Process Integration for Recovery of Recombinant Collagen Type I α1 from Corn Seed

Christopher M. Setina, Jason Haase, Charles E. Glatz

Department of Chemical and Biological Engineering

2114 Sweeney Hall

Iowa State University,

Ames IA 50011

Corresponding Author: Charles Glatz, cglatz@iastate.edu, 515-231-7465

A revised manuscript submitted for publication to Biotechnology Progress

August 7, 2015

ABSTRACT

Because of safety concerns and product consistency issues with the use of animal-derived collagen, several recombinant protein expression hosts have been considered for recombinant collagen corn seed. Full length, triple-helical, recombinant collagen (rCIα1) is expressed as a fusion with a foldon domain, which must later be removed. Here we have examined integration of purification and foldon removal by comparing advantages of removal before or after purification, using salt precipitation as the main purification step. Since expression levels in available maize lines are low, *Pichia*-produced recombinant collagens, both with and without foldon, were added to corn seed germ at the extraction step.

Salt precipitation an acidic corn seed extract yielded 100% of the collagen without foldon at >70% purity without the pepsin pretreatment. With pepsin pretreatment, yield was 94.0% with purity of 76.5%. Analysis of the protein molecular weight distribution of the pre- and post-treatment extracts showed that the corn proteins are largely resistant to pepsin proteolysis, explaining why little benefit was obtained by pepsin treatment. In the absence of pepsin treatment, the recovery of rCIα1 with foldon was still above 90% but the purity was only 44%. This still represented at ca. 13-fold purification with a 2.7-fold volume reduction which would reduce the pepsin requirement for post-recovery foldon cleavage.

KEYWORDS: collagen, corn, *Pichia*, transgenic, purification.

INTRODUCTION

Collagen and gelatin are widely used in the food, cosmetic, and pharmaceutical industries (1). The starting materials for collagen and its derived product gelatin are typically animal byproducts, which introduce product variability and the potential for transmission of infective and allergenic agents (1, 2). The ability to express collagen in the yeast *Pichia pastoris* with tailored degrees of proline hydroxylation demonstrates the potential to make structural proteins with correspondingly-tailored performance properties (3, 4, 5) specific to applications such as drug delivery and tissue engineering (3,6), without the variability and post-processing required of native collagens (6). Genetically modified amino acid sequences giving rise to collagen-like proteins further expands the possible functionalities (6). While recombinant collagen has now been expressed in a variety of hosts, purification of the collagen from some promising hosts has not been fully addressed and will be the focus of this paper.

Type I collagen (CIα1) is an assembly of three helical domains forming the structural heterotrimeric helix containing α1(I) and α2(I) chains (5). The triple helical procollagen molecule contains C- and N- telopeptide domains at the time of assembly, as well as N- and C-terminal peptides (7, 8). These non-helical sequences (i.e. the foldon) orient the chain to form intermolecular hydrogen bonds, stabilizing the helical form (9, 10). Stabilization of collagen at physiologically-relevant temperatures also requires expression of prolyl 4-hydroxylase (P4H) for the regioselective post-translational hydroxylation of proline residues (11, 12). Once synthesized in animal cells, the procollagen is secreted to the intercellular space before proteolytic processing of the C- and N- terminal peptides (13). After expression of recombinant collagen in other hosts, the terminal peptides can be removed by pepsin hydrolysis, which does not cleave triple-helical regions (14, 15, 16).

Formation of recombinant human triple-helical collagen and individual collagen chains as gelatins has been demonstrated in microbial expression systems known for high productivity of heterologous proteins. Several yeast genera including *Pichia pastoris* expressing animal P4H have shown the ability to perform the post-translational modifications and protein folding essential for the formation of the mature collagen molecule with its required physical properties (1, 3). *Pichia* possesses a tightly regulated methanol induced alcohol oxidase (AOX-1) promoter capable of co-expression of P4H with collagen chains in the endoplasmic reticulum where post-translational modifications occur (17, 18). Secretion of collagen chains to the fermentation medium by *Pichia* cells has been reported, while the intracellularly accumulated triple helical procollagen purification requires cell lysis (4, 19, 20). In tobacco, hydroxylation of both proline and lysine has been achieved and the capabilities of other hosts are reviewed by Shoseyov *et al.* (21)

Directed accumulation of recombinant proteins in corn grain provides storage stability (22) with product enrichment possible by targeting to specific tissues within the grain (23,24). Recombinant human Type-I collagen alpha 1 (rCIα1) has been expressed in transgenic corn and characterized as having about 1% hydroxyproline content or one tenth the value of fully prolyl-hydroxylated human CIα1. Further improvements in transgenic lines aimed at bringing accumulation up to commercially acceptable levels (5 g protein/kg grain) and hydroxylation levels up to those of native collagen coupled with the low cost, ease of scale-up and the capability of long term storage of corn seed should make the use of corn expression hosts attractive if low-cost recovery and purification can be achieved (7, 25).

Native corn seed proteins' solubility in aqueous buffers increases above pH 3.0 (23). Precipitation from a corn extract by sodium chloride addition following low pH extraction has

been used to reduce host cell protein (HCP) levels in recovering recombinant aprotinin in the supernatant (26). Collagen is suitable for low pH extraction as its solubility increases below pH 4. Solubility decreases with increases in salt concentrations up to 4% w/v (8, 27). Purifications of low expression level rCIα1 and a smaller fragment from corn extracted in an acidic buffer (pH 1.8) using hydrophobic interaction (HIC), ion exchange and gel filtration chromatography has demonstrated that high purity fractions of full-length collagen or the fragment can be obtained (7, 8).

Acidic extraction, preceded by pepsin digestion to facilitate release from the tissue, has been used to extract collagen from fish skins (27). Pepsin addition, followed by a series of salt precipitations at low and neutral pH were used to purify a recombinant collagen from tobacco (28). Collagen-like proteins were recovered from *E. coli* broths using a combination of pepsin degradation of HCP and acid precipitation (29).

Pepsin addition is also used for recombinant collagen recovery from *Pichia* in order to remove the foldon region and improve ease of recovery by hydrolysis of HCP. Cleaveage also provides the form ready for assembly into fibers (30). Both hydroxylated and non-hydroxylated forms of rCIα1 have proven to be resistant to pepsin (15) when treatment takes place below their triple helix melting temperatures, making pepsin addition an attractive processing option for simplifying purification from host cell proteins.

This work addresses 1) optimal pH for rCIα1 extraction from whole milled corn seed, 2) the sodium chloride concentration providing the best purity and yield for precipitation of rCIα1 from a corn seed matrix, 3) the effect of pepsin digestion of HCP on rCIα1 yield and purity in the precipitate, and 4) the effect of deferring foldon cleavage until after recovery. Because current

transgenic corn lines have low rCIα1 accumulation levels, an experimental model has been developed in which *Pichia*-derived rP-CIα1 has been spiked at 0.25 mg/mL into the extraction buffer, equivalent to complete extraction of a recombinant protein expressed at the level of 1 mg/g of dried solid. This accumulation level has been approached (e.g. porcine α-lactalbumin has been expressed at 0.4 mg/g in transgenic corn seed (31)) and would be a reasonable target for a lower-valued protein. Spiking the extraction buffer with rP-CIα, using the forms with and without the foldon, 1prior to corn extraction was designed to mimic potential interactions within the soluble matrix that corn-expressed rCIα1 would undergo. The corn line chosen, which currently expresses low levels of rCIα1, provides the properties and concentration of native corn components characteristic of a potential transgenic variety. A comparison of this spiking approach with recovery from truly transgenic grain validated this approach for β-glucuronidase (GUS) in canola seed. (32). Some limitations of the model exist: the spiked rP-CIα1 enters the system as part of a fully soluble fraction; hence, interactions within the insoluble matrix (prior to extraction) are not represented in this model. The influences of foldon presence and a corn host presenting a potentially different susceptibility to pepsin cleavage of HCP have not been reported previously.

MATERIALS AND METHODS

Materials

Reagent grade sodium chloride, phosphoric acid, hydrochloric acid, guanidine hydrochloride, trichloroacetic acid, SDS, glycerol, Tris hydrochloride, sodium hydroxide and Thermo–Pierce Bradford Total Protein Assay were obtained from Fisher Scientific (Hampton, NH). Dithiothreitol and iodoacetamide were obtained from BioRad (Hercules, CA).

Freeze-dried *Pichia pastoris* cell mass expressing recombinant human collagen (rP-CIα1) was provided by the ISU Fermentation Lab and used as the source of CIα1 with foldon. Anti-foldon monoclonal antibody (74550) was provided by Fibrogen Inc. (San Francisco, CA). Goat-anti-rabbit IgG (H+L) monoclonal antibody conjugated to alkaline phosphatase was obtained from Invitrogen (Carlsbad, CA). CAPS buffer and FastRed TR/Naphthol AS-MX alkaline phosphatase indicator tablets were obtained from Sigma-Aldrich Chemical (St, Louis MO).

Transgenic whole corn with low-level expression of recombinant human Type I collagen (rCIa1) with 2% hydroxyproline content from a 2004 field trial was provided by ProdiGene, Inc. (College Station, TX) as part of a research collaboration with Fibrogen, Inc. (San Francisco, CA). Purified human collagen standard without foldon (rP-CIα1 at 61% purity and 1.23 mg/mL) was provided by Fibrogen after having been expressed in *Pichia pastoris*, cleaved with pepsin, and purified by acid-salt precipitation (stored at 4^{o}C) (3). The foldon was a protein trimerization peptide obtained from bacteriophage T4 fibritin protein. Purified human collagen without foldon (rP-CIα1) for spiking was produced by ISU Fermentation Lab by the same method. Amino acid analysis performed at Fibrogen identified 33% of the prolines as hydroxylated for a total hydroxylated proline amino acid content of 7.5%. The pepsin used for extract digestion was recombinant human pepsin (19,500 EU/mL, stored at -20^{o}C) provided by Fibrogen.

Cell Rehydration and Washing

Freeze-dried *Pichia* cells were rehydrated by combining approximately 30 grams of dried cells with 300 mL of 0.1 M phosphoric acid. The mixture was stirred for 24 hours at 4^{o}C. Rehydrated cells were centrifuged and washed 5 times with 3 g wash solution/g cells to reduce a

green color that developed upon freeze-drying. Cells were resuspended in the wash liquid and stirred (1 hr, 4°C), then separated from the wash liquid by centrifugation in a Sorvall RC6-Plus centrifuge (10,000 *g*, 15 min, 4°C. Thermo-Sorvall, Asheville NC). The supernatant was decanted and its mass determined. All liquids were kept for analysis. Washes 1-2 were 0.1 M phosphoric acid; Wash 3 was 0.5 M sodium chloride, 0.01 M hydrochloric acid; Washes 4-5 were 0.01 M hydrochloric acid. Removal of a green color impurity was determined by absorbance at 440 nm using a Varian Cary 50 Bio spectrophotometer (Palo Alto, CA).

Cell Disruption

A 15% (w/v) cell weight suspension of washed cells was combined with cold lysis buffer (pH 7.5, 0.1 M Tris HCl, 0.5 M NaCl) with final pH adjustment to 7.4 with 6 M NaOH. Tempered 0.5 mm glass beads (250 g, -20°C) and 100 g of cell suspension were added to a beadmill with ice water in the jacket (BeadBeater, Biospec Products, Bartlesville OK) and milled for 10 min, held for 10 min, and milled an additional 10 min. The lysate was decanted from the beads through a mesh, and the process was repeated until all the cell suspension was disrupted. The beads were washed with 50 mL additional cold lysis buffer after the final milling to remove residual lysate, and the decanted lysates were combined.

Collagen Stock Preparation

The source of the rP-CIα1 with foldon for precipitation studies was the *Pichia* lysate. The disrupted cell suspension was adjusted to pH 7.4 and centrifuged (10,000 *g*, 30 min, 4°C) to remove the solids. The supernatant was adjusted to pH 2 with hydrochloric acid before the rP-CIα1 with foldon was precipitated at 0.5 M NaCl by addition of 5 M NaCl with stirring (1 h, 4°C). The precipitate was recovered by centrifugation (10,000 *g*, 30 min, 4°C,) and was

resolubilized in 0.01 M hydrochloric acid, then centrifuged (29,400 *g*, 4°C, 30 min) to clarify. This resolubilized acid-salt precipitate was called 'RASP'. To determine if any rP-CIα1with foldon remained in the cell debris fraction, it was extracted a second time with 200 mL of 0.1 M phosphoric acid (1 h, 4°C), and centrifuged (10,000 *g* , 30 min, 4°C). The supernatant was decanted for analysis.

To determine if RASP purity could be improved by a second precipitation step, 1 mL of RASP was precipitated in a 1.7 mL microcentrifuge tube at 1.0 M NaCl by addition of 5 M NaCl and stirred (1 hr, 4°C). The precipitate was recovered by centrifugation (Eppendorf Model 5424 microcentrifuge, 10,000 *g*, 30 min, 4°C,) and was resolubilized in 0.01 M hydrochloric acid.

Ultrafiltration of Pichia-derived Collagen

To determine whether the green contaminant, which was evident on rehydration of the freeze-dried *Pichia,* could be removed by ultrafiltration, 10 mL of the rehydrated cell supernatant was loaded into an Amicon Centriplus (#YM-30. Millipore, Billerica MA) 30 kDa Centrifugal Ultrafilter and centrifuged (10,000 *g*, 15 min, 4°C). The color appeared in the permeate.

Thus, ultrafiltration of the RASP was performed using a Millipore TFF system fitted with a 50 cm^2 Pellicon XL filter cartridge (50 kDa MWCO Biomax PES membrane). The 76 mL of RASP were diafiltered at a crossflow rate of 38 mL/min with a transmembrane pressure between 12.5 and 14.5 psi by periodic addition to the retentate of a total 100 ml of 0.01 M hydrochloric acid (4°C), then continuing until the retentate volume was reduced to approximately 25 mL.

Milling and Defatting

Whole transgenic corn seed containing rCIa1 was cracked and milled using a Witt Corrugated Mill with slot setting calibrated for grit particles less than 0.021 inches (Witt

Corrugating Inc, Witchita, KS). The grits were ground using a household coffee grinder (Kitchen-Aid Co., St.Joseph, MO) into a fine powder. The whole ground corn was defatted in a 20% w/v mixture with hexane and mixed on ice for 1 hour. The supernatant was separated by centrifugation (15 min., 2000 *g*, 23ºC) in a Sorvall RC6-Plus centrifuge (Thermo-Sorvall, Asheville NC) before a second hexane extraction was performed, The material was allowed to air dry overnight at room temperature prior to storage at -20ºC.

Preparation of Collagen-Containing Extracts of Milled/Defatted Whole Corn

Simulated extracts of corn-expressed CIα1 with and without foldon were prepared as follows. The milled, defatted whole corn was extracted at 10% w/v at pH 2 (0.10 M phosphoric acid, 0.15 M sodium chloride, and 0.25 mg rP-CIα1 (without foldon) /mL or 0.25 mg ultrafiltered RASP (with foldon) /mL) for one hour at 24ºC, centrifuged (4,600 *g*, 10 min, 4ºC), and decanted. The residual solids were extracted a second time in fresh extractant, centrifuged (15,200 *g*, 15 min, 4ºC) and the supernatant decanted. The extracts from both steps were combined, filtered through a 0.45 µm sterile SFCA syringe filter, (Corning, Corning NY) and stored at 4ºC. For an initial screening of precipitation conditions without pepsin pretreatment, rP-CIα1 w/o foldon was used at a spiking level of 0.50 mg rP-CIα1/mL.

Pepsin Digestion

For pepsin pretreatment of corn host proteins, a 20 mL aliquot of the 0.25 mg/mL collagen-spiked extract was digested with 0.2 mL of thawed pepsin enzyme at 4ºC, giving extract concentrations of 195 EU /mL pepsin and 0.248 mg/mL collagen. For the control, 0.2 mL of unspiked extraction buffer replaced the pepsin. Both samples were incubated overnight at 4ºC (3).

Precipitation

In the initial screening of precipitation conditions, three levels of pH (2, 4, and 6) and NaCl concentrations (0, 0.5, 1.0, 1.5 M) were set by addition of 5 N NaOH and NaCl (5 M) to 1.0 mL of extract in 1.7 mL microcentrifuge tubes. Final collagen levels in the pH 2, 4 and 6 trials following NaCl addition were nominally 0.45, 0.40 and 0.35 mg/mL although the mass balance on precipitate and supernatant after the pH 6 trials showed losses of up to 50% of the collagen. The pepsin-pretreatment comparison followed the same procedure using pH 2 and NaCl concentrations of 0.0, 0.25, 0.5 and 1.0 M. Final collagen levels were 0.195 and 0.198 mg/mL for all pepsin-pretreated and control samples respectively. Additional NaCl concentration levels were included for the "with foldon comparison" to allow for potentially different solubility behavior. All tubes were mixed by end-to-end tumbling on a rotating shaker at 4^{o}C for 1 hr and centrifuged (14,700 *g*, 30 minutes, Eppendorf Model 5424 microcentrifuge). The supernatants were decanted, the pellets were redissolved in their original volume of unspiked extraction buffer, and the solution analyzed for content of collagen and total protein. All precipitations were replicated three times.

Total Protein Microplate Assay

The Bradford Coomassie Total Protein microplate assay from Thermo-Pierce was used to determine total protein concentrations relative to BSA standards.

HPLC-SEC Assay for Collagen Content

All samples were prepared by combining 200 µL of sample with 200 µL of 4M guanidine hydrochloride and analyzed by HPLC-SEC (1200 Series HPLC, Agilent Technologies; BioBasic SEC 300 columns (30x78 mm guard; 300x78 mm separation column); 1.0 mL/min flow rate of

2M guanidine HCl mobile phase , 30°C; elution detection 220 and 280 nm). Peak areas were integrated using the Agilent's ChemStation software after manual identification of peak boundaries. Collagen absorption is much stronger at 220 nm than at 280 nm. Therefore, peaks collected at 220 nm (with neglible 280 nm absorbance) were identified as collagen and peak area was calibrated with injections of the purified rP-CIα1 w/o foldon (33).

Pepsin Digestion for Foldon Removal

To verify the presence of the foldon sequence in *Pichia*-derived rP-CIα1 with foldon, 0.99 mL of ultrafiltered RASP was combined in a 1.7 mL microcentrifuge tube with 0.01 mL of pepsin at 4°C for an effective dose of 195 EU/mL. A 0.99 mL ultrafiltered RASP control was combined with 0.01 mL of 0.10 M phosphoric acid without pepsin. A pepsin control was prepared by combining 0.99 mL of 0.10 M phosphoric acid and 0.01 mL of pepsin. All samples were incubated for 1 hr at 24°C and sampled for analysis by HPLC-SEC from which the distribution of collagen with and without foldon could be determined based on the retention time and relative absorbances at 220 and 280 nm.

To evaluate pepsin's removal of foldon as a processing step, precipitates of whole corn extract containing the ultrafiltered RASP with 1.0 M NaCl were formed with 0.183 mg rP-CIα1 present. These were redissolved in 200 µL of 0.10 M phosphoric acid and 2 µL of pepsin (195 EU/mL final concentration) and incubated for 1 hr at 24°C. After pepsin treatment, 50 µL of 5 M NaCl was added, the tubes mixed on a rotating shaker at 4°C for 1 hr, then centrifuged (14,700 *g*, 30 min, 4°C. Eppendorf Model 5424 microcentrifuge). The supernatants were decanted and the pellets were redissolved in unspiked extraction buffer followed by HPLC-SEC.

SDS-PAGE

Samples were prepared by 1:2 dilution in a preparatory aqueous sample buffer (pH ~ 8.3) containing 2% SDS, 0.5% bromophenol blue, 10% glycerol, 5% β-mercaptoethanol, and 0.0625M tris HCl (Fisher Scientific) in 1.7 mL microcentrifuge tubes. Each sample was vortexed and boiled for 30 seconds prior to loading onto a precast tris HCl 7.5% or 4-15% gels (BioRad Inc.). The gel was run at 200V using a cold running buffer of 1.5% tris base, 7.2% glycine, 0.5% SDS in deionized water. Gels were fixed for 30 min in 40% (v/v) ethanol and 10% acetic acid, then stained for one hour in 1 g/L Coomassie Blue Stain G-250 in 40% (v/v) methanol, 10% (v/v) acetic acid in deionized water. The gels were destained in 40% (v/v) methanol, 10% (v/v) acetic acid in deionized water. Molecular weight standards ranging from 10 to 250 kD (Precision Plus Protein Standard #161-0363, BioRad) were run in an adjacent lane.

Western Blotting

Samples for Western Blot identification of rP-CIα1 with foldon sequence were prepared using SDS-PAGE. A pre-stained broad range standard (# 161-0373, BioRad) was used to assess transfer and position of protein during the blotting procedure. Prior to fixing the gel after SDS-PAGE, the gel was applied to a PVDF membrane and the proteins were transferred (15 V, 30 min, 24ºC) in 0.01 mM CAPS, 5 % (v/v) methanol, 0.05 % SDS buffer, using a BioRad Trans-Blot Semi-Dry Electrophoretic Transfer Cell. The blotted gel was recovered from the blotting procedure and stained using SDS-PAGE staining technique to determine the level of protein transfer. The PVDF membrane was blocked overnight at 4ºC in a blocking buffer containing phosphate buffered saline (PBS) and 0.1% Tween-20, with 2% dry milk solids added. The membrane was rinsed three times with washing buffer containing PBS and 0.05% Tween-20,

before incubating (1 h, 24°C) on a rocking shaker with 15 mL total volume of a rabbit anti-foldon primary antibody (# 74550, diluted 1:5,000 in blocking buffer).

A 15-mL volume of goat/anti-rabbit IgG (H+L) secondary antibody with an alkaline phosphatase conjugate (diluted 1:10,000 in blocking buffer) was added after the membrane was rinsed three times with washing buffer. The secondary antibody was incubated (1 h, 24°C) on a rocking shaker, then the membrane was rinsed three times in washing buffer before adding 10 mL of FastRed alkaline phosphatase indicator until a red color developed.

Two-Dimensional Gel Electrophoresis

Samples of ultrafiltered RASP were diluted to a concentration of 200 µg/mL of total protein in Destreak Rehydration Buffer (GE Biosciences, Uppsala Sweden) and 5 µL of Bio-Lyte 3/10 Ampholyte solution (BioRad, Hercules CA) were added for a total volume of 200 µL. BioRad 3/10 11 cm gel strips were hydrated with the protein solution then focused for 16 hrs using a Amersham Pharmacia Biosciences Ettan IPGphor11 (Piscataway NJ) isoelectric focusing system using the following program: 1 hr at 250 V; 3 hrs ramp from 250 to 8000 V; 1 hr at 8000 V; 16 hrs at 50 V. Each strip was removed and equilibrated for 15 min in 5 mL of Equilibration Buffer (0.375 M Tris, 2 % SDS and 2% glycerol) and 2 % dithiothreitol. They were individually rinsed with Equilibration Buffer and soaked for 15 minutes in 5 mL per strip of Equilibration Buffer and 2.5 % iodoacetamide. The strips were loaded into BioRad Criterion 4-10% Tris HCl gels with 0.5% BioRad overlay agarose and separated in the second dimension at 200V for 1 hour and stained using the identical staining procedure used for SDS-PAGE above. pI's were calculated from position on the 2-D gels, assuming a linear gradient between pH 3.0 and

10.0.Estimated pI's of the proteins of interest were calculated from amino acid sequence using Protcalc3 Protein Calculator v3.3 (CDPutnam, Scripps Research Institute, LaJolla CA).

Statistical Analysis

Two full-factorial ANOVA statistical analyses were performed using JMP Software (SAS). The variables tested in the first were pH, sodium chloride concentration and digestion for their effect on purity, yield and purification factor. The error estimate for a 95% confidence interval was based on pooled standard deviation values with df = 4. Statistically significant groupings from ANOVA (Tukey's least square mean difference test) are identified by lettering in all Figures below. In the second design, the same parameters were used for testing the variables of NaCl concentration and foldon presence.

Calculations

From the collagen and total protein concentrations of the initial extracts and redissolved precipitates yield, Y, purity, P, and purification factor, PF, were calculated as follows:

$$Y = \frac{m_i}{m_f}(100\%) \qquad \text{(Eqn. 1)}$$

$$P = \frac{C_{collagen}}{C_{total\,protein}}\ (100\%) \qquad \text{(Eqn. 2)}$$

$$PF = \frac{P_i}{P_f} \qquad \text{(Eqn. 3)}$$

where m is mass, C is concentration, and i and f refer to initial and final values, respectively.

RESULTS AND DISCUSSION

Collagen Assay

Collagen can be seen (Figure 1) to elute separately from the HCPs in HPLC-SEC. Preparation and separation in guanidine hydrochloride dissociates the collagen molecule into monomers of approximately 90 kDa - still significantly larger in molecular weight than the HCPs. Thus, collagen recovery and purity can be determined from the separate chromatographic areas corresponding to collagen and HCP (15). The calculated concentration from the SEC peak area matched the measured addition to the extract within 2%.

Screening of rCIa1 w/o foldon Precipitation Conditions without Pepsin Pretreatment

Figure 2 illustrates the role of pH and salt concentration on collagen precipitation. At pH 6, little collagen was recovered. At pH 2 and 4, yields and purities were comparable but achieved at lower salt concentrations for pH 4. Decreased purification at the highest salt concentration is the combined result of lower collagen yield and higher HCP co-precipitation (confirmed qualitatively by SDS-PAGE, not shown). HCP present in precipitates from pH 2 and pH 4 samples were similar at the same NaCl levels, but increased with NaCl addition: 0.08, 0.13 and 0.17 mg/mL at 0.5, 1.0 and 1.5 M NaCl, respectively.

The best of the tested conditions for precipitation of rP-CIα1 from the corn extract was pH 4 and 0.5 M NaCl. Purification factors of 1.41$\pm$0.13 were achieved for pH 2 (1.0 M NaCl) and pH 4 (0.5 and 1.0 M NaCl). ANOVA analysis showed that both pH and NaCl were significant factors affecting the purity, yield and purification factors. However, pH 2 and 1.0 M NaCl were not statistically different from precipitations performed at pH 4 and 0.5 M and 1.0 M

NaCl as indicated by their common lettering (Figure 2). Therefore, pH 2 was selected for further trials as it is the condition used for extraction and thus does not require a pH adjustment step.

rCIa1 w/o foldon Precipitation Conditions with Pepsin Pretreatment

The possibility of improved purification using pepsin digestion of HCP in the extract was based on the hypothesis that pepsin digestion of susceptible HCP would increase their solubility, thus avoiding co-precipitation of HCP with collagen. Both total protein and collagen concentrations of the treated and untreated extracts were similar (0.198 and 0.195 mg collagen/mL and 1.58 mg/mL and 1.43 mg total protein /mL for treated and untreated, respectively). The lack of change in collagen content showed its expected resistance to pepsin hydrolysis. Purification factors of 5.39$\pm$1.53 for pepsin pretreated samples with 0.75 and 1.0 M NaCl, and slightly higher values of 5.53$\pm$0.57 for no pretreatment (at 0.75 and 1.0 M NaCl), were higher than the purification factors seen in the precipitation screening trials as a result of the higher initial collagen concentration and lower initial HCP content in the latter extract.

Comparison of Figures 3 (without pepsin pretreatment) and 4 (with pepsin) and the ANOVA analysis showed that the pretreatment did not improve purity or yield. Salt concentration remained the significant factor affecting precipitation. The 0.75 M NaCl level, not tested in the early set of experiments, proved optimal, giving a yield and purity of 94% and 76.5%, respectively.

The similar outcomes after pretreatment reflect the limited digestion of HCP. Evidence of the extent of hydrolysis is seen by SDS-PAGE (Figure 5), as well as a minor shift in area detected by HPLC-SEC (Figure 6).

That the precipitate at low salt concentrations is mainly HCPs suggested that a two-step addition of salt could result in an initial removal of the least soluble HCP, followed by an enriched product cut in the second step. Therefore, a precipitation at 0.25 M NaCl was followed by precipitation from the 0.25 M supernatant at 1.0 M NaCl. From an initial HCP content in the original extract of 1.22 mg/mL, HCP's were reduced to 0.0207 mg/mL in the 0.25 – 1.0 M NaCl fraction, in contrast to 0.056 mg/mL for precipitation directly at 1.0 M NaCl. Although purity increased to 95%, collagen yield dropped significantly to 89%. Of the HCP in the fractional cut, a wash step was able to remove 30%, raising the purity to 96.6% but reducing yield to 86%.

Recovery and Purification of rP-CIa1 with Foldon from Disrupted Pichia Cells

From the pH 7.4 extraction of 30 g of freeze-dried cells, a total of 89.9 mg of rP-CIa1 with and without foldon were recovered in the RASP (Table 1). An additional 75.4 mg of rP-CIa1 with and without foldon was lost to rehydration and wash liquids.

Subsequent ultrafiltration of the RASP using a Millipore TFF system and Pellicon 50 kDa MWCO membrane provided a purity of 69.6 % at a concentration of 0.61 mg total collagen/mL.

Identification of rP-CIa1 with Foldon via HPLC-SEC

Each of the recovered fractions was analyzed for rCIa1and rCIa1with foldon at 220 nm and for host proteins at 280 nm. With rP-CIa1with foldon showing absorbance at both 220 and 280 nm, the 220 nm absorbance area was used to determine mass, and the 280 nm signature was used to determine identity. The CIa1-containing region of rP-CIa1with foldon will be identical to that of the standard; therefore, the absorbance relationship to mass was assumed to be the same. The foldon peptide region contains 17% aromatic amino acid content, and is expected to

display significantly higher absorbance at 280 nm than rP-CIa1 without foldon, which contains 2.9%.

The RASP retentate from ultrafiltration showed two high molecular weight A_{220} peaks (Figure 7): one at 7 min, matching the elution time of the rP-CIa1 standard; and a larger peak at 6 min, which would have a higher molecular weight than the rP-CIa1. This peak at 6 min also had a corresponding A_{280} peak, thus both size and absorbance indicate this is rP-CIa1with foldon. Based on this assay, purity of the ultrafiltered RASP was 69.6% with 58.4 % of the collagen content having the foldon attached. No absorbance at 440 nm indicated removal of a green contaminant from this material.

Additional Confirmation of rP-CIa1with Foldon

Several experiments provided additional confirmation of the 7 min elution peak as being the rP-CIa1with foldon was obtained by pepsin digestion of the ultrafiltered RASP. Pepsin would cleave the foldon, but leave the rP-CIa1 intact. The addition of pepsin decreased rP-CIa1with foldon content by approximately 0.0318 mg with a corresponding increase in rP-CIa1 content of 0.0229 mg (Table 2). Other peaks present before digestion showed little change after digestion, but there was some increase in A_{220} beyond 10 min consistent with the presence of some degradation products.

Direct confirmation of the identity of rP-CIα1with foldon was made by Western Blotting. The ultrafiltered RASP material was separated by SDS-PAGE and compared to the purified rP-CIα1 (without foldon) reference standard, rP-CIα1 (without foldon) spiking material, and prestained molecular weight standards for transfer to the PVDF membrane. The presence of the foldon sequence was identified only in the ultrafiltered RASP at the approximate molecular

weight of 120 kDa corresponding to rP-CIα1 (Figure 8). The additional molecular weight of the foldon peptide could not be discriminated from non-foldon containing rP-CIα1 via SDS-PAGE during these trials. The rP-CIα1 high purity standard and rP-CIα1 used for comparison had been pretreated with pepsin to remove foldon, and foldon was not detected by Western blotting.

The full amino acid sequence of rCIα1with foldon expressed in corn has been reported including the helical regions, the C- and N-terminal telopeptides and the foldon domains (7). The pIs for those sequences, with and without foldon, were estimated using ProtCalc3 v3.3 online protein calculator to be pH 9.3 (with foldon) and pH 9.4 (without foldon). A 2-D separation using isoelectric focusing and SDS-PAGE of the ultrafiltered RASP produced two approximate pI values at the expected molecular weight of 120 kDa: pH 9.1 and pH 7.3. The difference in pI of those spots cannot be accounted for by either removal of the foldon or removal of the telopeptides. One would have to suspect some other post-translational modification. Confirmation of which spot includes the foldon was not pursued in this work.

Precipitation of rP-CIa1 with and without Foldon from Corn Extracts

Separate whole corn extracts were prepared containing rP-CIa1 either with or without foldon at levels of 0.25 mg/mL. HPLC-SEC analysis of the initial spiked extracts confirmed rP-CIa1 concentrations to be 0.247 mg/mL and 0.252 mg/mL, respectively. Average concentrations after addition of sodium chloride were 0.195 mg total collagens/mL.

Whole corn extracts spiked with rP-CIα1 showed little precipitation at 0 and 0.25 M sodium chloride concentrations. Yield quickly increased to >98 % as the sodium chloride concentration increased to 0.75 and 1.0 M. Whole corn extracts spiked with ultrafiltered RASP behaved very differently. Despite being clear and soluble at the time of sodium chloride

addition, approximately 60% of the collagen present formed a pellet during centrifugation after 0, 0.25 and 0.5 M sodium chloride was added. Yield improved to 91% in the presence of 0.75 M sodium chloride, but did not significantly increase with higher sodium chloride concentrations as indicated by common lettering (Figure 9). Both foldon presence and NaCl concentration influences were judged significant in the ANOVA analysis.

When precipitated with sodium chloride, purity of the ultrafiltered RASP samples started at approximately 22%, but did not increase to the same purity as the rP-CIα1 spiked samples (Figure 10). Increasing the sodium chloride concentration above 1.0 M did not improve purity of the ultrafiltered RASP samples. Both foldon presence and NaCl concentration were statistically significant in affecting purity, though foldon presence had less effect than NaCl concentration.

The results indicate that a high yield of rP-CIα1 with foldon in corn could be achieved by precipitation with sodium chloride concentrations above 0.75 M. However, the presence of the intact foldon sequence decreased the purity of the precipitate relative to pepsin-treated rP-CIα1. Below 0.75 M NaCl, 68.6% of the precipitated rP-CIα1 contained foldon --- a higher percentage than the ultrafiltered RASP (58.4%). At 0.75 and 1.0 M NaCl, the average percent of rP-CIα1 with foldon decreased to 59.5% as in excess of 90% of all rP-CIα1 was precipitated under those conditions.

Pepsin Digestion of Sodium Chloride Precipitated rP-CIα1 with Foldon

59.9% of the rP-CIα1 in the precipitate after 1.0 M sodium chloride addition remained with foldon. After pepsin addition, a second precipitation at 1.0 M sodium chloride was carried out to remove the pepsin. Calculated on a basis of the initial spiked extract, yield decreased from 94% (for the initial precipitation at 1.0 M sodium chloride) to 75.7%. Purity changed little

(44% to 43%), and rP-CIα1 with foldon content was reduced from 59.9 to 31.8 % of total collagen. The remaining rP-CIα1 with foldon following pepsin treatment precipitated preferentially. A longer period of enzyme action may be required for complete hydrolysis.

The benefit of maintaining the foldon sequence during precipitation may come in the form of volume reduction. The second precipitation with 1.0 M NaCl to remove the pepsin increased the collagen concentration from 0.25 mg/mL in the original extract to 0.669 mg/mL following resuspension in 0.01 M hydrochloric acid. This represents a volume reduction of 2.7-fold, along with a purification factor of ca. 13. A smaller pepsin addition to cleave the foldon sequences would be required for this volume. Additionally, supernatant concentrations from NaCl precipitations >0.75 M ranged between 0.04 and 0.07 mg/mL rP-CIα1 (with or without foldon). If this represents the solubility of rP-CIα1 at high NaCl concentrations, a starting concentration of >0.95 mg/mL would be required for >90% purity given the same host cell protein concentration.

CONCLUSIONS

- High recovery and purities of CIα1 without foldon from corn extracts were achieved at the extraction pH with optimal choice of salt concentration. A single precipitation gave complete recovery and greater than 70% purity. Fractional precipitation provided 95% purity with yield lowered to 89%.
- In contrast to other plant hosts, epsin pretreatment of the extract provided modest increase in purity at the cost of 6% yield loss and additional process complexity. The small impact of the pepsin pretreatment reflected the relatively small extent of hydrolysis of corn HCP.

- Delaying pepsin treatment until after precipitation of the foldon-intact form resulted in lower purity but still high yield showing that recovery of this form is feasible from the corn host. The resulting volume reduction would reduce pepsin requirements in a subsequent foldon cleavage step.

ACKNOWLEDGMENTS

The authors thank Dr Julio Baez and Sheri Almeda (Fibrogen, Inc., San Francisco, CA) for providing materials and helpful suggestions, Cheng Zhang for preparing defatted corn flour, Aisa Muhammud Pi-Remli for assisting in the lab, Xing Xu for preparing the gels in Figure 7, and Bill Colonna and John Strohl for providing *Pichia*-derived collagen. This work was supported by funding from the USDA CREES Grant Numbers. 2006-34496-17122, 2008-34496-19348, and 2009-34496-19899

REFERENCES

1. Olsen D, Yang C, Bodo M, Chang R, Leigh S, Baez J, Carmichael D, Perälä M, Hämäläinen E, Jarvinen M, Polarek J. Recombinant collagen and gelatin for drug delivery. *Advanced Drug Delivery Reviews*. 2003, 55:1547-1767.
2. Asghar A, Henrickson RL. Chemical, biochemical, functional and nutritional characteristics of collagen in food systems. *Adv. Food Res*. 1982, 28:231-372.
3. Baez J, Olsen D, Polarek JW. Recombinant microbial systems for the production of human collagen and gelatin. *App.. Microbiol. Biotechnol*. 2005, 69:245-252.
4. Werten MWT, Van Den Bosch TJ, Wind RD, Mooibroek H, De Wolf FA. High-yield secretion of recombinant gelatins by *Pichia pastoris*. *Yeast.* 1999, 15:1087-1096.

5. Perrett S, Merle C, Bernocco S, Berland P, Garrone R, Hulmes DJS, Theisen M, Ruggiero F. Unhydrolxylated triple helical collagen I produced in transgenic plants provides new clues on the role of hydroxyproline in collagen folding and fibril formation. *J. Biol. Chem*. 2001, 276:43693-32698.

6. Brodsky, B, Kaplan, D L. Shining light on collagen: Expressing collagen in plants. *Tissue Engineering Part A*. 2013, 19:1499-1501

7. Zhang C, Baez J, Pappu K, Glatz CE. Purification and characterization of a transgenic corn grain-derived recombinant collagen type I alpha 1. *Biotechnol. Prog*. 2009; 25:1660-1668.

8. Zhang C, Baez J, Glatz CE. Purification and characterization of a 44-kDa recombinant collagen I α 1 fragment from corn grain. *J. Agric. Food Chem*. 2009, 57:880-887.

9. Berisio R, De Simone A, Ruggiero A, Improta R, Vitagliano L. Role of side chains in collagen triple helix stabilization and partner recognition. *J. Peptide Sci*. 2009, 15:131-140.

10. Kar K, Amin P, Bryan M, Persikov A, Mohs A, Wang Y, Brodsky B. Self-association of collagen triple helix peptides into higher order structures. *J. Biol. Chem*. 2006, 281:33283-33290.

11. Geerlings T, De Boer A, Lunenborg M, Veenhuis M, Van der Klei I. A novel platform for the production of nonhydroxylated gelatins based on the methylotrophic yeast *Hansenula polymorpha*. *FEMS Yeast Res*. 2007, 7:1188-1196.

12. Kivirikko K, Myllylhapju J. Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biol*. 1998, 16:357-368.

13. Ricard-Blum S, Ruggiero, F. The collagen superfamily: from the extracellular matrix to the cell membrane. *Pathologie Biologie*. 2005, 52:430-442.

14. Olsen DR, Leigh SD, Chang R, McMullin H, Ong W, Tai E, Chisholm G, Birk DE, Berg RA, Hitzeman RA, Toman PD. Production of human type I collagen in yeast reveals unexpected new insights into the molecular assembly of collagen trimers. *J. Biol. Chem*. 2001, 276:24038-24043.

15. Pakkanen O, Hämäläinen ER, Kivirikko KI, Myllyharju J. 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. *J. Biol. Chem*. 2003, 278:32478-32483.

16. Nokelainen M, Tu HM, Vuorela A, Notbohm H, Kivirikko KI, Myllyharju J. High-level production of human type I collagen in the yeast *Pichia pastoris*. *Yeast*. 2001, 18:797-806.

17. De Bruin E, Werten M, Laane C, De Wolf F. Endogenous prolyl 4-hydroxylation in *Hansenula polymorpha* and its use for the production of hydroxylated recombinant gelatin. *FEMS Yeast Res*. 2002, 1:291-298.

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

19. Myllyharju J, Nokelainen M, Vuorela A, Kivirikko K. Expression of recombinant human collagens in the yeast *Pichia pastoris*. *Biochem. Soc. Trans*. 2000, 28;353-357.

20. Werten M, Wisselink W, van den Bosch T, De Bruin E, De Wolf F. Secreted production of a custom-designed, highly hydrophilic gelatin in *Pichia pastoris*. *Protein Eng*. 2001, 14:447-454.

21. Shoseyov O, Posen Y, Grynspan, F. Human recombinant Type I Collagen produced in plants. *Tissue Engineering Part A.* 2013**,** 19:1527-1533

22. Stoger E, Ma JKC, Fischer R, Christou P. Sowing the seeds of success: Pharmaceutical proteins from plants. *Current Opinion in Biotechnology*. 2005, 16:167-173.

23. Azzoni AR, Kusnadi AR, Miranda, EA, Nikolov ZL. Recombinant aprotinin produced in transgenic corn seed: Extraction and purification studies. *Biotechnol. Bioeng*. 2002, 80:268-276.

24. Zhang C, Fox S, Johnson L, Glatz C. Fractionation of transgenic corn seed by dry and wet milling to recover recombinant collagen-related proteins. *Biotechnol Prog.* 2009, 25:1396-1401.

25. Menkhaus T, Bai Y, Zhang C, Nikolov Z, Glatz C. Considerations for the recovery of recombinant proteins from plants. *Biotechnol. Prog*. 2004; 20:1001-1014.

26. Zhong Q, Xu L, Zhang C, Glatz, CE. Purification of recombinant aprotinin from transgenic corn germ fraction using ion exchange and hydrophobic interaction chromatography. *Appl. Microbiol. Biotechnol*. 2007, 76:607-613.

27. Woo J, Yu S, Cho S, Lee Y, Kim S. Extraction optimization and properties of collagen from yellowfin tuna (*Thunnus albacores*) dorsal skin. *Food Hydrocolloids*. 2008, 22:879-887.

28. Stein H, Wilensky M, Tsafrir Y, Rosenthal M, Amir R, Avraham T, Ofir K, Dgany O, Yayon A, and Shoseyov O. Production of bioactive, post-translationally modified,

heterotrimeric, human recombinant Type-I Collagen in transgenic tobacco. *Biomacromolecules*. 2009, 10:2640-2645

29. Peng YY, Stoichevska V, Madsen S, Howell L, Dumsday GJ, Werkmeister JA, Ramshaw JAM. A simple cost-effective methodology for large-scale purification of recombinant non-animal collagens. Appl. Microbiol. Biotechnol. 2014, 98:1807-1815
30. Ruggiero F, Exposito JY, Bournat P, Gruber V, Perret S, Comte J, Olagnier B, Garrone R, Theisen M. Triple helix assembly and processing of human collagen produced in transgenic tobacco plants. *FEBS Letters***.** 2000, 469:132-136.
31. Yang SH, Moran DL, Jia HW, Bicar EH, Lee M, Scott MP. Expression of a synthetic porcine α-lactalbumin gene in the kernels of transgenic maize. *Transgenic Res.* 2002, 11:11-20.
32. Zhang CM, Love RT, Jilka JM, Glatz CE. Genetic engineering strategies for purification of recombinant proteins from Canola by anion exchange chromatography: An example of β-glucuronidase. *Biotechnol. Prog*. 2001, 17:161-167.
33. Aspelund MA, Glatz CE. Purification of recombinant plant-made proteins from corn extracts by ultrafiltration. *J. Mem. Sci.* 2010**,** 353:103–110

Table 1. Content of rP-CIa1 with and without foldon and host cell protein content of the redissolved acid-salt precipitate (RASP) and the initial supernatant from *Pichia* cell lysate

Sample	rP-CIa1 w/ foldon (mg)	rP-CIa1 w/o foldon (mg)	Host Cell Protein (mg)	Total Protein (mg)	Purity[a], (%)	% rP-CIa1 w/ foldon of combined collagen
Supernatant	9.95	11.32	19.97	41.24	51.6	46.8 %
RASP	53.49	36.35	231.81	278.69	16.8	59.5 %

[a]Total rP-CIa1 as percent of Total Protein.

Table 2. Effect of pepsin on distribution of rP-CIα1with foldon, rP-CIα1, and other protein in RASP material

Sample	rP-CIa1 w/ foldon (mg)	rP-CIa1 (mg)	Other Protein (mg)
RASP	0.044207	0.067127	1.8746
Pepsin (195EU/mL)	------	------	0.5383
RASP + Pepsin (195 EU/mL)	0.012425	0.090011	2.4392

FIGURES

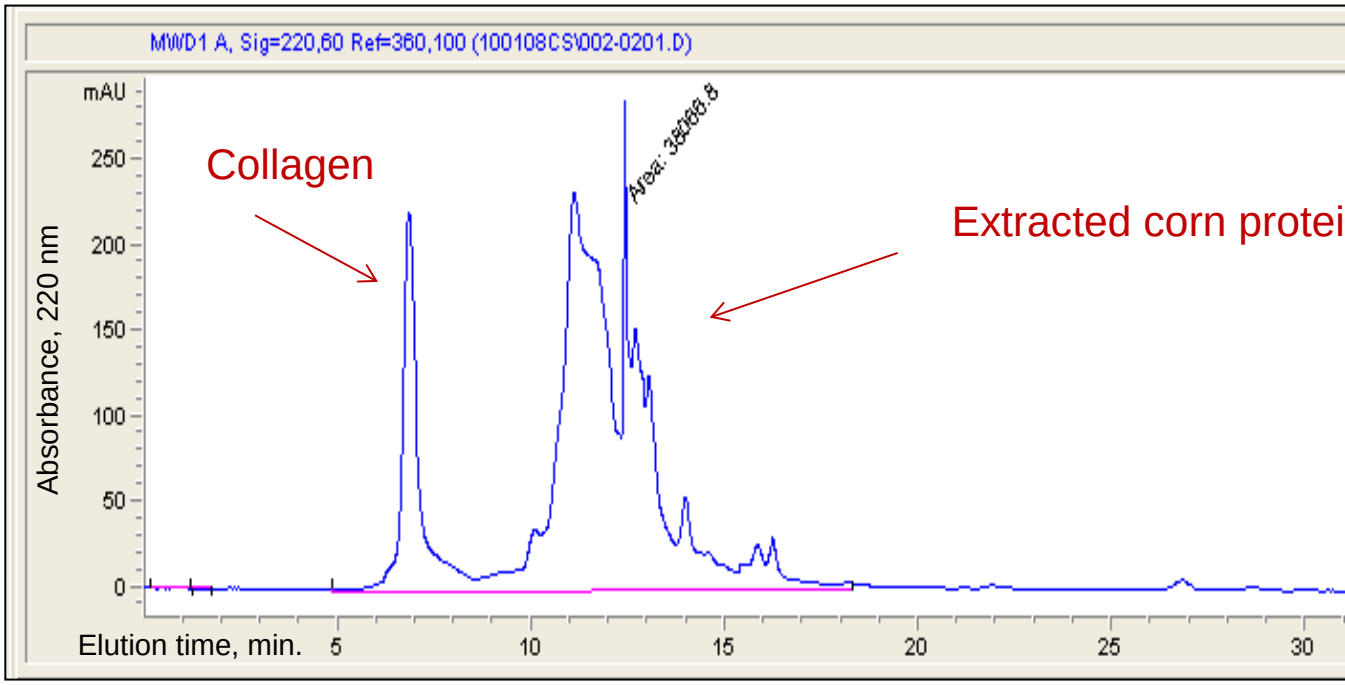

Figure 1: HPLC-SEC chromatogram of extract containing rP-CIα1collagen and corn proteins showing that the larger collagen molecule (~90 kDa) elutes quantitatively ahead of the smaller host proteins.

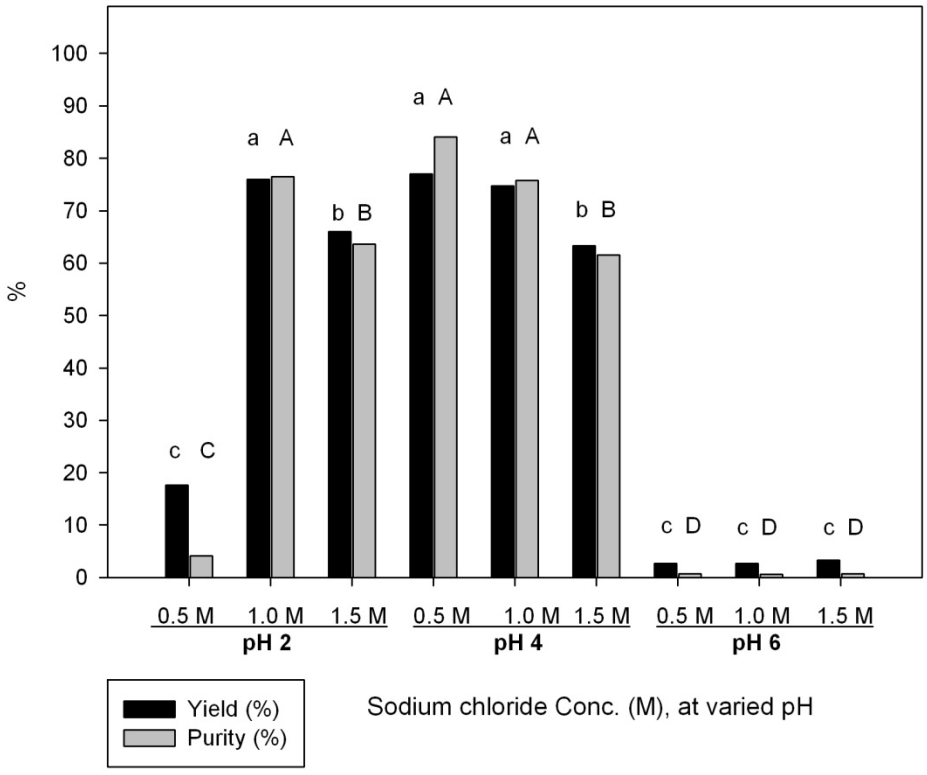

Figure 2. Reconstituted rP-CIα1 (w/o foldon) precipitate purity and yield over the range of precipitation conditions**.** Statistically significant (95% confidence level) groupings from ANOVA analysis for purity are given in upper case, yield groupings are in lower case.

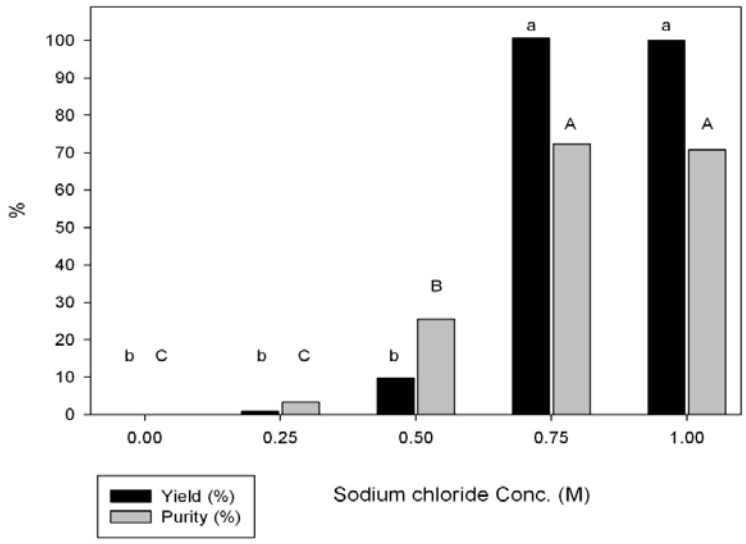

Figure 3. Reconstituted rP-CIα1 (w/o foldon) precipitate purity and yield after precipitation from corn extract with sodium chloride at pH 2 without pepsin pretreatment**.** Statistically significant groupings from ANOVA analysis for purity are given in upper case, yield groupings are in lower case.

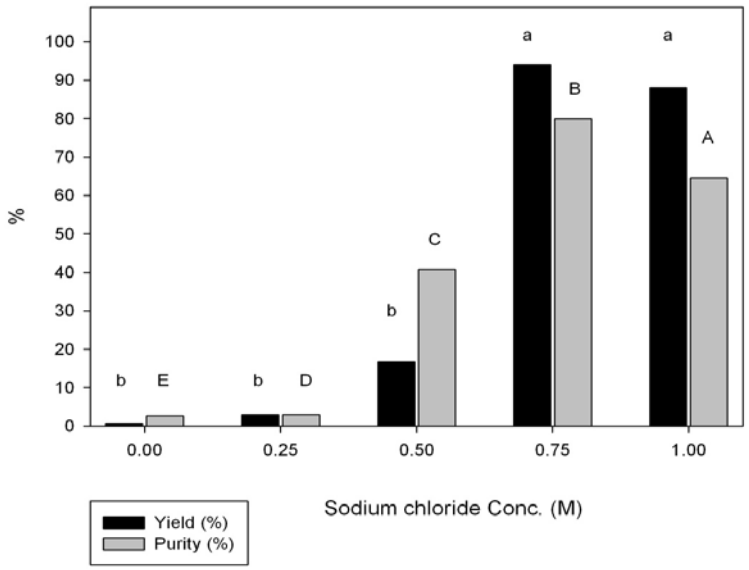

Figure 4. Reconstituted rP-CIα1 (w/o foldon) precipitate purity and yield after precipitation from corn extract with sodium chloride at pH 2 with pepsin pretreatment**.** Statistically significant groupings from ANOVA analysis for purity are given in upper case, yield groupings are in lower case.

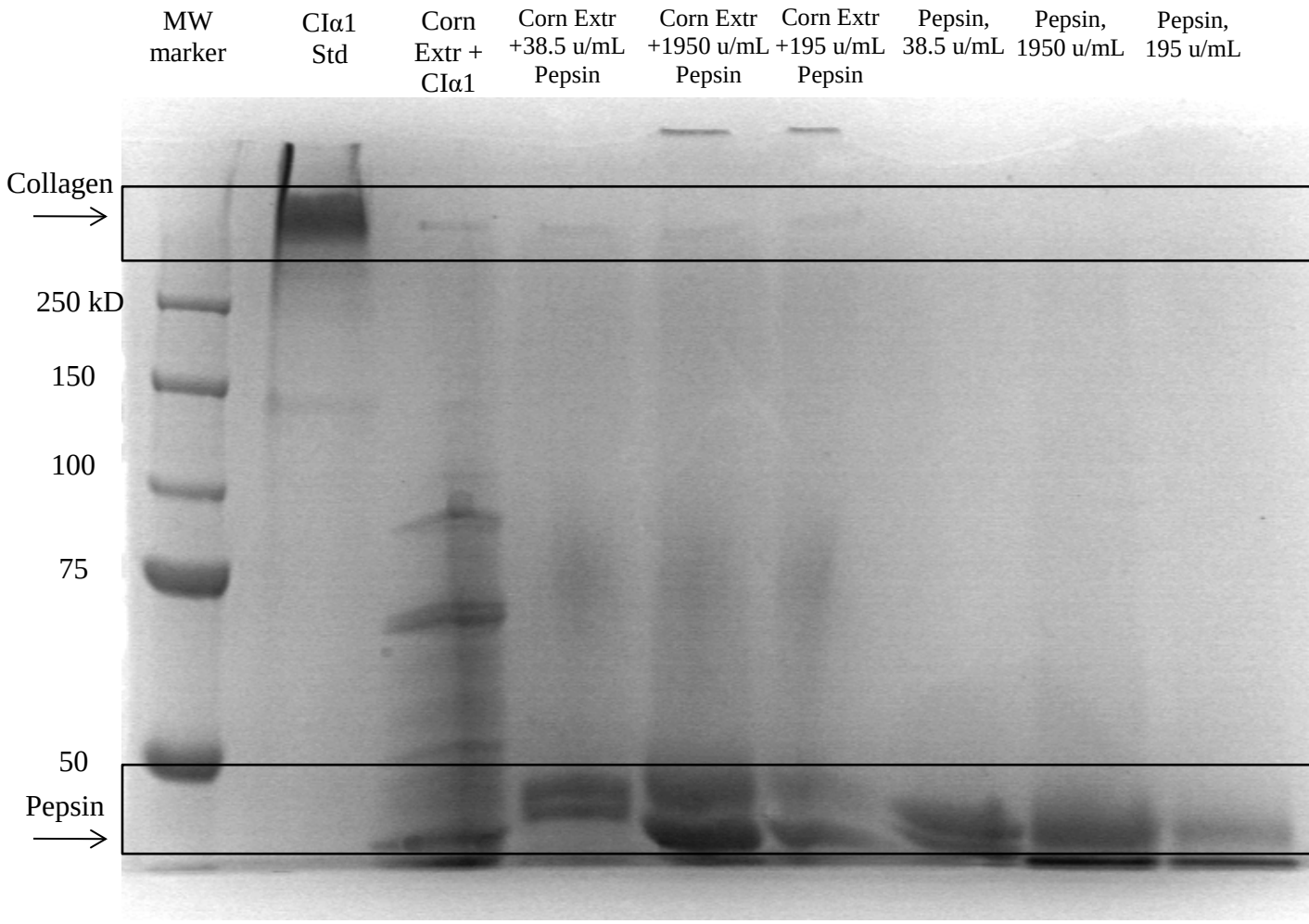

Figure 5. 2D-EP analysis of pepsin digests corn extracts spiked with rP-CIα1 (w/o foldon) at pepsin concentrations of 38.5, 195, 1950 EU/mL. Tris HCl 7.5% gel stained with Coomassie Blue G-250 shows pepsin degradation of HCP in the 50-100 kDa MW range. The same samples run on 4-15% tris HCl gradient gels show no detectable digestion of HCP.

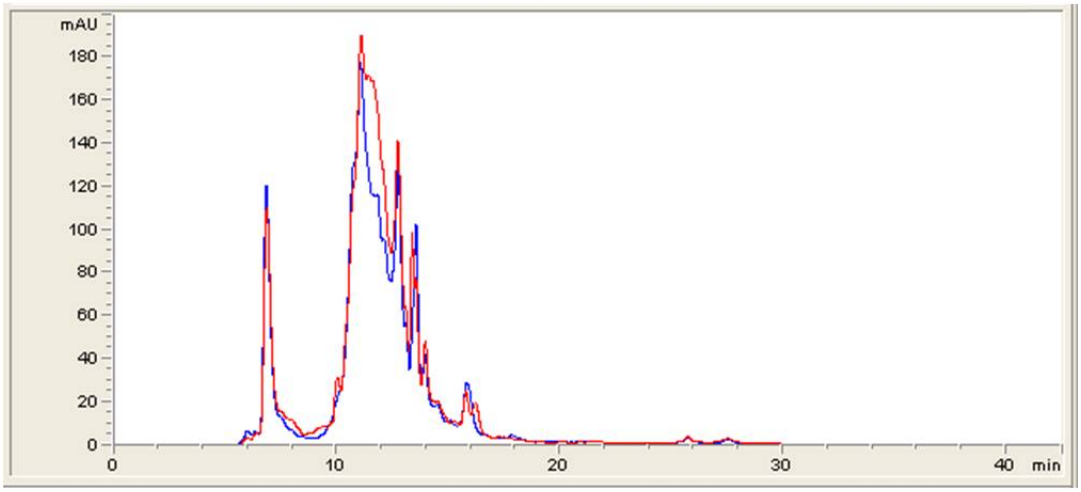

Figure 6. HPLC-SEC chromatogram of rP-CIα1 (w/o foldon) - spiked corn extract, with pepsin (red) and without pepsin (blue) digestion, showing the shift in the chromatographic area resulting from the addition of pepsin, with pepsin co-eluting among the corn proteins at 10.8 minutes, but otherwise little change in HCP profile

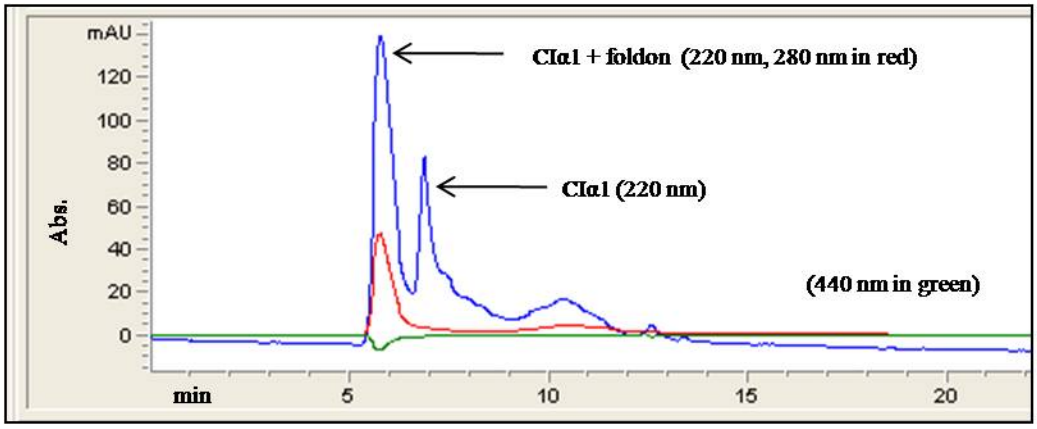

Figure 7. HPLC-SEC of ultrafiltered RASP showing peaks of rP-CIa1with and without foldon and absence of green contaminant.

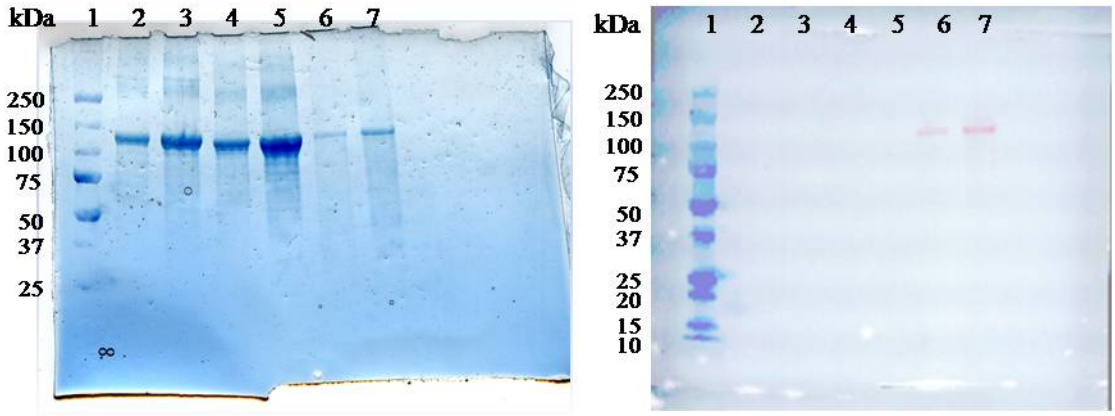

Figure 8: SDS-PAGE gel stained with Coomassie blue G-250 stain (left) and corresponding Western Blot PVDF membrane (right). 1) pre-stained MW Stds; 2) 10 µL high purity rP-CIα1 Std (1.2mg/mL); 3) 20 µL high purity rP-CIα1 Std (1.2 mg/mL); 4) 10 µL rP-CIα1 (2.43 mg/mL); 5) 20 µL rP-CIα1 (2.43 mg/mL); 6) 10 µL ultrafiltered RASP; 7) 20 µL ultrafiltered RASP. Red in Western blot indicates presence of rP-CIα1 with foldon in RASP samples.

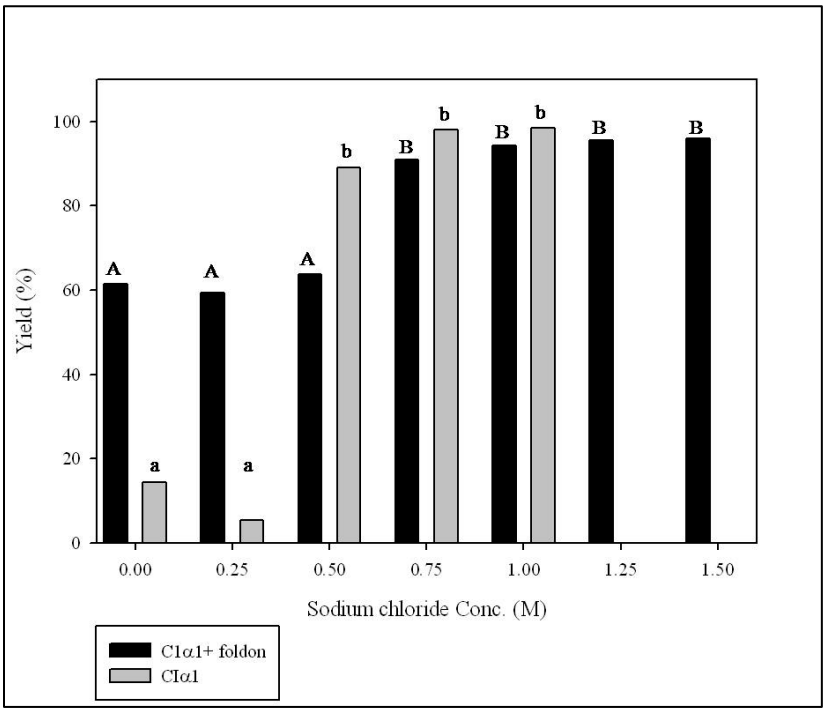

Figure 9: Yield after precipitation of whole corn extracts with sodium chloride when spiked with rP-CIα1 with foldon (ultrafiltered RASP) and rP-CIα1 w/o foldon . Statistically significant groupings for yield are indicated by letters.

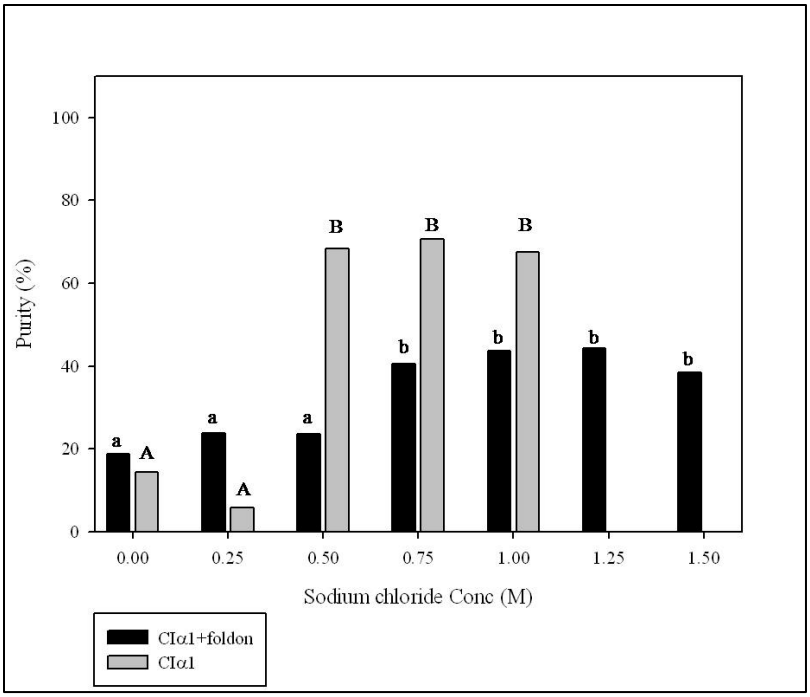

Figure 10: Purity after precipitation of whole corn extracts with sodium chloride when spiked with rP-CIα1 with foldon (ultrafiltered RASP) and rP-CIα1 w/o foldon. Statistically significant groupings from ANOVA analysis for purity with foldon are given in upper case, without foldon groupings are in lower case.