolyl 4-hydroxylase; the foldon domain was subsequently removed by pepsin treatment, which also digests non-triple helical collagen chains, whereas triple helical collagen molecules are resistant to it. The foldon domain was found to be very effective in chain assembly, as expression of the α(1)foldon or α(1)IIIfoldon chains gave about 2.5-3-fold the amount of pepsin-resistant type I or type III collagen homotrimers relative to those obtained using the authentic C-propeptides. In contrast, expression of chains with no oligomerization domain led to very low levels of pepsin-resistant molecules. Expression of α2(I)foldon chains gave no pepsin-resistant molecules at all, indicating that in addition to control at the level of the C-propeptide other restrictions at the level of the collagen domain exist that prevent the formation of stable [α2(I)]3 molecules. Co-expression of α1(I)foldon and α2(I)foldon chains led to an efficient assembly of heterotrimeric molecules, their amounts being about 2-fold those obtained with the authentic C-propeptides and the α(1) to α(2) ratio being 1.91 ± 0.31 (S.D.). As the foldon sequence contains no information for chain recognition, our data indicate that chain assembly is influenced not only by the C-terminal oligomerization domain but also by determinants present in the α chain domains.

The collagen superfamily of proteins includes more than 20 types of collagen and more than 15 additional proteins that have collagen-like domains. All collagen molecules consist of three polypeptide chains, called α chains, that are coiled around each other into a triple helix and contain the triplet sequence -Gly-X-Y-, in which the Y position amino acid is often 4-hydroxyproline. The most abundant collagens form fibrils and are therefore known as fibril-forming collagens, whereas others form other kinds of supramolecular structures. The molecules of the most abundant fibril-forming collagen, type I, consist of two α1(I) chains and one α2(I) chain, whereas the molecules of type III collagen are [α1(III)]3 homotrimers. In addition to the type I collagen heterotrimer, most tissues also contain a small amount of type I collagen with a chain composition of [α1(I)]αβ, known as the type I collagen homotrimer (for reviews, see Refs. 1–4).

The fibril-forming collagens are synthesized as procollagen molecules with N- and C-terminal propeptides. Chain assembly begins with association of the three C-propeptides through a process directed by their structures (2, 5). Renaturation experiments with individual type I collagen α chains that lack any propeptides have indicated, however, that the chains form both [α1(I)]2α2(II) heterotrimers and [α1(I)]3 homotrimers, although the process is very slow and the Tm values of the molecules formed are lower than those present in vivo (6, 7). Based on these findings and numerous subsequent studies carried out in a large variety of biological systems over a period of more than 25 years, the C-propeptides are now believed to be essential for correct chain recognition and to play a crucial role in chain assembly in vivo (2, 5, 8).

Most studies of collagen synthesis have used vertebrate cells, which usually possess sufficient levels of all the specific cotranslational and posttranslational enzymes needed for collagen processing. In recent years experiments have also been performed using recombinant expression in cultured insect cells (9–11), yeasts (12–14) and plants (15). These cells have been shown to assemble partially (13, 15) or fully (9–12, 14) hydroxylated recombinant collagen chains into molecules with partially (13, 15) or fully (9–12, 14) stable triple helices, provided that they co-express the two types of subunit of a recombinant prolyl 4-hydroxylase. This key enzyme of collagen synthesis is an α2β2 tetramer in vertebrates (16–18). Most of the triple helical collagen molecules assembled in insect cells, yeasts, and plants are not secreted (9–15) but accumulate within the endoplasmic reticulum (19). Studies of these systems have confirmed that the N-propeptides are not required for the assembly of triple helical molecules, as they can be swapped between collagen types (10, 11) or omitted (14, 15, 20).

A recent study reported, highly surprisingly, that assembly of triple helical recombinant type I collagen molecules from proα chains, a procollagen polypeptide chain; pCa chain, a procollagen polypeptide lacking the N-propeptide; αfoldon chain, a pCα chain with the C-propeptide replaced by the foldon domain.

This paper is available online at http://www.jbc.org

¶To whom correspondence should be addressed: Dept. of Medical Biochemistry and Molecular Biology, P. O. Box 5000, University of Oulu, FIN-90014 Oulu, Finland. Tel: 358-8-537-5740; Fax: 358-8-537-5810; E-mail: johanna.myllyharju@oulu.fi.

††The abbreviations used are: α chain, a collagen polypeptide chain; pCa chain, a procollagen polypeptide chain; pCα chain, a procollagen polypeptide lacking the N-propeptide; αfoldon chain, a pCα chain with the C-propeptide replaced by the foldon domain.
tially hydroxylated α chains in the yeast Saccharomyces cerevisiae does not even require the C-propeptides (20). The triple helical molecules had an α(1)I to α(2)I chain ratio of 5:1, however, suggesting that they consist of mixtures of [α(1)I]2α(2)I heterotrimer and [α(1)I]3 homotrimers (20).

In the present work we have studied whether fully hydroxylated recombinant human type I collagen α chains lacking C-propeptides can assemble effectively into triple helical molecules in the yeast Pichia pastoris when it co-expresses the two types of subunits of a recombinant human prolyl 4-hydroxylase (12, 14). As the collagen domains of the α chains of type III collagen, unlike those of type I, contain two cysteine residues at their C-terminal ends, which are involved in the formation of interchain disulfide bonds (see Ref. 21), we also determined whether type III collagen α chains lacking their C-propeptides would be assembled more effectively than those of type I. One major aspect studied here was whether the 30-kDa C-propeptides of the proc(1)I, proc(2)I, or proc(1)III chains could be replaced by foldon, a 29-amino acid peptide that is normally located at the C terminus of the polypeptide chains in the bacteriophagel T4 fibritin, a three-stranded α helical coiled-coil protein, and appears to be essential for the assembly of the bacteriophage T4 fibritin, a three-stranded α helical molecule with its own C-propeptides. Furthermore, co-expression of the α(1)Ifoldon and α(2)Ifoldon chains led to an even more efficient assembly of homotrimeric molecules with triple helical molecules than in the cases of chains expressed with their own C-propeptides. Furthermore, co-expression of the α(1)Ifoldon and α(2)Ifoldon chains led to the effective formation of type I collagen heterotrimer with the correct 2:1 chain ratio, indicating that the chain composition is determined not only by the C-terminal oligomerization domain but also by determinants present in the collagen domain of the polypeptide chains, at least in the case of very small oligomerization domains such as foldon.

MATERIALS AND METHODS

P. pastoris Expression Vectors and Generation of Recombinant Strains—The P. pastoris host strain yC300 (his4, arg4, ade1) and the expression vectors pBLADES and pBLARGIX (27) were gifts from Dr. James Cregg, (Keck Graduate Institute of Applied Life Sciences), and the vectors pPIC3K, pPICZB, and pPICZαA were from Invitrogen. The recombinant strains were generated by the electroporation method according to the manufacturer’s instructions (Invitrogen (32)). The recombinant strains were of the methanol utilization plus phenotype.

PCa(I)I, α(1)Ifoldon, and α(I)III Strains—A DNA for the β subunit of human prolyl 4-hydroxylase lacking the signal sequence and flanked by EcoRI restriction sites (12) was cloned into pPICzaA in frame with the S. cerevisiae α mating factor (αMF) pre-pro sequence. The expression cassette encoding the αMF-β polypeptide was digested from pPICzaAβ with BamHI-BgII and cloned into pBLADES. To generate a recombinant P. pastoris strain expressing human prolyl 4-hydroxylase α₂β₂ tetramers, the pBLARGIXα (14) and pBLADESα were linearized with HinII and Spol, respectively, and cotransformed into the yC300 strain. The resulting strain was named ArgαAdεβ.his.α I. P. pastoris expression cassette encoding human type I proc(1)I chains (procollagen chains lacking their N-propeptides) was digested from pPICZBcα(1)I (14) with Pmel-NotI and cloned into pPIC3K to generate pPICZKcα(1)I. A DNA coding for the 29-amino acid foldon domain of fibritin (22–24) (GenBank accession no. AAD42679) was generated by annealing the oligonucleotide Foldon-5′ (5′-AGCTTTTAGTCTCTGAAAGAAGATGGGCAAGTTAGCTCTGCAA-3′) with Foldon-5′-2 (5′-CCATTTTACGAACTGTTGCMCCCATCCCTGGAGGTCCTGAAATATA-3′), and Foldon-3′ (5′-AGATGGCGAATT-GGTATTTCTTCT TTCTACTACAACGATGAACG-3′) with Foldon-3′ (5′-GGCCGCTTATGCTGTGATAAAAGTAGAAGATAATCCATCG-3′). The oligonucleotides (Invitrogen) were designed so that HindIII and NotI overhangs (underlined) are created at the 5′ and 3′ ends of the annealed Foldon-5′ and Foldon-3′ fragments, respectively, and cohesive overhangs at the 3′ and 5′ ends. The foldon fragments were co-ligated into HindIII-NotI digested pBlueScript (Stratagene) to generate pBSfoldon. To replace the sequence coding for the C-propeptide of the procα(I)I chain with that coding for foldon, two fragments were generated by PCR, the first extending from an internal BamHI site in the proc(I)I cDNA and ending with the last amino acid of the C-telopeptide, and the second from the first codon of the cDNA to the NotI site following the stop codon. These fragments were co-ligated into BamHI-NotI digested pPIC3Kcα(1)I to generate pPIC3Kcα(1)I with that coding for foldon. To replace the C-propeptide, a fragment extending from the internal BamHI site of proc(I)I cDNA to the codon for the last amino acid of the C-telopeptide followed by a stop codon, and a NotI site was created by PCR and ligated into BamHI-NotI digested pPIC3Kcα(1)I to generate pPIC3Kcα(1)I with that coding for foldon. The pPIC3Kcα(1)I, pPIC3Kα(1)foldon, and pPIC3Kα(1) constructs were linearized with NotI and transformed into the ArgαAdεβ.his.α strain expressing recombinant human prolyl 4-hydroxylase tetramers. Schematic representations of the proc(I)I, α(I)foldon and α(I)I chains are shown in Fig. 1. PCa(1)II, α(2)Ifoldon, PCa(1)+PCa(2)I, and α(I)foldon + α(2)Ifoldon Strains—A P. pastoris expression cassette encoding human proc(I)I chains was digested from pBLADEIXcα(2)I (14) with Pmel-NotI and cloned into pPICZB to generate pPICZBcα(2)I. To replace the C-propeptide of proc(I)I with foldon, two fragments were generated by PCR, the first extending from an internal AorII site in the proc(I)I cDNA to the codon for the last amino acid of the C-telopeptide, and the second extending codon as above. These fragments were co-ligated into AorII-NotI digested pPICZBcα(2)I to generate pPICZBcα(2)I withfoldon. The constructs were linearized with Pmel and transformed into the ArgαAdεβ.his.α strain to generate the strains PCa(1)II and α(I)foldon. To study the expression of type I collagen heterotrimer, the linearized pPICZBcα(2)I and pPICZBcα(2)I were transformed into the above strains expressing PCa(1)I and α(I)foldon chains, respectively.

PCa(1)III, α(I)IIIfoldon, and α(I)I plus α(II)III Strains—To delete the sequence encoding the N-propeptide of proc(I)III chains, a PCR fragment extending from the Pmel site of the alcohol oxidase 5′ sequence of pPICZB to the end of the signal sequence of proc(I)III and followed directly by the sequence coding for the N-telopeptide of proc(I)III until an internal Ndel site was amplified using pPICZBproc(I)III (12) as a template and

![Fig. 1. Schematic representations of the various recombinant collagen polypeptide chains expressed in P. pastoris. A, a full-length procollagen (procα) chain; B, a procollagen chain lacking the N-propeptide (procα); C, a proc chain in which the C-propeptide is replaced by foldon (αfoldon); and D, a proc chain (α). The triple helical collagen domain is not drawn to scale. All chains contain the native signal peptide at their N termini (not shown).](http://www.jbc.org/)

Downloaded from http://www.jbc.org/ on April 26, 2019
A GC-rich oligonucleotide (with the nucleotides corresponding to the signal sequence and N-telopeptide shown in bold and italics, respectively, and the NdeI site underlined) as the reverse primer. The fragment was cloned into Pmel-NdeI digested pPICZBpro1(III) to generate pPICZBpCo1(III). To change the expression vectors, the pro1(III) and pC1(III) expression cassettes were digested from the PICZB vectors with Pmel-NdeI and cloned into pPIC3K to generate pPIC3Kpro1(III) and pPIC3KpCo1(III). To replace the C-propeptide of pC1(III) with fodont, two fragments were generated by PCR, the first extending from an internal AvrII site in the pC1(III) cDNA to the 3' end of the C-telopeptide sequence and the second encoding fodafone as described above. The fragments were co-ligated into AvrII-NdeI-digested pPIC3KpCo1(III) to generate pPIC3K1(III)foldon. To delete the C-propeptide, a fragment extending from the internal AvrII site to the 3' end of the C-telopeptide sequence followed by a stop codon, and a NotI site was created by PCR and ligated into AvrII-NotI-digested pPIC3KpCo1(III) to generate pPIC3K1(III)foldon. The pPIC3Kpro1(III), pPIC3KpCo1(III), pPIC3K1(III)foldon, and pPIC3K3K1(III)constructs were linearized with SfiI and transformed into the Arg1/His3 strain.

**Culture and Induction of P. pastoris Strains**—Cells were cultured in 25-ml shaker flasks in a buffered glycerol complex medium, pH 6.0, with 1 g/liter yeast extract and 2 g/liter peptone. Expression was induced in a buffered minimal methanol medium, pH 6.0, and methanol was added every 12 h to a final concentration of 0.5%. Amino acids were added up to 100 μg/liter as required.

**Analysis of the Recombinant Collagens**—Cells were harvested after a 60-h methanol induction at 30 °C, washed once, and suspended in cold (4 °C) 5% glycerol, 1 mM Pefabloc SC, and 50 mM sodium phosphate buffer, pH 7.4. The cells were broken by vortexing with glass beads, and the lysate was centrifuged at 10,000 g for 30 min. Aliquots of the soluble fractions were analyzed by SDS-PAGE under reducing conditions followed by Western blotting using a polyclonal antibody against type I (A) and after pepson digestion for 2 h at 22 °C (B and C) by SDS-PAGE under reducing conditions followed by Western blotting using a polyclonal antibody against type I collagen (A and C) or Coomassie Blue staining (B). The pCo1(III), α1(I)-foldon, and α1(III) chains are indicated by arrows.

Assembly of the chains into triple helical molecules was analyzed by digesting aliquots of the cell extracts with pepson. The triple helix of collagens is resistant to pepson, whereas non-triple helical chains and the propeptides of triple helical molecules are digested (28). Pepson-resistant polypeptides corresponding to the α1(I) chains were seen in samples from all three strains expressing pCo1(III), α1(I)-foldon, and α1(III) chains in Coomassie Blue-stained SDS-PAGE performed under reducing conditions (Fig. 2B, lanes 1–3) and in immunoblots (Fig. 2C, lanes 1–3). Likewise, bands corresponding to α1(III) chains were seen in pepson-digested samples from the strains expressing pro1(III), pCo1(III), α1(I)-foldon, and α1(III) chains in Coomassie Blue-stained SDS-PAGE carried out under reducing conditions (Fig. 3B, lanes 1–4).

The felden domain was found to be very efficient for chain assembly, as increased amounts of pepson-resistant α1(I) and α1(III) chains were seen in the samples from cells expressing the α1(I)-foldon and α1(III)foldon chains, respectively (Figs. 2B, C, and 3B). Densitometry of the bands in 11 individual samples indicated that the increase in the expression level of pepson-resistant type I collagen homotrimers in the strains expressing α1(I)-foldon chains was up to 3.0-fold that in strains expressing pCo1(III) chains at the highest levels. Similar experiments indicated a 2.4-fold increase in the expression level of pepson-resistant type I collagen homotrimers in strains expressing α1(III)foldon chains. The α1(I) and α1(III) chains lacking any oligomerization domain also became assembled into triple helical molecules (Figs. 2, 3, and 4), but their assembly levels were distinctly lower than those of the pCo1(III), α1(I)-foldon, pCo1(III), and α1(III)foldon chains (Figs. 2B, C, and 3B), and the chains lacking any oligomerization domain were susceptible to degradation (Figs. 2B, C, and 3B, lane 4).
The thermal stability of the type I collagen homotrimers assembled from the pCa1(I), α1(I)foldon, and α1(I) chains was analyzed by digestion with a mixture of trypsin and chymotrypsin after heating to various temperatures (28). The \( T_m \) of the recombinant type I collagen homotrimers expressed in all three strains was between 38 and 40 °C (Fig. 5A–C), as reported previously for human type III collagen produced in P. pastoris in shaker flasks (12).

Lack of Assembly of Any Stable \([α2(II)]_3\) Homotrimers—To study whether pCo2(I) chains with the C-propeptide replaced with foldon are able to form homotrimers, strains were generated that expressed pCo2(I) and α2(I)foldon chains in the presence of human prolyl 4-hydroxylase. When these strains were cultured, induced, and harvested as described above and analyzed by SDS-PAGE under reducing conditions followed by Western blotting, full-length pCo2(I) and α2(I)foldon chains were seen in the immunoblots of the soluble extracts (Fig. 6A, lanes 1 and 2). Several additional immunoreactive bands, corresponding to degradation products, were also present (Fig. 6A, lanes 1 and 2), however, and in the case of samples digested with pepsin at either 22 or 4 °C, no pepsin-resistant α2(I) chains were seen in either strain (Fig. 6, B and C, lanes 1 and 2).

Assembly of Stable Type I Collagen Heterotrimers—The C-propeptides are believed to have an essential role in the selective association of procollagen chains in a type-specific manner (2, 5, 8). The effect of replacement of the C-propeptides of both the pCa1(I) and pCa2(I) chains with foldon on the assembly of type I collagen heterotrimers was studied by transforming expression constructs encoding pCo2(I) and α2(I)foldon chains into the above described recombinant strains expressing pCo1(I) and α1(I)foldon chains, respectively. The strains were cultured and harvested as described above. Bands corresponding to the pCo1(I) and pCo2(I) chains and α1(I)foldon and α2(I)foldon chains were seen in the immunoblots of the recombinant strains (details not shown). Two pepsin-resistant polypeptides corresponding to the α1(I) and α2(I) chains were seen in samples from the two strains co-expressing pCa1(I) and
C-terminal Oligomerization Domains in Collagen Assembly

Fig. 7. Analysis of the assembly of type I collagen heterotrimers from modified proα1(I) and proα2(I) chains. The recombinant P. pastoris strains were induced and harvested and the cells broken as described in the legend for Fig. 2. Soluble fractions of cell lysates from strains co-expressing pcα1(I) and pcα2(I) chains (lane 1) or α1(I)foldon and α2(I)foldon chains (lane 2) were analyzed after digestion with pepsin by 8% SDS-PAGE under reducing conditions followed by Coomassie Blue staining. The arrows indicate α1(I) and α2(I) chains, respectively.

pcα2(I) chains (Fig. 7, lane 1) or α1(I)foldon and α2(I)foldon chains (Fig. 7, lane 2).

It has previously been reported that the pcα1(I) and pcα2(I) chains are assembled into type I collagen heterotrimers with the correct 2:1 chain ratio, 1.92 ± 0.13 (S.D.) (14). Densitometry of the pepsin-resistant α1(I) and α2(I) bands from 10 individual strains co-expressing α1(I)foldon and α2(I)foldon chains indicated that the α1(I) to α2(I) chain ratio was 1.91 ± 0.31 (S.D.) (Fig. 7, lane 2), and an increased quantity of pepsin-resistant α1(I) and α2(I) chains was also seen in these samples (Fig. 7). Densitometry of the bands in seven individual samples indicated that the increase in the expression level of type I collagen heterotrimers in these strains was 2.1-fold relative to that obtained with the strains co-expressing pcα1(I) and pcα2(I) chains. The Tm of the recombinant type I collagen heterotrimers produced in the strains co-expressing pcα1(I) and pcα2(I) chains or α1(I)foldon and α2(I)foldon chains was between 38 and 40 °C (data not shown).

DISCUSSION
Numerous studies carried out in a large variety of biological systems have demonstrated that the C-propeptides are essential for correct chain recognition and chain assembly in the procollagen molecules in vivo (2, 5, 8). These events are then followed by nucleation and alignment of the collagen domains of the proc chains, driven mainly by sequences present at the C-terminal end of these domains and propagation of the triple helix from the C terminus toward the N terminus (2, 5, 8, 29). Where the C-propeptides of the proα chains in the precursor forms of all three main fibril-forming collagens, types I–III, consist of about 245 amino acid residues (1, 30), the critical region in the C-propeptides required for correct chain recognition consists of a discontinuous sequence of 15 amino acids (8). No data are available, however, on any effective replacement of the C-propeptides by a smaller domain. One study performed in a cell-free translation system in the presence of semi-permeabilized cells indicated that the C-propeptides can be replaced with a transmembrane domain, albeit at a much lower efficiency (31).

The data reported here indicate that the short, 29-amino acid foldon sequence can effectively replace the large C-propeptides of the human pcα1(I), pcα2(I), and pcα1(III) chains in chain assembly in P. pastoris cells that also express the two subunits of human prolyl 4-hydroxylase. The foldon domains could subsequently be removed by pepsin treatment, leading to collagen molecules with stable triple helices. Expression of the α1(I)foldon and α1(III)foldon chains gave homotrimeric [α1(I)]3 and [α1(III)]3 collagen molecules in pepsin treatment, whereas co-expression of the α1(I)foldon and α2(I)foldon chains gave heterotrimeric collagen molecules with the correct 2:1 chain ratio. Replacement of the C-propeptides by the foldon sequence also increased the levels of the pepsin-resistant homotrimeric and heterotrimeric molecules by more than 2-fold. It therefore appears that the chains containing foldon also became assembled even more efficiently than those containing the C-propeptides, whereas the nonassembled chains were rapidly degraded.

Expression of α1 chains without any C-terminal oligomerization domain led to inefficient assembly of triple helical molecules, as the amounts of pepsin-resistant [α1(I)]3 and [α1(III)]3 homotrimers recovered were distinctly smaller than those obtained when the chains were expressed with an oligomerization domain. This result disagrees with a recent report of efficient assembly of type I collagen heterotrimers from partially hydroxylated α chains in the yeast S. cerevisiae (20). The reason for this difference is unknown, as the data reported for S. cerevisiae disagree not only with our data obtained with another yeast strain but also with those reported in numerous previous studies in a variety of biological systems (2, 5, 8).

Although the proα1(I) and proα2(I) chains can form both [proα1(I)]3proα2(I) heterotrimers and [proα1(I)]3 homotrimers, the former are preferred (2, 5). In agreement with this preference, co-expression of proα1(I) and proα2(I) chains in insect cells (10) and P. pastoris (14) led essentially to the formation of heterotrimers only, unless the proα1(I) chain was expressed in great excess (10). As the foldon sequence contains no information for chain recognition, it seemed possible that co-expression of α1(I)foldon and α2(I)foldon chains might lead to a mixture of pepsin-resistant [α1(I)]3 [α2(I)] heterotrimeric and [α1(I)]3 homotrimers. The ratio of α1(I) to α2(I) chains in the pepsin-resistant molecules was 1.91:1, however, indicating that essentially all the chains were present only in heterotrimers. Random trimer formation within the endoplasmic reticulum of the two types of chain expressed in a 1:1 ratio would give trimers [α1(I)]3, [α1(I)]2α2(I), [α1(I)]α2(I)2, and [α2(I)]3 in a ratio of 1:3:3:1. The α1(I)[α2(I)]3 and α2(I) trimers are thermally unstable (6, 7) and would thus be degraded, whereas the [α1(I)]3 homotrimer is stable, as shown here and previously (10, 14). In this case the samples would contain about 75% pepsin-resistant [α1(I)]3 [α2(I)] heterotrimers and 25% [α1(I)]3 homotrimers, and the ratio of α1(I) to α2(I) chains would be 3:1 rather than the observed 2:1. Our data thus indicate that even in the presence of the foldon sequence, the chain composition is influenced by determinants present in the α chain domains. The chain ratio obtained here differs distinctly from the α1(I) to α2(I) chain ratio of 5:1 obtained for pepsin-resistant molecules when α chains lacking any propeptides were expressed in S. cerevisiae (20). This difference may be due in part to a preferential degradation of the α2(I) chains in the absence of efficient trimer formation.

Renaturation experiments with α2(I) chains lacking any propeptides have demonstrated that they do form [α2(I)]3 homotrimers but that the Tm of such homotrimers is only 22–24 °C (6, 7). Because the foldon domain has been reported to promote marked stabilization of the triple helix formed by (Pro-Pro-Gly)10foldon, we also studied whether fully hydroxylated α2(I)foldon chains would form homotrimers that are converted to stable [α2(I)]3 molecules upon pepsin treatment. This possibility was supported by a report on in vitro translation experiments in a rabbit reticulocyte lysate system indicating that truncated proα2(I) chains with an internal deletion formed homotrimers with stable triple helices provided that the C-propeptide of the proα2(I) chains had been replaced by that of the proα1(III) chains (8). On the other hand, full-length proα2(I) chains in which the C-propeptide had been replaced
with that of the proc1(III) chains formed no homotrimeric molecules with stable triple helices in recombinant expression in insect cells (10). The present data clearly indicate that the proc2(II)domain chains formed no homotrimers with stable triple helices, a finding that supports the previous suggestion (10) that in addition to control at the level of the C-propeptide, additional restrictions may exist at the level of the collagen domain of the proc2(II) chains that prevent formation of stable \([\alpha2(II)]_3\) triple helices.

Acknowledgments—We thank Tanja Aatsinki, Raija Juntunen, Anne Kokko, Eeva Lehtimäki, Minna Siurua, and Anne Kurkela for expert technical assistance.

REFERENCES
1. Kadler, K. (1995) Protein Profile 2, 491–619
2. Prockop, D. J., and Kivirikko, K. I. (1995) Annu. Rev. Biochem. 64, 403–434
3. Bateman, J. F., Lamande, S. R., and Ramshaw, J. A. M. (1996) in Extracellular Matrix (Comper, W. D., ed) Vol. 2, pp. 22–67, Harwood, Amsterdam
4. Myllyharju, J., and Kivirikko, K. I. (1997) J. Biol. Chem. 272, 1318–1328
5. McLaughlin, S. H., and Bulleid, N. J. (1998) J. Mol. Biol. 279, 1081–1089
6. Tkocz, C., and Kühn, K. (1999) Eur. J. Biochem. 266, 454–462
7. Leikina, E., Mertts, M. V., and Leikin, S. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 1314–1318
8. Lees, J. F., Tsah, M., and Bulleid, N. J. (1997) EMBO J. 16, 908–916
9. Lamberg, A., Helaakoski, T., Myllyharju, J., Peltonen, S., Notohm, H., Pihlajaniemi, T., and Kivirikko, K. I. (1996) J. Biol. Chem. 271, 11888–11895
10. Myllyharju, J., Lamberg, A., Notohm, H., Fietzék, P. P., Pihlajaniemi, T., and Kivirikko, K. I. (1997) J. Biol. Chem. 272, 21824–21830
11. Nokelainen, M., Helaakoski, T., Myllyharju, J., Notohm, H., Pihlajaniemi, T., Fietzék, P. P., and Kivirikko, K. I. (1998) Matrix Biol. 16, 329–338
12. Vuorela, A., Myllyharju, J., Nissi, R., Pihlajaniemi, T., and Kivirikko, K. I. (1997) EMBO J. 16, 6702–6712
13. Toman, P. D., Chiashelm, G., McMullin, H., Gieres, L. M., Olsen, D. R., Kovach, R. J., Leigh, S. D., Feng, B. E., Chang, R., Daniels, G. A., Berg, R. A., and Hitzeman, R. A. (2000) J. Biol. Chem. 275, 23903–23909
14. Nokelainen, M., Tu, H., Vuorela, A., Notohm, H., Kivirikko, K. I., and Myllyharju, J. (2001) FEBS Lett. 511, 979–986
15. Merle, C., Perret, S., Lacour, T., Jonval, V., Hudaverdian, S., Garnere, R., Ruggiero, F., and Theisen, M. (2002) FEBS Lett. 511, 114–118
16. Kivirikko, K. I., and Myllyharju, J. (1996) Matrix Biol. 16, 357–368
17. Kivirikko, K. I., and Pihlajaniemi, T. (1998) Adv. Enzymol. Relat. Areas Mol. Biol. 72, 325–400
18. Myllyharju, J. (2003) Matrix Biol. 22, 15–24
19. Keizer-Gunnink, I., Vuorela, A., Myllyharju, J., Pihlajaniemi, T., Kivirikko, K. I., and Veenhuis, M. (2000) Matrix Biol. 19, 29–36
20. Olsen, D. R., Leigh, S. D., Chang, H., McMullin, H., Ong, W., Tai, E., Chiashelm, G., Birk, D. E., Berg, R. A., Hitzeman, R. A., and Toman, P. D. (2001) J. Biol. Chem. 276, 24038–24043
21. Bulleid, N. J., Wilson, R., and Lees, J. F. (1996) Biochem. J. 317, 195–202
22. Tao, Y., Strelkov, S. V., Mesyanzhinov, V. V., and Rossmann, M. G. (1997) Structure 5, 789–798
23. Letarov, A. V., Londer, Y. Y., Boudko, S. P., and Mesyanzhinov, V. V. (1999) Biochemistry (Moscow) 64, 817–823
24. Boudko, S. P., Londer, Y. Y., Letarov, A. V., Sernova, N. V., Engel, J., and Mesyanzhinov, V. V. (2002) Eur. J. Biochem. 269, 833–841
25. Frank, S., Kammerer, R. A., Meelhing, D., Schulthess, T., Landwehr, R., Mann, J., Guo, Y, Lustig, A., Bachinger, H. P., and Engel, J. (2001) J. Mol. Biol. 308, 1081–1089
26. Boudko, S., Frank, S., Kammerer, R. A., Stetefeld, J., Schulthess, T., Landwehr, R, Lustig, A., Bachinger, H. P., and Engel, J. (2002) J. Mol. Biol. 317, 459–470
27. Cereghino, G. P., Cereghino, J. L., Sunga, A. J., Johnson, M. A., Lim, M., Gleeson, M. A., and Cregg, J. M. (2001) Gene 263, 159–168
28. Bruckner, P., and Prockop, D. J. (1981) Anal. Biochem. 110, 360–368
29. Engel, J., and Prockop, D. J. (1991) Annu. Rev. Biophys. Biophys. Chem. 20, 137–152
30. Don, A. S., and Myers, J. C. (1987) Mol. Biol. 193, 127–143
31. Bulleid, N. J., Dally, J. A., and Lees, J. F. (1997) EMBO J. 16, 6684–6701
32. Invitrogen (2002) Manual of Methods for Expression of Recombinant Proteins in Pichia pastoris, Version M, Invitrogen, Carlsbad, CA
Assembly of Stable Human Type I and III Collagen Molecules from Hydroxylated Recombinant Chains in the Yeast *Pichia pastoris*: EFFECT OF AN ENGINEERED C-TERMINAL OLIGOMERIZATION DOMAIN FOLDON

Outi Pakkanen, Eija-Riitta Hämäläinen, Kari I. Kivirikko and Johanna Myllyharju

*J. Biol. Chem.* 2003, 278:32478-32483.
doi: 10.1074/jbc.M304405200 originally published online June 12, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M304405200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 31 references, 8 of which can be accessed free at http://www.jbc.org/content/278/34/32478.full.html#ref-list-1