2010

Alternative routes for purification of collagen type I α1 from a whole corn extract

Christopher Michael Setina
Iowa State University

Alternative routes for purification of collagen type I α1 from a whole corn extract

by

Christopher M. Setina

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Chemical Engineering

Program of Study Committee:

Charles E. Glatz (Major Professor)

Laura Jarboe

Marvin Paul Scott

Iowa State University

Ames, Iowa

2010

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

ABSTRACT

Transgenic corn is capable of producing and storing Type I α1 human collagen (r-CIα1) containing a sequence that assists assembly (foldon) and the mature cleaved form (without foldon). Extracts of ground corn grain spiked with *Pichia*-sourced collagen (rP-CIα1), with and without foldon, were precipitated with NaCl to purify collagen from host cell proteins. The low acid solubility of host cell proteins (HCP's) along with the acid stability of collagen favored extraction at pH 2. NaCl concentrations >0.75 M provided complete precipitation of rP-CIα1 without foldon and >70% purity. A fractional cut to remove HCP's with 0.25 M NaCl and precipitation of rP-CIα1 with 1.0 M NaCl increased the purity to 95% at the expense of yield. The presence of foldon negatively impacted yield and purity results, 94% and 44% respectively, and these did not improve with increased NaCl concentration. Pepsin added to cleave foldon and host cell proteins prior to precipitation had no effect on yield or purity. Cleavage after precipitation reduced the volume to be treated with pepsin , and subsequent reprecipitation of the partially converted rP-CIα1 improved purity to 67%. The purity of rP-CIα1 with foldon recovered from *Pichia* extracts was lower (16.8%) than when recovered from corn extracts, but maintained yield >90% and foldon content of approximately 60%.

CHAPTER 1.

GENERAL INTRODUCTION

1.1 Introduction

Motivation

Collagen and its denatured form, gelatin, are proteins which are widely used by the food, cosmetic and medical device industries. Both are extracted from bone, skins and other connective tissues of animals processed by the food industry, and have been widely considered for over 125 years, and by the FDA, as GRAS (Generally Recognized As Safe) for use in consumer products. Their melting points make gelatins useful in oral drug delivery (1, 2); the fiber's structural form makes it useful in injectable physiological fillers and implantable medical devices (3). The matrices formed by these fibers provide a three dimensional substrate for the *in vitro* cell growth critical to wound healing and tissue regeneration (4, 5, 6).

In 1992 and 1993, the FDA issued its first guidance letters alerting the food and medical device industries to the potential transmission, from animal-derived materials to humans, of prion diseases which result in chronic wasting encephalopathies: Bovine Spongiform Encephalopathy (BSE) in cows, and Creutzfeldt-Jacob Disease (CJD) in humans . Early interaction with the FDA by the gelatin industry had identified process conditions capable of inactivating potential prion inclusion in the final product (7). However, the 1996 confirmation of a rapid onset CJD variant in humans who had consumed contaminated meat in Great Britain, led to a series of control measures designed to reduce infectivity risk. Official guidelines were developed and completed in

1997 by the FDA, as well as the European Commission, instructing the Animal and Plant Health Inspection Safety (APHIS), USDA Veterinary Service, and manufacturers to systematically eliminate the use of bovine- and ovine-derived products from nine European countries, and identified twenty-one additional countries as having significant risk. The guidance document further identified specific tissues by a hierarchy of their potential infectivity, with the goal of removing high risk material of any source from animal feeds and the human food supply (7, 8, 9).

Transmission to other animal species broadened the categorization from a bovine and ovine disease to a general infectious agent, Transmissible Spongiform Encephalopathy (TSE). Guidance documents were updated regulating collagen and gelatin sourced for injection, opthalmics, and implantation, that imposed limits on collagen and gelatin's origins regardless of the source animal's country of origin (10). For the first time, USDA veterinarians enforced animal tracking, required animals to be ambulatory prior to slaughter with increased ante- and post-mortem inspection, and revised abattoir and mechanical separation methods to reduce nerve tissue cross-contamination from vertebral columns (8). Despite the identification of a molecular marker for the disease-form proteins and development of a generalized TSE assay, gelatin and collagen extracted from bovine and porcine bone or skins became regulated as a potentially infective animal material (11).

When harvested from animals, several types of collagen exist. Variations in amino acid sequence and the level of prolyl hydroxylation result in different physical characteristics and make any extracted product of animal origin highly heterogeneous (12, 13). Thus, concern exists that epitopes caused by variation among collagen types

and animal species may induce immune reactions upon repeated introduction to human patients (14, 15). Medical application demands exceed thousands of tons annually. Therefore, variations in physical properties and reliance upon an animal-based supply chain make tailored forms of collagen and gelatins with specific melting point temperatures, dissolution pH, viscosity and gel strength candidates for recombinant protein production. Fermentation of Type I human collagens in yeasts has proven these expression systems capable of the post-translational modifications required for specific collagen formation. Co-expression of prolyl 4 hydroxylase modifies proline in the helix forming regions to form hydroxylproline, essential to hydroxylation patterns comparable to animal-origin Type I molecules (16, 17).

Recent advances in molecular plant technologies have brought new capability to low-cost production of heterologous proteins (18). Plant host systems have correctly expressed, folded and modified proteins (19), and directed transport and storage to specific structures in the seed (20) including recombinant human Type I α1 collagen. Some degree of initial purification can be achieved for some plant hosts, e.g. corn, by separation of component tissues using standard wet and dry milling procedures and equipment (21). In a 1998 economic analysis of recombinant β-glucuronidase recovered from 8% dry weight content in transgenic corn, ion exchange and hydrophobic interaction chromatography produced 83% purity material at an estimated cost of $43,000/kg using 1998 cost structures (22). To be successful, recombinant collagen and its related gelatins must be comparable in purity and price to animal-derived products, less than $50/kg. This price reflects their roles as commodities and pharmaceutical ingredients --- not actives. To achieve this, co-product extraction combining traditional

corn derivatives with transgenic protein product may be a solution to offset refining costs (23).

Research Objectives

Through a research collaboration with Fibrogen and ProdiGene, our lab obtained transgenic corn seed containing low expression levels of Type I α 1 collagen. An evaluation of a low-cost initial processing step to extract and purify intact collagen was undertaken. To represent collagen levels expected from further expression system development, Type I α1 collagen purified from recombinant *Pichia* cells, either with or without foldon, was spiked at concentrations of 0.25 mg/mL into the extraction buffer. Extraction pH, sodium chloride (i.e. precipitant) concentration and the effect of pepsin treatment prior to precipitation were evaluated to determine their effects on the purity and yield of precipitated collagen. Extraction pH's of 2, 4 and 6 were tested. Sodium chloride concentrations ranging from 0 to 1.5 M were evaluated for their ability to precipitate collagen in the extract while minimizing host cell protein inclusion. Upon addition of the precipitant, Type I α1 collagen concentrations were 0.20 mg/mL. Supplementing the precipitation studies, a proteolysis step employing pepsin was used in two ways: 1) to remove the foldon sequence present in pro-collagen to evolve the mature collagen form,; and 2) to determine whether any collateral hydrolysis of host cell proteins would influence the selectivity and yield of the precipitation of collagen.   In short, the objective of this work was to yield a relatively high purity Type I α1 collagen from whole corn extract in a few low-cost processing steps.

1.2 Thesis Organization

This thesis includes four chapters. This first chapter includes a general literature review which covers the structure of Type I α1 collagen, how its properties will affect its purification from host cell proteins, separation techniques and approaches to collagen extraction and purification, and the specific susceptibility of pro-collagen to pepsin enzymatic digestion and the mature collagen form which results.

The second chapter examines the effectiveness of precipitation of a purified form of Type I α1 collagen from a low pH transgenic corn extract containing spiked collagen. ANOVA analysis of three variables (pH, sodium chloride concentration and preparative pepsin digestion of the whole corn extract) was evaluated for significance in precipitating collagen from the extract as a low-cost initial purification step.

The third chapter specifically examines the precipitation of a procollagen form which contains a foldon peptide sequence from a transgenic corn extract. Collagen with the foldon sequence is expected to represent a large portion of the expressed form in transgenic corn. At pH 2, the ability to precipitate Type I α1 with foldon collagen from a transgenic corn extract containing added procollagen was evaluated along with the subsequent pepsin digestion of the purified collagen with foldon was performed.

The fourth and final chapter of this thesis provides general conclusions on precipitation and purification of collagen from a whole corn extract, and its relevance to low-cost initial separation procedures when collagen is expressed by different expression systems in the presence of other host cells proteins.

1.3   Literature Review

Protein Solubility and Precipitation

The solubility of proteins in aqueous solutions can be manipulated for use in bioprocesses where precipitation provides a low-cost means of purification and concentration from complex solutions. Several factors can affect solubility including protein concentration, pH, temperature, salt strength and ionic species present in the aqueous buffers.

Protein solubility is generally recognized to improve with increases in temperature, up to a point where the energy of the system denatures the protein's tertiary structure. Therefore, increasing the temperature to improve solubility is not a preferred method for improving active protein solubility. Correct folding of a protein is essential for activity, making the maintenance of low temperatures during processing a necessity (24). pH affects the level of protonation of amino acid side chains in contact with the aqueous system. The overall charge interaction which results between the protein and the solvent contributes to the protein's stability in that solvent. When the pH is adjusted to the protein's specific pI, a net charge of '0' results . At the pI, hydration of the protein's surface is smallest, destabilizing it within the liquid (25). To exploit the effects of pI, estimated pI values can be calculated from a known amino acid sequence and represent the charge of the primary structure. These estimated pI values represent a fully denatured form and generally provide good correlation with experimental pI results. pI shifts result from protein modifications incurred in biological processes or from charge interactions of partially denatured samples (26).

The presence of salts changes the hydration interactions between the protein and the solvent by altering the electrostatic distribution over the protein surface, and therefore it's interaction with the solvent. In aqueous systems, this changes in the ordering of water molecules incrementally with salt concentration. Salting-in (solubilization) or salting-out (precipitation) can be induced by the change of preferential hydration level of the protein surface (27, 28). Ion effects tend to follow the Hofmeister series (Table 1.) which lists cations and anions in order of their preferred hydration interactions and salt-binding to the protein surface (24, 29). A higher order within the Hofmeister series indicates a greater ability for ions to reduce preferential hydration, destabilizing the protein and causing precipitation (30, 31).

Table 1. Cations and Anions of the Hofmeister Series

anions: $SO_4^{2-} > H_2PO_4^- > CH_3CO_2^- > Cl^- > Br^- > I^- > NO_3^- > SCN^-$

cations: $Li^+ > Na^+ > K^+ > NH^{4+} > Mg^{2+}$

Collagen Structure

Collagen is a fiber-forming protein assembled from three subchains, each containing a domain with repeated triplet amino acid sequences of Gly-X-Y. Proline and hydroxyproline are often present in the X and Y positions of this sequence respectively. In human collagen, these three amino acids represent approximately 22 percent of all residues. The high content of these amino acids and the Y-position occupied by hydroxyproline produces a conformational pucker which lends to the bend formation of

the secondary structure and exposes the 4-hydroxyl subgroup of the hydroxyproline residue (13, 32).

During translation, this sequence is assembled as Gly-X-Y with proline predominantly at both X and Y positions (17). However, it doesn't remain entirely a proline-based sequence. In Type I α1 collagen expressed by *Pichia,* the γ-carbon of the Y-position proline is hydroxylated in the endoplasmic reticulum to 4-hydroxyproline after assembly of the primary structure, and when co-expressed with the non-heme iron enzyme prolyl-4-hydroxylase (17, 33, 34, 35).

Additional sequences which do not conform to the Gly-X-Y triplet are also present in the procollagen form. The N-terminus contains an N-propeptide and an N-telopeptide which remain after initial post-translational cleavage of a signal sequence. Similarly, the C-terminus of the helical region contains a C-telopeptide, and C-propeptide which contains a foldon sequence (36, 37). Helical association of procollagen chains is governed by prolyl hydroxylation of multiple prolines at a nucleation region of the C-terminus, and proceeds down the helix to the N-terminus. The proximity of the winding C-terminal residues assists in the alignment of the next residue pairing and self-assembly of the helical trimer. The telopeptide residues associate with adjacent strands to form intramolecular bonds prior to propeptides cleavage (32, 38). Ordering of the glycine to the interior is stabilized by additional hydrogen bonding of hydroxyprolines to the adjacent chain and final association of the trimer occurs. Formation of the helix from C- to N- terminus with interaction by molecular chaperones precludes incorrect pairing that would destabilize the helix (38, 39). Hydroxylation of proline in the Y-position and the

resulting hydrogen bond formation between monomers improves stability which is measurable by melting point temperature (16, 39).

Final processing of the procollagen into the final mature collagen involves cleavage of the C- and N-terminal propeptides, leaving a stable helical trimer which retains the C- and N- telopeptides, prior to secretion from animal cells to the extracellular spaces and final fibril assembly (13, 38). When human recombinant collagen is expressed in plants, this processing is also observed to occur to varying degrees producing procollagen and mature collagen (40,41).

Pepsin and Its Role in Procollagen Maturation

Pepsin is a member of the aspartyl protease family of enzymes which has endopeptidase proteolytic capability and is found in the gastric lining of animals. The most common industrially available form consists of Pepsin A, which is derived from porcine intestinal mucosa and is used in food processing (42). Its characteristic structure of β-sheets, common within the family of aspartyl proteases, results in folding of the enzyme into a two-lobe formation. At low pH's, most of the carboxyl groups are protonated, stabilizing the lobed structure via hydrogen bonding and placing the catalytic region in the correct conformation of the fold for association with its substrate. This effect on folding explains the maximum activity of the enzyme occurring at pH 2 (43,44).

Aspartic acid residues of the catalytic region are located at relative pepsin positions of Asp32 and Asp215, which cleave peptide bonds of substrate proteins by the presence of highly electronegative oxygen atoms at the free carboxyl ends in low pH conditions. Nucleophilic attack of the target substrate peptide bond occurs due to the

close distance between the two free carboxyl groups of Asp32 and Asp 215 within the active site, resulting in the separation of the carboxyl and amino ends of the target protein cleavage site (45). Aromatic and hydrophobic amino acid residue pairings are preferred by pepsin for binding within its seven residue catalytic active site (46).

The procollagen form of Type I α1 triple helix is processed by pepsin to form the mature collagen molecule by removal of non-helical residues within the C- and N-terminal telopeptides. The foldon sequence attached at to the C-terminus is also removed. The monomer subunit structure within the helical region has a high content of glycine and proline, which are hydrophobic residues. However, their stabilization within the helical trimer confers pepsin resistance to the remaining undenatured Type I α1 molecule, while monomers not stabilized by the helical formation remain subject to pepsin activity (34, 47).   Pepsin's smaller molecular weight of 35 kDa is easily separated from 98 kDa collagen and can be detected by chromatographic methods which separate based upon size including HPLC-SEC and SDS-PAGE (48).

1.4 References

1. US Food and Drug Adminstration. "Database of Select Committee on GRAS Substances (SCOGS) Reviews: Gelatin". **1975.** Accessed June 14 2010 (Link - http://www.accessdata.fda.gov/scripts/fcn/fcnDetailNavigation.cfm?rpt=scogsListing&id=14

2. Sikorski Z. *Chemical and functional properties of food proteins.* CRC Press LLC. Boca Raton, FL. **2001**.

3. US Food and Drug Adminstration. "Wrinkle Relief: Injectable Cosmetic Fillers". **06/26/2008.** Accessed June 14 2010 (Link -

    http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm049349.htm)

4. Meyers MA, Chen PY, Lin AYM, Seki Y. Biological materials: structure and mechanical properties. *Progr. Mater. Sci*. **2008**; 53, 1: 1-206.

5. Abraham LC, Zuena E, Perez-Ramirez B, Kaplan DL. Guide to collagen characterization for biomaterial studies . *J. Biomed. Mater. Res., Part B – Applied Biomaterials*. **2008;** 87B, 1: 264-285.

6. Ruszczak Z, Schwartz RA. Collagen uses in dermatology- an update. *Dermatology*. **1999**; 199: 285-289.

7. US Food and Drug Administration . "Guidance for Industry: The Sourcing and Processing of Gelatin to Reduce the Potential Risk Posed by Bovine Spongiform Encephalopathy (BSE) in FDA-Regulated Products for Human Use". Sept 1997. Accessed June 14 2010 (Link:

    http://www.fda.gov/RegulatoryInformation/Guidances/ucm125182.htm).

8. US Food and Drug Adminstration. "Animal Disposition/Food Safety: Bovine Spongiform Encephalopathy (BSE): Key Points for the Public Health Veterinarian". 03/23/09.

9. European Commission Decision 2000/418/EC: Regulating the use of material presenting risks as regards transmissible spongiform encephalopathies and amending Decision 94/474/EC. *Offical Journal of the European Communities*. **6/30/2000**.

10. US Food and Drug Adminstration. "Guidance for FDA Reviewers and Industry Medical Devices Containing Materials Derived from Animal Sources - Except for In Vitro Diagnostic Devices". **Nov 6, 1998**. Accessed June 14 2010 (Link - http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm073810.htm)

11. Dormont D. Prions, BSE and food. *Int. J. Food Microbiol*. **2002**; 78: 181-189.

12. Olsen D, Yang CL, Bodo M, Chang R, Leigh S, Baez J, Carmichael D, Perälä M, Hämäläinen ER, Jarvinen M, Polarek J. Recombinant collagen and gelatin for drug delivery. *Adv. Drug Delivery Rev*. **2003;** 55: 1547-1567.

13. Ricard-Blum S, Ruggiero, F. The collagen superfamily: from the extracellular matrix to the cell membrane. *Pathologie Biologie*. **2005**; 53: 430-442.

14. Sakaguchi M, Inouye S. Systemic allergic reactions to gelatin included in vaccines as a stabilizer. *Japanese J. of Infect. Diseases*. 2000; 53, 5: 189-195.

15. Nandakumar, KS. Pathogenic antibody recognition of cartilage. *Cell Tissue Res*. **2010;** 339: 213–220.

16. Baez J, Olsen D, Polarek JW. Recombinant microbial systems for the production of human collagen and gelatin. *Appl. Microbiol. Biotechnol*. **2005**; 69: 245-252.

17. Kivirikko K, Myllyharju. Prolyl 4-hydroxylases and their proteins disulfide isomerase subunit. *Matrix Biol*. **1998**; 16: 357-368.

18. Menkhaus TJ, Bai Y, Zhang CM, Nikolov ZL, Glatz CE. Considerations for the recovery of recombinant proteins from plants. *Biotechnol. Prog*. **2004**; 20: 1001-1014.

19. Zhong QX, Gu ZR, Glatz CE. Extraction of recombinant dog gastric lipase from transgenic corn seed. *J. Agric. Food Chem*. **2006**; 54: 8086-8092.

20. Paraman I, Fox SR, Aspelund MT, Glatz CE, Johnson LA. Recovering corn germ enriched in recombinant protein by wet-fractionation. *Bioresour. Technol*. **2010**; 101: 239-244.

21. 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.

22. Evangelista RL, Kusnadi AR, Howard JA, Nikolov ZL. Process and economic evaluation of the extraction and purification of recombinant beta-glucuronidase from transgenic corn. *Biotechnol. Prog*. **1998**; 14: 607-614.

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

24. Arakawa T, Timasheff SN. Mechanism of protein salting in and salting out by divalent-cation salts : balance between hydration and salt binding. *Biochem*istry. **1984**; 23: 5912-5923.

25. Boström M, Tavares FW, Finet S, Skouri-Panet F, Tardieu A, Ninham BW. Why forces between proteins follow different Hofmeister series for pH above and below pI. *Biophys Chem.* **2005**; 117: 217-224.

26. Gianazza E. Isoelectric focusing as a tool for the investigation of post-translational processing and chemical modifications of proteins. *J. Chromatogr. A*. **1995**; 705: 67-87.

27. Ruckenstein E, Shulgin IL. Effect of salts and organic additives on the solubility of proteins in aqueous solutions. *Adv. Colloid Interface Sci*. **2006**; 123: 97-103.

28. Arakawa T, Timasheff SN.   Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry*. **1982**; 21: 6545-6552.

29. Dumetz AC, Snellinger-O'Brien AM, Ann M., Kaler EW, Lenhoff AM. Patterns of protein - protein interactions in salt solutions and implications for protein crystallization. *Protein Sci*. **2007**; 16: 1867-1877.

30. Curtis RA, Lue L.   A molecular approach to bioseparations: Protein-protein and protein-salt interactions. *Chem. Eng. Sci.*   **2006**; 61: 907-923.

31. Watanabe EO, Popova E, Miranda EA, Maurer G, Pessoa PD. Phase equilibria for salt-induced lysozyme precipitation: Effect of salt type and temperature. *Fluid Phase Equilib***. 2009**; 281: 32-39.

32. 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.

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

34. 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.

35. Shoulders MD, Raines RT. Collagen Structure and Stability. *Ann. Rev. Biochem*. **2009**; 78: 929-958.

36. 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.

37. 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.

38. Koide T, Nagata K. Collagen biosynthesis. *Top. Curr. Chem*. **2005**; 247: 85-114.

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

40. Baez J, Olsen D, Polarek JW. Recombinant microbial systems for the production of human collagen and gelatin. *Appl. Microbiol. Biotechnol*. **2005**; 69: 245-252.

41. 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.

42. Fruton JS. A history of pepsin and related enzymes. *Quarterly Rev. Biol*. **2002**; 77: 127-147.

43. Kageyama T. Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development. *Cell Mol. Life Sci.* **2002**; 59: 288-306.

44. Chitpinityol S, Crabbe MJC. Chymosin and aspartic proteinases. *Food Chem.* **1998**; 61: 395-418.

45. Dunn BM. Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chem. Rev*. **2002**; 102: 4431-4458.

46. Kageyama H, Ueda H, Tezuka T, Ogasawara A, Narita Y, Kageyama T, Ichinose M. Differences in the P1' substrate specificities of pepsin A and chymosin. *J. Biochem*. **2010**; 147: 167-174.

47. Myllyharju J, Nokelainen M, Vuorela A, Kivirikko KI. Expression of recombinant human type I-III collagens in the yeast *Pichia pastoris*. *Biochem. Soc. Transactions*. **2000**; 28: 353-357.

48. Jones RGA, Landon J. Enhanced pepsin digestion: a novel process for purifying antibody F(ab ')$_2$ fragments in high yield from serum. *J. Immunol. Methods*. **2002**; 263: 57-74.

CHAPTER 2.

RECOVERY OF COLLAGEN TYPE I α1 FROM A WHOLE-CORN EXTRACT USING PEPSIN DIGESTION AND PRECIPITATION

A paper to be submitted to *Biotechnology Progress*

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

2.1 Abstract

Because of safety concerns and product consistency issues with animal-derived collagen and its derived product form, gelatin, there is commercial interest in producing recombinant forms of collagen expressed in a non-animal host to meet the demands of the food, cosmetic and pharmaceutical markets. We describe a recovery strategy for collagen spiked into corn seed extracts that is appropriate to the need to produce these proteins at low cost to meet market demands. The method was based on one developed by FibroGen for recovery from a *Pichia* host. An initial enzymatic digestion of host cell proteins by pepsin, to which undenatured triple-helical collagen without a foldon sequence is resistant, was followed by addition of sodium chloride to selectively precipitate spiked yeast-derived recombinant Type I α1 (rP-CIα1) collagen from acidic extracts of corn grain. Spiking with rP-CIα1 was done to achieve collagen levels in the range that have been observed in transgenic corn lines developed for high expression of other heterologous proteins.

Salt precipitation at pH 2 of a corn seed extract containing 0.20 mg rCIα1 /mL solution (with rP-CIα1 representing 14% of the total soluble protein) at 0.75 M and 1.0 M sodium chloride yielded 100% of the collagen at >70% purity without the pepsin pretreatment. With pepsin pretreatment, the highest yield (94.0%) of rCIα1 was in the precipitate at 0.75 M sodium chloride. The collagen purity at 0.75 M sodium chloride was 76.5%; hence, under these conditions the improvement from the pepsin pretreatment would not likely justify the additional cost. 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.

2.2 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). Recently, recombinant technology has made possible production via microbial fermentation with high titers of collagen. The ability to express collagen and gelatin in *Pichia pastoris* with tailored degrees of proline hydroxylation demonstrates the potential to make structural proteins with tailored performance properties (3, 4, 5) specific to applications such as drug delivery and capsule production (3).

Directed expression to corn grain provides storage stability (6) and additional product enrichment can be obtained by targeting to specific tissues within the grain (7, 8).

Recombinant human Type-I collagen alpha 1 (rCIα1) has been expressed in transgenic corn and characterized as having 1.23% hydroxyproline content, 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 will make the use of corn expression hosts attractive if low-cost recovery and purification can be achieved (9, 10).

Native corn protein solubility in aqueous buffers increases above pH 3.0 (7). 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 (11). 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 (12, 13). Purification of a low expression level rCIα1 from corn extracted in an acidic buffer (pH 1.8) using hydrophobic interaction (HIC), ion exchange and gel filtration chromatography has already demonstrated that high purity fractions of a collagen fragment can be obtained (9, 12).

Acidic extraction, preceded by pepsin digestion to facilitate release from the tissue, has been used to extract collagen from fish skins (13). Pepsin addition is used for recovery from *Pichia* in order to complete post-translational modification and improve purity by partial hydrolysis of HCP's. The triple helical structure of collagen is tightly packed and stabilized by strand interactions of the hydroxyl proline residues, into a form that is resistant to pepsin activity (7, 14). When initially expressed, the procollagen molecule contains C- and N- terminal telopeptides and propeptides to aid in forming the

correct helical structure. Subsequent cleavage of these sequences, which can be accomplished by pepsin addition (15), provides the form ready for assembly into fibers (16). Both hydroxylated and non-hydroxylated forms of rCIα1 have proven to be resistant to pepsin (15) below their melting temperatures, making pepsin addition a processing option for simplifying separation 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, and 3) the effect of pepsin digestion of HCP on rCIα1 yield and purity in the precipitate.   Because current rCIα1 expression levels are low, an experimental model has been developed in which *Pichia*-derived rP-CIα1 has been spiked at 0.25 mg/mL into the extraction buffer, corresponding to complete extraction of a recombinant protein expressed at the level of 1 mg/g of dried solid. This level has been approached (e.g. porcine α-lactalbumin has been expressed at 0.4 mg/g in transgenic corn seed (17)) and would be a reasonable target for a lower-valued protein.   Spiking the buffer with rP-CIα1prior 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, accounts for the nature and concentration of native corn components characteristic of a potential transgenic variety. A comparison of this spiking approach with recovery from truly transgenic gain validated this approach for β-glucuronidase (GUS) in canola seed. (18). Some limitations of the model exist: the spiked rP-CIα1 enters the system as part of a fully soluble fraction. Interactions within the insoluble matris (prior to extraction) are not represented in this model.

2.3 Materials and Methods

Materials

Reagent grade sodium chloride, phosphoric acid, hydrochloric acid, guanidine hydrochloride, and Thermo–Pierce Bradford Total Protein Assay were obtained from Fisher Scientific (Hampton, NH). 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). Recombinant human collagen was expressed in *Pichia pastoris* (rP-CIα1) in a 50 L fermentor (ISU Fermentation Facility), released from the cells by ball milling, and recovered by precipitation after a pepsin digest of the extract following procedures developed by Fibrogen (3). 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°C) provided by Fibrogen.

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*, 23°C) 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 -20°C.

Extraction and Spiking

For an initial screening of precipitation conditions without pepsin pretreatment, 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 0.50 mg rP-CIα1/mL) for one hour at 24°C, centrifuged (4,600 *g*, 10 min, 4°C), and decanted. The pellet was resuspended for a second extraction, centrifuged (15,200 *g*, 15 min, 4°C) and the supernatant was decanted. The combined extracts were filtered (0.45 μm SFCA syringe filter) and stored at 4°C. For the additional precipitation experiments that included comparison of pepsin pre-treatment, the same extraction procedure was followed but the rP-CIα1 level was reduced to 0.25 mg/mL.

Pepsin Digestion

For pepsin pretreatment, a 20 mL aliquot of the 0.25 mg/mL collagen-spiked extract was digested with 0.2 mL of thawed pepsin enzyme at 4°C, 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 4°C (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. All tubes were mixed by end-to-end tumbling on a rotating shaker at 4^oC 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^oC; elution detection 220 and 280 nm). Peak areas were integrated using the Agilent's ChemStation software after manual identification of peak boundaries. Collagen is detectable at 220 nm,

but not at 280 nm absorbance. Therefore, peaks collected at 220nm (without 280 nm absorbance) were identified as collagen and peak area was calibrated with injections of the purified rP-CIα1 (18).

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% and 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 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.

Statistical Analysis

A full-factorial ANOVA statistical analysis was performed using JMP Software (SAS). The variables tested were pH, sodium chloride concentration and digestion for their effect on purity, yield and purification factor, determined using a 95% confidence interval determined from pooled standard deviation values with df = 4. Statistically significant groupings from ANOVA (Tukey's least square mean difference test) were identified by lettering in all Figures.

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.

2.4 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 (14). The calculated

concentration from the SEC peak area matched the measured addition to the extract

within 2%.
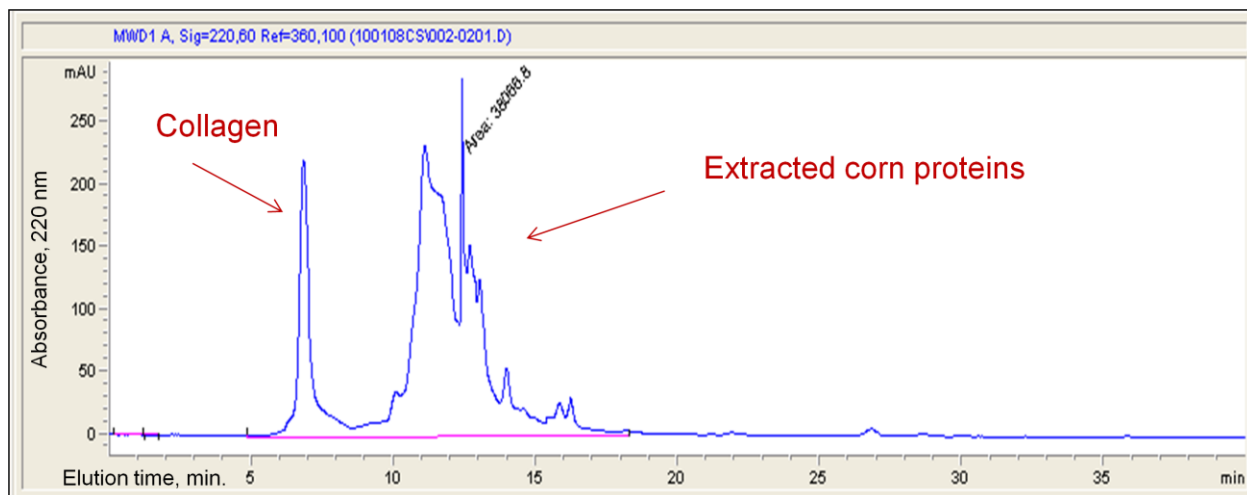
Figure 1: HPLC-SEC chromatogram of extract containing rP-CIα1collagen and corn
proteins. The larger collagen molecule (~90 kDa) elutes ahead of the smaller corn
proteins.

Screening of 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. 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.
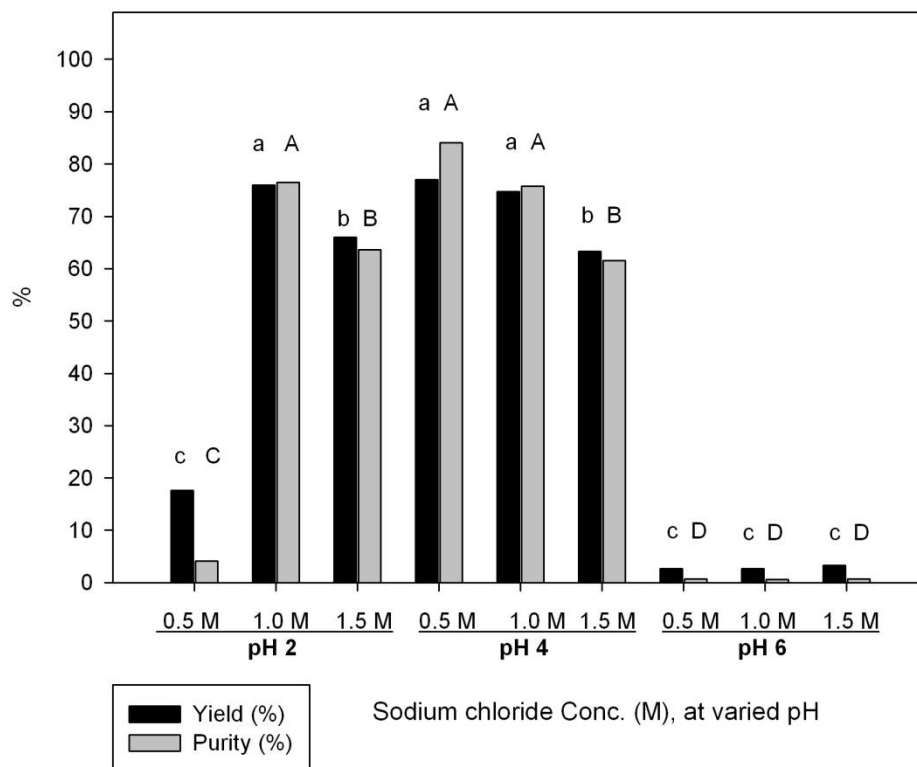
Figure 2. Reconstituted precipitate collagen purity and yield over the range of precipitation conditions. Statistically significant groupings from ANOVA analysis for purity are given in upper case, yield groupings are in lower case.

The best of the tested conditions for precipitation of rP-CIα1 from the corn extract was pH 4 and 0.5 M NaCl. The purity result can also be seen qualitatively by SDS-PAGE (Figure 3). Purification factors of 1.41 ± 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 (see Appendix). 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.
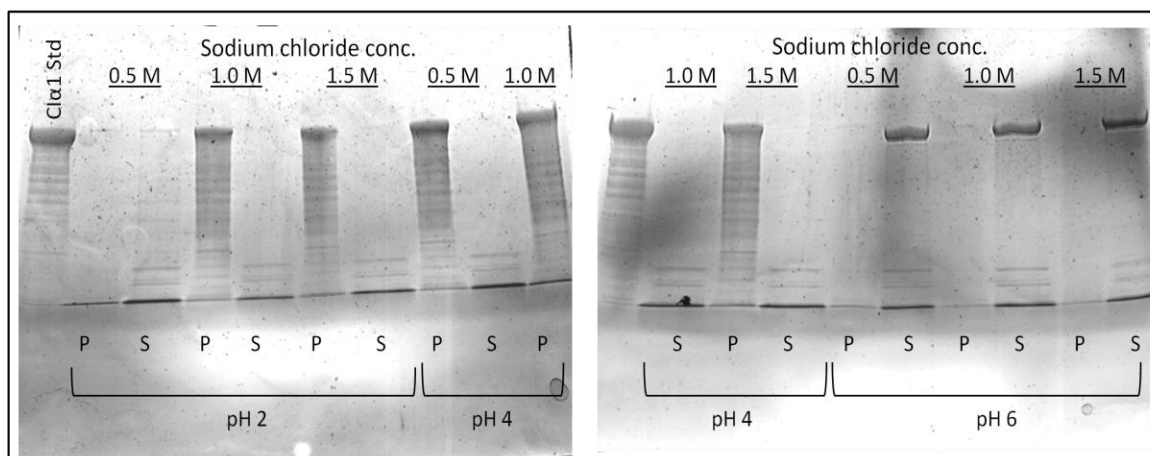
Figure 3. SDS PAGE gels of rP-CIα1-spiked whole corn extract, pH adjusted to 2, 4, and 6 prior to precipitation with sodium chloride.   S = supernatant, P= precipitant.

Precipitation 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, 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 ± 1.53 for pepsin pretreated samples with 0.75 and 1.0 M NaCl, and slightly higher values of 5.53 ± 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.
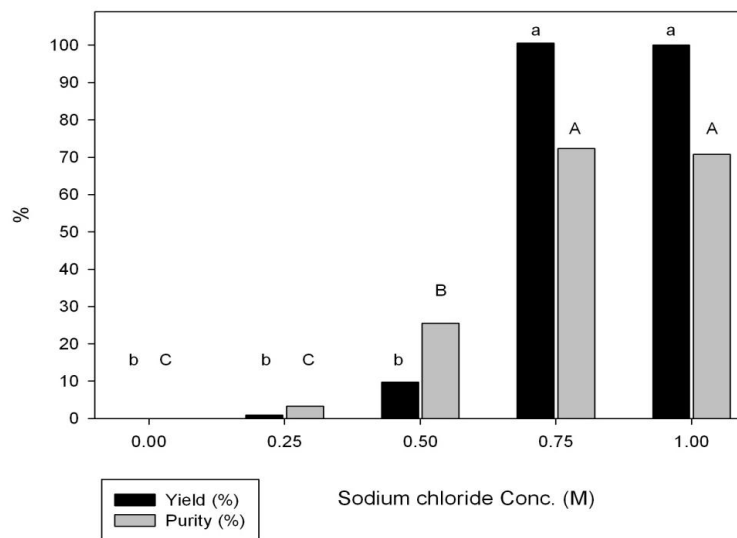
Figure 4.   Reconstituted precipitate collagen purity and yield after precipitation
with sodium chloride at pH 2 without pepsin pretreatment. Statistically significant
groupings from ANOVA analysis for purity are given in upper case, yield
groupings are in lower case.
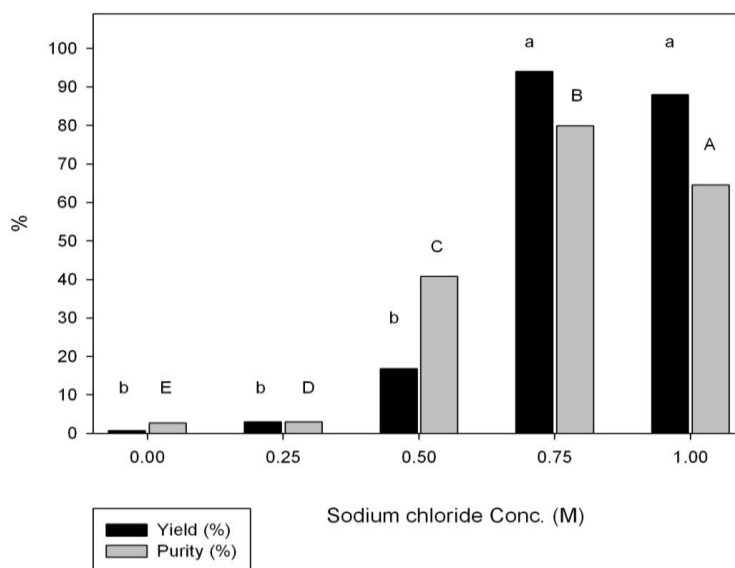
Figure 5. Reconstituted precipitate collagen purity and yield after pepsin
pretreatment and precipitation with sodium chloride at pH 2 . Statistically
significant groupings from ANOVA analysis for purity are given in upper case,
yield groupings are in lower case.

Comparison of Figures 4 and 5 and the ANOVA analysis showed that the pretreatment did not improve purity or yield. Salt concentration, however, was statistically significant (see Appendix). 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. Recovery of rP-CIα1 in the pellet is clearly seen by comparison of HPLC-SEC chromatograms of the original extract (Figure 1) and the precipitated product (Figure 6).
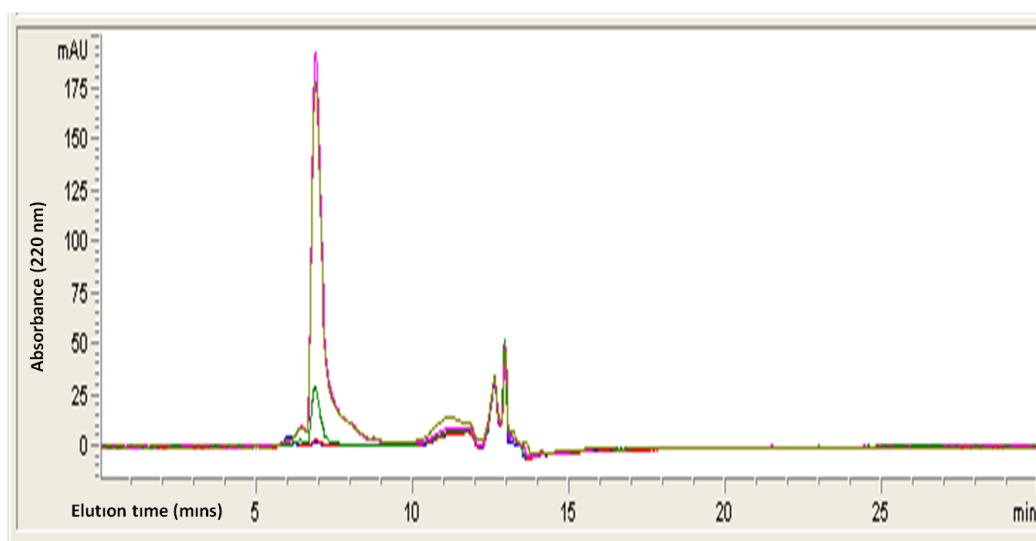
Figure 6. HPLC-SEC chromatogram of resuspended, pepsin-pretreated, pH 2 precipitates, showing an increase in rP-CIα1 at 7 min with increasing sodium chloride concentration.

The similar outcomes after pretreatment reflect the limited digestion of HCP. Evidence of a this hydrolysis is seen by SDS-PAGE (Figure 7), as well as a minor shift in area detected by HPLC-SEC (Figure 8).
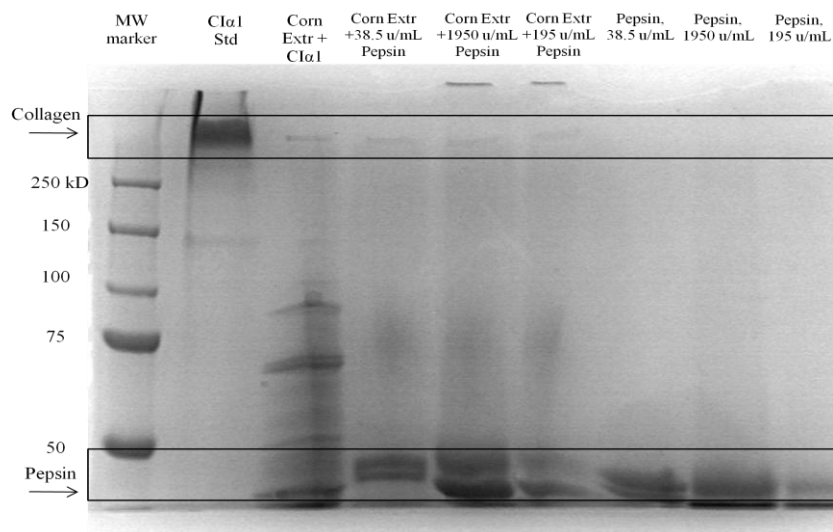
Figure 7. Digestion of corn extracts spiked with rP-CIα1 hydroxylated collagen by pepsin at concentrations of 38.5, 195, 1950 EU/mL. Above is a tris HCl 7.5% gel stained with Coomassie Blue G-250 and shows pepsin degradation of HCP in the 50-100 kDa MW range. The same samples run on 4-15% tris HCl gradient gels show no detectable digestion of HCP.
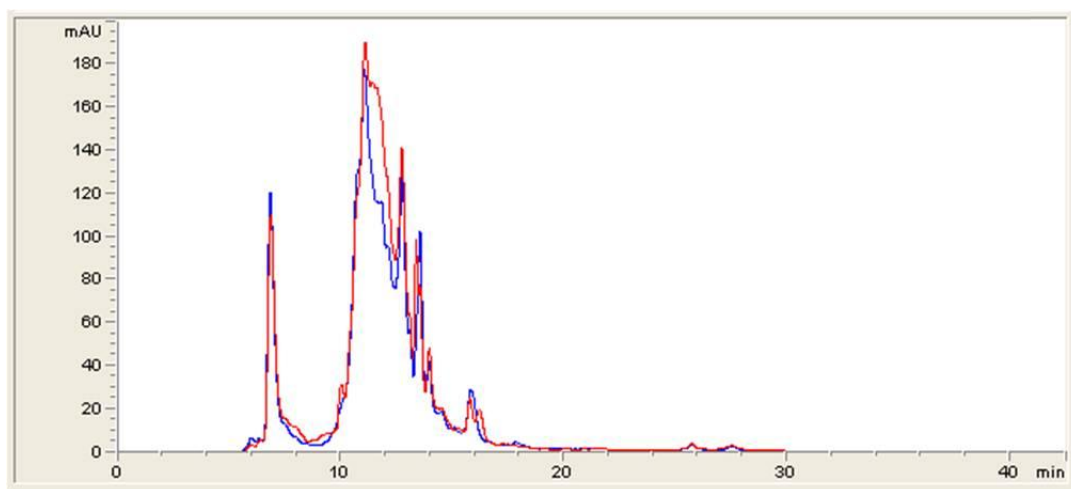
Figure 8. HPLC-SEC chromatogram of collagen spiked corn extract, with pepsin (red) and without pepsin (blue) digestion, showing the shift in the chromatographic area resulting from the addition of pepsin, with pepsin co-eluting among the corn proteins at 10.8 minutes, but otherwise little change in HCP profile

Figure 6 also reveals co-precipitation of HCP at low salt concentration, where collagen remains in solution. This 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%.

2.5 Conclusion

Carrying out both extraction and salt precipitation at pH 2 provided complete collagen recovery at greater than 70% purity. Replacing a single precipitation step with a fractional cut increased purity to 95% at the cost of an 11% yield loss. A pepsin 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 achieved in the step.

2.6 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) 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.

7) 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.

8) 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.

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

10) 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.

11) 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.

12) 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.

13) 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.

14) Berisio R, De Simone A, Ruggiero A, Improta R, Vitagliano L. Role of side chains in collagen triple helix stabilization and partner recognition. *J. Pept. Sci*. **2008**; 15: 131-140.

15) Pakkanen O, Hämäläinen ER, Kvirikko 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) 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.

17) 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.

18) 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 beta-glucuronidase. *Biotechnol. Prog*. **2001**; 17: 161-167.

CHAPTER 3.

RECOVERY OF COLLAGEN TYPE I α1 WITH FOLDON SEQUENCE FROM AN EXTRACT OF WHOLE COR*N AND PICHIA*

Christopher M. Setina and Charles E. Glatz

Department of Chemical and Biological Engineering, Iowa State University, Ames IA 50011

3.1 Abstract

Recombinant Type I α1 collagen (rCIα1) has been expressed in several hosts with an added C-terminal foldon which provides for assembly of the required triple helical structure. Pepsin can then be used to cleave the foldon. Here we examine the possibility of corn as the expression host, while comparing the merits of purification before foldon cleavage. To do so, we simulate recombinant expression by addition of *Pichia*-expressed rP-CIα1 with foldon and compare purity and yield of this protein recovered by precipitation with sodium chloride. Comparison is also made to recovery from *Pichia* extracts, as well as results for rP-CIα1 without foldon.

Precipitation from an acidic corn extract of rP-CIα1spiked to 0.20 mg/mL concentration produced yields greater than 90% for NaCl levels >0.75 M in both rP-CIα1 with foldon and non-foldon containing extracts. A maximum purity of 44% was obtained for the former at >0.75 M NaCl and 67% purity for the latter at >0.5 M NaCl. While the foldon-containing precipitate was not as pure, the ca. 13-fold purification and reconstitution with 2.7-fold volume reduction, relative to the original extract, allowed subsequent foldon cleavage with reduced pepsin requirement.

In contrast to the recovery from corn extract, precipitation of rP-CIα1 with 1.0 M NaCl directly from *Pichia* lysate achieved a purity of 16.8% and yield of 31.2%. Precipitating this material a second time increased the purity to 68.9% purity, but a green contaminant could not be removed. Ultrafiltration (50 kDa MWCO) was used to remove the green coloration of the initial precipitate, while increasing purity from 16.8% to 70%.

KEYWORDS: recombinant collagen, corn, *Pichia*, precipitation, purification, recovery.

3.2 Introduction

Collagens are a family of fiber-forming proteins commonly used in food, cosmetic and pharmaceutical applications, primarily for their gelation and matrix properties (1, 2). Type I collagen is the most prevalent, being the only type which is used in the production of gelatins (3). However, natural sources of collagen are animal connective tissues which carry an inherent infectivity risk (4). Stricter regulation on the sourcing, processing and traceability of animal-derived materials makes production of non-animal collagens commercially desirable (5).

Type I collagen is an assembly of three helical domains forming the structural heterotrimeric or homotrimeric helix from α1(I) or α2(I) chains.(6) The triple helical procollagen molecule contains C- and N- telopeptide domains at the time of assembly, as well as a C-terminal foldon peptide (7, 8). These sequences orient the monomers to form intermolecular hydrogen bonds, stabilizing the helical form (9, 10) . Stabilization of collagen also requires expression of prolyl 4-hydroxylase (P4H) for 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 in other hosts, the foldon can be removed by pepsin hydrolysis which does not cleave correctly folded triple-helical regions (14, 15, 16).

Formation of recombinant human triple-helical collagen and individual gelatins has been demonstrated in microbial expression systems known for high productivity of heterologous proteins. Several yeast genera including *Pichia pastoris* have shown the ability to perform the post-translational modifications and protein folding important for

formation of the mature collagen molecule with its required physical properties (17, 18). *Pichia* possesses a tightly regulated methanol induced alcohol oxidase (AOX-1) promoter capable of co-expression of P4H in the endoplasmic reticulum where post-translational modifications occur (19, 20). Secretion of gelatin to the fermentation medium by *Pichia* cells has been reported, while procollagen purification requires cell lysis (21, 22, 23). Lysis conditions affect the resulting host cell release profile (23).

Production of transgenic collagen in plants, such as corn, offers low cost commercial scalability and stabile storage in seed form for convenient process scheduling. Expression of pharmaceutically relevant proteins requiring post-translational modifications has been demonstrated in tobacco and corn, with promoters able to direct protein to specific tissues that are separable using existing milling technology prior to cell lysis (24, 25). Collagen's solubility in acidic solutions adds further benefit to separation from plant host cell proteins which exhibit limited solubility at low pH. Therefore, low pH extraction and salt precipitation of collagen in a simple step can provide relatively high purity (26, 27).

A model system of a low rCIα1expressing transgenic corn and an extraction liquid spiked with Pichia-derived rP-CIα1 was used to represent the possible interactions with the native corn constituents, and those constituent's nature and concentration in the soluble matrix. The spiking rP-CIα1 only represents the soluble form. When expressed in corn seed, rCIα1 is resident in unsolubilized tissue and therefore all effects of resolubilizing from that solid matrix cannot be determined (28).

3.3 Materials and Methods

Materials

Reagent grade sodium chloride, phosphoric acid, hydrochloric acid, guanidine hydrochloride, trichloroacetic acid, SDS, glycerol, Tris hydrochloride, and sodium hydroxide 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) were 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) from a 2004 field trial was provided by ProdiGene Inc. (College Station, TX) as part of a research collaboration with Fibrogen (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°C). Purified human collagen without foldon (rP-CIα1) for spiking was produced by ISU Fermentation Lab by the same method. Recombinant human pepsin (19,500 EU/mL, stored at -20°C) was also provided by Fibrogen Inc. (San Francisco, CA).

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°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 green color 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^oC) 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^oC). The precipitate was recovered by centrifugation (10,000 g, 30 min, 4^oC,) and was resolubilized in 0.01 M hydrochloric acid, then centrifuged (29,400 g, 4^oC, 30 min) to clarify. This resolubilized acid-salt precipitate was called 'RASP'. To determine if any rP-CIα1 with 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^oC), and centrifuged (10,000 g, 30 min, 4^oC). 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^oC). The precipitate was recovered by centrifugation (Eppendorf Model 5424 microcentrifuge, 10,000 g, 30 min, 4^oC,) 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 of Whole Corn

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 resulting grits were ground into a fine powder using a household coffee grinder (Kitchen-Aid Co., St.Joseph, MO). The whole ground corn was defatted in a 20% w/v mixture with hexane, mixed on ice for 1 h. The supernatant was separated by centrifugation (2000 g,15 min, , 23°C) before a second hexane extraction was performed. The material was allowed to air dry overnight at room temperature prior to storage at -20°C.

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 24°C, centrifuged (4,600

g, 10 min, 4°C), and decanted. The residual solids were extracted a second time in fresh extractant, centrifuged (15,200 *g*, 15 min, 4°C) 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 4°C.

Precipitation of Collagen from Corn Extracts

Collagen-spiked corn extracts at the extract pH of 2 were precipitated in 1.7 mL microcentrifuge tubes with NaCl concentrations of 0.0, 0.25, 0.5, 1.0, 1.25 and 1.5 M. Collagen levels were 0.195 mg/mL for all samples. The tubes were mixed on a rotating shaker (1 h, 4°C) and centrifuged (14,700 *g,* 30 min, Eppendorf Model 5424 microcentrifuge). The supernatants were decanted, the pellets were redissolved in unspiked extraction buffer, and the solution analyzed for content of collagen and total protein. All precipitations were replicated three times.

HPLC-SEC Assay for Collagen and Total Protein

Samples for analysis 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 220nm (with negligible 280 nm absorbance) were identified as collagen and peak area was calibrated with injections of the purified rP-CIα1 (without foldon).

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 loaded onto a precast tris HCl 4-15% gel (BioRad Inc.). The gel was run at 200 V 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 alongside.

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, 24°C) 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 4°C 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.

Pepsin Digestion

Pepsin was used for two purposes: to verify the presence of the foldon sequence in *Pichia*-derived rP-CIα1, and to remove foldon to provide a purified rP-CIα1. To verify foldon presence, 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.

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.

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

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 were calculated from amino acid sequence using Protcalc3 Protein Calculator v3.3

(CDPutnam, Scripps Research Institute, LaJolla CA).

Statistical Analysis

A full-factorial ANOVA statistical analysis was performed using JMP Software (SAS). The variables tested were sodium chloride concentration and foldon presence for its effect on purity and yield, determined using a 95% confidence interval determined from pooled standard deviation values with df = 4. Statistically significant groupings from ANOVA (Tukey's lease square mean difference test) were identified by lettering in all Figures.

Calculations

Yield and purity of collagen and total protein concentrations of the whole corn and *Pichia* extracts and redissolved precipitates 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.

3.4 Results And Discussion

Washing Pichia Cells

Purification of rCIa1 with foldon from *Pichia* requires separation from the host cell proteins. A strong green color was present immediately upon rehydration of the freeze-dried cell mass, which persisted through washing, lysis and purification steps in earlier trials. A scan from 200 to 800 nm was performed on the wash liquids using a Varian scanning spectrophotometer. Maximum absorbance was found at 440 nm (Figure 1). This wavelength was used to quantify the green contaminant's presence throughout the washing process.
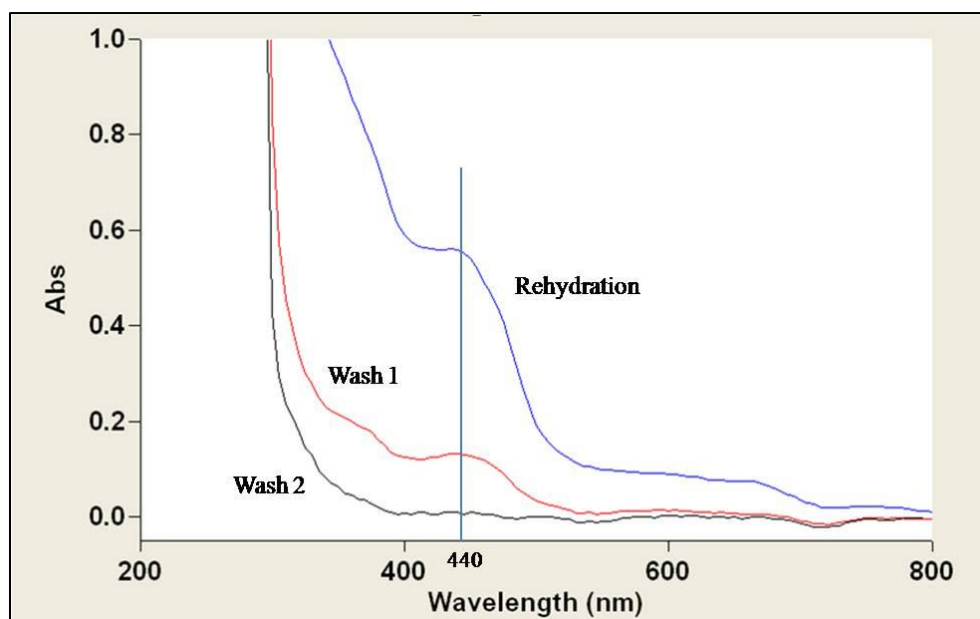

Figure 1. Spectrophotometric scan from 200 to 800 nm of the first three wash steps of the freeze dried *Pichia* cells.

As the number of wash cycles increased, the amount of green color in the supernatant decreased, despite a residual green coloration of the wet cell mass. The green

color was tracked at 440 nm for the rehydration supernatant and over the subsequent 6 wash cycles. In consideration that the green colored wet cell mass might contain membrane bound green proteins, Wash 3 incorporated sodium chloride in an attempt to remove further color from the wet cell mass. Upon addition of sodium chloride, an increase in supernatant color was measured (Figure 2).
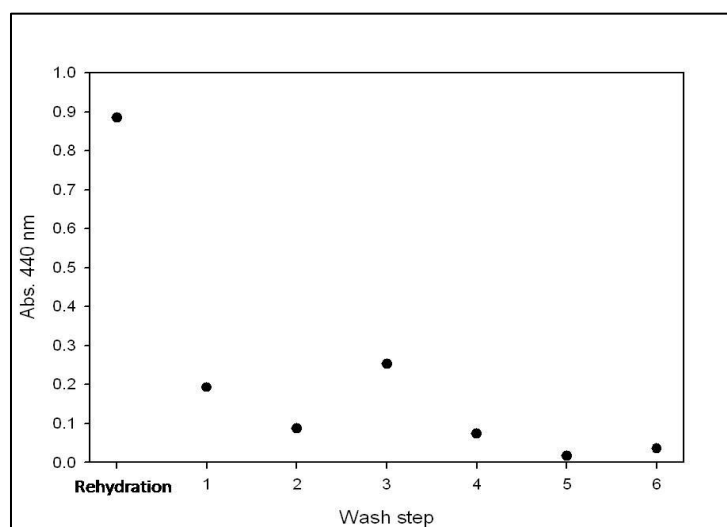
Figure 2. Absorbance values (440 nm) for subsequent washes of the rehydrated cell mass.

All wash volumes were analyzed by HPLC-SEC to determine total rP-CIa1 collagen and host protein content. The initial rehydration and wash removed significant amounts of rP-CIa1 and other proteins (Table 1). The presence of NaCl in Wash 3 did not appear to increase the removal of other proteins, but did increase the amount of rP-CIa1 collagen removed with the wash liquid.   It was not determined if the presence of rP-CIa1 in the Wash 3 liquid was the result of cell lysis. With a reduction in green color ($A^{440} < 0.04$) in Wash 6 liquid, the cells were resuspended in lysis buffer for disruption.

Table 1.   Total of rP-CIa1 and rP-CIa1 w/ foldon and Host Protein Contents of Cell Wash Supernatants

Wash Step	Total Vol (mL)	Total rP-CIa1 (mg)	Host Proteins (mg)
Rehydration (0)	246.51	73.942	11194.14
Wash 1	250.2	0.000	2180.72
Wash 2	265.14	0.000	686.13
Wash3	246.96	1.421	545.14
Wash 4	278.8	0.000	331.50
Wash 5	245.63	0.000	132.53
Wash 6	243.14	0.000	158.66

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 2). An additional 75.4 mg of rP-CIa1 with and without foldon was lost to Rehydration and Wash 3 liquids (Table 1).

Table 2. rP-CIa1 With and Without Foldon and Host Cell Protein Content of the RASP and 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, (%)	% 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 portion of RASP was reprecipitated in an attempt to improve purity. The second precipitation at 1 mL volume of the RASP increased purity to 68.9%. However, yield dropped to 43.9% and the pellet recovered from the second precipitation step was green in color. SDS-PAGE showed that the RASP contained a protein with molecular

weight similar to a rP-CIα1 standard at approximately 120 kDa, as well as a large amount contaminant proteins with molecular weight less than 37 kDa (Figure 3) making ultrafiltration a better option for concentrating and improving the RASP purity.

Ultrafiltration of the RASP using a Millipore TFF system and Pellicon 50 kDa MWCO significantly improved purity from 16.8% to 69.6 % in addition to concentrating the material to 0.61 mg total collagen/mL. However, 8.61 mg of total collagen were lost to the permeate representing a 9.6% yield loss. Previous studies showed complete retention of rP-CIa1with a 100 kDa MWCO membrane (29). Thus, presence of collagen in the 50 kDa permeate suggests a possible loss of membrane integrity so the specific purity and retention outcomes may be misleading. We did not seek to remedy this because our purpose was merely to prepare material for addition to the extracts. The 0.61 mg/mL retentate (ultrafiltered RASP) was tested to confirm the presence of rP-CIa1with foldon and was used for spiking the extraction buffer for the whole corn.
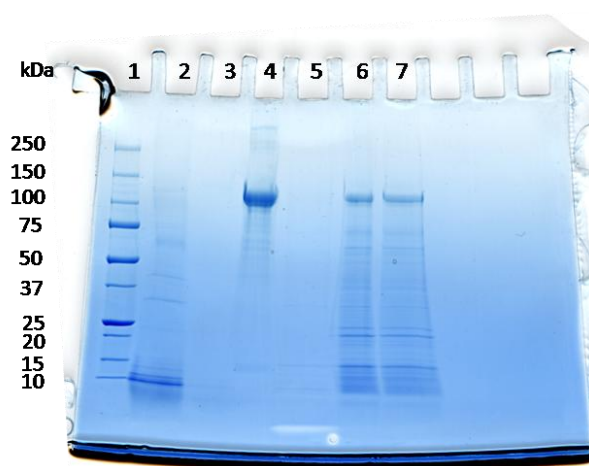

Figure 3.   Coomassie Blue G-250 Stained gel.   Sample wells are 1. Mol Wt Std, 2. Rehydration liquid, 3. Wash 1, 4. rP-CIα1 Std 0.12 mg,   5. supernatant, 6. RASP, 7. second extract of *Pichia* cell debris.

Identification of rP-CIa1 with Foldon via HPLC-SEC

Each of the recovered fractions was analyzed for rCIa1, 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: 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 the green contaminant from this material (Figure 4).
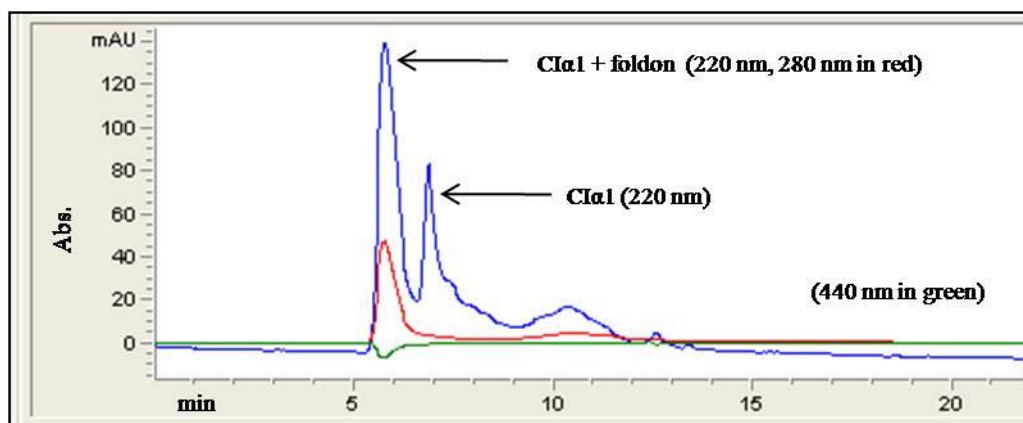
Figure 4.   HPLC-SEC (Abs. 220nm) of ultrafiltered RASP extracted from freeze-dried *Pichia* showing peaks of rP-CIa1with and without foldon

Identification of rP-CIa1with Foldon via Pepsin Digestion

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 3).   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.

Table 3. Effect of Pepsin on Content 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

Western Blotting

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 5). 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 comparision were pretreated with pepsin to remove foldon, and foldon was not detected by Western blotting.
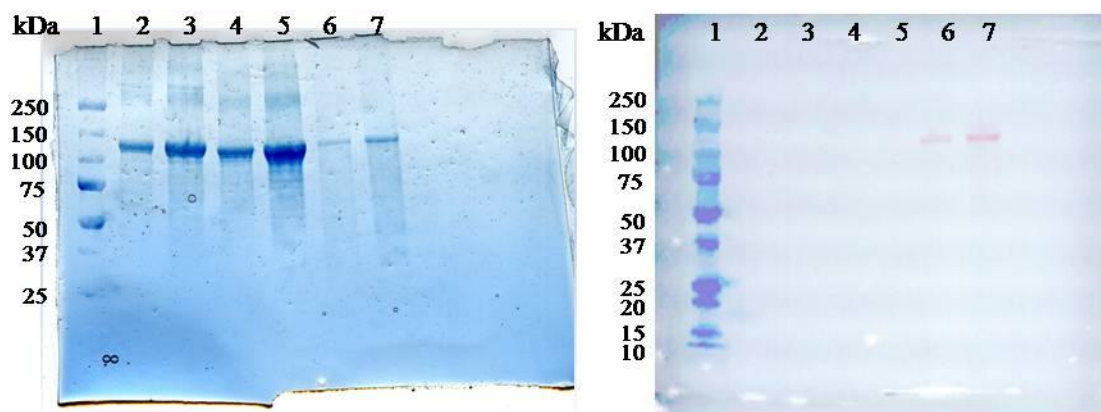
Figure 5: Recovered SDS-PAGE gel stained with Coomassie blue G-250 stain (left) and corresponding Western Blot PVDF membrane (right). 1) pre-stained MW Stds; 2) 10 µL high purity rP-CIα1 Std (1.2mg/mL); 3) 20 µL high purity rP-CIα1 Std (1.2 mg/mL); 4) 10 µL rP-CIα1 (2.43 mg/mL); 5) 20 µL rP-CIα1 (2.43 mg/mL); 6) 10 µL ultrafiltered RASP; 7) 20 µL ultrafiltered RASP. Red color indicates alkaline phosphatase activity, which was conjugated to the secondary antibody.

Two Dimensional Gel Electrophoresis

The full amino acid sequence of rCIα1with foldon expressed in corn has been reported by 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 pI values at the expected molecular weight of 120 kDa: pH 9.1 and pH 7.3.
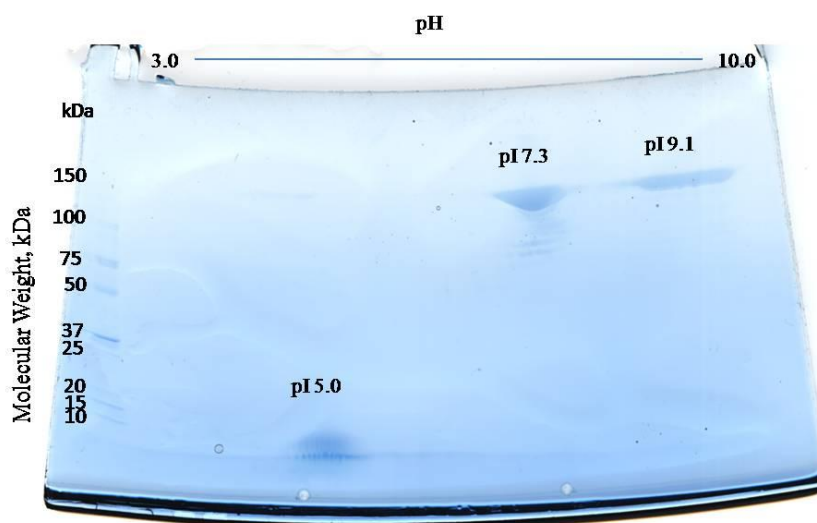

Figure 6: 2D gel separation of 200 µg load of ultrafiltered RASP extracted from *Pichia*, identifying molecular weight and pI.

Given the wide band at pI 9.1 (Figure 6), this spot could be either form of the rP-CIα1. The HPLC-SEC measure of relative amounts of components would argue for the pI 5.0 material being the host cell protein and both of the larger spots being one form of

collagen or another. The difference in pI of those spots can not 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 7). Both foldon presence and NaCl concentration significantly affected yield (see Appendix).

Figure 7: Yields after precipitation with sodium chloride of rP-CIα1 with foldon (ultrafiltered RASP) and without foldon spiked whole corn extracts.   Statistically significant groupings from ANOVA analysis for yield with foldon are given in upper case, without foldon groupings are in lower case.

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. Increasing the sodium chloride concentration above 1.0 M did not improve purity of the ultrafiltered RASP samples (Figure 8). Both foldon presence and NaCl concentration were statistically significant in effect on purity, though foldon presence had less effect than NaCl concentration (see Appendix).
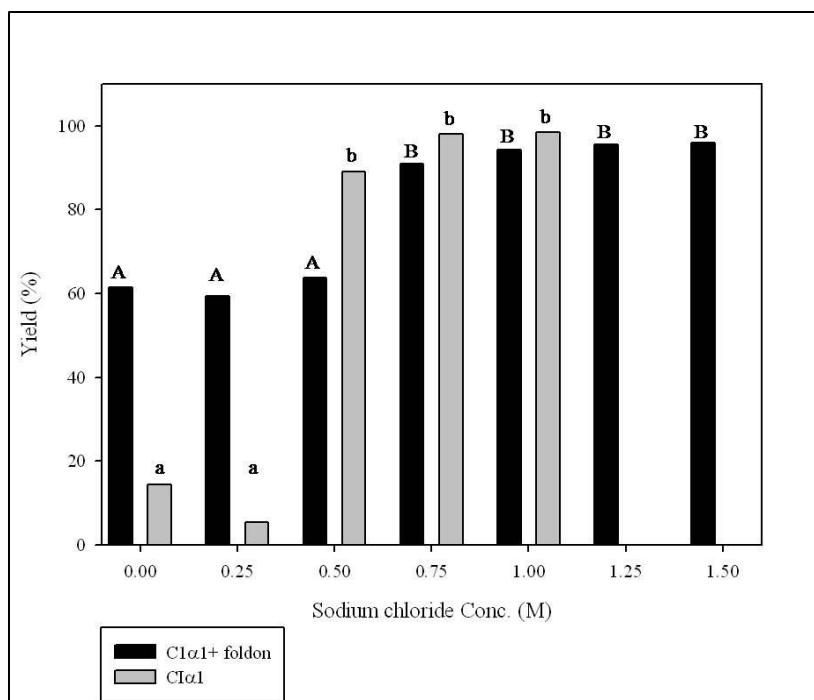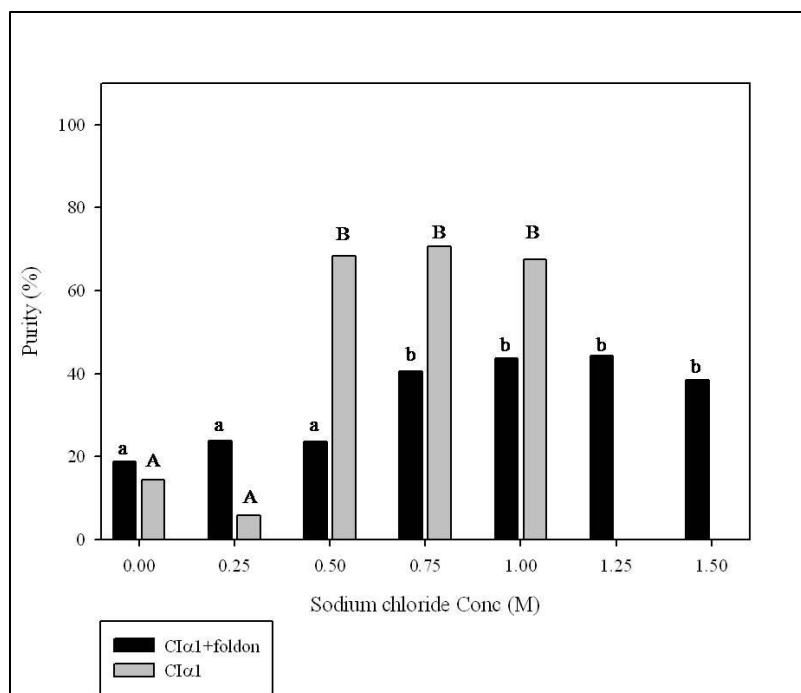
Figure 8: Purity after precipitation with sodium chloride of rP-CIα1 with foldon (ultrafiltered RASP) and without foldon spiked whole corn extracts. Statistically significant groupings from ANOVA analysis for purity with foldon are given in upper case, without foldon groupings are in lower case.

The results indicate that a high yield of rP-CIα1 with foldon expressed 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 59.9 to 31.8 % of total collagen. The remaining 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 purification factor of ca. 13, and a volume reduction of 2.7-fold. A smaller pepsin addition to cleave the foldon sequences would be required for this volume, which can be removed by subsequent 1.0 M sodium chloride precipitation. 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.

3.5 Conclusion

rP-CIa1 collagen containing the foldon sequence can be purified from freeze-dried *Pichia pastoris* cells to approximately 70% when ultrafiltration at 50 kDa MWCO is employed.   The ultrafiltered RASP material was confirmed to contain the foldon sequence by Western Blotting and its susceptibility to removal by pepsin digestion. Precipitation of rP-CIa1 with ca. 59% foldon intact from the corn extract had equivalent yields to precipitation of the rP-CIa1 without foldon for sodium chloride concentrations ranging from 0.75 to 1.5 M. However, in contrast to the without foldon form, about 60% of the rP-CIa1 containing foldon material precipitated at sodium chloride concentrations less than 0.5 M and this precipitate was ca. 15% richer in the "with foldon" form.   The foldon sequence significantly decreased the purity of the final precipitate for all sodium chloride concentrations. Sodium chloride concentrations greater than 0.75 M achieved the highest purities (44%).   Therefore, it is expected that a precipitation step can successfully remove most of the rP-CIa1 containing the foldon sequence from a corn extract, without an initial pepsin digestion. However, the purity is likely to be significantly less when a single precipitation step is employed. Cleavage of the foldon after an initial sodium chloride precipitation reduced the amount of pepsin required for processing.

3.6 References

1. Ruszczak Z, Schwartz RA. Collagen uses in dermatology- an update. *Dermatology*. **1999**; 199: 285-289.

2. Ku S, Li W, Dulin W, Donahue F, Cadeb D, Benameur H, Hutchison K. Performance qualification of a new hypromellose capsule: Part I. Comparative evaluation of physical, mechanical and processability quality attributes of Vcaps Plus®, Quali-V® and gelatin capsules. *Int. J. Pharm*. **2010**; 386: 30-41.

3. Sikorski Z. *Chemical and functional properties of food proteins*.   CRC Press LLC. Boca Raton, FL. **2001**.

4. Dormont D. Prions, BSE and food. *Int. J. Food Micro*. **2002**; 78: 181-189.

5. US Food and Drug Administration. "Guidance for Industry: The Sourcing and Processing of Gelatin to Reduce the Potential Risk Posed by Bovine Spongiform Encephalopathy (BSE) in FDA-Regulated Products for Human Use". Sept 1977. Accessed June 14 2010 (Link: http://www.fda.gov/RegulatoryInformation/Guidances/ucm125182.htm).

6. Perret S, Merle C, Bernocco S, Berland P, Garrone R, Hulmes DJS, Theisen M, Ruggiero F. Unhydroxylated 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: 43693-43698.

7. Zhang C, Baez J, Pappu Kameshwari, 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**: 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. Olsen D, Yang CL, Bodo M, Chang R, Leigh S, Baez J, Carmichael D, Perala M, Hämäläinen ER, Jarvinen M, Polarek J. Recombinant collagen and gelatin for drug delivery. *Adv. Drug Delivery Rev.* **2003;** 55: 1547-1567.

18. Baez J, Olsen D, Polarek JW. Recombinant microbial systems for the production of human collagen and gelatin. *Appl. Microbiol. Biotechnol*. **2005**; 69: 245-252.

19. 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.

20. 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.

21. 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.

22. Werten M, Van den Bosch T, Wind R, Mooibroek H, De Wolf F. High-yield secretion of recombinant gelatins by *Pichia pastoris*.   *Yeast*. **1999;** 15: 1087-1096.

23. 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.

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

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

26. 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.

27. 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.

28. Setina C, Haase J, Glatz CE. Recovery of collagen Type I α1 from a whole-corn extracting using pepsin digestion and precipitation. (An unsubmitted paper) **2010**.

29. Aspelund MT, Glatz CE. Purification of recombinant plant-made proteins from corn extracts by ultrafiltration. *J. Membr. Sci*. **2010**; 353: 103-110.

CHAPTER 4

GENERAL CONCLUSION

Recombinant Type I α1human collagen (rCIα1) produced via microbial and plant expression systems can be purified from other components in the host cell extract using few simple steps. A model system where whole ground corn was extracted using buffer spiked with *Pichia*-derived rP-CIα1 (with and without foldon) simulated eventual corn grain expression levels and also minimized the extraction of host cell proteins (HCP's). Precipitation of rP-CIα1 without foldon from spiked extracts indicated that pH 2 produced the best yields while limiting HCP inclusion. From these extracts, rP-CIα1 without foldon could be completely precipitated with a purity >70% when >0.75 M NaCl was present. It was possible to improve purity to >95% when a fractional cut precipitation at 0.25 M NaCl was used to remove HCP's prior to rP-CIα1 precipitation with 1.0 M NaCl. Additional washing of the 1.0 M NaCl precipitate further increased purity to 96.5%, but both the fractional cut and the precipitate washing steps decreased yield.

When corn extracts containing rP-CIα1 "with foldon" were precipitated with salt, the foldon collagen preferentially precipitated between 0 and 0. 5 M NaCl. Above 0.5 M NaCl, "with foldon" precipitate yields were similar to non-foldon precipitates, but had much lower purities. Hydrolysis with pepsin then provided for removal of the foldon. When carried out on the spiked whole corn extracts prior to precipitation, the enzyme provided little benefit in yield or purity.

The fractional cut that improved purity of non-foldon extracts with a 0.25 M NaCl precipitation (prior to addition of 1.0 M NaCl to precipitate collagen) would decrease yield and purity for extracts spiked with rP-CIα1 with foldon. If fractional cut were used with corn extracts "with foldon", addition of 0.25 M NaCl would precipitate 60% of rP-CIα1 with foldon prior to the 1.0 M NaCl second precipitation. By comparison, high concentrations of HCP's in *Pichia* lysates also caused low precipitate purity of 16.8%. Ultrafiltration was necessary to increase purity to 70% and remove the green contaminant that a second precipitation step could not.

A recommended process therefore involves extraction at pH 2 and precipitation with ≥0.75 M NaCl, before redissolving the pellet in < 20% of the original extract volume. Removal of foldon necessitates pepsin addition. Concentrating rP-CIα1 in corn extracts prior to precipitation will reduce yield losses and improve purity.

APPENDIX

ANOVA ANALYSIS - EFFECTS TEST RESULTS

Table 1.   Tukey's least square mean difference effects test results for pH and NaCl concentration's effect on yield and purity

OUTPUT: YIELD

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
pH	2	2	10022.193	523.079	<.0001
NaCl Conc.	2	2	8666.245	452.3093	<.0001
pH*NaCl Conc.	4	4	6867.472	179.2138	<.0001

OUTPUT: PURITY

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
pH	2	2	0.70375417	246.9313	<.0001
NaCl Conc.	2	2	0.56552083	198.4284	<.0001
pH*NaCl Conc.	4	4	0.41282885	72.4261	<.0001

Table 2. Tukey's least square mean difference effects test results for pepsin addition (enzyme) and NaCl concentration's effect on yield and purity

OUTPUT: YIELD

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Enzyme	1	1	0.0000296	0.0039	0.9512
NaCl Conc	4	4	4.8534916	157.9875	<.0001
Enzyme*NaCl Conc	4	4	0.0694253	2.2599	0.1081

OUTPUT: PURITY

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Enzyme	1	1	0.0003522	0.4205	0.5259
NaCl Conc	4	4	2.1592758	644.5755	<.0001
Enzyme*NaCl Conc	4	4	0.0317465	9.4768	0.0004

Table 3. Tukey's least square mean difference effects test results for foldon presence (collagen type) and NaCl concentration's effect on yield and purity

OUTPUT: YIELD

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
collagen type	1	1	3360.667	210.4318	<.0001
NaCl Conc	4	4	22525.669	352.6174	<.0001
collagen type*NaCl Conc	4	4	7484.582	117.1638	<.0001

OUTPUT: PURITY

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
collagen type	1	1	59.724	6.6567	0.0179
NaCl Conc	4	4	10194.361	284.0573	<.0001
collagen type*NaCl Conc	4	4	4075.174	113.5513	<.0001