3 1st Generation PAA - Pro_mGlu_n Diblock*

(* contains material reprinted from Ref.1 with permission from the American Chemical Society)

3.1 Prior Work

This chapter focuses on our initial attempt to produce a novel surface-active poly(amino acid), or PAA, designed to form brush layers on aluminum oxide surfaces. Prior work has indicated that through recombinant DNA technology, PAA's can be produced biosynthetically with a higher level of control than current polymer synthesis methods.[2,3,4] The 20 naturally occurring amino acids possess a wide range of physio-chemical properties, allowing them to be used in the design of PAA's that interact with specific surfaces and/or solvents. The most common brush-forming polymers are diblock copolymers, which possess two different moieties referred to as the *anchor* and *tail* blocks. Figure 3.1 shows the brush layer formation by a diblock copolymer.

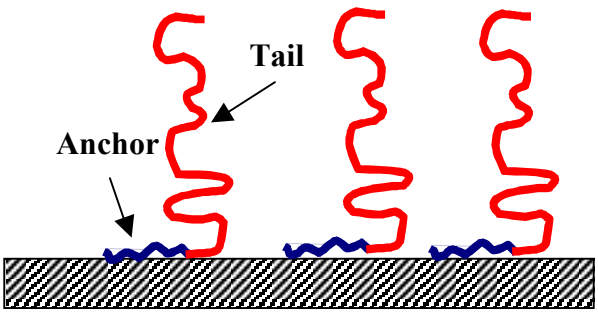

Figure 3.1 Schematic of brush-forming diblock copolymers.

In order to determine the best molecular composition with which to design our initial diblock PAA, homopolymer adsorption experiments were conducted.[5] Commercially available PAA homopolymers were used to identify potential candidates for the *anchor* and *tail blocks*. This experiment was conducted by suspending aluminum oxide particles in aqueous solutions with varying PAA homopolymer concentrations. The amount of polymer adsorbed was determined by comparing initial and final polymer concentrations in the supernatant. Polymer concentrations were determined using capillary zone electrophoresis (data not shown).

From this work, it was found that the acidic homopolymers of aspartate and glutmate adsorbed strongly in the pH range 7-9. This was to be expected since under these conditions the aluminum oxide surface has a net positive charge and these PAA's have a strong negative charge.[5] It was also found that homopolymers of proline did not

adsorb onto the aluminum oxide surface and were sufficiently soluble in water under the same conditions. During competitive adsorption experiments, weakly bound poly-proline was actually displaced by poly-glutamate. As a result of this work, it was determined that poly-glutmate and poly-proline were prime candidates for the anchor block and tail block, respectively.

3.2 Synthetic Gene Design

A gene was designed to encode for a PAA diblock copolymer with the composition Pro_mGlu_n (where m=40 and n=10). At this composition, the anchor block monomers comprise ~20mol% of the total polymer. It has been demonstrated theoretically and experimentally that optimum brush layer formation occurs at ~5mol% anchor monomer.[6] Due to the inherent limitations of the conventional recombinant DNA methods used, we were unable to attain this goal with our first attempt at designing a surface-active PAA.

In the bacterial genetic code, there are mulitple codons encoding for each amino acid. However, organisms sometimes prefer to use some codons more often than others.[7] Therefore, the coding sequence for Pro_mGlu_n was generated by randomly selecting from these 'high usage codons' to optimize protein translation. DNA restriction enzyme recognition sites flanking the sequence encoding for Pro_mGlu_n were included to facilitate the cloning of the synthetic gene into the chosen expression vector, pThioHisA. A schematic showing this cloning process is given in Figure 3.2.

Once the recombinant plasmid containing the Pro_mGlu_n gene was constructed, it was transformed into HS996 *E. coli*. Transformant colonies were selected on LB agar containing ampicillin and were used to inoculate small cultures from which plasmid could be isolated for further analysis. DNA sequencing was necessary to check for the presence and correct orientation of the synthetic gene.

DNA sequencing analysis determined that most transformant colonies contained a gene insert encoding for $Pro_{32}Glu_{10}$. The results of the DNA sequence analysis and the corresponding amino acid sequence is given in Figure 3.3 (where P = Pro; E = Glu).

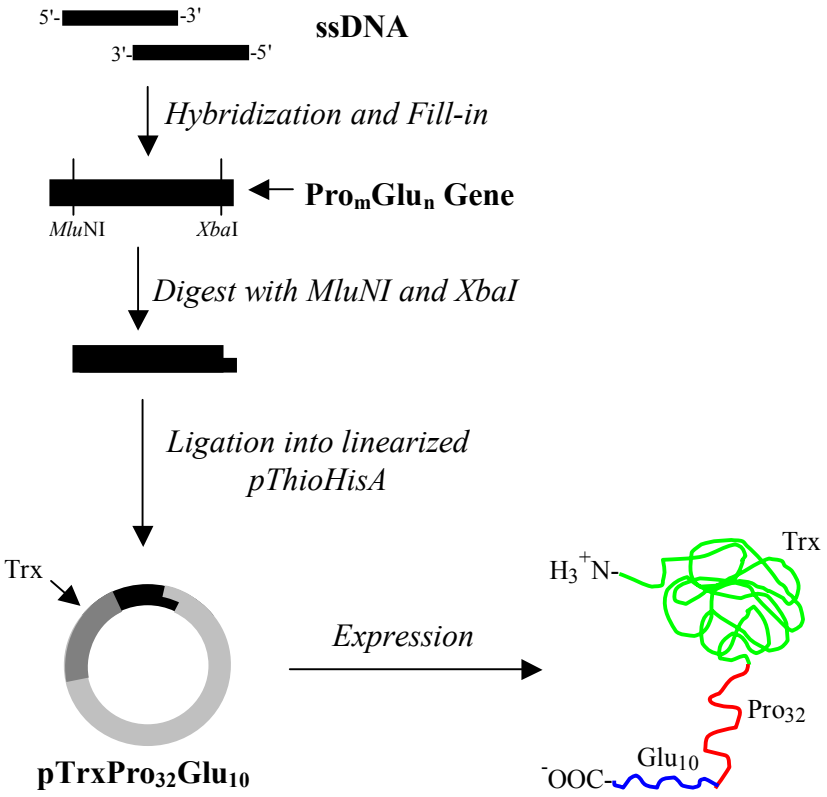

Figure 3.2 Schematic showing the construction of the Pro$_m$Glu$_n$ gene and the restriction enzyme cloning of the gene into a linearized expression vector forming the recombinant vector, pTrxPro$_{32}$Glu$_{10}$. (Adapted from Ref. 1; reprinted from Ref. 1 with permission from the American Chemical Society).

cct ccc cct cct cct cct cct cct cct cct cct cct cct cca cct ccg ccg cca ccc ccg cct cca cca cca cca cca cca cca

P P P P P P P P P P P P P P P P P P P P P P P P P P P P

cca cca cca cca gaa gaa gaa gaa gaa gaa gaa gaa gaa gaa

P P P P E E E E E E E E E E

Figure 3.3 DNA sequence of synthetic gene insert encoding for Pro$_{32}$Glu$_{10}$. (P=Pro; E=Glu).

The proline block is shorter than expected as a result of 'slippage' during gene construction. Due to long stretches of repeating nucleotides, the single-stranded oligonucleotides used to construct the double-stranded DNA molecule did not anneal perfectly in line, or 'slipped', resulting in a shorter proline block. It was decided that a construct consisting of a polypeptide sequence of $Pro_{32}Glu_{10}$ was sufficient for our initial studies and we continued on to the expression experiments. The resulting recombinant expression vector was denoted as $pTrxPro_{32}Glu_{10}$.

3.3 Expression System

It is important here to discuss the key aspects of the expression system we have chosen for producing our PAA construct. The expression vector pThioHisA (Invitrogen) was chosen due to its ability to aid in the expression, detection and purification of recombinant proteins. Using this expression system, the PAA construct is expressed with an amino-terminal His-Patch Thioredoxin (HP-Thioredoxin or Trx) fusion partner. The fusion protein product will then have the sequence: H_3^+N-Trx-$Pro_{32}Glu_{10}$-COO^-. This recombinant form of thioredoxin has been engineered to contain a 'histidine patch' which has been shown to have high affinity for divalent cations. Therefore, these fusions can be easily purified using metal chelating resins.[8] Trx is also beneficial in that it aids in the solubility of the fusion proteins, as well as in their detection by offering an epitope that is recognized by commercially available antibodies, such as Anti-Thio™ (Invitrogen). After isolation of the fusion protein, the Trx fusion partner is designed to be removed by enzymatic cleavage using EKMax™ (Invitrogen).

In addition, the chosen expression system provides a high level of fusion protein expression driven by the *trc* promoter. Isopropyl-β-*D*-thiogalactopyranoside (IPTG) is used to induce overexpression of the fusion protein. It has been shown that the fusion protein can be expressed to approximately 40% of the total cellular protein.[9] Hence, pThioHisA is a powerful expression vector that can be used to produce large amounts of soluble fusion protein which can be easily detected and purified.

3.4 Expression Experiments

Initial expression experiments were conducted at the shaker flask scale (~50 mL). Cultures were induced with IPTG at mid-log phase and incubated for 4-5 hrs before the cells were harvested by centrifugation. Cells were lysed and extract samples were analyzed by SDS-PAGE and Western blotting (Figure 3.4).

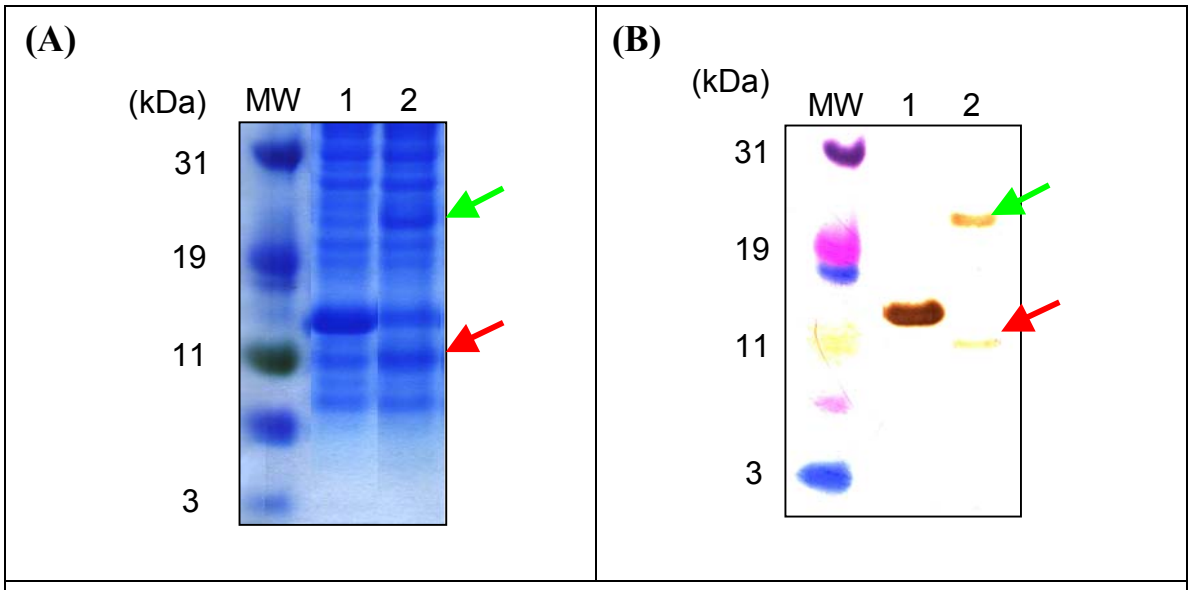

Figure 3.4 Trx-$Pro_{32}Glu_{10}$ expression analysis. (A) SDS-PAGE (NuPAGE® Novex 10% Bis-Tris polyacrylamide gels visualized by staining with Colloidal Blue Staining Kit). (B) Western blot (on PVDF membrane, using Anti-Thio™ antibody): Lane MW, MultiMark Multi-Colored Standard; Lane 1, cell extract from a positive expression control; Lane 2, cell extract containing Trx-$Pro_{32}Glu_{10}$ (top arrow).

The band in Lane 1 of the SDS-PAGE gel at ~13 kDa and its corresponding positive signal in Lane 1 of the Western blot, confirms the successful expression of a control protein also containing a Trx domain. In Lane 2 of the Western blot, it can be seen that two protein bands are detected. The top band represents a protein with a molecular weight (MW) of ~25 kDa. The lower band represents a protein with a MW of ~12 kDa. Trx has a MW of ~12 kDa and $Pro_{32}Glu_{10}$ is expected to have a MW of ~4.4 kDa. Therefore, the total fusion protein (Trx-$Pro_{32}Glu_{10}$) will have a MW of ~16.4 kDa. The top band was subsequently identified by amino acid analysis (data not shown) to be the fusion protein even though its apparent MW is higher than the expected value, and the lower band was identified as Trx without the $Pro_{32}Glu_{10}$.

It is not definitively known why the contaminating Trx protein is present in the cell extract. One theory is that a contaminating pThioHisA vector is present within the

expression strain that does not contain the $Pro_{32}Glu_{10}$ gene. Therefore, it expresses Trx without a fusion partner. This contaminating protein does not significantly interfere with the production of our PAA and can be subsequently removed in the downstream purification of the fusion protein.

Another theory is that the fusion protein is being digested by some unidentified protease. There are examples in the literature that suggest that protease activity increases during overexpression of a recombinant protein in *E. coli.*[10] For example, Matsuo *et al.* showed that OmpT, an outer membrane protease, and a homologue, OmpP, which is located on the commonly used F' episome, can complicate recombinant protein expression in some strains of *E. coli* by their protease activity.[11] The substrate specificity of these proteases is not fully understood, but it is believed that they preferentially cleave between two consecutive basic residues.[12] A corresponding site was not found in the fusion protein sequence. However, further work with the OmpP protease has shown that additional residues around the cleavage site may affect proteolysis.[13] Therefore, it can not be determined if proteolysis is occurring in this case, but the literature shows that it is a possibility.

The size discrepancy between the expected and apparent fusion protein MW also requires an explanation. During SDS-PAGE, SDS binds proteins, causing them to denature and giving them roughly the same charge-to-mass ratio.[14] The negatively charged protein-SDS complexes migrate through the polyacrylamide gel towards the positive cathode and are separated solely on their size. However, the $Pro_{32}Glu_{10}$ sequence likely does not bind SDS as effectively as 'normal' proteins causing it to migrate slower, resulting in a higher apparent MW.

Initial protein expression experiments proved to be efficient and relatively simple. For further analysis, culture growth was scaled-up to larger shaker flasks (500 mL). Larger cultures were grown using a New Brunswick BioFlo 3000 fermentor (2 L). Using this system, growth conditions (pH, temperature, mixing rate, aeration, etc.) could be monitored and controlled for optimum cell growth and protein expression. Once large amounts of protein were produced, several purification steps were required to isolate the Trx-$Pro_{32}Glu_{10}$ fusion protein from the native proteins expressed within the bacterial cell.

3.5 Purification

Cells harvested from expression cultures were lysed and the cell extract was obtained by centrifuging out the cell debris. The cell extract consists mostly of a complicated mixture of proteins. The knowledge of a protein's physiochemical properties can be exploited to aid in its purification. Fusion tags, like HP-Thioredoxin, are often used to supply the protein of interest with a specific trait that can be utilized to purify the protein in just one or two steps.

3.5.1 Immobilized Metal Affinity Chromatography

As previously mentioned, the 'histidine patch' on the Trx protein has a high affinity for divalent cations. This property is exploited to purify Trx fusions on resins containing immobilized divalent metal cations. This type of purification is known as immobilized metal affinity chromatography, or IMAC, and was used as the first step to purify our Trx-$Pro_{32}Glu_{10}$ product.

Cell extract in 20 mM sodium phosphate, 0.5 M NaCl, pH 7.8 was loaded onto a chromatography column containing a resin charged with Ni^{+2} ions. Most proteins did not interact with the column resin and simply passed straight through. Some proteins bound weakly and these could be washed away by lowering the buffer pH to 6. The Trx-$Pro_{32}Glu_{10}$ bound strongly to the column resin and could be subsequently eluted off by lowering the pH to 4. Hence, the Trx-$Pro_{32}Glu_{10}$ product was purified away from most of the proteins in the cell extract. Protein elution from the column was monitored by UV absorbance at 280 nm and peaks were collected with an automated fraction collector. The chromatogram in Figure 3.5A shows the elution profile of this purification step.

The SDS-PAGE gel in Figure 3.5B shows the protein content of each of the three main elution peaks seen in the chromatogram. As expected, the contaminating Trx protein coeluted with our Trx-$Pro_{32}Glu_{10}$ product from the IMAC column in the pH4 fraction. By comparing the gel lanes, many proteins have been purified away. However, further purification was required to remove additional proteins, as well as the Trx contaminant.

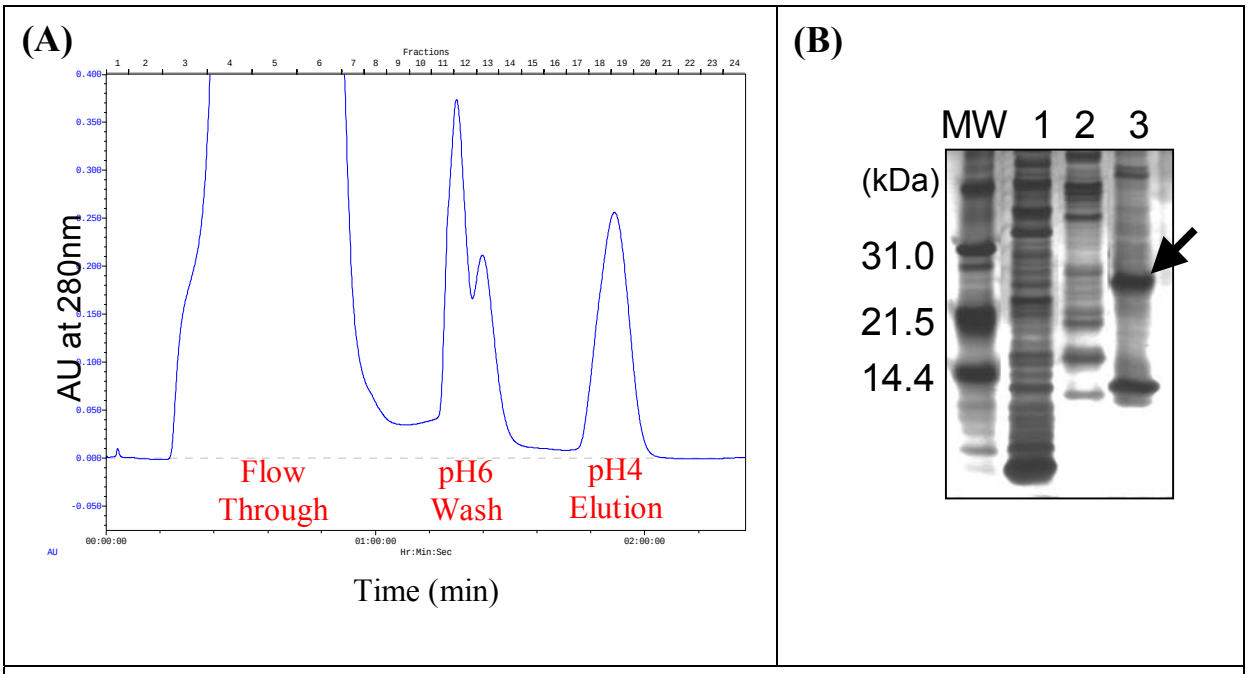

Figure 3.5 IMAC purification of the Trx-Pro$_{32}$Glu$_{10}$ product. (A) Chromatogram showing elution profile from Chelating Sepharose™ Fast Flow (Amersham Biosciences) with a pH step gradient. (B) SDS-PAGE analysis (NuPAGE® Novex 10% Bis-Tris polyacrylamide gels visualized by staining with silver) of column fractions: Lane MW, Silver Stain SDS-PAGE standards, Low Range ; Lane 1, Flow Through; Lane 2, pH6 Wash; Lane 3, pH4 Elution. Trx-Pro$_{32}$Glu$_{10}$ product indicated by arrow.

3.5.2 Anion Exchange Chromatography

During anion exchange chromatography, a positively charged resin is used to bind negatively charged proteins. The resin used here was Diethylaminoethyl (DEAE) Sepharose™ Fast Flow from Amersham Biosciences. Negatively charged proteins bind to this resin with different strengths, depending on their own charge strength. Bound proteins are eluted by increasing the salt concentration gradually, which decreases the ionic interactions.

Protein from the IMAC pH 4 fraction (including the Trx-Pro$_{32}$Glu$_{10}$ product) was loaded onto a DEAE Sepharose™ Fast Flow column in 20 mM Tris, pH 8.0. The pH is important since proteins must have a negative charge to interact with the column. Protein was eluted with a linear gradient of increasing NaCl concentration from 0 to 0.35 M. Again, protein elution was monitored by UV absorbance at 280 nm and peaks were collected with an automated fraction collector. The chromatogram in Figure 3.6A shows the elution profile of this purification step.

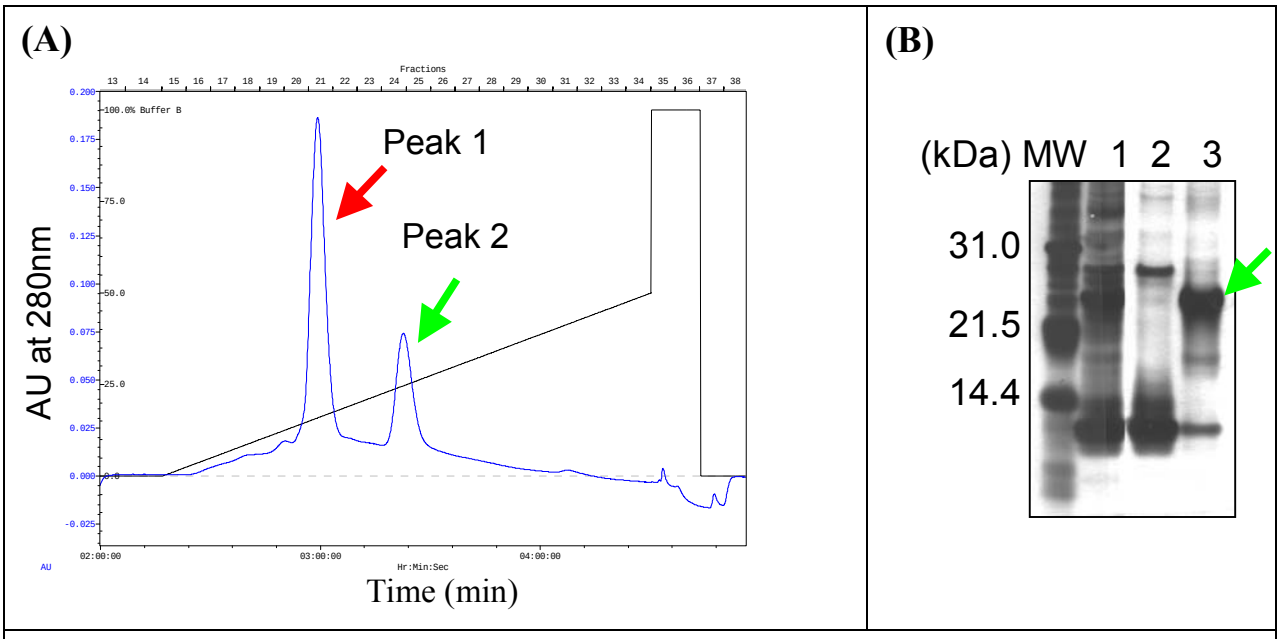

Figure 3.6 Ion exchange purification of the Trx-Pro$_{32}$Glu$_{10}$ product. (A) Chromatogram showing elution profile from DEAE Sepharose™ Fast Flow (Amersham Biosciences) with increasing NaCl gradient from 0-0.35 M. (B) SDS-PAGE analysis (NuPAGE® Novex 10% Bis-Tris polyacrylamide gels visualized by staining with silver) of column fractions: Lane MW, Silver Stain SDS-PAGE standards, Low Range; Lane 1, column feed; Lane 2, Peak 1 (Trx); Lane 3, Peak 2 (Trx-Pro$_{32}$Glu$_{10}$ product, indicated by arrow).

SDS-PAGE was used to analyze the two main protein peaks seen in the chromatogram (Figure 3.6B). The gel shows that we were successfully able to separate the contaminating Trx from the Trx-Pro$_{32}$Glu$_{10}$ product (indicated by arrow). The fraction containing the Trx-Pro$_{32}$Glu$_{10}$ product was determined to be sufficiently pure and no further purification was deemed necessary.

The samples containing the Trx-Pro$_{32}$Glu$_{10}$ product and the Trx contaminant were subsequently dialyzed separately against deionized water and lyophilized to dryness. Samples of each were then submitted to the Biomolecular Research Facility (Charlottesville, VA) for amino acid compositional analysis. As expected, the analysis concluded that the contaminant was indeed HP-Thioredoxin without the Pro$_{32}$Glu$_{10}$ insert (data not shown). The analysis also confirmed that the Trx-Pro$_{32}$Glu$_{10}$ sample had the correct amino acid composition.

3.6 Enzymatic Cleavage

To remove the Trx fusion tag, an enzymatic recognition site was engineered between the tag and the $Pro_{32}Glu_{10}$ sequence. Enterokinase is a highly specific enzyme that cleaves at the C-terminus of the peptide sequence Asp-Asp-Asp-Asp-Lys. Invitrogen offers a recombinant version of the catalytic subunit of enterokinase, known as EKMax™, which was used during this work.

Digestion of Trx-$Pro_{32}Glu_{10}$ was conducted using varying amounts of EKMax™ and incubation times to determine the optimum conditions. The digestion reaction was monitored by SDS-PAGE (Figure 3.7). Peptide fragments appear in some reaction samples indicating partial digestion of the Trx. A band corresponding to the $Pro_{32}Glu_{10}$ fragment may not appear because it is unclear whether it will stain with silver staining.

The gel also shows that a considerable amount of enzyme is required to digest the fusion protein completely. Even with 8 units of enzyme, some Trx-$Pro_{32}Glu_{10}$ remained uncleaved. Experiments were also conducted at longer incubation times with similar results. Literature suggests that proline residues directly downstream of the EKMax™ recognition site might prevent efficient digestion.[15]

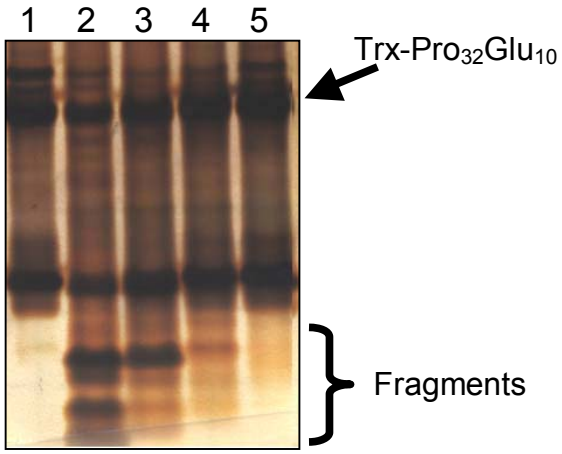

Figure 3.7 SDS-PAGE analysis of enzymatic digestion of Trx-$Pro_{32}Glu_{10}$ with varying amounts of EKMax™ after 16 hrs at 37°C. (NuPAGE® Novex 10% Bis-Tris polyacrylamide gels visualized by staining with silver): Lane 1, 0 u; Lane 2, 8 u; Lane 3, 4 u; Lane 4, 1 u; Lane 5, 0.1 u. Undigested Trx-$Pro_{32}Glu_{10}$ indicated by arrow. (u = units of enzyme).

3.7 Complications and Improvements

Our first attempt at producing a $Pro_{32}Glu_{10}$ diblock was met with several complications. The most glaring of these was the inability to efficiently cleave the Trx fusion partner from $Pro_{32}Glu_{10}$. A large amount of EKMax™ would have been required to produce sufficient amounts of $Pro_{32}Glu_{10}$. In addition to being inefficient, EKMax™ is very expensive. Literature shows that these types of fusion protein cleavage problems are notorious in recombinant protein expression. Therefore, other methods to release the Trx fusion partner were investigated.

Moving away from enzymatic digestion, we attempted to use chemical cleavage as a method to release our PAA construct from the fusion partner. Two of the most common techniques for chemically cleaving proteins use cyanogen bromide (CNBr) and hydroxylamine. CNBr cleaves proteins at Met-X bonds[16] and hydroxylamine cleaves Asn-Gly bonds.[17] In our research, chemical cleavage with hydroxylamine was chosen for its simplicity and its relatively mild reaction conditions. In addition, Asn-Gly pairings are relatively uncommon, making hydroxylamine cleavage more selective and a better method for cleaving fusion partners from proteins of interest than the CNBr method.

In order to utilize hydroxylamine as a means of releasing the Trx fusion tag, the engineering of a cleavage site, an Asn-Gly paring between the Trx and $Pro_{32}Glu_{10}$ domains, was required. It was also decided to add an extra glycine residue to increase flexibility around the chemical cleavage site. DNA restriction enzyme sites were also included to enable the cloning of the gene into pThioHisA. A schematic showing this cloning process is given in Figure 3.8.

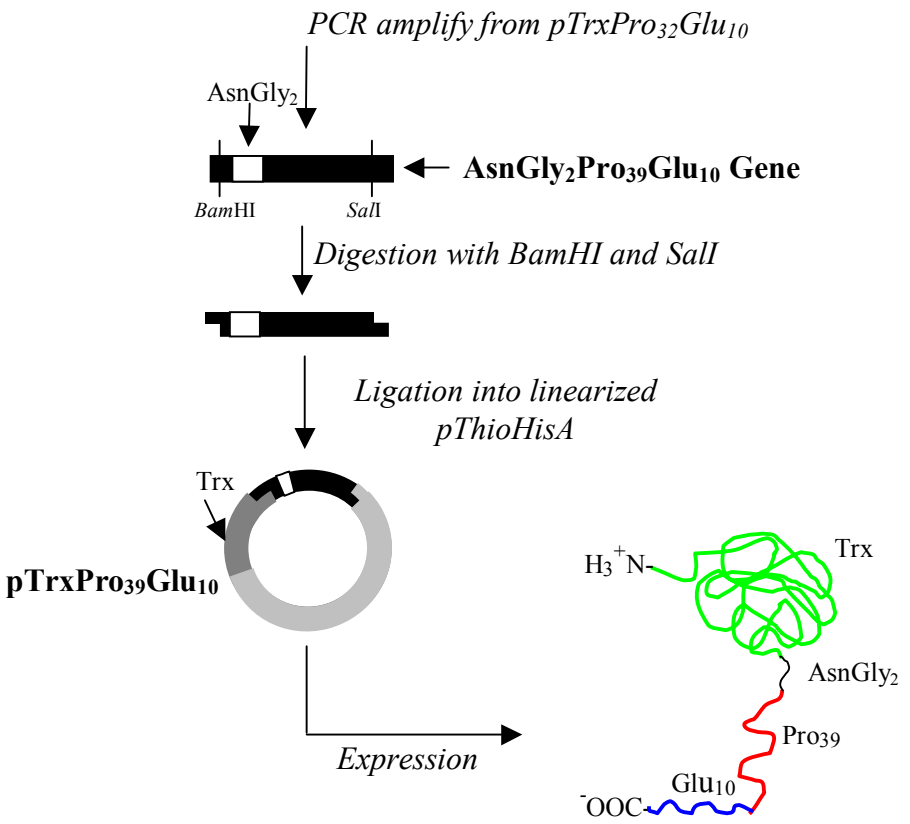

Figure 3.8 Schematic showing the subcloning of the Pro_mGlu_n gene into a linearized expression vector forming $pTrxPro_{39}Glu_{10}$. Note the incorporation of the peptide sequence: $AsnGly_2$ between the Trx and Pro_mGlu_n genes to facilitate fusion tag removal via hydroxylamine chemical cleavage. **(**Adapted from Ref. 1; reprinted from Ref. 1 with permission from the American Chemical Society).

This gene was constructed by amplifying the original $Pro_{32}Glu_{10}$ gene out of the $pTrxPro_{32}Glu_{10}$ vector using the polymerase chain reaction (PCR) with primers that included additional bases coding for the Asn-Gly-Gly sequence and restriction sites. The PCR product was restriction cloned back into pThioHisA and the new recombinant plasmid was transformed into TOP10F' *E. coli* and isolated, as previously, for DNA sequence analysis. The analysis results gave a DNA sequence encoding for a fusion protein product with the sequence: H_3^+N-Trx-Asn-Gly_2-$Pro_{39}Glu_{10}$-COO^- (see Section 3.13). The poly-proline tail block of the new PAA is actually longer than the initial template gene. This may be another example of 'slippage' that occurred during the PCR amplification. Fortunately, the longer block is more desirable for our brush forming requirement. The new recombinant expression vector was denoted as $pTrxPro_{39}Glu_{10}$.

3.8 Expression and Purification of the Trx-$Pro_{39}Glu_{10}$ Product

The new fusion product, which will be referred to as Trx-$Pro_{39}Glu_{10}$, was expressed and purified in the same manner as the first fusion, Trx-$Pro_{32}Glu_{10}$. Cell extract from a large-scale expression was loaded onto an IMAC column and the Trx-$Pro_{39}Glu_{10}$ product was eluted in the pH 4 fraction. Protein from this fraction was loaded onto a DEAE column to further purify the product using ion exchange chromatography (IEC). The gel in Figure 3.9 shows the progression of the purification of Trx-$Pro_{39}Glu_{10}$ using the two chromatographic techniques described.

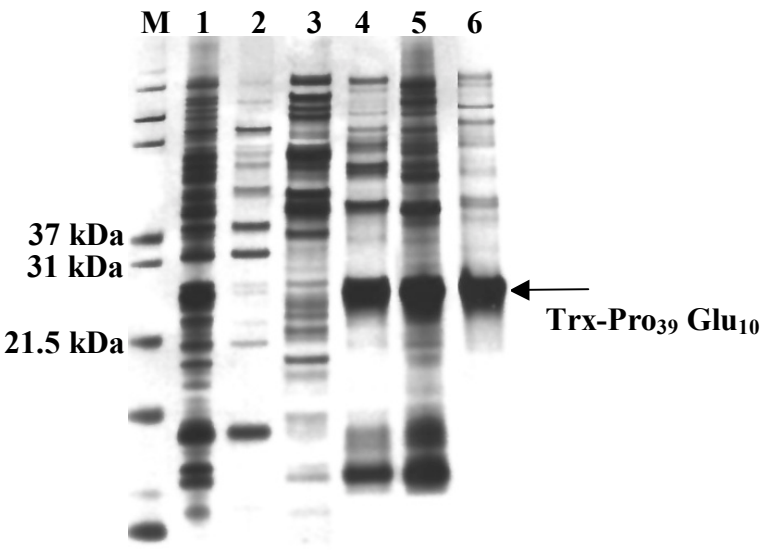

Figure 3.9 Silver stained SDS-PAGE gel showing purification of Trx-$Pro_{39}Glu_{10}$ using IMAC and IEC. Lane M, Mark12™ Unstained Standard; Lane 1, IMAC feed; Lane 2, IMAC flow-through; Lane 3, IMAC pH6 Wash; Lane 4, IMAC pH4 Elution; Lane 5, IEC feed; Lane 6, IEC eluate. The Trx-$Pro_{39}Glu_{10}$ product is indicated by arrow. (Adapted from Ref. 1; reprinted from Ref. 1 with permission from the American Chemical Society).

In Lane 6 of the gel in Figure 3.9, it can be seen that the Trx-$Pro_{39}Glu_{10}$ product (arrow) is relatively pure. The product migrates at ~28 kDa. This is higher than expected (~17 kDa) for the same reasons discussed previously for the Trx-$Pro_{32}Glu_{10}$ product ($AsnGly_2Pro_{39}Glu_{10}$ has a MW of ~5.3 kDa). The fraction containing the purified product was subsequently desalted by gel filtration and lyophilized to dryness. The dried Trx-$Pro_{39}Glu_{10}$ product was then re-suspended in deionized water to a concentration of ~5 mg/mL. The final expression yield was determined to be ~40 mg Trx-$Pro_{39}Glu_{10}$ product

per liter of culture. The purity of the Trx-$Pro_{39}Glu_{10}$ was determined by densitometry conducted on SDS-PAGE gels stained with Coomassie Brilliant Blue R250. As shown in Figure 3.10, it was determined that the polished Trx- $Pro_{39}Glu_{10}$ product was >98% pure.

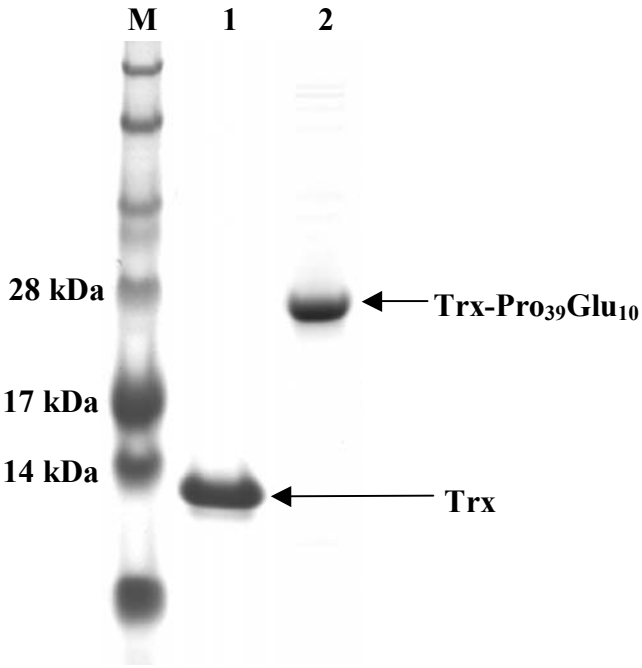

Figure 3.10. Product Purity Analysis. SDS-PAGE (NuPAGE® Novex 10% Bis-Tris polyacrylamide gels visualized by staining with Coomassie Brilliant Blue R250) used for purity analysis of products by densitometry. Lane M, SeeBlue® Plus2 Pre-Stained Protein Standard; Lane 1, polished Trx product (>98% pure); Lane 2, polished Trx-$Pro_{39}Glu_{10}$ product (>98% pure). **(**Adapted from Ref. 1; reprinted from Ref. 1 with permission from the American Chemical Society).

3.9 Confirmation of the Trx-$Pro_{39}Glu_{10}$ Product Identity

After the Trx-$Pro_{39}Glu_{10}$ was purified, it was necessary to confirm its identity. This was accomplished using two common biochemical characterization techniques: amino acid compositional analysis and MALDI-TOF mass spectrometry. The results are discussed below.

3.9.1 Amino Acid Compositional Analysis

The Trx-$Pro_{39}Glu_{10}$ composition was confirmed by amino acid compositional analysis. A lyophilized sample of the Trx-$Pro_{39}Glu_{10}$ product was submitted to Commonwealth Biotechnologies, Inc. (Richmond, VA) as well as a sample blotted on

PVDF. The analysis results agreed well with the expected composition based on the amino acid sequence (see Section 3.13). The results are tabulated in Table 3.1.

Residue	Dried Sample residues/mol	PVDF Blot residues/mol	Average residues/mol	Predicted residues/mol	Difference
As(x)	16	16	16	16	0
Gl(x)	17	15	16	14	-2
Ser	6	5	5	5	0
His	3	3	3	3	0
Gly	13	14	13	15	2
Thr	6	5	6	6	0
Ala	13	13	13	12	-1
Arg	4	2	3	1	-2
Tyr	3	2	2	2	0
Val	7	6	6	5	-1
Met	2	1	1	2	1
Phe	4	4	4	4	0
Ile	8	9	9	9	0
Leu	14	13	13	13	0
Lys	9	10	9	10	1
Pro	44	52	48	44	-4

Table 3.1 Results of amino acid compositional analysis conducted on the Trx-Pro$_{39}$Glu$_{10}$ product. As(x) = Asp + Asn; Gl(x) = Glu + Gln; Difference = (Predicted)-(Average).

There is some variability in results, but it is within acceptable limits. Because of the nature of our product, proline and glutamate were the amino acids of interest. The expected moles proline per mole Trx-Pro$_{39}$Glu$_{10}$ was detected, along with the expected amounts of Glx (Glu + Gln). Therefore, this analysis confirmed that the isolated product has the expected amino acid composition.

3.9.2 Product Polishing Using HP-IEC

To effectively analyze the Trx-Pro$_{39}$Glu$_{10}$ product by mass spectrometry (covered in next Section 3.9.3), an additional purification step was necessary. To accomplish this, high performance ion exchange chromatography (HP-IEC) was employed. After purification using the IMAC/IEC steps discussed in Section 3.8, a sample of the fusion product was loaded onto a Vydac DEAE 301 VHP column. Separation was obtained by using an increasing salt gradient and fractions were collected. A chromatogram showing the elution profile at 280 nm is shown in Figure 3.11.

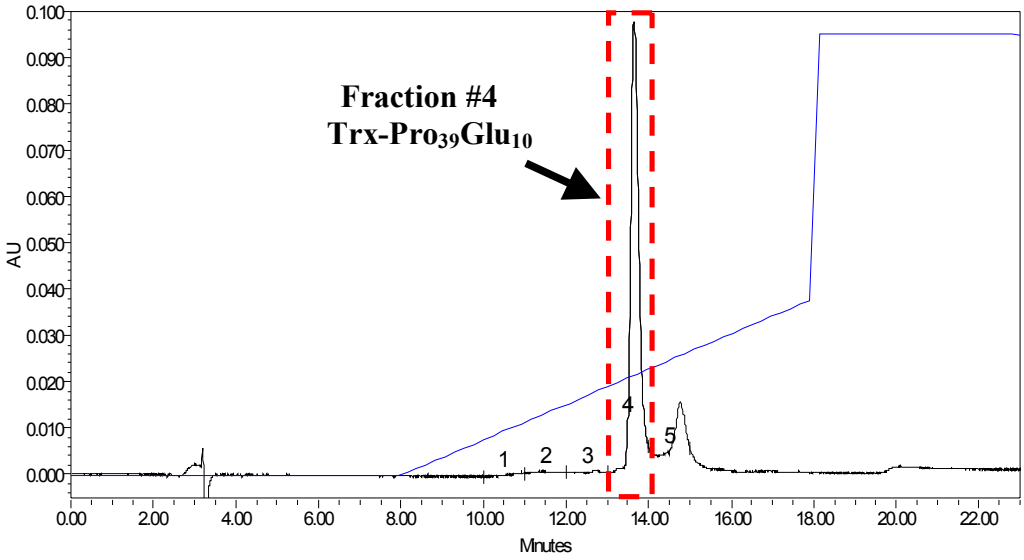

Figure 3.11 Chromatogram showing the HP-IEC purification of the Trx-Pro$_{39}$Glu$_{10}$ product. Amount loaded: 100 µL. Column: Vydac DEAE 301VHP with Vydac CPF10 prefilter. 5 µm, 900 Å, 7.5 mm ID x 50 mm L. Conditions: 24°C, load at 0.5 mL/min., UV absorbance at 280 nm. Mobile phase: A = 7.5 mM Tris-HCl in 25% acetonitrile, pH 8.0. B = 500 mM NaCl in Buffer A. Gradient: Linear 0 to 40%B over 10 min at 1 mL/min. Then held at 100%B for 4 min. Collection: 10-20 min, 1 min/tube.

The strong peak observed in Fraction #4 (outlined in box) is assumed to represent the highly purified Trx-Pro$_{39}$Glu$_{10}$ product. The fusion will absorb UV at 280 nm due to residues in the Trx domain (Trp, Tyr and Cys). The confirmation of the identity of this peak was obtained by MALDI-TOF MS and is discussed in the next section.

3.9.3 MALDI-TOF MS Analysis of the Trx-Pro$_{39}$Glu$_{10}$ Product

Mass spectrometry was used to confirm that the molecular weight of the purified Trx-Pro$_{39}$Glu$_{10}$ product agreed with the expected value calculated from the amino acid sequence. A sample of HP-IEC Fraction #4 (see Figure 3.11) was analyzed by MALDI-TOF Mass Spectrometry using an Applied Biosystems Voyager DE Pro in linear positive ion detection mode. The results are shown below in Figure 3.12.

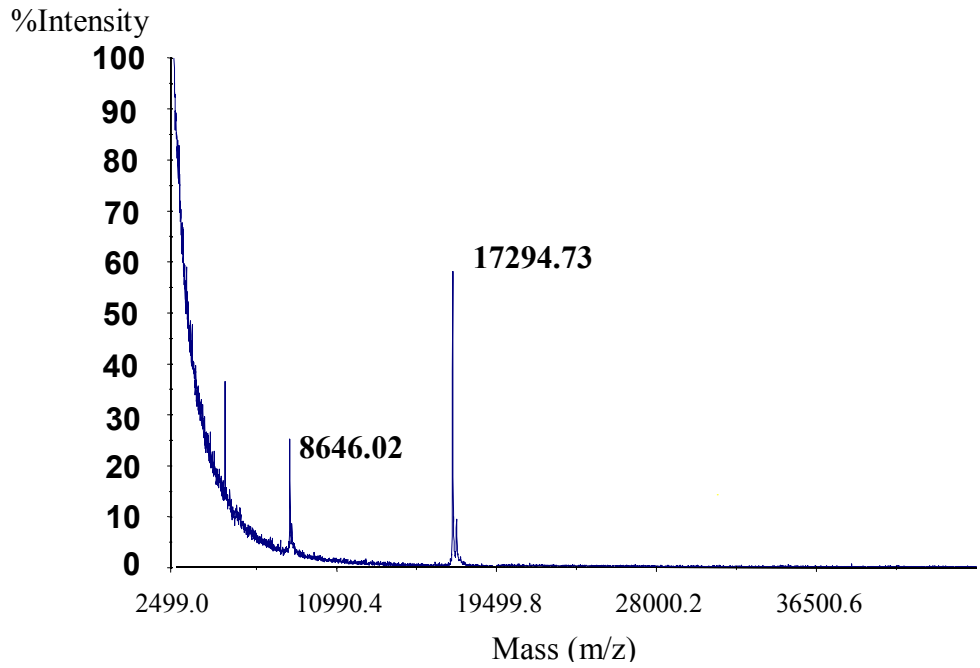

Figure 3.12 Results of MALDI-TOF MS analysis of the purified Trx-Pro$_{39}$Glu$_{10}$ product. Sample was diluted to 0.5 mg/mL then mixed with 1:1 with Sinapic Acid (Fluka) matrix solution (10 mg/mL Sinapic Acid in 50% acetonitrile/0.05% TFA). The plates were spotted with 1 uL of sample/matrix, and allowed to air dry. A minimum of 400 shots were taken to obtain the spectrum.

A strong signal was obtained at 17,294.73 m/z, along with the corresponding doubly charged ion at 8,646.02 m/z (Figure 3.12). This was in good agreement with the calculated $[M+H]^+$ m/z of 17,275.71 (MW_{avg}).

Identity	MALDI Results	Predicted[a] $[M+H]^+$, MW_{avg}	% Diff[b]
Trx-Pro$_{39}$Glu$_{10}$	**17,294.73 m/z**	**17,275.71 m/z**	**0.1%**

Table 3.2 Results from MALDI-TOF MS analysis of purified the Trx-Pro$_{39}$Glu$_{10}$ product.

[a] Molecular weight calculated using PeptideMass software
[b] Difference relative to predicted value

Therefore, using MALDI-TOF MS, we were able to confirm the molecular weight of our product to be ~17.3 kDa, as expected. This also supports our claim that we were seeing abnormal migration of our PAA products during SDS-PAGE (see Sections 3.4 and 3.8).

3.9.4 Conclusions from the Characterization Studies

From the results of the amino acid compositional analysis and MALDI-TOF MS, as well as the DNA sequence results of the pTrx-$Pro_{39}Glu_{10}$ vector, we confirmed that the highly purified product was in fact the target PAA fusion: a His-Patch Thioredoxin protein containing a C-terminal $Pro_{39}Glu_{10}$ polypeptide.

3.10 Isolation of the $Pro_{39}Glu_{10}$ Product

Once we had confirmed the successful production of purified Trx-$Pro_{39}Glu_{10}$, the Trx could be removed and the desired $Pro_{39}Glu_{10}$ product could be isolated. The following sections discuss the hydroxylamine chemical cleavage experiments, as well as the isolation of the $Pro_{39}Glu_{10}$ product.

3.10.1 Hydroxylamine Chemical Cleavage

The hydroxylamine cleavage reaction was first described by Bornstein and Balian[17] and its effectiveness as a method to cleave fusion proteins was first shown by Canova-Davis *et al.*[18] Cleavage of the Trx-$Pro_{39}Glu_{10}$ was carried out in a buffer containing 2 M hydroxylamineHCl, 2 M guanidineHCl and 0.2 M K_2CO_3 at pH9.0. The guanidine was included to act as a denaturant which has been show to improve yields.[19] The reaction mixture was incubated at 45°C for various time lengths. To stop the chemical cleavage, reaction components were dialyzed away against 20 mM Tris, pH9.0. A basic pH was used to ensure that the glutamate residues were ionized, thus promoting solubility of the $Pro_{39}Glu_{10}$ cleavage product. The cleavage reaction was monitored by SDS-PAGE analysis (Figure 3.13).

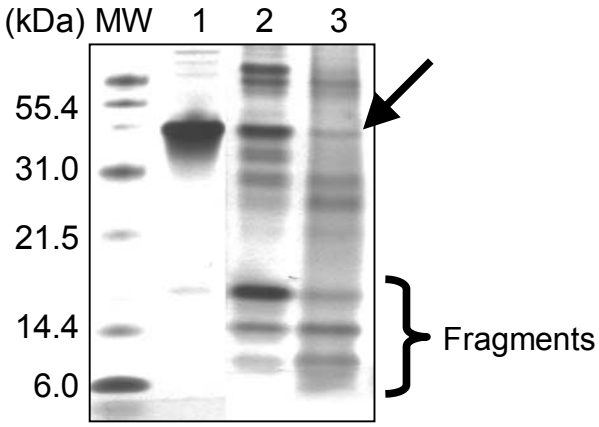

Figure 3.13 SDS-PAGE analysis of chemical cleavage of the Trx-$Pro_{39}Glu_{10}$ product with hydroxylamine. (NuPAGE® Novex 10% Bis-Tris polyacrylamide gels visualized by staining with silver): Lane MW, Mark12™ Unstained Standard; Lane 1, t=0 hrs; Lane 2, t=4 hrs; Lane 3, t= 24hrs. Uncleaved Trx-$Pro_{39}Glu_{10}$ indicated by arrow.

The gel in Figure 3.13 shows that the cleavage reaction progressed efficiently. After 4 hrs, the intensity of the Trx-$Pro_{39}Glu_{10}$ band (arrow) had decreased. The band at ~12 kDa suggests the appearance of the released Trx fusion partner. After 24 hrs, complete cleavage has occurred with very little fusion remaining, as well as the cleavage of the released Trx tag. Trx contains one internal Asn-Gly bond and additional bonds that are mildly susceptible to cleavage by hydroxylamine.[17,20] The location of the bonds that are susceptible to cleavage are shown in Table 3.3. Assuming 100% cleavage of these bonds, 5 fragments are expected, including the $Pro_{39}Glu_{10}$ product. The sequences and sizes of these fragments are given in Table 3.4. As previously discussed, a band representing $Pro_{39}Glu_{10}$ will not appear in Figure 3.13 as a result of its inability to be visualized with silver-staining. Therefore, we had to use other methods to locate and isolate the desired product.

MSD↓KIIHLTDDSFDTDVLKADGAILVDFWAHWCGPCKMIAPILDGIADGYQGK
LTVAKLNIDHNPGTAPKYGIRGIPTLLLFKN↓GEVAATKVGALSKGQLKEFLDAN
↓LAGSGSN↓GGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPEEEEEEEEEE

Table 3.3 Amino acid sequence of the Trx-$Pro_{39}Glu_{10}$ product and the location of the peptide bonds susceptible to cleavage by hydroxylamine (↓).

Peptide Fragment	Amino Acid Sequence	$[M+H]^+$ MW_{avg}
$Pro_{39}Glu_{10}$	GGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPEEEEEEEEEE	5211.83
Trx Fragment #1	MSD	352.38
Trx Fragment #2	KIIHLTDDSFDTDVLKADGAILVDFWAHWCGPCKMIAPILDGIAD GYQGKLTVAKLNIDHNPGTAPKYGIRGIPTLLLFKN	8834.29
Trx Fragment #3	GEVAATKVGALSKGQLKEFLDAN	2347.67
Trx Fragment #4	LAGSGSN	605.62

Table 3.4 Peptide fragments resulting from 100% cleavage of Trx-$Pro_{39}Glu_{10}$ by hydroxylamine. MW's were calculated using PeptideMass software.

3.10.2 Isolation of the $Pro_{39}Glu_{10}$ Product Using RP-HPLC

Reversed-phase high performance liquid chromatography (RP-HPLC) was used to isolate the $Pro_{39}Glu_{10}$ once it had been cleaved from the Trx fusion tag. A sample of the Trx-$Pro_{39}Glu_{10}$ fusion was cleaved with hydroxylamine for 20 hrs, stopped with 3vol of 2% TFA and loaded directly onto the RP column. The peptide fragments from the cleavage reaction were separated using an increasing acetonitrile gradient and fractions were collected. Chromatograms showing the elution profiles at 214 and 280 nm are given in Figure 3.14.

The $Pro_{39}Glu_{10}$ product does not contain amino acids that absorb UV at 280 nm (Trp, Tyr and Cys). However, due to the high absorptive energy of the peptide bond that connects each residue, $Pro_{39}Glu_{10}$ will absorb strongly at 214 nm. Therefore, we can identify the $Pro_{39}Glu_{10}$ product by comparing the chromatograms seen in Figure 3.14 and looking for peaks that appear in the 214 nm trace but not in the 280 nm trace. Using this method, the peak eluted at 13.1 min becomes a prime candidate for containing the $Pro_{39}Glu_{10}$ product (compare boxes highlighted on both traces). This peak was collected in Fraction #14. To confirm that this peak contained purified $Pro_{39}Glu_{10}$, we used MALDI-TOF MS and the results are presented in the next section.

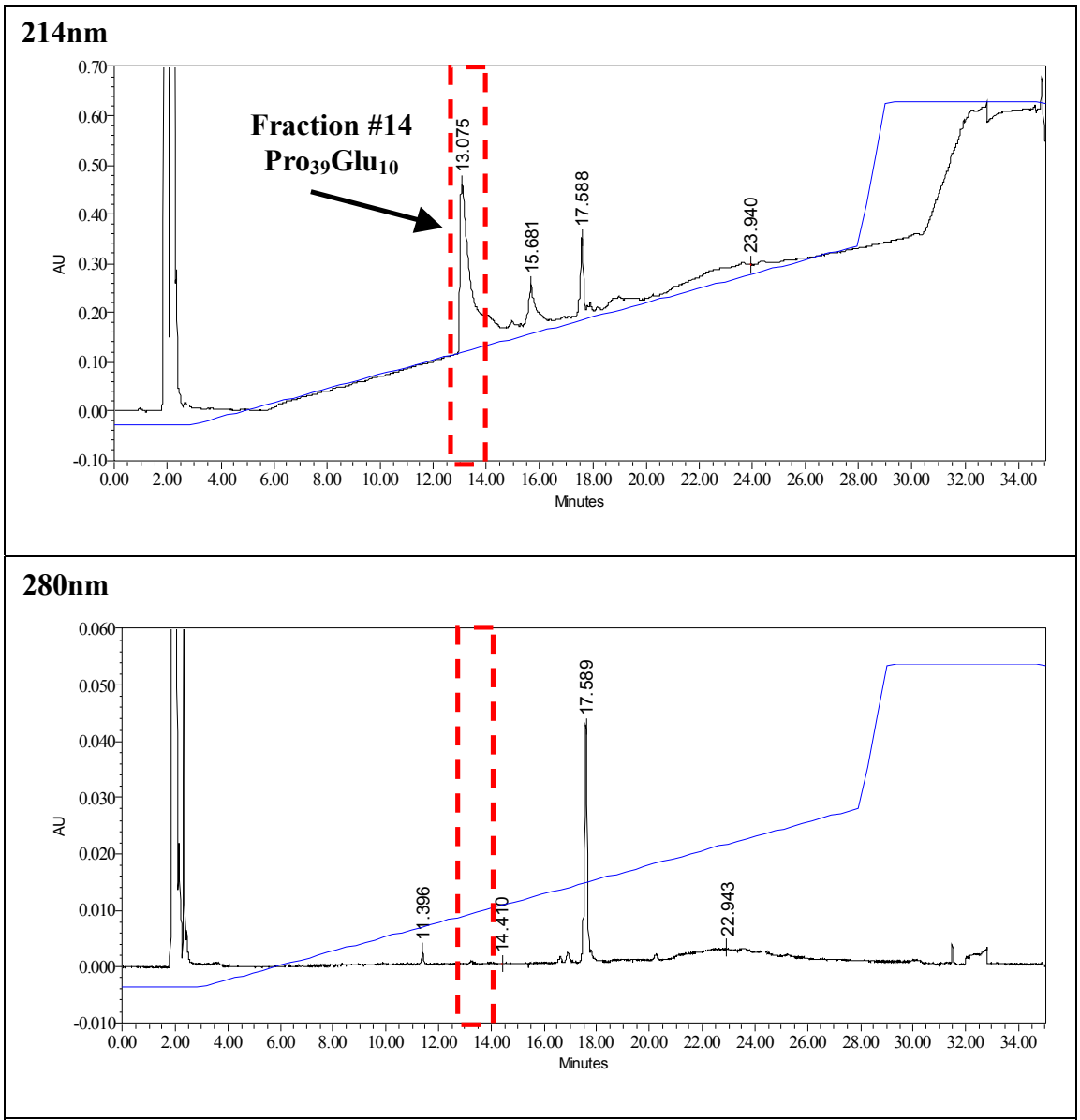

Figure 3.14 Chromatograms showing RP-HPLC purification of the $Pro_{39}Glu_{10}$ product from the hydroxylamine cleavage reaction. Sample: Hydroxylamine digest (t=20 hrs) of Trx-$Pro_{39}Glu_{10}$. Dried fusion was reconstituted in 40 µL hydroxylamine reaction buffer at 2.5 mg/mL and incubated 20 hrs at 45°C. Cleavage reaction was stopped with the addition of 3vol (120µL) of 2% TFA and cooled to -80°C. The final protein concentration was 0.625 mg/ml. Amount loaded: 75 µL. Column: Supelco Biowide C8 with Vydac CPF10 prefilter. 5 µm, 300 Å, 4.6 mm ID x 15 cm L. Conditions: 35°C, 1.0 mL/min., absorbance at 214 nm (top trace) and 280 nm (bottom trace). Mobile phase: A = 0.1% Formic Acid. B = 0.1% Formic Acid in acetonitrile. Gradient: Linear 5 to 55%B over 25 min. Then to 95%B over 1 min. Held at 95%B for 6 min. Collection: 0-25 min, 1 min/tube.

3.10.3 MALDI-TOF MS Analysis of the Isolated $Pro_{39}Glu_{10}$ Product

Mass spectrometry was used to confirm that Fraction #14 collected in Section 3.10.2 contained the isolated $Pro_{39}Glu_{10}$ product. A sample of the purified product was analyzed by MALDI-TOF Mass Spectrometry using an Applied Biosystems Voyager DE Pro in linear positive ion detection mode. The results are shown below in Figure 3.15

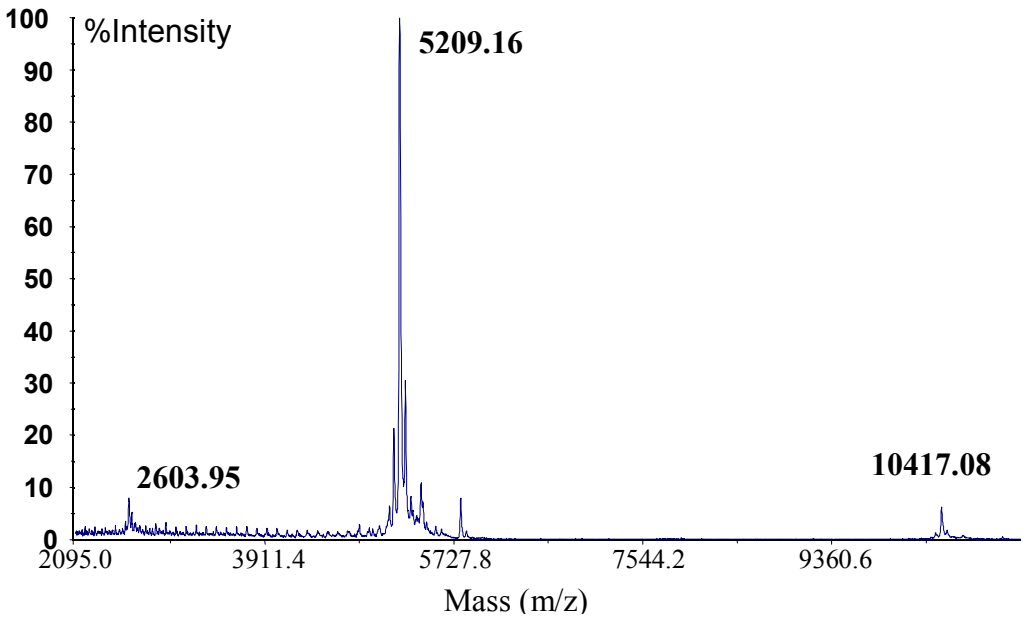

Figure 3.15 Results of MALDI-TOF MS analysis of the isolated $Pro_{39}Glu_{10}$ product. HPLC fraction was dried down and reconstituted with 100 µL deionized water. The sample was then mixed 1:1 with Sinapic Acid (Fluka) matrix solution (10 mg/mL Sinapic Acid in 50% acetonitrile/0.05% TFA). The plates were spotted with 1 uL of sample/matrix, and allowed to air dry. A minimum of 400 shots were taken to obtain the spectrum.

A strong signal was obtained at 5,209.16 m/z, along with the corresponding doubly charged ion at 2,603.95 m/z (Figure 3.15). This was in good agreement with the calculated $[M+H]^+$ m/z of 5,211.83 m/z (MW_{avg}). The peak at 10,417.08 m/z likely corresponds to a dimer of the $Pro_{39}Glu_{10}$ that forms during the sample preparation and ionization process. It should also be noted that to the left of the main peak, there are a series of smaller peaks that appear at 96 m/z intervals. Interestingly, 96 Da is the residue mass of proline. We believe that with increasing exposure to the laser (during the desorption/ionization step), proline residues are being broken off the N-terminus, one at a time (post-source decay). In a sense, we are sequencing the polypeptide and confirming the presence of the proline block.

Identity	MALDI Results	Predicted[a] $[M+H]^+$, MW_{avg}	% Diff[b]
$Pro_{39}Glu_{10}$	5,209.16 m/z	5,211.83 m/z	0.05%

Table 3.5 Results from MALDI-TOF MS analysis of the purified $Pro_{39}Glu_{10}$ product.

[a] Molecular weight calculated using PeptideMass software
[b] Difference relative to predicted value

The MS analysis of the HPLC fraction confirms that we have purified the $Pro_{39}Glu_{10}$ product. We have also validated hydroxylamine chemical cleavage as an efficient method for the removal of fusion tags. More importantly, these results indicate that we have succeeded in producing and isolating a specifically designed PAA.

3.11 Conclusions

The work presented in this chapter describes the design, production and isolation of a unique PAA. The successful completion of this work was important because it confirmed the feasibility of the following:

- The biosynthetic production of a unique PAA specifically designed for brush-forming applications.
- Using hydroxylamine cleavage as a simple method for the removal of fusion tags.

The successful biosynthetic production and isolation of the $Pro_{39}Glu_{10}$ product is a significant accomplishment. This is because this is the first biosynthetically produced PAA specifically designed for brush-forming applications.

Unfortunately, we were unable to produce sufficient amounts of a pure Pro_mGlu_n construct for surface adsorption studies to characterize its ability to form brush layers. This was due to limited production and purification facilities at the present time. In the future, we hope get access to the necessary equipment to scale-up the production of the $Pro_{39}Glu_{10}$ product for surface adsorption studies.

As an extension the work presented in this chapter, we became interested in the brush-forming abilities of a PAA consisting of: an anchor block, a linker region and a 'capping domain'. Trx-$Pro_{39}Glu_{10}$ acts as a convenient model for this system where the proline block becomes the linker region that tethers the capping domain (Trx) to the adsorbing acidic anchor block. This work is covered in the Chapter 4.

During the Pro_mGlu_n work it also became evident that a different cloning approach was necessary to make an effective surface-active PAA. For sufficient brush formation, much higher molecular weight PAA's are required than the conventional recombinant DNA methods used in this chapter can provide. Therefore, a new cloning strategy was required and subsequently developed to correct for some of the limitations that were present in this first generation PAA construct. This work is presented in Chapters 5 and 6.

3.12 Materials and Methods

Materials. HS996, TOP10, One Shot® TOP10F' Chemically Competent *E. coli* strains, His-Patch ThioFusion™ Expression System, NuPAGE® Novex 10% Bis-Tris polyacrylamide gels, Anti-Thio™ antibody, EKMax™, LB Broth Base, UltraPure™ agarose, UltraPure™ ethidium bromide solution, MultiMark® Multi-Colored Protein Standard, Mark12™ Unstained Protein Standard, SeeBlue® Plus2 Pre-Stained Protein Standard, Colloidal Blue Staining Kit, synthetic oligonucleotides and DNA primers were purchased from Invitrogen™ (Carlsbad, CA). *Taq* PCR Core Kit, QIAprep® Spin Miniprep and QIAquick® Gel Extraction/PCR Purification kits were purchased from Qiagen, Inc. (Valencia, CA). LB Agar, CelLytic™ B Bacterial Cell Lysis Extraction Reagent, Coomassie Brilliant Blue R250 and Goat Anti-Mouse HRP-conjugate antibody were obtained from Sigma (St. Louis, MO). Ampicillin, isopropyl β-*D*-thiogalactopyranoside (IPTG) and 0.2 μm pre-sterilized syringe filters were purchased from VWR International (So. Plainfield, NJ). Silver Stain SDS-PAGE Protein Standard (Low Range) and Immun-Blot™ PVDF membrane were purchased from Biorad (Hercules, CA). Chelating Sepharose™ Fast Flow, DEAE Sepharose™ Fast Flow and Sephadex™ G-25 Coarse were obtained from Amersham Biosciences (Piscataway, NJ) for chromatography applications. Metal Enhanced DAB Substrate Kit, SnakeSkin® Pleated Dialysis Tubing (3.5 and 10 kDa molecular weight cutoff) and Slide-A-Lyzer® Dialysis Cassettes (3.5 kDa molecular weight cutoff) were purchased from Pierce (Rockford, IL). All restriction enzymes, T4 DNA Ligase, 100 bp DNA Ladder and 1 kbp DNA Ladder were all purchased from New England Biolabs (Beverly, MA).

General Methods. The procedures for the manipulation of DNA, transformation, cell growth, protein expression and analysis were adapted from published literature[21, 22] or from instructions provided by manufacturers. All reagents for the manipulation of DNA were sterilized and DNAse/RNAse free. Enzymatic manipulations of DNA were conducted in reagent buffers supplied by the manufacturer. PCR amplification and DNA extension reactions were performed in an Omnigene thermal cycler from Hybaid (United Kingdom). Automated DNA sequence analysis was performed on a Perkin-Elmer ABI Prism model 377 DNA Sequencer at Cleveland Genomics, Inc. (Cleveland, OH). Cells were lysed using a Tekmar Sonic Disruptor with a microtip sonicator (Mason, OH). Product analysis was conducted by SDS-PAGE using pre-cast NuPAGE Novex 10% Bis-Tris polyacrylamide gels on an Xcell Surelock Mini-Cell apparatus from Invitrogen™. The gels were visualized using a silver staining protocol adapted from published literature[23] or by staining with Coomassie Blue. Polyacrylamide and agarose gel images were captured using a Microtek ScanMaker X6EL scanner and a Polaroid GelCam, respectively. Polypeptide product and DNA concentrations were calculated by optical density (OD) using a Milton Roy Spectronic 1201 UV spectrophotometer (Ivyland, PA). All chromatographic separations were conducted on a BioRad BioLogic DuoFlow system with 280 and 214 nm detection.

Construction of Pro_mGlu_n gene (Figure 3.2). Two single-stranded oligonucleotides of 107 bases each with complimentary 3' ends and an ultimate coding capacity of $Pro_{40}Glu_{10}$ were purchased. To construct the gene, the sense and antisense oligonucleotides were mixed and then heated to 95°C. The mixture was cooled down slowly to 5°C below the estimated melting temperature (T_m) of the oligonucleotides and held for 5 min to anneal the complimentary bases. A solution of dNTPs was added, followed by the addition of *Taq* DNA Polymerase to begin the 'fill-in' reaction. Agarose electrophoresis showed that the average product size was ~180 base-pairs, as desired (data not shown). A *Mlu*NI site was included at the 5' end of the synthetic gene, just upstream of the proline codons and an *Xba*I site at the 3' end, downstream of the glutamate codons. The product was then digested with *Mlu*NI and *Xba*I. The *Mlu*NI digest resulted in a blunt end at the 5' end. The *Xba*I digest allowed for convenient insertion into the vector, pThioHisA. This vector was sequentially digested with *Asp*718I and Mung Bean Nuclease to form a blunt end, followed by digestion with *Xba*I. The Pro_mGlu_n gene was ligated into the linearized vector with T4 DNA Ligase and transformed into HS996 *E. coli*. Transformants were selected on LB agar containing 100 μg/mL ampicillin. Thirty colonies were picked and plasmids were isolated from each using a standard alkaline lysis protocol.[21] The recombinant plasmids were digested with *Sma*I and *Xba*I to identify clones with inserts of the correct size. Plasmids from 2 positive clones were purified and were submitted for sequence analysis. The results indicated a sequence containing a gene encoding for the amino acid sequence $H_3^+N\text{-}Pro_{32}Glu_{10}\text{-}COO^-$ fused to the C-terminus of the Trx fusion protein (Trx-$Pro_{32}Glu_{10}$). Hence, this plasmid was denoted

as pTrxPro$_{32}$Glu$_{10}$. Stocks were made of the positive clone by adding glycerol to a final volume of 15% and these were stored at -80°C.

Re-engineering of Pro$_m$Glu$_n$ Gene (Figure 3.8). The Pro$_{32}$Glu$_{10}$ gene was PCR subcloned to include an Asn-Gly-Gly tripeptide between the C-terminus of the thioredoxin fusion protein (Trx) and the N-terminus of the Pro$_m$Glu$_n$ polypeptide. This tripeptide sequence was designed to provide the option to remove the HP-Thioredoxin fusion tag by chemical cleavage with hydroxylamine, which cleaves the peptide bond between Asn-Gly.[17,18] The Pro$_m$Glu$_n$ gene was PCR amplified from the recombinant vector using the following primer set: [Forward: 5'-attaggatccaacggtggtcctccccctcctcctcct-3'; Reverse: 5'-gtcgactctagagctattcttcttc-3']. The amplifying primers also contained restriction enzymes sites for *Bam*HI and *Sal*I for convenient cloning into the expression vector, pThioHisA. Using *Taq* DNA polymerase and an annealing temperature of 62°C, a new double-stranded Pro$_m$Glu$_n$ gene was amplified. After purification, the PCR product and expression vector were digested with *Bam*HI and *Sal*I overnight (~16 hrs) at 37°C and ligated in a 3:1 insert:vector molar ratio with T4 DNA Ligase overnight (~16 hrs) at 16°C. The ligation reaction was transformed into One Shot® TOP10F' Chemically Competent *E. coli* and selected on LB agar plates containing 100 µg/mL ampicillin. Colonies were selected and used to inoculate 5 mL of LB Broth containing 100 µg/mL ampicillin and incubated for 16 hours at 37°C. Plasmids were isolated using the QIAprep® Spin Miniprep kit. Plasmids from several colonies were submitted for DNA sequencing (Cleveland Genomics, Inc.) using Trx sequencing primers to validate gene sequence and orientation. The analysis yielded the following sequence downstream of the Trx gene: 5'- aac ggt ggt cct ccc cct cct cct cct cct ccc cct cct cct cct cct cct cct cct cct cct cca cct ccg ccg cca ccc ccg cct cca cca cca cca cca cca cca cca cca cca gaa gaa gaa gaa gaa gaa gaa gaa gaa gaa gaa tag -3'. This gene encodes for the amino acid sequence H_3^+N-AsnGly$_2$Pro$_{39}$Glu$_{10}$-COO$^-$ which is fused to the C-terminus of the Trx fusion protein (Trx-Pro$_{39}$Glu$_{10}$). Hence, this plasmid was denoted as pTrxPro$_{39}$Glu$_{10}$. Stocks were made of the positive clone by adding glycerol to a final volume of 15% and these were stored at -80°C.

Product Expression. In a 250 mL culture flask, 50 mL of LB Broth containing 100 µg/mL ampicillin was inoculated with stock culture containing the recombinant vector (pTrxPro$_{32}$Glu$_{10}$ or pTrxPro$_{39}$Glu$_{10}$). The culture was incubated overnight at 37°C with shaking at 200 rpm. The culture was pelleted in a 50 mL centrifuge tube at 3000 X ***g*** for 10 min at 4°C. The supernatant was removed and the pellet was re-suspended in 10 mL of fresh LB media. Four 1 L culture flasks with 250 mL of fresh LB media containing 100 µg/mL ampicillin were inoculated with 2 mL of this cell suspension. The culture was incubated at 37°C with shaking at 200 rpm until mid-log phase (OD_{600} ~ 0.5). Expression of product was induced by the addition of IPTG to a final concentration of 1 mM, and the cultures were incubated for an additional 5-6 hours. The cells were harvested by centrifuging at 3000 X *g* for 10 min at 4°C, and the pellets were stored at -80°C. This procedure was repeated for the expression of HP-thioredoxin from a TOP10 strain containing the pThioHisA control expression vector that is available with the pThioHis expression kit from Invitrogen™.

Cell Lysis. Approximately 1 mg of wet cell pellet was re-suspended in 15 mL of CelLytic™ B Bacterial Cell Lysis Extraction Reagent. The cell suspension was sonicated with a Tekmar Sonic Disruptor with a microtip sonicator (Mason, OH) with four 10 sec bursts while keeping the samples on ice. The samples were then frozen at -80°C followed by a rapid thaw in a 42°C water bath. This sonication/freeze/thaw procedure was repeated 4 times. After the final round, an appropriate volume of 5X IMAC Binding Buffer was added to make a 1X solution (20 mM sodium phosphate, 0.5 M NaCl, pH 7.8). Cell debris was centrifuged out at 13,000 X g for 15 min at 4°C. The soluble cell extract containing the product was collected and cell debris pellet was discarded.

Immobilized Metal Affinity Chromatography Purification. Purification of the product from the soluble cell extract was accomplished by immobilized metal affinity chromatography (IMAC) using a column packed with Chelating Sepharose™ Fast Flow (8 cm x 2.5 cm i.d.) charged with Ni^{2+} (100 mM $NiCl_2$) and equilibrated in 1X IMAC Binding Buffer. Cell extract in 1X IMAC Binding Buffer was loaded onto the column at a flow rate of 0.8 cm/min. The column effluent was monitored at 280 or 214 nm. The column was washed with 2 column volumes (CV) of 1X IMAC Binding Buffer. Weakly bound proteins were eluted with 4 CV of 20 mM sodium phosphate, 0.5 M NaCl pH 6. The Trx fusion product was eluted with 4

CV of 20 mM sodium phosphate, 0.5 M NaCl pH 4. The pH of the eluate sample was adjusted to pH~7 with 4 M NaOH. Following elution, the column resin was re-equilibrated with 4 CV of 1x IMAC Binding Buffer. All steps were conducted at a linear flow rate of 0.8 cm/min. The fractions containing the product were pooled together. Product concentration was estimated by measuring the OD at 280 nm (assuming an extinction coefficient of 0.834 L/g·cm calculated based on the amino acid sequence of the product[24]). Samples were stored at -80°C. This procedure was also repeated for the purification of HP-thioredoxin from the control expression vector. The products were dialyzed against deionized water using Snakeskin® Pleated Dialysis Tubing (3.5 kDa molecular weight cutoff), lyophilized to dryness, and stored at -80°C.

Ion Exchange Chromatography Purification. Ion exchange chromatography (IEC) was used to separate the Trx fusion product from other proteins that eluted from the IMAC column at pH 4. The dried sample was re-suspended in 1X IEC Loading Buffer: 20 mM Tris-HCl, pH 8.0 to a concentration of ~0.50 mg/mL. This solution was loaded onto a column packed with Diethylaminoethyl (DEAE) Sepharose™ Fast Flow (8 cm x 2.5 cm i.d.) that had been equilibrated with 1X IEC Loading Buffer. Following sample loading, the column was washed with 2 CV of 1X IEC Loading Buffer. Bound proteins were eluted with a linear gradient from 0-35% 20 mM TrisHCl, 1 M NaCl pH 8.0 over 10 CV at 0.5 cm/min. The column resin was regenerated with 3 CV of 20 mM sodium phosphate, 4 M NaCl pH 8.0, followed by re-equilibration with 1X IEC Loading Buffer. All steps with the exception of the salt gradient were conducted at a linear flow rate of 1.2 cm/min. The column effluent was monitored at 280 or 214 nm. The fractions containing the product were pooled and concentration was measured by OD at 280 nm. The purified sample was then dialyzed and lyophilized as above.

Product Polishing. Dried, purified product was re-suspended in deionized water to a concentration of ~1 mg/mL. The solution was loaded (0.25 CV) onto a column packed with Sephadex™ G-25C desalting resin. The column effluent was monitored at 280 or 214 nm. The fractions containing the desalted product were pooled together and the concentration was measured by OD at 280 nm. The sample was frozen at -80°C and lyophilized to dryness. The product was then re-suspended in deionized water to a concentration of ~5 mg/mL. The purity and identity of the products were confirmed by SDS-PAGE and amino acid analysis.

Product Purity Analysis. Product purity was determined using densitometry of SDS-PAGE gels. The purified sample was loaded onto an SDS-PAGE gel and stained with Coomassie Brilliant Blue R250. The gel was then scanned using a Biorad GS-800 Imaging Densitometer. Using the accompanying software Quantity One®, the lane containing the product sample was analyzed for stained protein bands. The relative intensity of the bands were determined and compared to obtain the purity of the sample.

Amino Acid Compositional Analysis. Product samples were characterized by amino acid compositional analysis. Approximately 10 µg of purified product sample was loaded onto 3 lanes of a SDS-PAGE gel. After electrophoretic separation, the sample was transferred to PVDF membrane for 1.5 hrs at 30 volts. After the transfer, the membrane was washed extensively with deionized water. The immobilized product was stained with a 0.025% Coomassie Brilliant Blue R250 solution in 40% methanol for 5 min. The blot was de-stained in 50% methanol, followed by a deionized water wash. The blot was then dried and the bands corresponding to the product were cut out. The bands were submitted to Commonwealth Biotechnologies, Inc. (Richmond, VA) for amino acid compositional analysis. Lyophilized protein samples were also submitted.

Enzymatic Cleavage. EKMax™ (Invitrogen™) was used to remove the HP-Thioredoxin fusion partner. The procedure was supplied by the manufacture. 20 µg of $TrxPro_{32}Glu_{10}$ product at a concentration of 1 mg/mL was added to several digestion reaction mixtures containing various amounts of EKMax™ (0.0, 0.001, 0.01, 0.1, 1.0, 4.0 and 8.0 units) to determine optimum conditions. Reactions were incubated at 37°C and samples were taken at 0 hrs, 20 hrs and 40 hrs. The digestion reactions were monitored by SDS-PAGE and visualized by staining with silver or Colloidal Coomassie Blue.

Hydroxylamine Chemical Cleavage. Chemical cleavage procedure was adapted from published literature.[25] $TrxPro_{39}Glu_{10}$ product was re-suspended in Reaction Buffer (2 M hydroxylamineHCl, 2 M guanidineHCl, 0.2 M K_2CO_3 pH 9.0). The reaction mixture was incubated at 45°C for a specified amount

of time. To stop the reaction, the reaction components were dialyzed away using Slide-A-Lyzer® Dialysis Cassettes (3.5 kDa molecular weight cutoff) against 20 mM Tris pH 9.0 at 4°C for 24 hrs or stopped with 3vol of 2% TFA. The protein samples were collected and stored at -80°C. A control reaction was run simultaneously exactly as above without the inclusion of hydroxylamineHCl in the Reaction Buffer. The cleavage reactions were monitored by SDS-PAGE and visualized by staining with silver or Colloidal Coomassie Blue.

MALDI-TOF Mass Spectrometry. Samples were analyzed by MALDI-TOF Mass Spectrometry using an Applied Biosystems Voyager DE Pro (Nebraska Center for Mass Spectrometry at the University of Nebraska-Lincoln). Spectra were acquired manually in linear positive ion detection mode, with an accelerating voltage of 25 kV and a 300 nsec extraction delay time. A minimum of 400 shots per samples were used to obtain each final spectrum. Spectra y-axis represents %Intensity normalized to strongest peak. Spectra are presented without any smoothing. Before analysis, samples were mixed 1:1 with Sinapic Acid (Fluka) matrix solution (10 mg/mL Sinapic Acid in 50% acetonitrile/0.05% TFA). 1 µL of sample/matrix was spotted onto a 100-well stainless steel sample plate (ABI, Framingham, MA) and dried in ambient conditions. The instrument was calibrated using the following protein standards (Sigma, MS-CAL1): bovine insulin, equine Cytochrome C, equine apomyoglobin and rabbit muscle aldolase.

HP-IEC Purification Trx-$Pro_{39}Glu_{10}$ Product. HPLC analysis of all protein samples was obtained on a Waters Alliance® HPLC system with: cooled autosampler chamber and column heater; a Waters 2996 photodiode array detector; and a Waters Fraction Collector. High performance ion exchange chromatography (HP-IEC) was performed on a Vydac DEAE 301VHP column (DEAE stationary phase, 5 µm particle size, 900 Å pore size, L = 5 cm, ID = 7.5 mm, column temperature = 24°C) with a Vydac CPF10 column prefilter. The mobile phases used were A: 7.5 mM Tris-HCl in 25% Tris-HCl, pH 8.0 in 25% acetonitrile, and B: 500 mM NaCl in Buffer A. Gradient programs are given in individual chromatograms. Fractions were collected at 1 min intervals, frozen, then vacuum dried in a Labconco Centrivap. Dried samples were reconstituted in water prior to further analysis by MALD-TOF MS.

RP-HPLC Purification $Pro_{39}Glu_{10}$ Product. HPLC analysis of all protein samples was obtained on a Waters Alliance® HPLC system with: cooled autosampler chamber and column heater; a Waters 2996 photodiode array detector; and a Waters Fraction Collector. High performance reversed-phase chromatography (RP-HPLC) was performed on a Supelco BIO Wide HPLC column (C8 stationary phase, 5 µm particle size, 300 Å pore size, L = 15 cm, ID = 4.6 mm, 1 mL/min flowrate, column temperature = 35°C) with a Vydac CPF10 prefilter. The mobile phases used were A: 0.1% formic acid in water, and B: 0.1% formic acid in acetonitrile. Gradient programs are given in the individual chromatograms. Fractions were collected at 1 min intervals. Collected fractions were frozen, and then vacuum dried in a Labconco Centrivap®. Dried samples were reconstituted in 100 µL of water prior to further analysis by MALDI-TOF MS.

3.13 Supplemental Information

3.13.1 DNA Sequencing Results

The results from the DNA sequencing (conducted by Cleveland Genomics, Inc.) of the plasmids relevant to the work presented in this chapter are given below. The one-letter nucleotide abbreviations are used and the sequences are written in order of 5' → 3'.

Sequencing results for:	pTrx-$Pro_{39}Glu_{10}$
Sequencing primer used:	Trx Reverse (Invitrogen)

```
   1  anaaggnnaa acngccccnt tttgaaaaaa natggnnggg ncccgtnttt
  51  tttaannnan ntttnaanng nnnaattttn ntnggcncnn nggccngggn
 101  atttgtnggn ttaagttttt tttntttang ntttgngggg gnnnaaaaaa
 151  ntttttgag agaaaatntt nnnttgggnn ggggnttttt tttgnnaact
 201  gaangnnaat tagggcaggt ngggccattc tggnttnatt tntgggnana
 251  cttgnngggt ncgnnccaaa atgntcggtc cgantttng anggaattng
 301  tgggggganat nagggcanac tggccgtttc naaaantnga cattgatcac
 351  aacccgggcn ntgcgccgaa atnatggcaa tccgttgnta tcccgaantc
 401  tgctgctgtc aaaaacggtg aagtggcggc aaccaaagtg ggtgcactgt
 451  ttnaaggtca gttgaaagag ttccttcgac gntaacctgg ccggctgtgg
 501  atccaacggt ggtcctcccc ctcctcctcc tcctcctccc cctcctcctc
 551  ctcctcctcc tcctcctcct cctccacctc cgccgccacc cccgcctcca
 601  ccaccaccac caccaccacc accaccacca gaagaagaag aagaagaaga
 651  agaagaagaa tagctctaga gtcgacctgc agtaatcgta cagggtagta
 701  caaataaaaa aggcacgtca gatgacgtgc ttttttcttg tgagcngtaa
 751  gntcnaaggg
```

Asn-Gly-Gly	bases 505-513
Pro_{39}	bases 514-630
Glu_{10}	bases 631-660
Stop codon	bases 661-663

Table 3.6 DNA sequencing results for pTrx-$Pro_{39}Glu_{10}$. (DNA sequences shown are reverse compliment of raw data).

Sequencing results for:	pThioHisA
Sequencing primer used:	Trx Forward (Invitrogen)

```
  1  ccctttnann nnncntngga gatgacgatg acaaggtacc atgggagctc
 51  gagatcttcg aattccgcgg ccgcaggcct ctagagtcga cctgcagtaa
101  tcgtacaggg tagtacaaat aaaaaaggca cgtcagatga cgtgcctttt
151  ttcttgtgag cagtaagctt ggcactggcc gtcgttttac aacgtcgtga
201  ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc
251  ctttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc
301  caacagttgc gcagcctgaa tggcgaatgg cgcctgatgc ggtattttct
351  ccttacgcat ctgtgcggta tttcacaccg catatatggt gcactctcag
401  tacaatctgc tctgatgccg catagttaag ccagccccga cacccgccaa
451  cacccgctga cgcgccctga cgggcttgtc tgctcccggc atccgcttac
501  agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc
551  gtcatcaccg aaacgcgcga gacgaaaggg cctcgtgata cgcctatttt
601  tataggttaa tgtcatgata ataatgggtt tcttanacgt tcaggtggca
```

EK site (DDDDK)	bases 21-35
Stop codon	bases 111-113

Table 3.7 DNA sequencing results for pThioHisA.

3.13.2 Amino Acid Sequences

The amino acid sequences of the polypeptides relevant to the work presented in this chapter are given in Table 3.8. The one-letter amino acid abbreviations are used and the sequences are written in order of amino terminus → carboxy terminus. The sequences for the Trx domain of Trx-$Pro_{39}Glu_{10}$ and for HP-Thioredoxin (control) were obtained from Invitrogen Corp.

Polypeptide	Amino Acid Sequence
Trx-$Pro_{39}Glu_{10}$	MSDKIIHLTDDSFDTDVLKADGAILVDFWAHWCGPC KMIAPILDGIADGYQGKLTVAKLNIDHNPGTAPKYGI RGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA GSG**S**NGGPPPPPPPPPPPPPPPPPPPPPPPPPPPPP PPPPEEEEEEEEEE
HP-Thioredoxin (control)	MSDKIIHLTDDSFDTDVLKADGAILVDFWAHWCGPC KMIAPILDGIADGYQGKTVAKLNIDHNPGTAPKYGIR GIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGS GDDDDKVPWEPEIFEFRGRRPLESTCSNRTG

Table 3.8 Amino acid sequences of polypeptides relevant to Chapter 3.

3.14 References

[1] Tulpar A, Henderson DB, Mao M, Caba B, Davis RM, Van Cott KE and Ducker WA. "Unnatural proteins for the control of surface forces". *Langmuir*. (accepted for publication).

[2] Cappello J, Crissman J, Dorman M, Mikolajczak M, Textor G, Marquet M and Ferrari F. "Genetic engineering of structural protein polymers". *Biotechnol. Progr.* (1990); 6: 198-202.

[3] McMillan RA, Lee TAT and Conticello VP. "Rapid assembly of synthetic genes encoding protein polymers". *Macromolecules.* (1999); 32: 3643-3648.

[4] McGrath KP, Tirrell DA, Kawai M, Mason TL and Fournier MJ. "Chemical and biosynthetic approaches to the production of novel polypeptide materials" *Biotechnol. Progr.*(1990); 6: 188-192.

[5] Krsmanovic JL. Ph.D. Thesis, Virginia Polytechnic Institute & State University, Virginia, USA, (2003).

[6] Wu DT, Yokoyama A and Setterquist RL. "An experimental study on the effect of adsorbing and non-adsorbing block sizes on diblock copolymer adsorption". *Polym. J.* (1991); 23(5): 709-714.

[7] Zhang SP, Zubay G and Goldman E. "Low-usage codons in *Escherichia coli*, yeast, fruit fly and primates". *Gene.* (1991); 105(1): 61-72.

[8] Lu Z, Diblasio-Smith EA, Grant KL, Warne NW, LaVallie ER, Collins-Racie LA, Follettie MT, Williamson MJ and McCoy JM. "Histidine patch thioredoxins". *J. Biol. Chem.* (1996); 271(9): 5059-5065.

[9] Lavallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF and McCoy JM. "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *Escherichia coli* cytoplasm". *Bio-Technol.* (1993); 11(2): 187-193.

[10] Jordan GL and Harcum SW. "Characterization of up-regulated proteases in an industrial recombinant *Escherichia coli* fermentation". *J. Ind. Microbiol. Biot.* (2002); 28:74-80.

[11] Matsuo E, Sampei G, Mizobuchi K and Ito K. "The plasmid F OmpP protease, a homologue of OmpT, as a potential obstacle to *E. coli*-based protein production". *FEBS Lett.* (1999); 461: 6-8.

[12] Kaufmann A, Stierhof YD and Henning U. "New outer membrane-associated protease of *Escherichia coli* K-12". *J. Bacteriol.* (1994). 176: 359-367.

[13] Striebel HM and Kalousek F. "Eukaryotic precursor proteins are processed by *Escherichia coli* outer membrane protein OmpP". *Eur.J. Biochem.* (1999); 262: 832-839.

[14] Stryer L. *Biochemistry 4th edition*. W.H. Freeman and Company, New York, (1995).

[15] Hosfield T and Lu Q "Influence of the amino acid residue downstream of $(Asp)_4Lys$ on enterokinase cleavage of a fusion protein". *Anal. Biochem.* (1999); 269: 10-16.

[16] Smith BJ "Chemical cleavage of proteins at methionyl residues". *The Protein Protocol Handbook.* Humana Press, Inc. Totowa, NJ, (1996).

[17] Bornstein P and Balian G. "Cleavage at Asn-Gly bonds". *Method. Enzymol.* (1977); 47: 132-145.

[18] Canova-Davis E, Eng M, Mukku V, Reifsnyder DH, Olson CV, and Ling VT. "Chemical heterogeneity as a result of hydroxylamine cleavage of a fusion protein of human insulin-like growth factor I". *Biochem. J.* (1992); 278: 207-213.

[19] Kwong MY and Harris RJ. "Identification of succinimide sites in proteins by N-terminal sequence analysis after alkaline hydroxylamine cleavage". *Protein Sci.* (1994); 3: 147-149.

[20] Hiller Y, Bayer EA and Wilchek M. "Studies on the biotin-binding site of avidin. Minimized fragments that bind biotin". *Biochem. J.* (1991); 278: 573-585.

[21] Sambrook J, Fritsch EF and Maniatis T. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, (1992).

[22] Ausubel FM. *Current Protocols in Molecular Biology*. J. Wiley, New York, NY, (1997).

[23] Hochstrasser DF, Patchornik A and Merril CR. "Development of polyacrylamide gels that improve the separation of proteins and their detection by silver staining". *Anal. Biochem.* (1988); 173: 424-435.

[24] Gill SC, Von Hippel PH. "Calculation of protein extinction coefficient from amino acid sequence data". *Anal. Biochem.* (1989); 182(2): 319-326.

[25] Walker, JM. *The Protein Protocols Handbook*. Humana Press Inc.: Totowa, NJ, (1996).