1-24-2007

Expression, Purification, and Characterization of the SIAA M79A Protein

Brian Basden

EXPRESSION, PURIFICATION, AND CHARACTERIZATION OF THE SIAA M79A PROTEIN

by

BRIAN J BASDEN

Under the Direction of Dabney White Dixon

ABSTRACT

Some pathogenic bacteria derive significant amounts of iron heme from their hosts. In this study we investigated SiaA, a heme binding protein from *Streptococcus pyogenes*. The wildtype methionine79 putative axial ligand was mutated to alanine. SiaA M79A was expressed in *E. coli* in three production runs, lysed by sonication or French press, and purified by fast protein liquid chromatography (FPLC). Nickel affinity FPLC was found to give much purer SiaA when 30 mM imidazole was added to the binding buffer. The protocol using extensive sonication resulted in SiaA weighing 30464 Da. The protocol using French press resulted in SiaA weighting 33358 Da. Despite the difference in masses, the two forms of SiaA interacted with heme similarly.

INDEX WORDS: Heme, SiaA, *Streptococcus pyogenes*, heme binding proteins, Axial ligand, Methionine, Alanine, Fast protein liquid chromatography FPLC, Nickel affinity, Western Blot, SDS-PAGE, Homology modeling, ATP binding cassette, ABC transporters, Periplasmic binding proteins, PBP, Bacterial iron acquisition, Protein cleavage.

EXPRESSION, PURIFICATION, AND CHARACTERIZATION OF THE SIAA M79A PROTEIN

by

BRIAN J BASDEN

An Honors Thesis Submitted in Partial Fulfillment of the

Requirements for Graduation with Undergraduate Research Honors

in the College of Arts and Sciences

Georgia State University

2006

EXPRESSION, PURIFICATION, AND CHARACTERIZATION OF THE SIAA M79A PROTEIN

by

BRIAN J BASDEN

Honors Thesis Director: Dr. Dabney W. Dixon
Honors Program Director: Dr. Robert Sattelmeyer

Electronic Version Approved:

Honors Program
College of Arts and Sciences
Georgia State University
December 2006

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Dixon for allowing me to pursue this project in her research group. Dr. Dixon's input was invaluable throughout the research and writing process of this thesis. I would also like to thank Dr. Dixon for her continued support towards the completion of my Chemistry degree and towards my future plans. I would like to thank the entire Dixon laboratory, especially Brian Sook, Thanh Quach, and Marianna Libkind, who provided their assistance to the experimental aspects of this project. I thank Dr. Zehava Eichenbaum and co-workers including Griselle Montañez, Elizabeth Cunha, and Courtney Meier from Georgia State University as well as Dr. Mario Rivera from the University of Kansas. I also thank Gene Drago, Tanu Soni, and Drew Kohlhorst from Georgia State University.

TABLE OF CONTENTS

LIST OF TABLES

Introduction

Heme is a flat, planar molecule with a single iron atom in the middle. Heme is a common molecule used by proteins to carry out a variety of cellular tasks. In humans, and all mammals, heme is a part of the protein hemoglobin, which is responsible for binding oxygen from air in the lungs. Without iron, mammals would not be able to breathe. Iron is also an important nutrient for pathogenic bacteria; in fact, most pathogenic bacteria need iron to survive while in human hosts (Lei et al., 2002; Bates et al., 2003; Lei et al., 2003; Liu & Lei, 2005; Nygaard et al., 2006). Pathogenic bacteria are known to use heme binding proteins to acquire heme, and thus iron, in the human host (Lei et al., 2002; Bates et al., 2003; Lei et al., 2003; Liu & Lei, 2005; Nygaard et al., 2006).

Therefore, iron acquisition proteins of pathogenic microbes are potential targets for virulence reducing therapies. SiaA is a heme binding protein in *Streptococcus pyogenes* and it is part of an ATP binding cassette (ABC) transporter system. *S. pyogenes* employs an ABC transporter system labeled SiaABC that binds and imports heme (Lei et al., 2002; Bates et al., 2003; Lei et al., 2003; Liu & Lei, 2005; Nygaard et al., 2006). The probable axial ligands for SiaA are Met79 and His229. In this study, wildtype (WT) SiaA was mutated by replacing the native residue, Met79, with Ala79 to give the M79A SiaA mutant. M79A SiaA was expressed, purified, and studied.

ABC transporters. ABC transporters are part of the ABC-ATPase superfamily of proteins (Jones & George, 2004). Members of this superfamily are present in all organisms, and also comprise a significant proportion of the proteome. Genome analysis of *E. coli* showed at least 80 ABC proteins, which corresponds to about 5% of the genome.

ABC transporters occupy the ABC-ATPase superfamily along with non-transporters (Jones & George, 2004). Non-transporters are cytosolic and are generally used for DNA related functions. Transporters are mainly used by eukaryotes as exporters and prokaryotes typically use them as importers (Locher & Borths, 2004).

ABC proteins utilize the energy from ATP to perform or contribute to a variety of cellular functions including nutrient acquisition, resistance to xenotoxins, cell division, and bacterial immunity (Jones & George, 2004). Specifically, ABC transporters utilize the energy of ATP to transport substrates across lipid membranes.

ABC transporters are distinguished by their characteristic core structure of four domains (Jones & George, 2004). Two of the four domains are alpha-helical transmembrane domains (TMDs). The remaining two domains are cytosolic ABCs, or nucleotide-binding domains (NBDs). The TMDs form the channel through which solutes selectively pass. ABC exporters possess TMDs with a total of 12 α-helices while the TMDs of ABC importers vary from 8 to 20 α-helices (Locher, 2004).

ABC transporters, like all ABC proteins, are identified by two highly conserved regions in the NBDs (Jones & George, 2004). These two regions are called the Walker A and B consensus motifs. The Walker A motif (G-X-X-G-X-G-K-S/T) forms a phosphate binding loop, or a P-loop, aided by small glycine residues. The P-loop binds ATP by electrostatic interactions with the negatively charged tri-phosphate region. The Walker B motif (θ-θ-θ-θ-D), where 'θ' represents hydrophobic amino acids, forms a buried β-strand whose aspartate orients the active site through hydrogen bonds. Despite the

establishment of general structural and mechanistic details from biochemical studies and X-ray crystallography of specific ABC transporters, the question of how ATP hydrolysis specifically mechanistically transduces solutes across the membrane by ATP transporters is one of active investigation (Locher & Borths, 2004).

Most bacterial ABC importers possess binding proteins which deliver the substrate to the TMDs (Locher & Borths, 2004). Specific ABC transporter systems with their cognate binding proteins will be discussed.

The Isd gene cluster: Heme uptake in *Staphylococcus aureus*. The gram positive pathogen *S. aureus* is a potentially fatal microbe. *S. aureus* causes bacteremia and endocarditis (Skaar & Schneewind, 2004) and continues to evolve into virulent and fatal drug resistant strains.

Iron inside the human body is found largely in hemoproteins and iron-sulfur proteins (Skaar & Schneewind, 2004). Hemoproteins are the most abundant source, with 60-75% of iron located in hemoglobin; myoglobin accounts for an additional 8-10%. However, hemoproteins are not the only available source of iron for *S. aureus*. The plasma protein transferrin can serve as a sole iron source. Siderophores, low molecular weight chelating molecules, are thought to enable *S. aureus* to bind the iron from transferrin. Multiple siderophores have been documented in *S. aureus*, though no specific protein receptor has been conclusively identified (Skaar & Schneewind, 2004).

S. aureus, like *S. pyogenes*, utilizes an extracellular protein assembly to scavenge heme from these iron stores. *S. aureus* uses the iron-regulated surface determinant (Isd) protein assembly. Because the Isd system scavenges for iron, Isd is only transcribed when concentrations of iron are low.

Isd is comprised of five transcriptional subunits: *isdA, isdB, isdCDEFsrtBisdG, isdH,* and *isdI*. Four of these ten protein products, IsdA, IsdB, IsdC, and IsdH, are cell wall anchored. IsdDEF encodes for a transporter while IsdG and IsdI represent two cytosolic proteins. All nine Isd proteins bind heme. Finally, sortase B (SrtB) is also included in the Isd gene cluster.

IsdB, a cell wall linked protein, binds hemoglobin. This putative hemoglobin receptor has a intrachromosonal paralog, IsdH, with 85% sequence similarity. IsdH has two regions, D1 and D2, which bind hemoglobin-haptoglobin.

The individual inactivation of *isdA* or *isdF* decreased the amount of radiolabled heme-iron that was imported into *S. aureus*, providing evidence for Isd's heme acquisition role. However, inactivation of either *isdA* or *isdF* did not stop the importation of heme-iron, implying additional heme acquisition routes.

IsdE is related to periplasmic binding and heme binding proteins that donate heme to the permease in the ABC transporter. Recent work has not been able to obtain definite spectroscopic evidence for the heme axial ligands in this protein (Mack et al., 2004).

The Isd system depends upon sortases for protein placement. Sortases catalyze an amide linkage between the C-terminal end of proteins and a crossbridge of peptidoglycan. Of the four cell wall anchored protein products, sortase A links three: IsdABH. Sortase B links only IsdC. Over twenty proteins in *S. aureus* are cell wall bound by sortase A. Sortase B is thought to only function on IsdC.

Inactivation of *srtA* in mouse infection models markedly reduces the virulence of *S. aureus*, and renders cells incapable of utilizing heme to support bacterial growth

(Mazmanian et al., 2001; Jonsson et al., 2003). However, since sortase A links at least 20 proteins to the peptidoglycan layer, all of the anti-virulence effect of *srtA* inactivation cannot be attributed to disruption of Isd. *srtB* inactivation, leaving IsdC unbound extracellularly, reduces virulence marginally in comparison to *srtA* inactivation. The observation that *srtB* inactivation reduces virulence leads to the conclusion that reduced heme intake reduces the ability of *S. aureus* to reproduce due to limited iron concentrations.

The Btu gene cluster: Cobalamin uptake. Vitamin B_{12} is a cofactor used by organisms for reactions such as methyl transfers and ribonucleotide reduction (Karpowich et al., 2003). While some bacteria biosynthesize B_{12} using about 30 steps, most bacteria and all eukaryotes import B_{12} through transporters.

Btu (B twelve uptake) is an *E. coli* protein assembly that imports vitamin B_{12} (Karpowich et al., 2003). Btu includes the ABC transporter BtuCD, its associative periplasmic binding protein (PBP) BtuF, and the outer membrane transporter BtuB. The PBP BtuF binds B_{12} for delivery to the transmembrane (TM) dimer $BtuC_2$ (Locher, 2004). BtuF is thought to use electrostatic interactions between its conserved glutamates and the conserved arginines in BtuC for association (Locher & Borths, 2004). BtuD also dimerizes *in vivo* and each monomer contains a NBD of the ABC transporter. To complete the ABC transporter, $BtuD_2$ and $BtuC_2$ associate yielding the tetramer $BtuC_2D_2$.

The complete dimer of BtuD is needed for hydrolysis of ATP, thus ATPase activity is dependent upon dimerization of NBDs. Following hydrolysis, conformational changes are transmitted via various loops in the conserved NBD region to another conserved region: the BtuD – BtuC interface. It is presumed, therefore, that conformational changes starting in the NBD region transmit themselves to the TMDs, thus clearing the way for the import of the selected solute. Interestingly, it is thought that the ATPase ability of the NBDs is triggered when BtuF delivers vitamin B_{12} to the BtuC transmembrane dimer (Locher & Borths, 2004). Current research on ABC transporters focuses on these and other mechanistic complexities of solute transport.

BtuF has two globular domains (I and II) connected via a rigid α-helix, α-6 (Karpowich et al., 2003). Each globular domain has a secondary structure that alternates between α-helixes and parallel β–sheets. Domain I, starting at the N-terminus, has an α-helix, α-5, bordering the rigid interdomain α-6. This partial helical bundle contacts the B_{12} at its dimethylbenzimidazole (DMB) and propionamide groups.

BtuF binds B_{12} at domain II differently than other proteins with regions homologous to domain II. While domains homologous to domain II in methionine synthase (MetE) and methylmalonyl-CoA mutase (MMCM) use a conserved histidine to displace the DMB group from binding at the cobalt atom, domain II in BtuF binds on the opposite face, leaving the DMB group bound.

The B_{12} binding site of BtuF is acidic while the protein itself is basic. About 60% of the B_{12} is inside the binding cleft, held in by 11 residues that make direct contacts with B_{12} in both of the crystal structure monomers. Domain I contacts B_{12} 50 times with six residues while domain II uses 5 residues to contact B_{12} 39 times. BtuF makes fewer contacts with B_{12} than either MetE or MMCM, both of which make over 150 contacts compared to 89 for BtuF. Fewer contacts in BtuF is consistent with its function: BtuF delivers B_{12} to a transporter while MetE and MMCM immobilize B_{12} for catalytic use.

The Sia gene cluster: Heme uptake in *Streptococcus pyogenes.* *S. pyogenes*, also known as Group A Streptococcus (GAS), is a Gram positive pathogenic bacterium which colonizes the throat and skin (Cunningham, 2000). *S. pyogenes* causes pharyngitis, streptococcal toxic shock syndrome, impetigo, and scarlet fever; its role in increasing host susceptibility to other diseases has been investigated.

S. pyogenes demonstrates iron-dependent growth (Eichenbaum et al., 1996a), and parasitizes animal hosts where most iron is bound in heme and non-heme proteins (Lei et al., 2003). Hemoproteins including hemoglobin, myoglobin, and catalase can serve as iron donors for *S. pyogenes* (Eichenbaum et al., 1996b). Since iron is a necessary element for this pathogen, iron acquisition proteins in *S. pyogenes* serve as potential targets to limit full virulence.

S. pyogenes acquires heme via an ABC transporter, SiaABC (Bates et al., 2003). SiaABC is encoded by three genes (*spy1793-spy1795*) in an operon that also contains *shp* (*spy 1796*), *shr* (*spy1798*), and five other genes (Bates et al., 2003; Lei et al., 2002; Liu & Lei, 2005; Nygaard et al., 2006). Expression of the *sia* operon is negatively regulated by iron; thus, *S. pyogenes* produces more iron acquisition proteins when iron concentrations are low (Bates et al., 2003).

SiaA (*spy1795*) is anchored to the cell membrane with a lipoprotein tail (Bates et al., 2003; Lei et al., 2003). SiaA binds hemoglobin and heme (Bates et al., 2003; Lei et al., 2002; Lei et al., 2003). SiaA can not bind myoglobin (Bates et al., 2003) yet *S. pyogenes* can acquire iron from myoglobin (Eichenbaum et al., 1996b). These two observations give rise to two possibilities: either SiaABC is not the exclusive iron acquisition route in *S. pyogenes*, or additional proteins bind the heme from myoglobin and deliver it to SiaA. Isolation and analysis of the lipoprotein constituents of fifteen putative ABC transporters in *S. pyogenes* suggests SiaABC is the only heme ABC transporter in GAS (Lei et al., 2003). Twelve of the lipoproteins did not show chromophores, while the other two showed no homology with iron or heme transporters (Lei et al., 2003). Therefore, it is probable that other proteins bind the heme from myoglobin and deliver it to SiaA.

SiaB (*spy 1794*) contains the TMDs and consists of extended hydrophobic regions. SiaC is the ATPase and contains the two NBDs. Each NBD contains the Walker A and B consensus motifs, which transform the energy of ATP into conformational changes to transport heme (Jones & George, 2004).

Shp, also found on the *sia* operon, binds heme and delivers it to SiaA based on the observation that holo-Shp transferred heme almost irreversibly to apo-SiaA (Liu & Lei, 2005). Shr binds a variety of heme-containing proteins including hemoglobin, myoglobin, and hemoglobin-haptoglobin yet its role, if any, with SiaA has not yet been determined (Bates et al., 2003).

ShuT: One protein in the heme uptake pathway from *Shigella dysenteriae.* ShuT is a periplasmic heme binding protein in *S. dysenteriae* (Eakanunkul et al., 2005). It likely works in concert with ShuU, the presumptive cytoplasmic permease, and ShuV, the presumptive ATPase to form an ABC transporter system. ShuT shares sequence homology with BtuF, described above, and FhuD, an iron-hydroxamate PBP. Most PBPs contain two globular domains. The interface between the two globular domains forms

the active site, and the globular domains are connected by a hinge region. Three groups of PBPs have been characterized based upon differences in the hinge region. Group I has three interdomain connections, Group II has two interdomain connections, and Group III has one interdomain connection (α-helical). The proteins BtuF and FhuD are part of Group III PBPs and sequence homology suggests that ShuT is also a Group III PBP. Specifically, ShuT shares important aromatic residues with BtuF whose function is ligand stabilization.

Studies on SiaA. As outlined above, previous work in the Eichenbaum (Bates et al., 2003), Musser and Lei (Lei et al., 2002; Lei et al., 2003; Liu & Lei, 2005; Nygaard et al., 2006), and Dixon laboratories have shown that SiaA is a heme protein, whose probable axial ligands are Met79 and His229. The goal of this study was to isolate the protein in which the Met79 ligand had been mutated to Ala. This work was successful, and we report the first isolation and initial characterization of this protein. We have also observed that various techniques to break open the cells give different protein constructs; the details of these studies are reported.

Experimental

3D-JIGSAW and RasMol. The protein sequence of WT SiaA was submitted to the 3D-JIGSAW website (http://www.bmm.icnet.uk/servers/3djigsaw/) in interactive mode. A return email led the user to multiple homologous structures. The structure for BtuF (IN4D) was selected and the coordinates for BtuF with WT SiaA superimposed were emailed in response. The graphical coordinates in the email (starting with the line ATOM 1) were copied and pasted into Microsoft Office Word 2003 and saved as a txt file. The file was renamed as a pdb file and opened with RasMol 2.7.2.2.1 (originally developed by Roger Sayle).

Luria-Bertani (LB) medium preparation. A batch of LB medium was made from 30 g Bacto® Tryptone (Becton Dickinson, Sparks, MD), 30 g NaCl, 15 g Bacto® Yeast Extract (Becton Dickinson), and deionized water to make a total of 3 L. The solution was pH adjusted to 7.0 and autoclaved on a liquid cycle for 15 min. Ampicillin (3 mL of 100 mg/mL) was added to the 3 L of autoclaved medium.

Inoculation of LB medium and growth of M79A transfected *E. coli*. LB medium (5 mL) was added to each of six 15 mL falcon tubes. An M79A mutant was prepared using a Stratagene® QuikChange® site-directed mutagenesis kit by following the manufacturer's instructions (Amrita Nargund and Z. Eichenbaum). The M79A cryovial (stored at - 80 °C) was dabbed with the blunt end of an inoculating loop, and swirled into a 15 mL falcon tube. This process was repeated with a different inoculating loop for each falcon tube. The falcon tube lids were secured by tape in a position that allowed aeration and were incubated for 16 h at 37°C with shaking to form the aliquots. The OD_{660} of the aliquots were taken and incubation continued until the OD_{660} reached about 1. The remaining LB media was divided into six 1 L beakers (500 mL each) and inoculated with an aliquot each. Beakers were incubated at 37°C until OD_{660} readings reached 0.5. A 2% L-arabinose solution was made by dissolving 2 g L-arabinose (Acros Organics, Morris Plains, NJ) into 100 mL of 18 MΩ water delivered from a Barnstead NANOpure Diamond® purifier. The LB medium was induced to 0.02% L-arabinose by adding 5 mL

of 2% L-arabinose to each beaker. Beakers were incubated at 37°C for 4 h. The LB medium was centrifuged at 8000 *x* g and 4°C for 10 min and pellet was divided into ten 2 mL cryovials for storage at -80°C.

The procedure was run again on a 2 L scale.

Growth with δ-aminolevulinic acid (ALA) and ferrous sulfate. A stock solution of $FeSO_4$ (0.109 M) was prepared by dissolving 0.303 g $FeSO_4$ (Aldrich Chemical Company, Milwaukee, WI) in 10.0 mL nanopure water. A stock solution (0.100 M) of ALA (MP Biomedicals, Inc., Solon, OH) was prepared by dissolving 0.0838 g ALA in 5.0 mL of nanopure water (Marianna Libkind). LB medium (20 mL) was inoculated with the M79A strain and incubated at 39°C for 17 h in a 50 mL falcon tube. Erlenmeyer flasks, two of 250 mL volume, each containing 100 mL LB medium were each inoculated with 9 mL of the 20 mL solution. The two resultant 109 mL LB solutions were shaken at 200 rpm for 3 hours at 37°C. At the three hour mark both solutions were induced to a 0.02% solution of L-arabinose by adding 1.1 mL of a 2% L-arabinose stock solution each. Additionally at the three hour mark, one of the 109 mL solutions was induced to a final concentration of 0.1 μM ALA and 0.1 μM $FeSO_4$ by adding 0.109 mL of the ALA stock solution and 0.100 mL of the $FeSO_4$ stock solution. The solutions were incubated for an additional 4 h at 200 rpm and 37°C. Portions (30 mL) of both the solutions were French pressed for two cycles each on a SimAminco French Pressure Cell Press on high ratio at a display pressure of 960 and an effective cell pressure of 15,000 psi. A UV-Vis spectrum of both solutions was obtained.

Fermentation growth (9 L). Concentrated (10X) LB medium was prepared by dissolving 90 g Bacto® Tryptone (Becton Dickinson), 90 g NaCl, 45 g Bacto® Yeast Extract (Becton Dickinson), and deionized water to make a total of 900 mL. The solution was pH adjusted to 7.0 and autoclaved on a liquid cycle for 15 min. Ampicillin (9 mL of 100 mg/mL) was added to this medium. Diluted (1X) LB medium (900 mL), prepared according to previous LB medium procedures, was inoculated with the M79A strain and grown for 14 h with shaking at 37°C and 200 rpm. The inoculated 1X 900 mL solution and the 10X LB medium were provided to Dr. Gene Drago (Biology Department, Georgia State University) for a 9 L fermentation run. A Sartorius C30-2 fermentor (Sartorius AG, Goettingen, Germany), with water sterilized in the vessel, was loaded with the sterilized 10X solution which was diluted up to approximately 8 L. The inoculated 1X 900 mL solution was dabbed with a sterile loop and swabbed on a plate (and incubated overnight to verify that only *E. coli* was in the culture). The inoculated 1X 900 mL was delivered to the C30-2 to complete the 9 L. The C30-2 was set to a dissolved oxygen set point of 15% with a temperature of 37°C and a pH set point of 7.0. The pH was stabilized with H_2SO_4 and NaOH. The rpm maximum was 200 rpm, and the airflow maximum was 3.5 L/min. The airflow maximum was proportional-integral-derivative (PID) controlled. The 9 L run was incubated for 4 h whereupon the 9 L run was induced to a final concentration of 0.02% L-arabinose, 0.1 μM ALA, and 0.1 μM $FeSO_4$ by adding 0.151 g ALA (MP Biomedicals), 0.250 g $FeSO_4$ (Aldrich), and 9 mL 2.0% L-arabinose, and grown for an additional 3 h. A final volume of about 9.4 L was recovered and was centrifuged at 8000 *x* g and 4°C for 10 min and the pellet (approximately 60 g) was collected and frozen in 15 mL Nalgene cryovials at -80°C.

Sonication and suspension of pellet into buffer from 3 L run. Extraction buffer was made by adding 2.42 g Tris base (Fisher Chemicals, Fair Lawn, NJ), 5.84 g NaCl

(Aldrich Chemical Company, Milwaukee, WI), and 1 mL Triton X-100 (Aldrich) with 18 MΩ water up to 1 L to give a solution of 20mM Tris base, 100 mM NaCl, and 0.1% Triton X-100; the extraction buffer was then pH adjusted to 8. The ten cell pellets were combined into a 50 mL solution of extraction buffer with two Complete Mini EDTA-free protease inhibitor tablets (Roche Diagnostics, Indianapolis, IN). The 50 mL solution was divided into five 10 mL fractions. Each 10 mL fraction was sonicated on ice by a Branson Sonifier 450 for six 30 s pulsing intervals with a duty cycle of 85% with a displayed output of 20 – 25%. Sonicated samples were centrifuged for 20 min at 17,000 rpm and 4°C. The supernatant was decanted and filtered through an Acrodisc® 25 mm syringe filter with a 0.45 nm Supor® membrane and part number 4614 (Pall Corporation, East Hills, NY). The pellet was discarded.

Sonication of the 2 L run. The cell pellet from the 2 L run was suspended in 30 mL extraction buffer with two Complete Mini EDTA-free protease inhibitor tablets (Roche Diagnostics). The 30 mL solution was divided into three 50 mL falcon tubes of 10 mL each. The 10 mL fractions were sonicated on ice with a Branson Sonifier 450. Sonication was performed with 5 cycles of 2 minutes each with pulsing intervals (duty cycle) with resting periods of 1-2 minutes in between. A duty cycle of 85% was used with a recorded output in between 40 – 60%. The solution (30 mL) was centrifuged and filtered in the same method as the sonication of the 3 L run.

Lysis by French press of a portion of the 9 L run. The cell pellet (6 g) from the 9 L run was suspended in 60 mL of extraction buffer. The 60 mL solution was divided into two 50 mL falcon tubes to give two 30 mL samples; each 30 mL sample had one Complete Mini EDTA-free protease inhibitor tablet (Roche Diagnostics) dissolved in solution. The solutions were put on ice and kept on ice while not being actively French pressed. A large (35 mL maximum volume) Thermo French press cell, kept at 4°C, was loaded with each sample (30 mL) individually and each sample was French pressed twice. A SimAminco French pressure cell press was used on high setting and adjusted to a display psi of 1280 yielding an effective psi of 20000. Samples were collected in their original falcon tubes kept in a small beaker of ice. The solution (60 mL) was centrifuged and filtered in the same method as the sonication of the 3 L run. The large Thermo French pressed was cleaned with deionized water and ethanol and air dried at room temperature before re-refrigeration.

Nickel affinity fast protein liquid chromatography (FPLC). Binding buffer (50 mM potassium phosphate and 250 mM NaCl at pH 7.4) was made from adding 6.8 g KH_2PO_4 (Fisher), 5.5 g K_2HPO_4 (EM Science, Gibbstown, NJ), 14.61 g NaCl (Aldrich), and 18 MΩ water up to 1 L and adjusting the pH to 7.4. Elution buffer was made from binding buffer plus 0.5 M imidazole (EMD Chemicals Inc., Gibbstown, NJ). Charging solution (100 mL) was prepared by dissolving 2.37 g $NiCl_2$ (Fischer Scientific, Fair Lawn, NJ) into nanopure water and making the solution up to 100 mL. Stripping buffer was prepared by 18.61 g of disodium-EDTA (Fischer) in 1 L of binding buffer to give a concentration of 0.05 M EDTA. Binding buffer, elution buffer, charging solution, and stripping buffer were all filtered though 0.45 μL disposable Nalgene filter units with 50 mm diameter membranes and 45 mm necks. A HiTrap® Chelating HP (GE Healthcare Bio-sciences) column (5 mL) was washed with 5 column volumes (CV) of nanopure water, 5 CV of stripping buffer, 5 CV of nanopure water, 4 CV charging solution (0.1 M $NiCl_2$), 12 CV nanopure water, 5 CV of elution buffer, and 5 CV of binding buffer to

complete the stripping and charging process. An Amersham Biosciences A920 FPLC (GE Healthcare Bio-sciences, Piscataway, NJ) superloop was loaded with the filtered supernatant from the sonicated and centrifuged sample from the 3 L run. Pumps A and B were washed with water. The 5 mL HiTrap® Chelating HP (GE Healthcare Bio-sciences) column was washed with 5 CV of water. The column was equilibrated with 5 CV of binding buffer. The sample was injected into the column, flushed with 6 CV of binding buffer, and eluted with 10 CV of a gradient starting with 100% binding buffer and ending with a concentration of 100% elution buffer. Fractions 1 – 33 were collected.

The same process was repeated for the 2 L run except that 30 mM imidazole (EMD Chemicals) was added to the binding buffer.

The original method with no imidazole in the binding buffer was applied for 6 g of cell pellet from the 9 L run that was French pressed.

SDS-PAGE of nickel affinity FPLC fractions. Two 0.75 mm gels were prepared by adding 15% resolving gel (Sambrook & Russel, 2001) and a thin layer of 2-propanol (Fischer) to sit for 30 min. The 2-propanol was removed carefully with deionized water, the gels were dried, and 5% stacking gel (Sambrook & Russel, 2001) was added on top of the resolving gels and combs were placed on top of both gels. After 45 min the combs were removed and the gels were loaded into a Mini PROTEAN® 3 Cell (Bio-Rad). Tris-glycine-SDS 1X (500 mL) was made by adding 50 mL of Tris-glycine SDS 10X (Fisher) with 18 MΩ water up to 1 L. Tris-glycine-SDS 1X (500 mL) was poured into the Mini PROTEAN® 3 Cell. Mixed Laemmli sample buffer was made by adding 1 part 2-mercaptoethanol (Bio-Rad) with 19 equivalents of Laemmli sample buffer (Bio-Rad). Five µL was taken from each of the affinity FPLC fractions 10-18 and was combined with 10 µL of mixed Laemmli sample buffer. Samples were centrifuged for 15 sec on a VWR mini-centrifuge C-1200 and heated for 5 min at 103°C. Five µL of Precision Plus Protein® molecular weight marker standard (Bio-Rad) was added to lane one. Heated samples were loaded into the remaining gel lanes. The gel was ran for at 60 – 80 V until the blue Laemmli sample buffer band was 1 cm below the lanes, at which point the voltage was increased to 100 – 120 V until the dye reached the bottom of the gel. Destaining solution was made by combing 5 parts methanol (Fischer) to 1 part glacial acetic acid (EMD Chemicals, Gibbstown, NJ) and 4 parts nanopure water. A Coomassie® Blue staining solution was prepared by dissolving 0.25 g of Coomassie® Brilliant Blue R-250 into 100 mL of destaining solution. The gels were removed and stained in the staining solution for 30 min and destained for 1 h. The gels were rinsed with deionized water and stored in a 20% glycerol (Fischer) solution.

Combining nickel affinity FPLC fractions and exchanging buffer to remove salt. A solution of 10 mM Tris-HCl pH 8 was prepared by adding 1 part 1.0 M Tris-HCl Buffer 1 M (Fischer) to 99 parts nanopure water then adjusting the pH to 8.0. Fractions 16, 17, 18, and 19 from Nickel Affinity FPLC were combined and placed into a Amicron® Ultra 15 5,000 MWCO (Millipore, Billerica, MA) filter tube. The tube was centrifuged at 6000 rev/min for 15 min at 4°C. Approximately 6 mL of 10 mM Tris-HCl was added to the remaining protein solution and the tube was centrifuged in the same manner again. This process was repeated once more, and yielded 2 mL of protein suspended in 10 mM Tris-HCl pH 8.0.

Lowry assay and enterokinase cleavage. A Modified Lowry Protein Assay (Pierce, Rockford, IL) was used following the manufacturer's instructions to determine the

protein concentration of the 2 mL solution. The 2 mL solution was treated with an EnterokinaseMax™ (EKMax™) digestion (Invitrogen, Carlsbad, CA). Manufacturer's instructions were followed except that 0.1 µL of EKMax™ was used per 20 µg of protein instead of the recommended range of 1 – 4 µL (B. Sook, personal communication). The 2 mL solution was combined with 525 µL 10X EKMax™ buffer, 17.5 µL of EKMax™, and nanopure water up to 5.25 mL and stored at 4°C for 16 h.

Buffer exchange and anion exchange FPLC. Using the same buffer exchange method previously described, the EKMax™ digestion solution was exchanged into 10 mM Tris-HCl pH 8.0 resulting in a solution of 5 mL. The FPLC was fitted with a HiTrap™ Q HP (GE Healthcare Bio-sciences) column and washed with water along with pumps A and B. The column was washed with five CV of anion exchange start buffer (10 mM Tris-HCl pH 8.0), five CV of anion exchange elution buffer (anion exchange start buffer and 1 M NaCl), and then equilibrated with seven CV of anion exchange start buffer. The 5 mL protein solution was loaded into the superloop and injected into the column. Elution of the protein was performed by running five CV of anion exchange start buffer through the column followed with a 10 CV linear gradient. The gradient began with 100% anion exchange start buffer and ended with 100% anion exchange elution buffer. Fractions 1 – 30 were collected.

For the 2 L run, the same process was repeated except that a HiTrap™ DEAE FF (5 mL) column was used, resulting in weak anion exchange FPLC, and the gradient was started at 7% elution buffer. For the repurification of the 2 L run, the original method was applied except that a HiTrap™ DEAE FF (5 mL) column was used and the elution buffer was anion exchange start buffer and 0.4 M NaCl.

For the 6 g of cell pellet from the 9 L run, the original method was applied except that a HiTrap™ DEAE FF (5 mL) column was used to give weak anion exchange FPLC and the elution buffer was reduced 0.4 M NaCl and start buffer.

SDS PAGE of anion exchange FPLC fractions and buffer exchange. SDS-PAGE was performed. Lane one was filled with standard while lanes 2 – 10 were filled with fractions 12 – 18, a control, and fraction 2, respectively. Fraction 13 was exchanged to a 10 mM Tris-HCl pH 8 solution.

The same buffer exchange process was repeated with fractions 15 – 18 from Figure 15 except that the 10 mM Tris-HCl pH 7.5 was used.

Fractions 17 – 21 from Figure 20 were transferred into 10 mM Tris-HCl pH 8.

Mass spectrometry. Matrix-assisted laser desorption/ionization (MALDI) experiments were performed using an ABI Voyager-DE™ Pro Workstation (Applied Biosystems, Warrington, UK) MALDI reflectron time-of-flight spectrometer. The mass range scanned was 10000 – 50000 Da in positive mode. The protein samples (approximately 20 µL) were transferred into nanopure water and run with sinipinic acid (3,5-dimethoxy-4-hydroxycinnamic acid) as the matrix. Sinipinic acid was prepared at 10 mg/mL in 30% acetonitrile with 0.1% trifluoroacetic acid. The sample solution was mixed with the matrix solution in and spotted on the MALDI target plate. Throughout the text, peaks are rounded to the nearest unit.

Western blot. SDS-PAGE was run in the same manner as described previously, except that a 1.0 mm gel was used instead of a 0.75 mm gel. All protein solutions except the MW lanes were combined with two parts mixed Laemmli sample buffer. Lane 1 (MW) was loaded with the Precision Plus Protein® molecular weight marker standard. Lane 2

was loaded with 10 µL of protein from combined fractions 17 – 21 (Figure 20). The protein for lanes 3 and 4 were provided by Brian Sook and contained 10 µL of a dilute SiaA WT solution and 0.3 µL of a concentrated SiaA WT solution, respectively. Lane 5 was left empty. Lane 6 (MW) was loaded with the Precision Plus Protein® molecular weight marker standard. Lane 7 was loaded with 1 µL of protein from combined fractions 17 – 21 (Figure 20). Lane 8 was loaded with 5 µL of protein from combined fractions 17 – 21 (Figure 20). Lanes 9 and 10 were loaded identically to lanes 3 and 4, respectively. After electrophoresis, the gel was cut down lane 5; lanes 1 – 4 were stained and destained as described previously. A solution (1 L) of 10X CAPS was made by dissolving 22.13 g CAPS (3-cyclohexylamino-1-propane sulfonic acid, Fischer) into 800 mL nanopure water, adjusting the pH to 11.0 with NaOH, and diluting up to 1 L with nanopure water. A solution of 1X CAPS (2 L) was prepared by combining 200 mL 10X CAPS, 200 mL MeOH, and 1.6 L nanopure water. The portion of the gel containing lanes 6 – 10, 1 sheet of Invitrogen nitrocellulose membrane (0.45 µm pore size, LC2001), and two pieces blotting paper were soaked in 1X CAPS for 10 min, taking care that the gel did not touch the nitrocellulose membrane. A plastic clamp with two sponges was submerged in 1X CAPS and the apparatus was clamped together in the following order: sponge, blotting paper, membrane, gel, blotting paper, and sponge. The apparatus was loaded into a transfer chamber filled with 1X CAPS such that the gel was closest to the negative electrode and the membrane was closest to the positive electrode. The transfer chamber was kept in a tub of ice inside a 4°C refrigerator during the run of 1 h at 70 V. PBS-Tween (3 L) was prepared by combining 24 g NaCl (Fischer), 0.6 g KH_2PO_4 (Fischer), 3.45 g Na_2HPO_4 (Sigma), 0.6 g KCl (Fischer), and 0.6 g sodium azide. A solution (50 mL) of blocking solution, PBS-Tween and 5% skim milk, was prepared by dissolving 2.5 g Carnation Instant Nonfat dry milk into 50 mL PBS-Tween. The primary antibody solution was made by diluting 3 µL of the monoclonal IgG antibody Anti Xpress N-Term (Invitrogen) into 15 mL of PBS-Tween. The secondary antibody solution was made by diluting 1.5 µL of the secondary antibody (Anti-Mouse IgG, A-5153 from goat, Sigma) into 15 mL of PBS-Tween. A BCIP stock solution was prepared by dissolving 5 mg of BCIP (5-bromo-4-chloro-3-indolyl phosphate, Sigma) in 1 mL dimethyl formamide (Fischer). A stock solution of NBT (nitro-blue tetrazolium, Sigma, 1 mg/mL) was prepared by dissolving 5 mg NBT into 5 mL of nanopure water. The alkaline phosphatase (AP) substrate mix (10 mL) was prepared by combining 9.5 mL 50 mM Tris-HCl pH 7.4, 40 µL of 1 M $MgCl_2$, 0.5 mL of NBT stock, and 0.2 mL of BCIP stock. The membrane was soaked in blocking solution for 1 h, and washed with PBS-Tween three times. The membrane was incubated on a rotary shaker with the monoclonal IgG primary antibody (Anti Xpress N-Term, Invitrogen) diluted in PBS-Tween for 1 h. The primary antibody solution was decanted and the membrane was rinsed with nanopure water once and PBS-Tween three times. The membrane was incubated on a rotary shaker with the secondary antibody for 1 h. The secondary antibody solution was decanted and the membrane was rinsed with nanopure water once and PBS-Tween three times. The membrane was washed on a rotary shaker for 1 h in 50 mM Tris-HCl pH 10. The Tris-HCl solution was decanted, and the AP substrate mix was added and the membrane was allowed to develop for about 1 h. The membrane was flushed with nanopure water to stop development and dried between two pieces of filter paper in the dark.

Reduction with dithionite. The spectrum of a solution of EK cleaved M79A SiaA (10 mM Tris-HCl, pH 8.0, plastic cuvette) was taken on a Cary 50 Bio UV-Visible spectrometer. The solution was gently bubbled with nitrogen via syringe for approximately 20 min to remove dissolved oxygen. The solution (about 1 mL) was transferred to a mini Eppendorf tube (1.7 mL) and reduced by adding a few grains of sodium dithionite. The solution was inverted 5 times and the spectrum was taken immediately.

Generation of cleavage possibilities. The complete construct of WT SiaA and M79A SiaA were submitted individually to the ExPASy FindPept tool (http://ca.expasy.org/tools/findpept.html). The sequences were entered along with the observed MALDI molecular ion peak for each protein. The query for M79A SiaA cleavage possibilities was executed with the molecular weight of 30464.38 Da (Figure 16). The query for WT SiaA cleavage possibilities was executed with the molecular weight of 30528.8 Da (Figure 27). The queries were filtered by choosing monoisotopic masses with a positive charge and with a mass tolerance of ± 5 Da from the observed weight.

Results and Discussion

Homology modeling.

Because the three-dimensional (3D) structure of SiaA is as yet unresolved, homology modeling was used to create a model of the protein. A homology model is the superposition of the residues from a protein of unknown 3D structure into the dimensional coordinates of a protein with a known 3D structure. The 3D-JIGSAW program was used. 3D-JIGSAW works by taking the input sequence and searching the Protein Database (PDB) for homologous proteins. If the user submits the protein sequence in interactive mode, and if multiple homologues are identified, then the user can choose from the different homologous structures to visualize the submitted protein sequence. The sequence of WT SiaA gave multiple hits in 3D-JIGSAW. The protein showing the highest homology was the cobalamin binding protein BtuF. WT SiaA was superimposed into the protein structure of BtuF (Borths et al., 2002) in IN4D (ligand unbound) structure and visualized with RasMol (Figure 3). The five residues that are possible axial ligands in Figure 3 are one of three amino acids: histidine, tyrosine, or methionine.

Growth and isolation of M79A SiaA.

Three production runs of *E. coli* were completed: a 3 L run, a 2 L run, and a 9 L fermentation run. In the 9 L run, the plate showed only *E. coli* growth, and approximately 9.4 L of growth were recovered from the to give about 60 g of cell pellet after centrifugation.

An SDS-PAGE gel (not shown) on the cell pellet from the 3 L run verified over-expression of protein. A nickel affinity FPLC chromatogram was run starting with a binding buffer of 50 mM potassium phosphate and 250 mM NaCl at pH 7.4. The majority of unbound protein appeared to be eluted after about 50 mL (10 CV) of binding buffer. Elution buffer (binding buffer plus 0.5 M imidazole) was then added, ramping up to 100% elution buffer after another 50 mL. Figure 4 shows significant protein eluting in an unresolved double peak at 117 and 129 min (a buffer range of about 0.05 M – 0.25 M

imidazole). By the time the eluent had reached a concentration of 0.25 M imidazole, nearly all proteins had eluted.

SDS-PAGE of the fractions 10 – 18, Figure 5, indicated that only three fractions (16, 17, and 18) contained significant amounts of SiaA. Fraction 15 was the first fraction that gave a SiaA band, while the earlier fractions (10 – 14) were primarily contaminant proteins. By fraction 16, only four light contaminant protein bands were observed. One contaminant was heavier than SiaA (33 kD) while three were lighter. Lanes 17 and 18 appeared identical to lane 16 qualitatively on SDS-PAGE, while the quantity of protein was decreasing with 16 > 17 > 18. By lane 18, only the SiaA band was visible.

For the 3 L run, SiaA from 1.5 L was purified through the nickel affinity column step. A Lowry assay, with the standard curve shown in Figure 6, was performed on a portion of the Ni affinity fractions that had been combined and exchanged into 10 mM Tris-HCl buffer (2.0 mL total volume). A protein concentration of 1.75 mg/mL was calculated. Thus, the total mass of the M79A SiaA that was isolated was 3.5 mg. Given a total growth culture 1.5 L and a yield of 3.5 mg, M79A was produced at approximately 2.3 mg/L of growth culture.

After EKMax™ digestion, strong anion exchange FPLC (Figure 7) with a gradient of NaCl in the buffer was used to purify the protein. The SiaA eluted at approximately 97 min, at an NaCl concentration of approximately 0.4 M NaCl. A post anion exchange SDS-PAGE (Figure 8) showed strong anion exchange FPLC was successful in removing the contaminant proteins found in solution after nickel affinity FPLC. The post-anion exchange SDS-PAGE gel also indicated that the EKMax™ digestion was successful. Lane EK shows a sample that was taken from EKMax™ digestion immediately after the addition of enterokinase and frozen; bands representing both cleaved (33 kD) and uncleaved (35 kD) SiaA are visible.

Figure 9 shows the UV-Vis spectrum of the oxidized EK cleaved M79A SiaA. The UV-Vis spectrum shows a much smaller Soret/280 ratio than the spectrum of the SiaA forms without EK cleavage (Figure 24). This is consistent with loss of heme during the EK cleavage and subsequent purification steps. In general, it seemed that heme was lost during manipulation and purification of M79A SiaA. Figure 9 also shows the protein after treatment with dithionite. The Soret is red-shifted and the α and β bands are evident at 552 nm and 529 nm, respectively.

The nickel affinity chromatogram from the 2 L run of SiaA (Figure 10) shows a much smaller total elution peak area for M79A SiaA than is seen in the 3 L run (Figure 4). Besides the volume, the only difference between the runs is that the 2 L run of SiaA had 30 mM imidazole in the binding buffer. SiaA is the main peak in Figure 10 (900 mAU) in the 2 L run of SiaA and contaminants represent the shoulder (350 mAU). The contaminants are the main peak in Figure 4 (3500 mAU) in the 3 L run and SiaA represents the shoulder (1200 mAU). The SDS-PAGE separation of the contaminant proteins are seen in Figure 11 (lanes 12 – 14) for the 2 L run of SiaA and Figure 5 (lanes 12 – 14) for the 3 L run of SiaA. In both cases SiaA elutes at a higher imidazole concentration that the contaminant proteins, but the amount of contaminant proteins is decreased dramatically upon the addition of 30 mM imidazole in the binding buffer (2 L run).

At the end of the purification for the 2 L run, the M79A protein appeared pure. Only one band was observable in SDS-PAGE (Figure 15). M79A was probably isolated

most effectively due to combination of a Ni affinity binding buffer with 30 mM imidazole, and two cycles of weak anion exchange FPLC (Figures 12 and 14).

The final solution of M79A SiaA from the 9 L run was analyzed to predict the amount of holoprotein purified. Assuming an extinction coefficient of 1.0×10^5 $mol^{-1}cm^{-1}$ at 400 nm, the final protein solution of the 9 L run yielded 0.61 mg/L of holo-M79A SiaA. This calculation is gives only the amount of holoprotein because apoprotein does not absorb at 400 nm. The final M79A solution from the 9 L run was also found to contain a protein weighing 25 kD (Figures 21 and 23). However, because the 25 kD protein band did not develop on a Western blot (Figure 22) using the Invitrogen antibody to the N-terminal portion of the construct, the 25 kD fragment is not likely to be a portion of SiaA.

ALA was added to the medium of a 100 mL run of M79A SiaA with the expectation that it would increase the production of heme, and therefore perhaps result in more holoprotein, as has been observed for other heme proteins. However, in this run, ALA did not stimulate the production of heme (Figure 25).

Forms of SiaA.

Three procedures were used to break open the cells. Procedure A involved 6 cycles of 30 sec each with the protein kept on ice. Procedure B involved 5 cycles of 2 min each with the protein kept on ice. Procedure C involved French pressure lysis of the cells. MALDI mass spectrometry of the protein from Procedure B showed a protein of 30464 Da (Figure 16). This was approximately 5094 Da lighter than the weight expected from the sequence. Possible cleavage points in the sequence that would result in a mass of 30464 Da are shown in Figure 28. Procedure C gave a protein of 33358 Da, 2000 Da lighter than the expected mass of 35558 Da.

Cleavage could have been induced in Procedure B either by the vigorous sonication itself, by proteases which were released by the sonication, or by increased activity of the proteases at the higher temperatures that resulted from the more extensive sonication. It is likely that sonication itself did not cause the cleavage of the protein down to 30464 Da because the shorter form of the protein (30471 Da) was also observed as a minor product in the solution in which a French press was used to break open the cells.

Cleavage might be a process that is a function of storage time. If the protein is stored at 4°C for an extended period of time, cleavage might occur either from proteases in solution or via auto-cleavage. The effect of cleavage due to storage may explain the differences between Figure 13 and Figure 15. In Figure 13, after weak anion exchange FPLC, three proteins are present, one each at approximately 35 kD, 33 kD, and 30 kD. However, after a repurification with another weak anion exchange FPLC, only one protein is present in Figure 15; it is at approximately 30 kD. Although Procedure B resulted in a shorter form of the protein than Procedure C by 2894 Da (30464 Da vs. 33358 Da), the UV-Vis spectra of these two forms of the protein are almost identical in the 350 – 700 nm region (Figure 24).

Shorter forms of the protein (30529 Da) were also observed when the WT SiaA was isolated with vigorous sonication (Figure 27). Possible cleavage points are shown in Figure 29. Possible cleavage points for the M79A SiaA solution from Procedure B are shown in Figure 28. Comparing Figure 28 with Figure 29 indicates that three cleavage possibilities are shared, resulting in three likely candidates for the isolated protein:

residues 16 to 282, 30 to 296, and 45 to 312. Of these, the third (45 to 312) seems unlikely because homology modeling (Figure 26) shows that the N-terminal cleaved portion would be deep inside the protein. Were this tightly packed portion of the protein to be removed it is unlikely that the protein would remain folded. The similarity of the short and long forms of the M79A protein in the UV-Vis spectra (Figure 24) indicate that the short form is very similar to the long form near the heme, thus ruling out significant changes to the globular domains that support the heme pocket. The cleavage points of sequences 16 to 282 and 30 to 296 are outside the area of homology modeling, and thus no choice can be made between them at this point.

Conclusions

M79A SiaA was successfully produced and purified to homogeneity. Encouragingly, it retains enough heme for spectroscopic and denaturation studies. However, it does appear to lose heme during purification.

Unexpectedly, initial nickel affinity FPLC resulted in the elution of a high concentration of proteins close to His-tagged SiaA. The addition of 30 mM imidazole in the binding buffer (2 L run) resulted in many fewer contaminant proteins eluting close to SiaA. The 2 L run also underwent two cycles of weak anion exchange FPLC; the first was done with an elution buffer with 1.0 M NaCl eluent and the second was performed with an elution buffer of 0.4 M NaCl eluent. This gave the purest form of SiaA.

For the 3 L run, EKMax^{TM} digestion and anion exchange FPLC resulted in cleaved and purified fractions of SiaA. The EK cleaved SiaA was successfully reduced, with α and β bands evident in the UV spectrum.

Different purification protocols resulted in SiaA proteins that differed in molecular mass by a few thousand Daltons. Multiple variables could influence cleavage of SiaA: rigorous sonication, temperature, proteases, and storage time. The observation of shorter versions of SiaA in both the rigorous sonication and French press protocols leads us to conclude that sonication *per se* does not cleave the protein. Despite a range in sizes of purified SiaA, the UV spectra of the different forms are similar, suggesting that cleavage does not alter the heme binding pocket. The spectral observations, as well as comparison of the cleavage forms of M79A with WT SiaA, and homology modeling, indicate that the cleavage sites are outside of the well-folded core of the protein.

Reference List

Bates, C. S., Montanez, G. E., Woods, C. R., Vincent, R. M., and Eichenbaum, Z. (2003). Identification and characterization of a *Streptococcus pyogenes* operon involved in binding of hemoproteins and acquisition of iron. *Infect. Immun.* **71**, 1042-1055.

Borths, E. L., Locher, K. P., Lee, A. T., and Rees, D. C. (2002). The structure of *Escherichia coli* BtuF and binding to its cognate ATP binding cassette transporter. *Proc. Natl. Acad. Sci. USA* **99**, 16642-16647.

Cunningham, M. W. (2000). Pathogenesis of group A streptococcal infections. *Clin. Microbiol. Rev.* **13**, 470-511.

Eakanunkul, S., Lukat-Rodgers, G. S., Sumithran, S., Ghosh, A., Rodgers, K. R., Dawson, J. H., and Wilks, A. (2005). Characterization of the periplasmic heme-binding protein ShuT from the heme uptake system of *Shigella dysenteriae*. *Biochemistry* **44**, 13179-13191.

Eichenbaum, Z., Green, B. D., and Scott, J. R. (1996a). Iron starvation causes release from the group A streptococcus of the ADP-ribosylating protein called plasmin receptor or surface glyceraldehyde-3-phosphate-dehydrogenase. *Infect. Immun.* **64**, 1956-1960.

Eichenbaum, Z., Muller, E., Morse, S. A., and Scott, J. R. (1996b). Acquisition of iron from host proteins by the group A *Streptococcus*. *Infect. Immun.* **64**, 5428-5429.

Jones, P. M. and George, A. M. (2004). The ABC transporter structure and mechanism: Perspectives on recent research. *Cell Mol. Life Sci.* **61**, 682-699.

Jonsson, I. M., Mazmanian, S. K., Schneewind, O., Bremell, T., and Tarkowski, A. (2003). The role of *Staphylococcus aureus* sortase A and sortase B in murine arthritis. *Microbes Infect.* **5**, 775-780.

Karpowich, N. K., Huang, H. H., Smith, P. C., and Hunt, J. F. (2003). Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. *J. Biol. Chem.* **278**, 8429-8434.

Lei, B. F., Liu, M. Y., Prater, C. I., Kala, S. V., Deleo, F. R., and Musser, J. M. (2003). Identification and characterization of HtsA, a second heme-binding protein made by *Streptococcus pyogenes*. *Infect. Immun.* **71**, 5962-5969.

Lei, B. F., Smoot, L. M., Menning, H. M., Voyich, J. M., Kala, S. V., Deleo, F. R., Reid, S. D., and Musser, J. M. (2002). Identification and characterization of a novel heme-associated cell surface protein made by *Streptococcus pyogenes*. *Infect. Immun.* **70**, 4494-4500.

Liu, M. Y. and Lei, B. F. (2005). Heme transfer from streptococcal cell surface protein Shp to HtsA of transporter HtsABC. *Infect. Immun.* **73**, 5086-5092.

Locher, K. P. (2004). Structure and mechanism of ABC transporters. *Curr. Opin. Struct. Biol.* **14**, 426-431.

Locher, K. P. and Borths, E. (2004). ABC transporter architecture and mechanism: Implications from the crystal structures of BtuCD and BtuF. *FEBS Lett.* **564**, 264-268.

Mack, J., Vermeiren, C., Heinrichs, D. E., and Stillman, M. J. (2004). *In vivo* heme scavenging by *Staphylococcus aureus* IsdC and IsdE proteins. *Biochem. Biophys. Res. Commun.* **320**, 781-788.

Mazmanian, S. K., Hung, I. T., and Schneewind, O. (2001). Sortase-catalysed anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. *Mol. Microbiol.* **40**, 1049-1057.

Nygaard, T. K., Blouin, G. C., Liu, M. Y., Fukumura, M., Olson, J. S., Fabian, M., Dooley, D. M., and Lei, B. F. (2006). The mechanism of direct heme transfer from the streptococcal cell surface protein shp to HtsA of the HtsABC transporter. *J. Biol. Chem.* **281**, 20761-20771.

Sambrook, J. and Russel, D. W. (2001). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory Press.

Skaar, E. P. and Schneewind, O. (2004). Iron-regulated surface determinants (Isd) of *Staphylococcus aureus*: stealing iron from heme. *Microbes Infect.* **6**, 390-397.

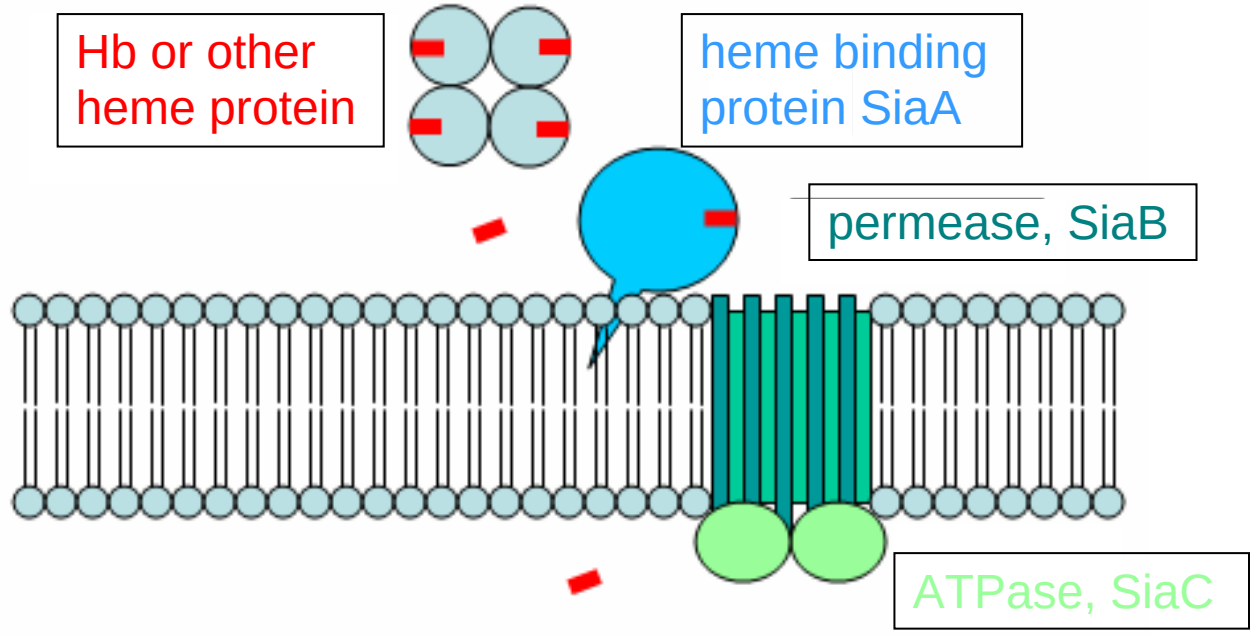

Figure 1: Representative diagram showing the heme protein hemoglobin (Hb), SiaA, SiaB, and SiaC. Heme binds SiaA, which is bound to the plasma membrane, and heme is transported through the ABC transporter system of SiaB and SiaC across the plasma membrane.

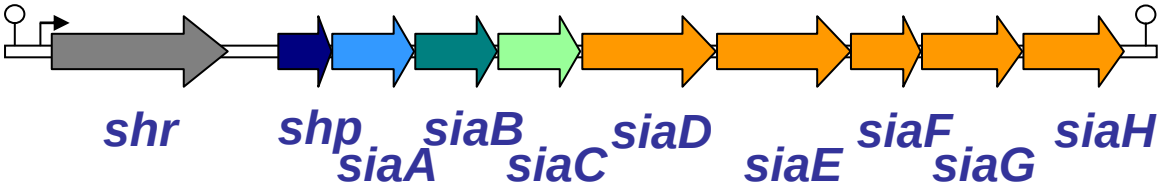

Figure 2: SiaA gene cluster.

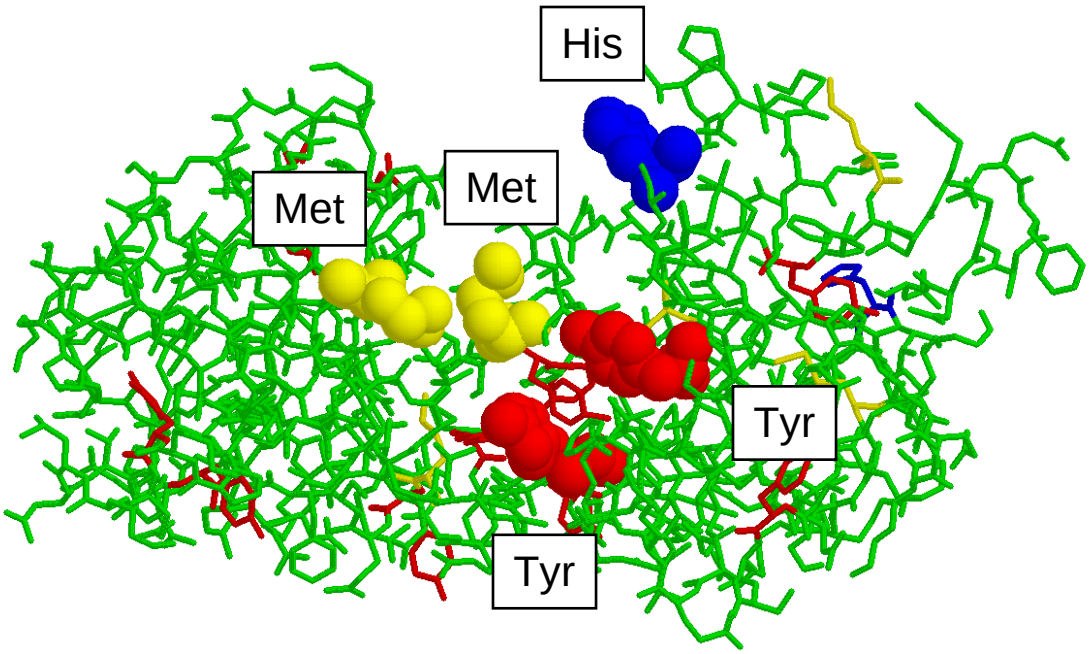

Figure 3: SiaA WT threaded into 1N4D (BtuF ligand unbound) (Borths et al., 2002). Histidine, methionine, and tyrosine residues colored blue, yellow, and red, respectively. The histidine, methionine, and tyrosine residues in the pocket are displayed using space filling models. Visualized with RasMol 2.7.2.2.1.

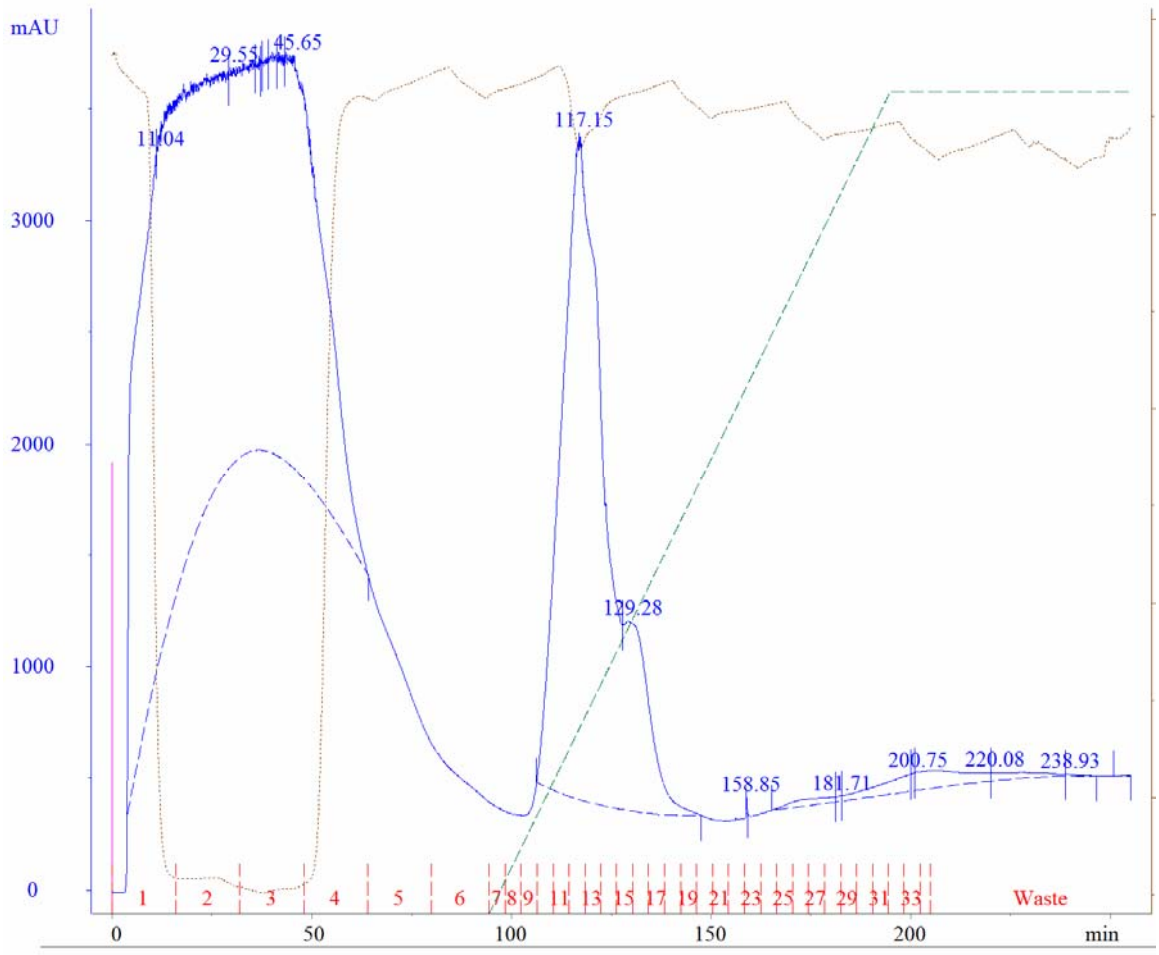

Figure 4: Nickel affinity FPLC chromatogram performed on sonicated, centrifuged, and filtered supernatant from 1.5 L of run 1 (3 L). Dotted line starting at fraction 7 indicates a linearly increasing concentration of elution buffer (binding buffer with 0.5 M imidazole). SiaA elutes at 129 min, at approximately 0.15 M imidazole. Fractions 10 – 19 were collected and purified further.

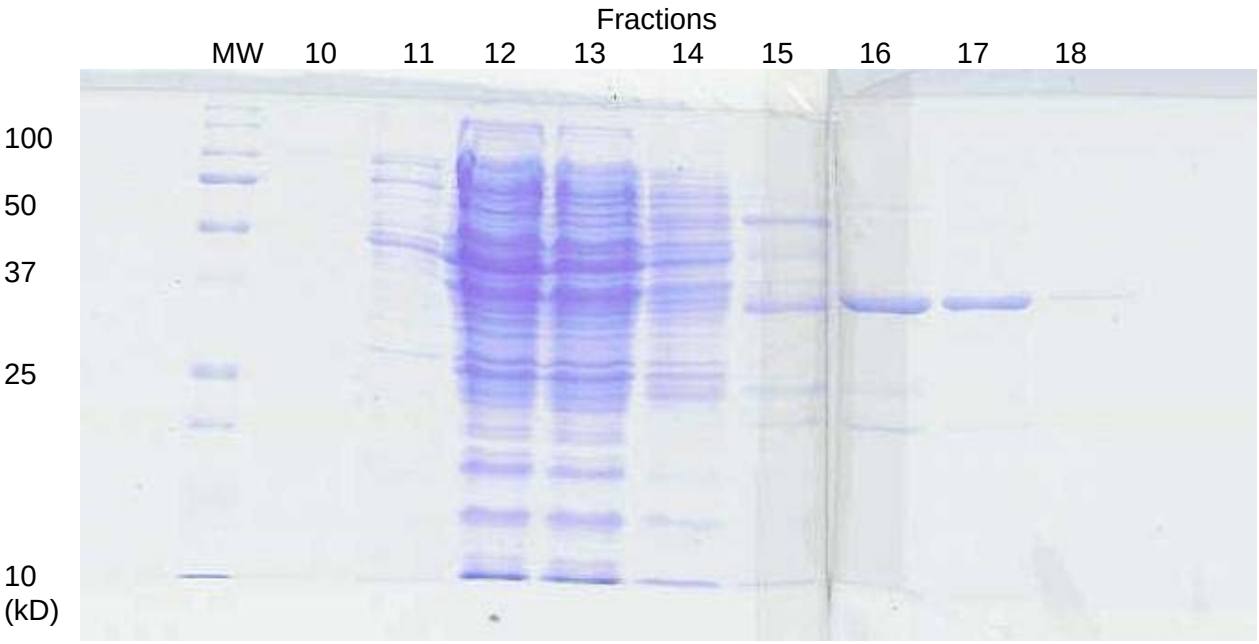

Figure 5: SDS-PAGE of nickel affinity FPLC fractions 10 – 18 from 3 L run (Figure 4). The MW lane contains the Precision Plus Protein® molecular weight marker standard. Fractions 11 – 15 contain contaminant proteins while lanes 16 – 18 contain mostly SiaA.

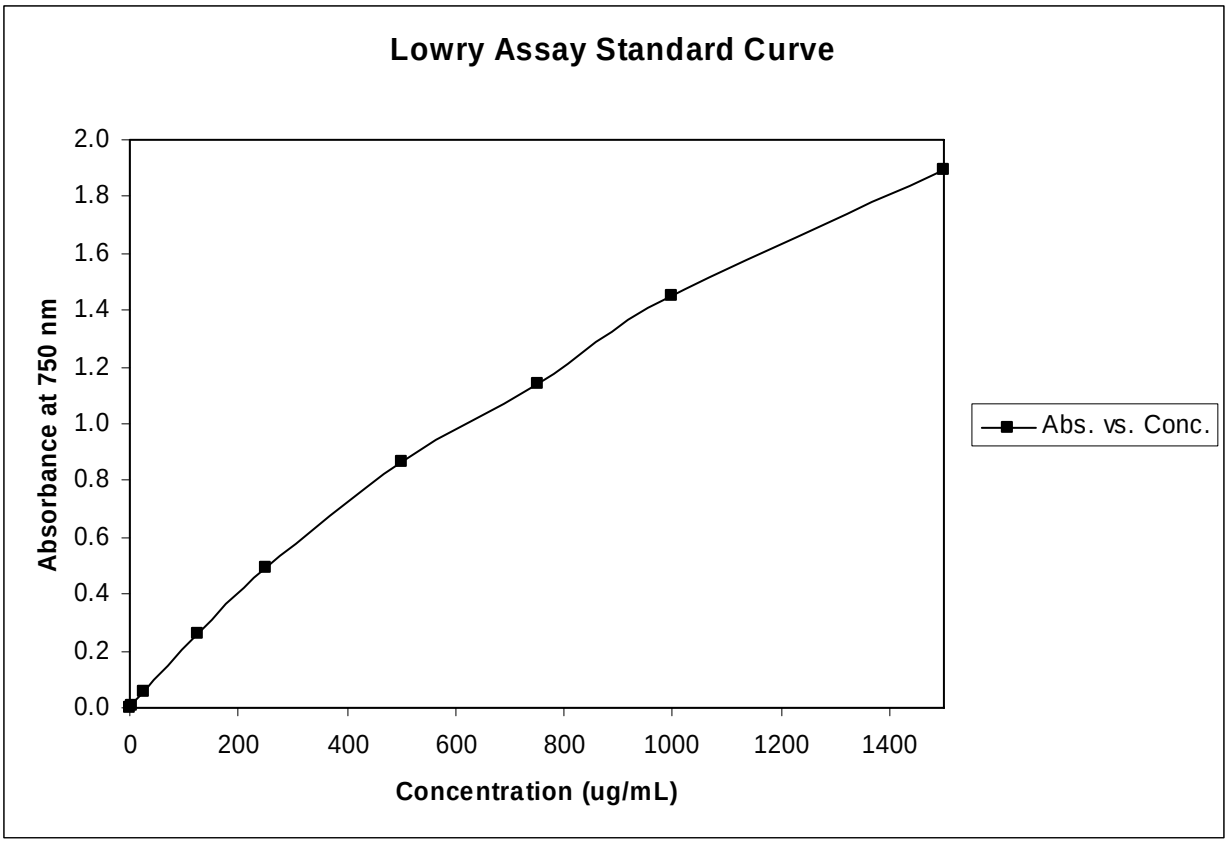

Figure 6: Lowry Assay standard curve of the absorbance at 750 nm vs. the concentrations of bovine serum albumin. The absorbance from a Lowry assay performed on combined fractions 16 – 19 from Figure 5 gave a protein concentration for SiaA of 1.75 mg/mL.

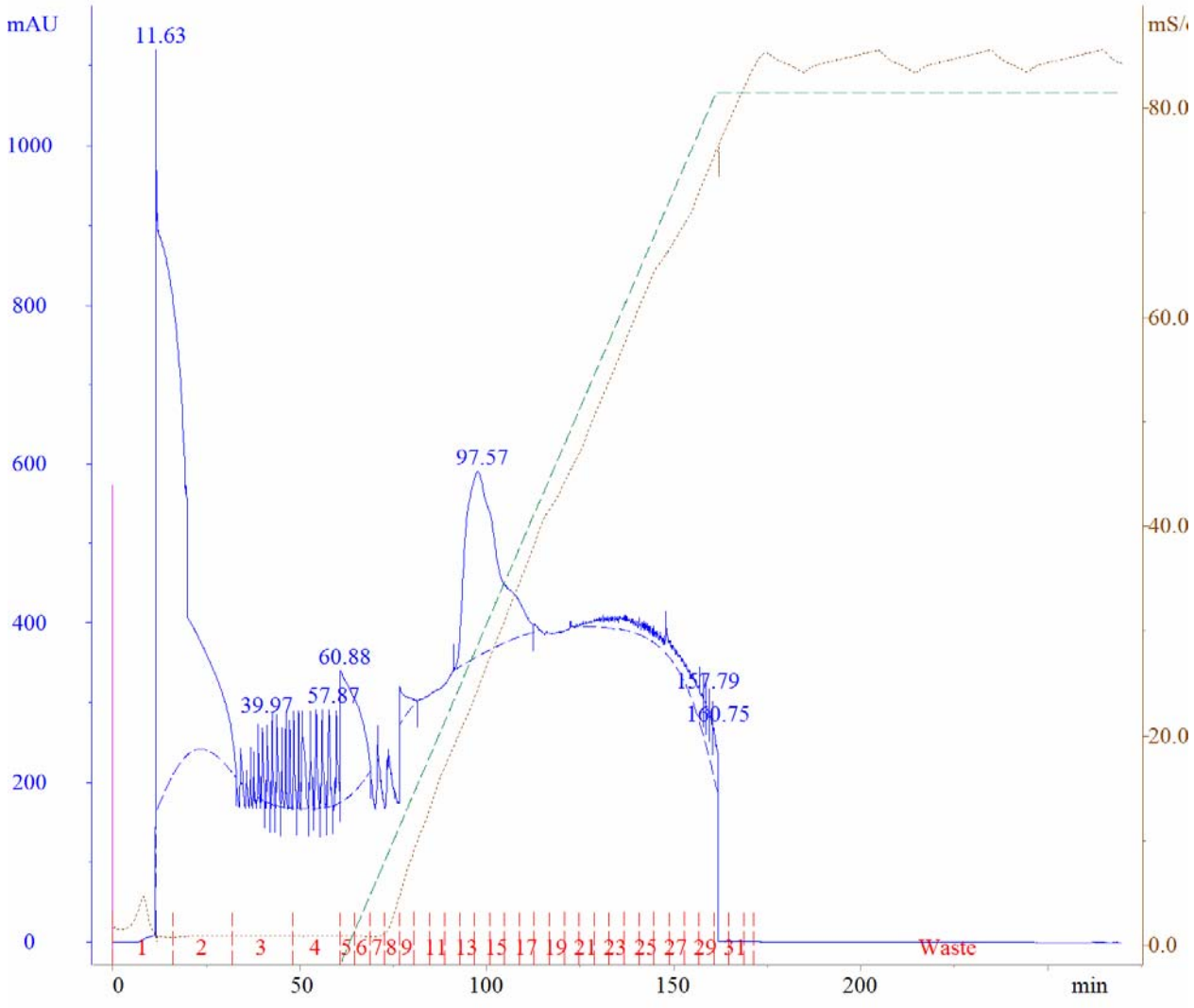

Figure 7: Strong anion exchange FPLC of EKMax™ digestion and buffer exchange solution using HiTrap Q HP (5 mL) column. Dotted line starting at fraction 5 indicates a linearly increasing concentration of NaCl in the elution buffer, ending with 1.0 M NaCl at fraction 29. SiaA starts eluting at approximately 0.32 M NaCl (97 min).

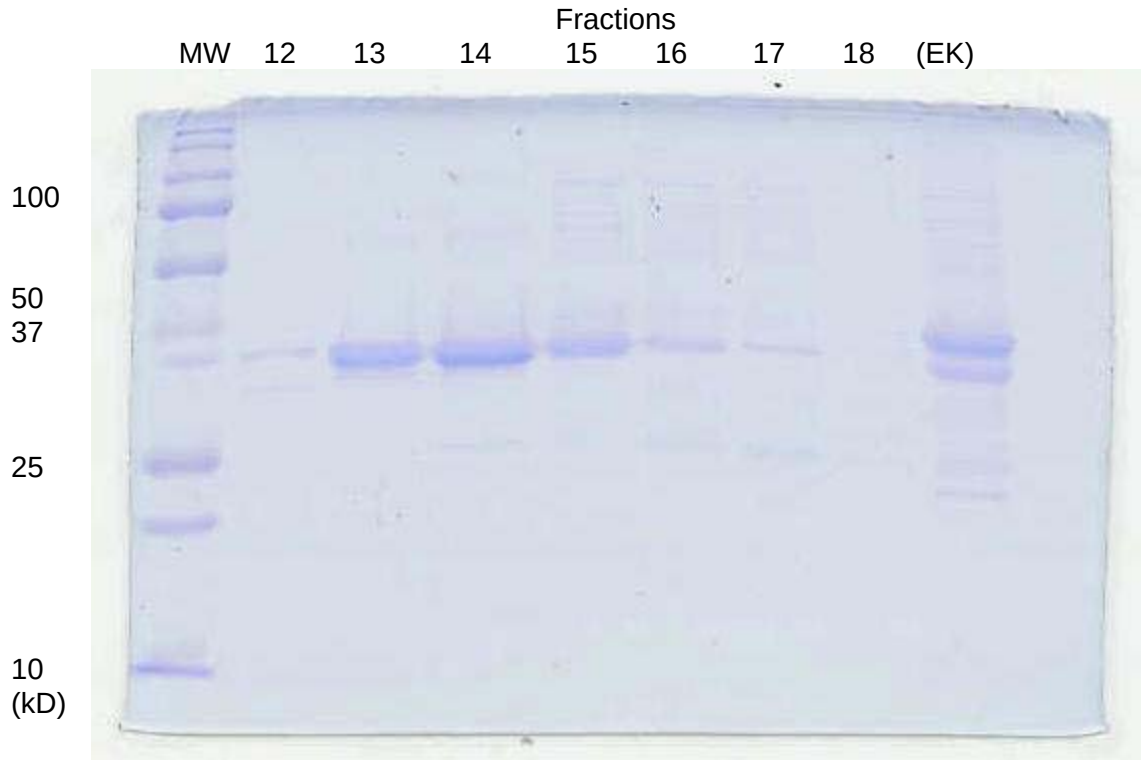

Figure 8: SDS-PAGE of strong anion exchange FPLC fractions (Figure 7). The first lane (MW) contains the Precision Plus Protein® molecular weight marker standard. Subsequent lanes contain anion exchange FPLC fractions 12 – 18. The final lane (EK) shows a sample that was taken from EKMax™ digestion immediately after the addition of enterokinase and frozen; it shows partial cleavage (the full length protein has a molecular weight of 35 kD).

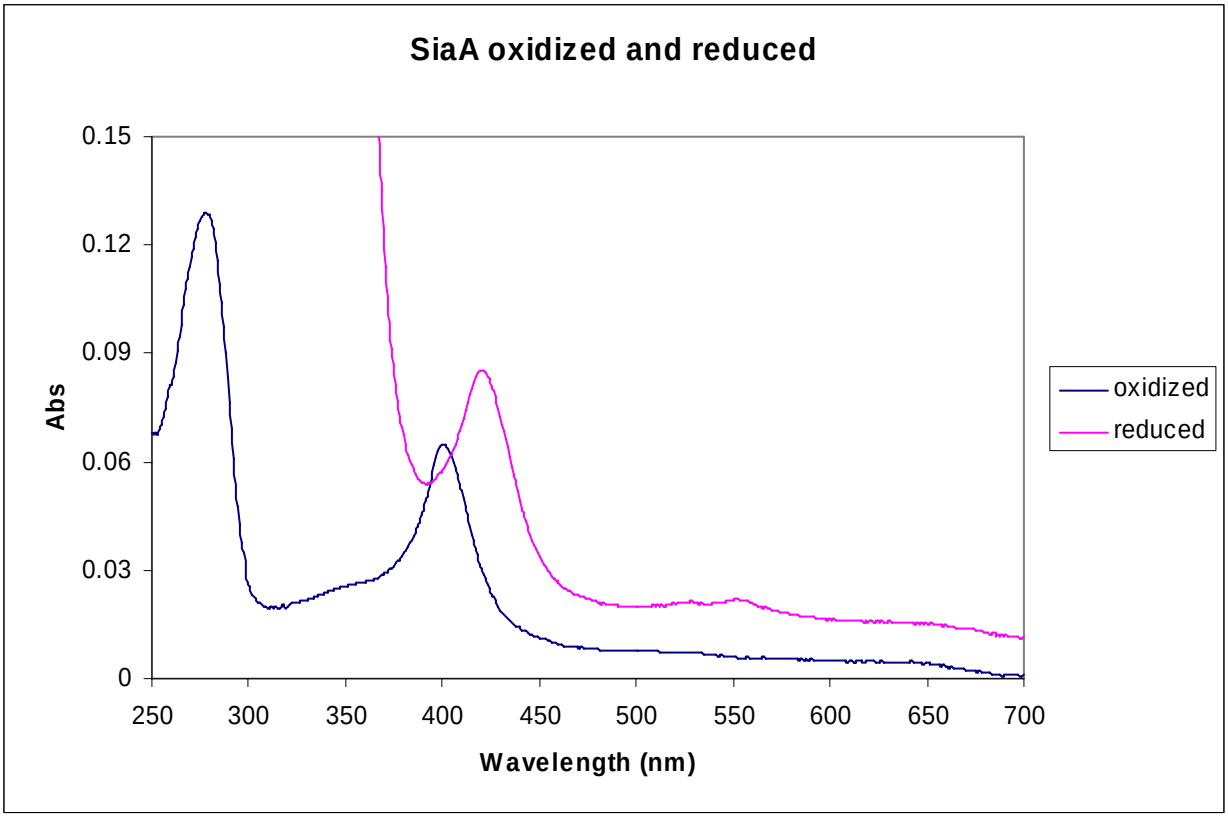

Figure 9: UV-Vis spectral overlay of reduced and oxidized SiaA (EK cleaved). Upon reduction with dithionite, the Soret is red-shifted and the α and β bands are evident at 552 nm and 529 nm, respectively.

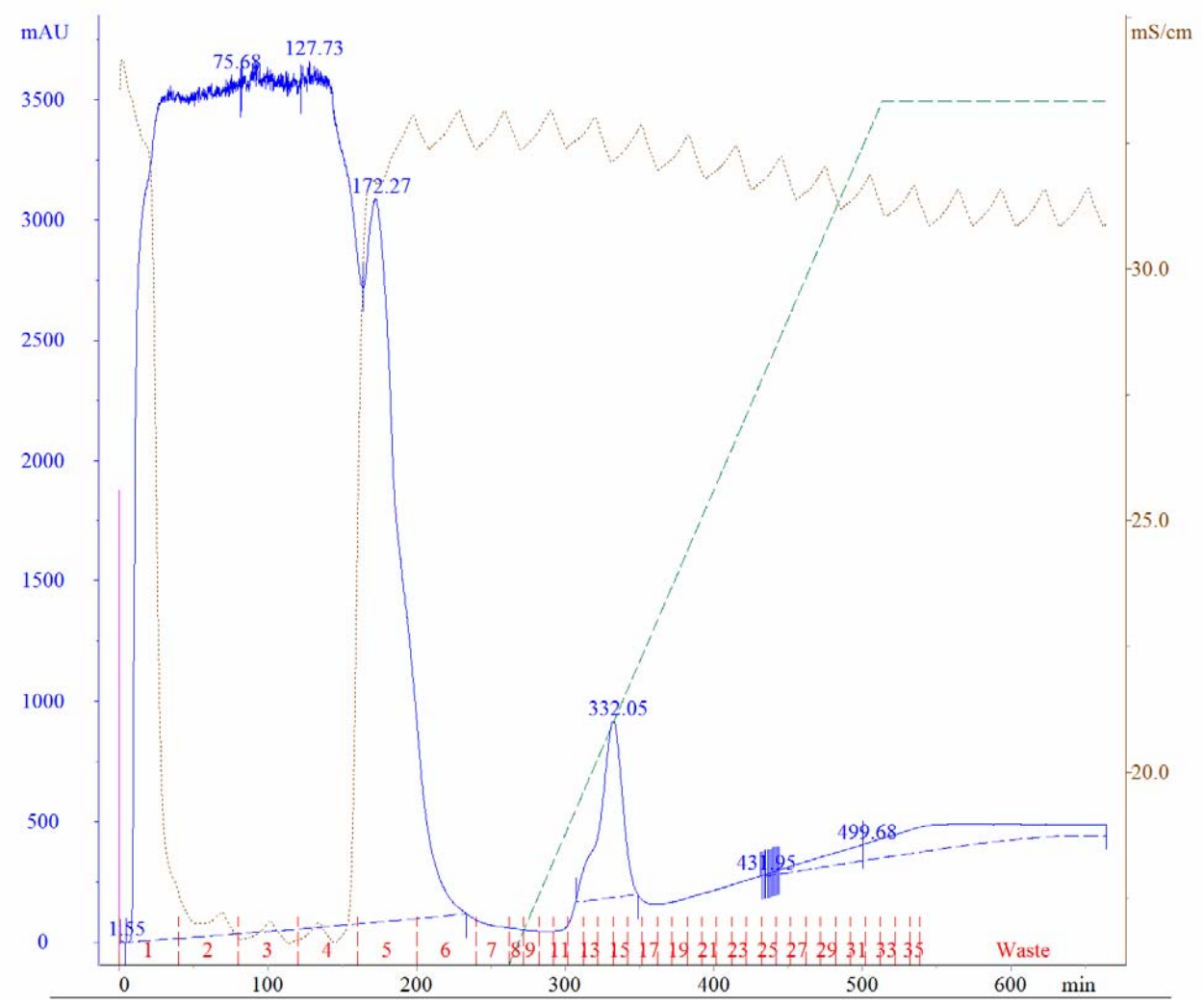

Figure 10: Ni affinity FPLC chromatogram with 30 mM imidazole in starting buffer (2 L run). Dotted line starting at fraction 8 indicates a linearly increasing concentration of elution buffer (binding buffer with 0.5 M imidazole). SiaA elutes at 332 min, at approximately 0.15 M imidazole. Fractions 12 – 17 were collected.

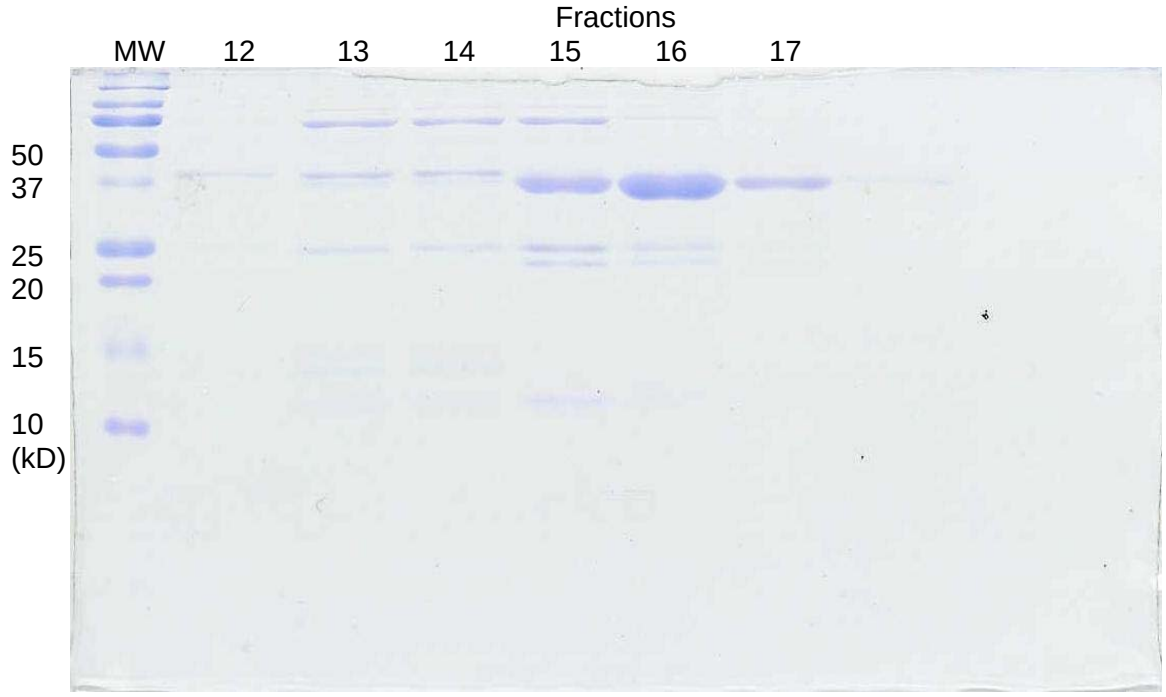

Figure 11: SDS-PAGE of fractions from Ni affinity column with 30 mM imidazole in starting buffer (Figure 10). The first lane (MW) contains the Precision Plus Protein® molecular weight marker standard. Fractions 12 – 14 show contaminant proteins while fractions 15 – 17 show SiaA. Fraction 16 shows the clearest and most concentrated SiaA band.

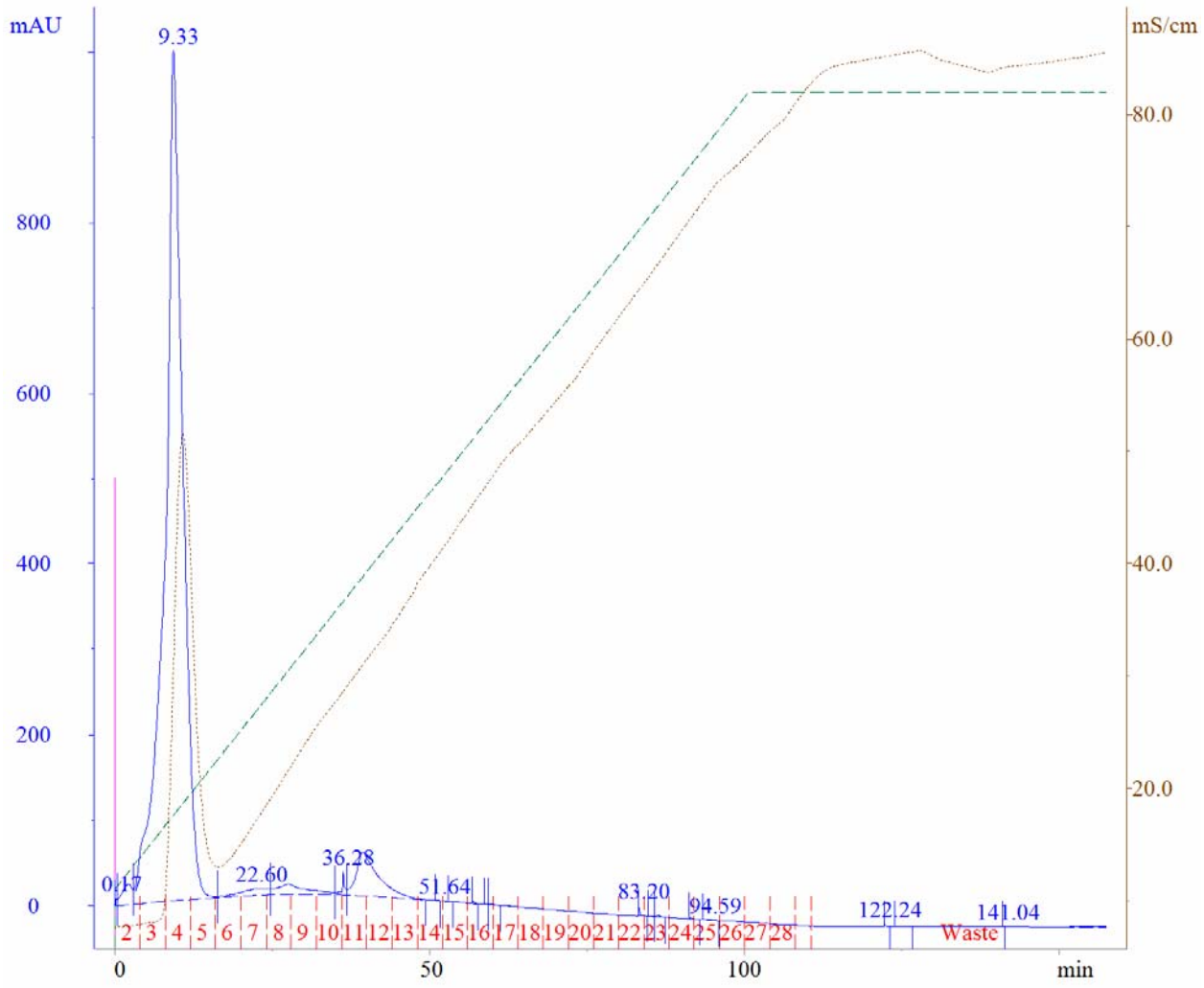

Figure 12: Weak anion exchange chromatography of fractions 14 – 17 from Figure 11. A DEAE FF column was used with 10 mM Tris-HCl pH 8.0 buffer. The linear gradient (indicated by the dashed line) started at 7% and ended at 100% 1 M NaCl in the same buffer. SiaA eluted at 9 min. The tube that was supposed to contain fraction 5 was empty. Fractions 2, 3, 4, 6, and 7 were collected.

Figure 13: SDS-PAGE of fractions 2, 3, 4, 6, and 7 from Figure 12. The first lane (MW) contains the Precision Plus Protein® molecular weight marker standard. Fraction 4 shows three protein bands, one at approximately 35 kD, another at 33 kD and one at 30 kD. The other lanes contain no visible protein.

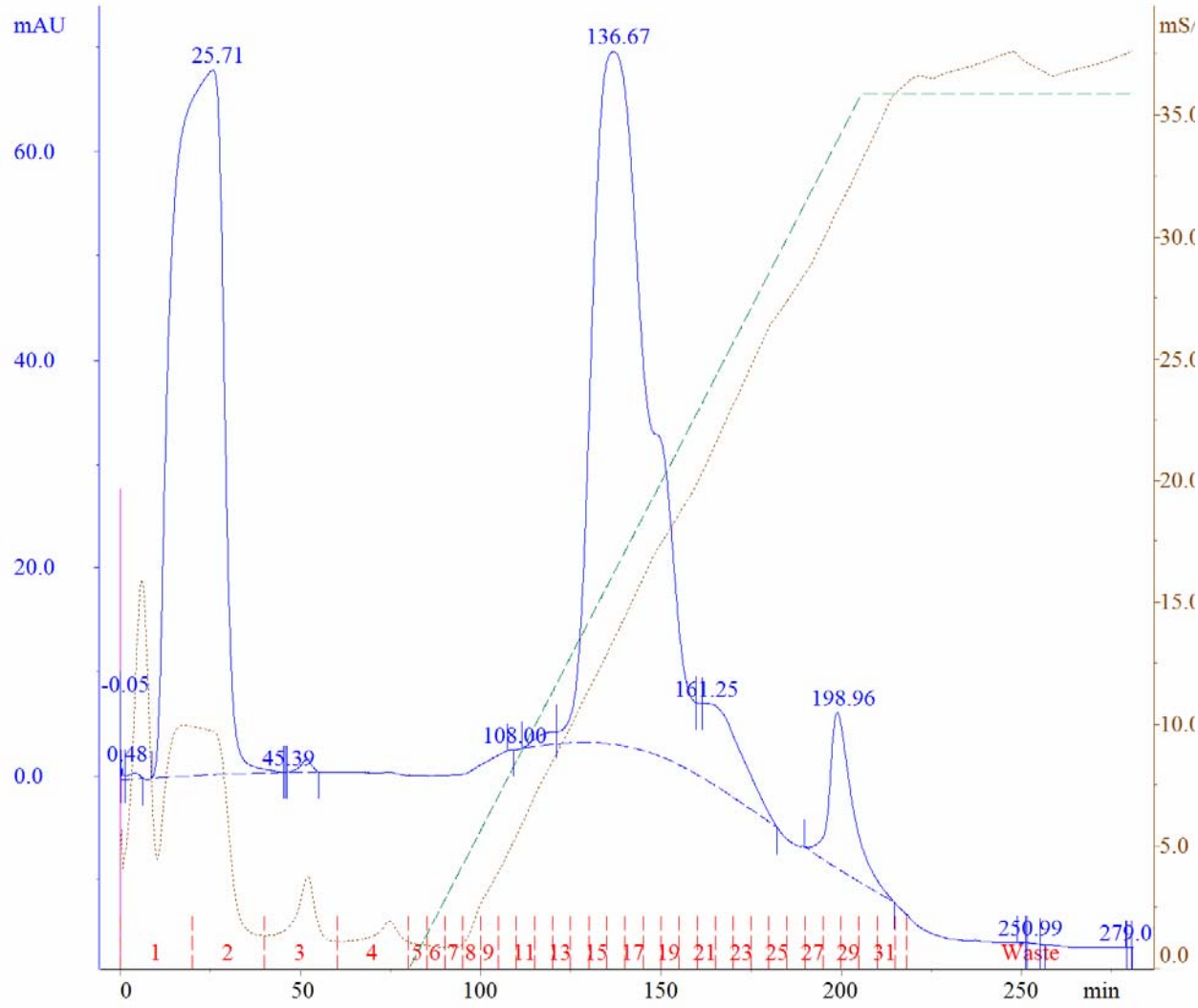

Figure 14: Repurification of fraction 4 in Figure 13 with weak anion exchange FPLC using 10 mM Tris-HCl pH 8.0 buffer as starting buffer and 0.4 M NaCl as eluting buffer with a DEAE FF column (5 mL). The gradient of eluting buffer started at fraction 5 and continued until fraction 30. The main peak eluted at 136 min and approximately 0.18 M NaCl (fractions 15 – 17). Fractions corresponding to the main peak and minor peak at 198 min (fractions 28 – 30) were collected for SDS-PAGE.

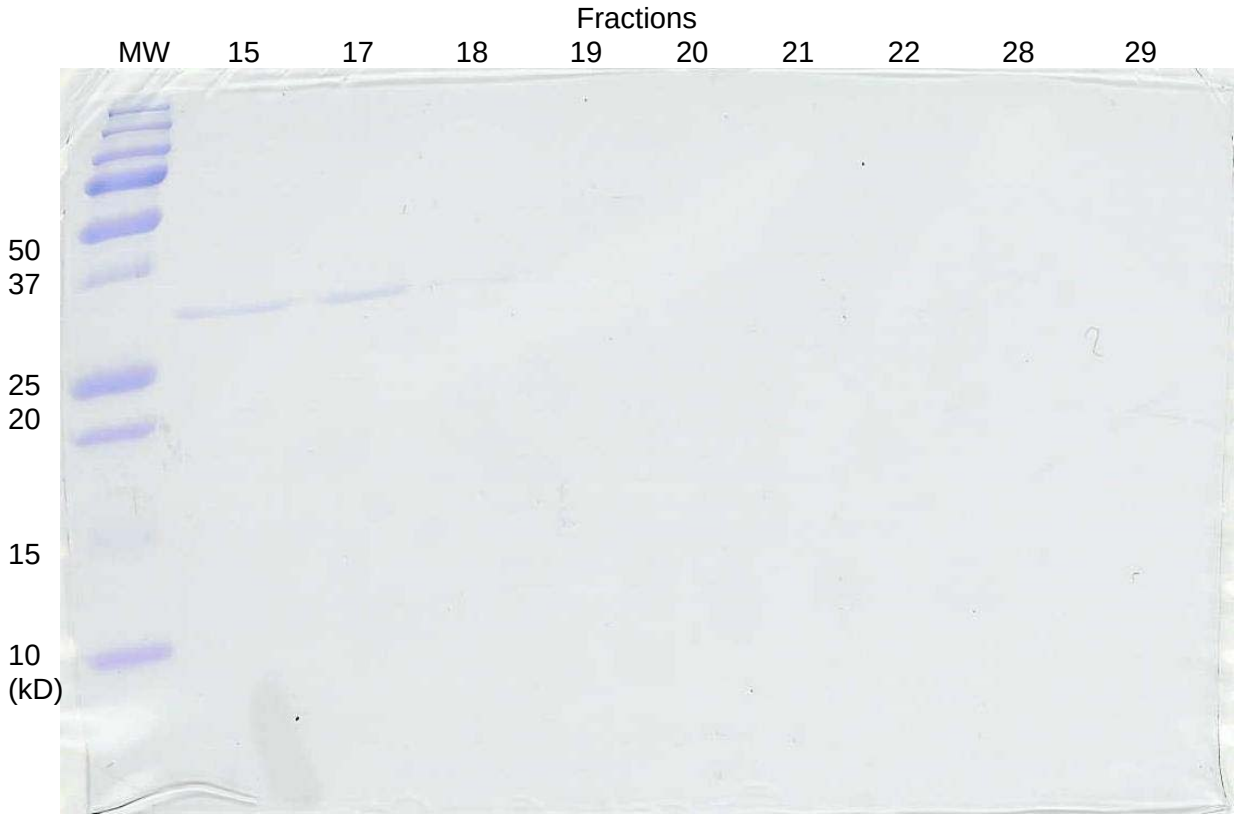

Figure 15: SDS-PAGE of fractions 15, 17 – 22, 28, and 29 from Figure 14. The first lane (MW) contains the Precision Plus Protein® molecular weight marker standard. Fractions 15, 17, and 18 display bands at approximately 30 kD. No other bands are visible.

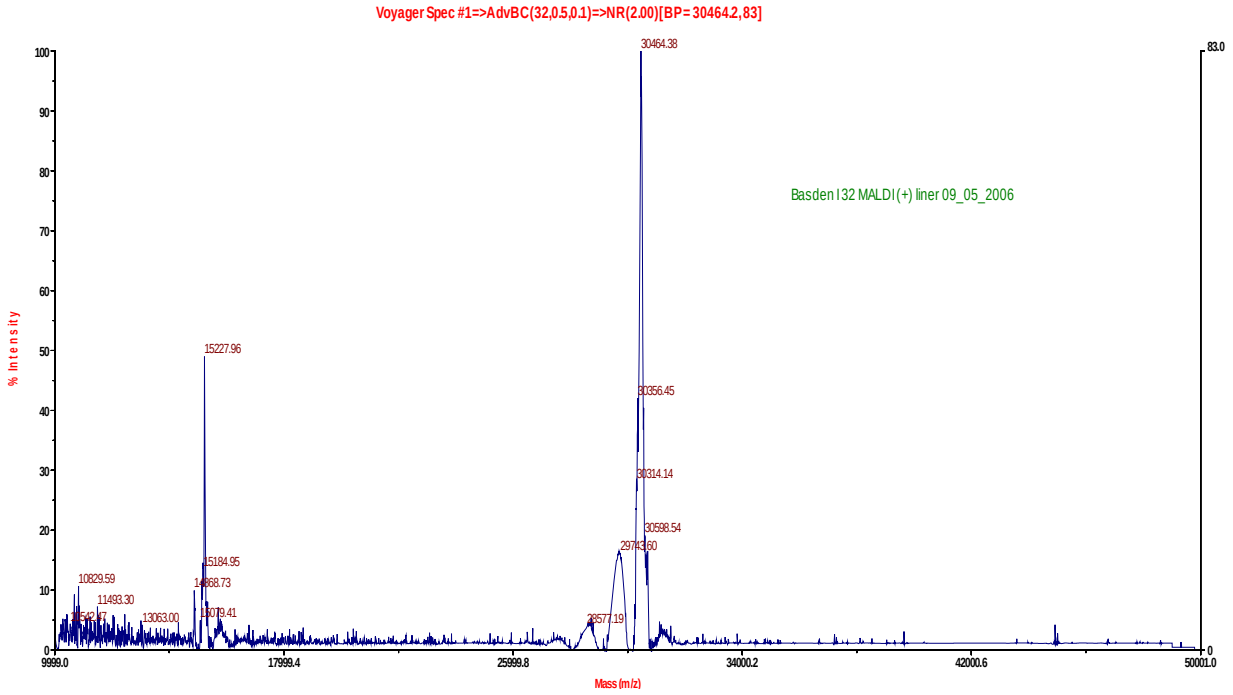

Figure 16: MALDI mass spectrometry of combined fractions 15, 16, 17, and 18 from Figure 15. The main molecular ion peak is at 30464 Da.

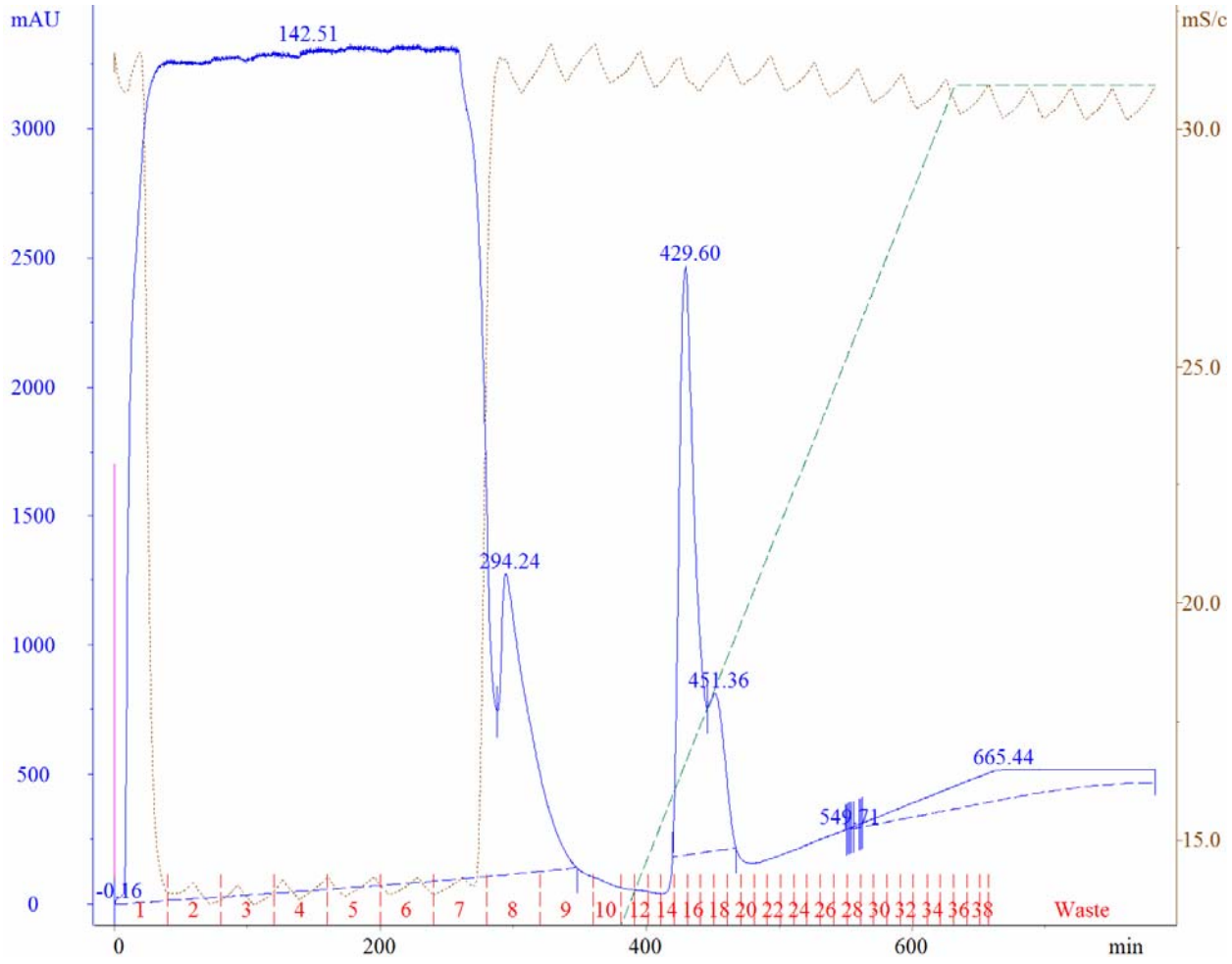

Figure 17: Ni Affinity FPLC run with no imidazole in starting buffer (6 g of cell pellet from 9 L run). Dotted line starting at fraction 10 indicates a linearly increasing concentration of elution buffer (binding buffer with 0.5 M imidazole). The peak at 429 min is not SiaA as shown by SDS PAGE (Figure 18). SiaA elutes at 451 min, at approximately 0.23 M imidazole (fractions 17 – 20).

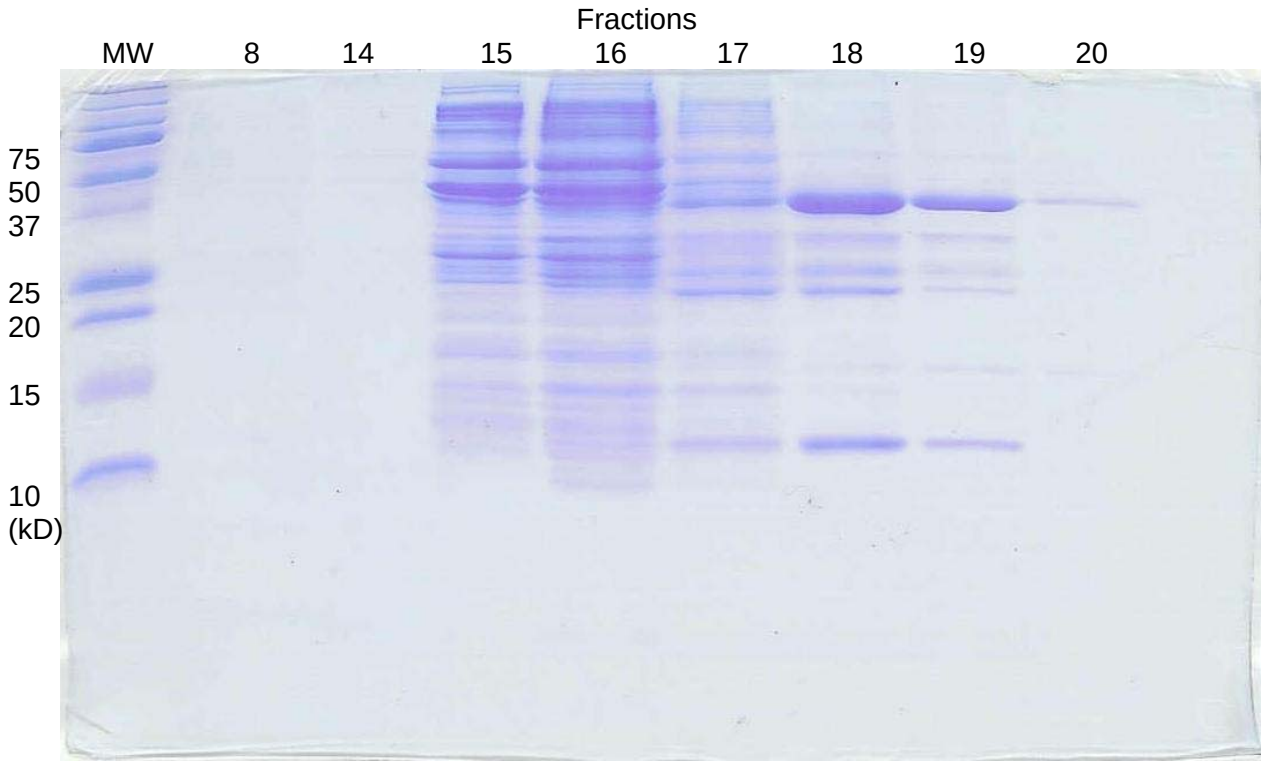

Figure 18: SDS-PAGE of Ni affinity FPLC run with no imidazole in starting buffer. The first lane (MW) contains the Precision Plus Protein® molecular weight marker standard. Lanes 15 – 17 are mostly contaminant proteins. Significant amounts of SiaA, along with contaminant proteins, are seen in fractions 18 and 19.

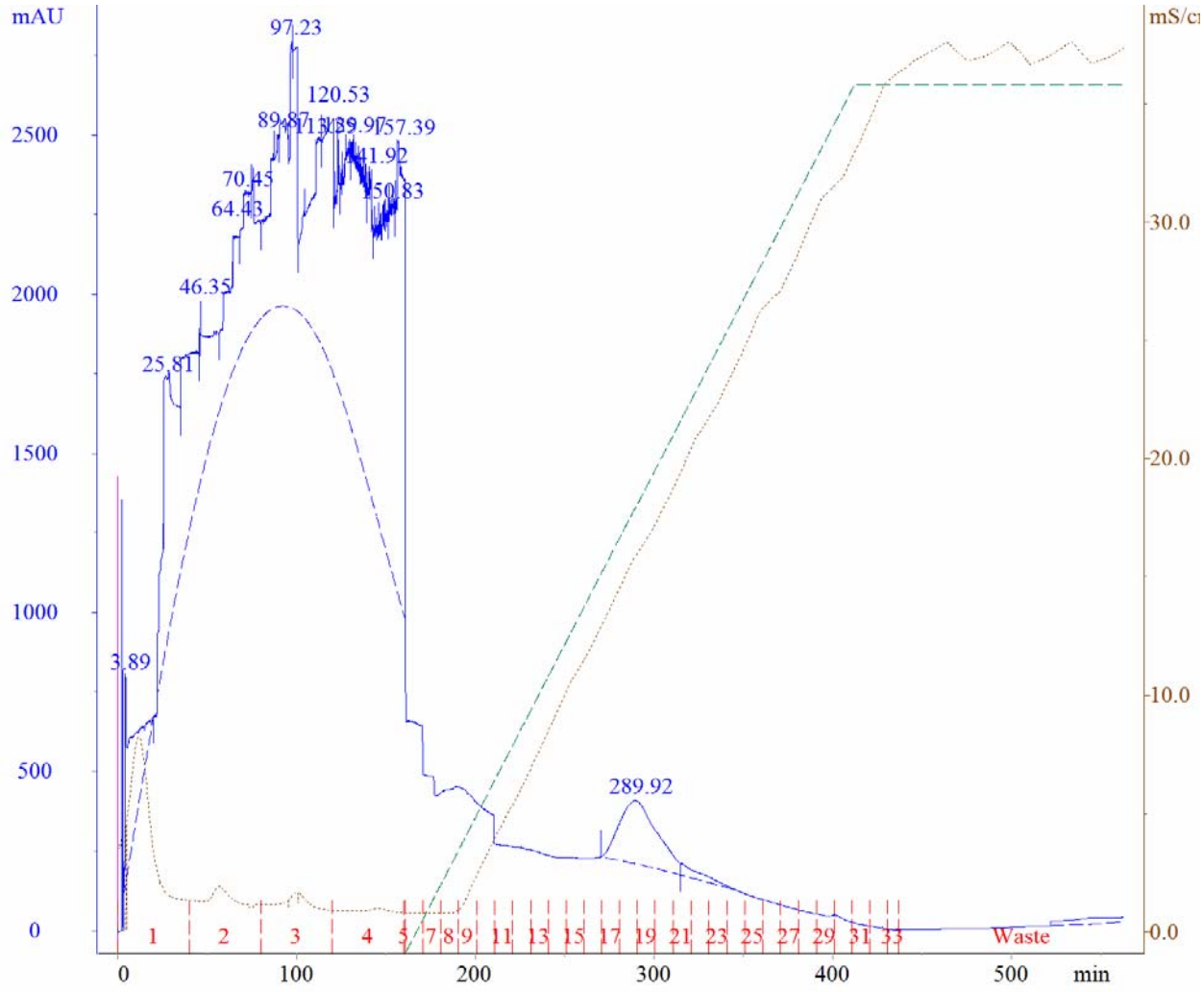

Figure 19: Weak anion exchange of fractions 18 – 20 in Figure 18 using 10 mM Tris-HCl pH 8.0 buffer as starting buffer and 0.4 M NaCl as eluting buffer with a DEAE FF column (5 mL). The gradient of eluting buffer started at fraction 5 and continued until fraction 30. The main peak eluted at approximately 0.18 M NaCl (fractions 15 – 17). Fractions corresponding to the main peak and minor peak (fractions 28 – 30) were collected for SDS-PAGE.

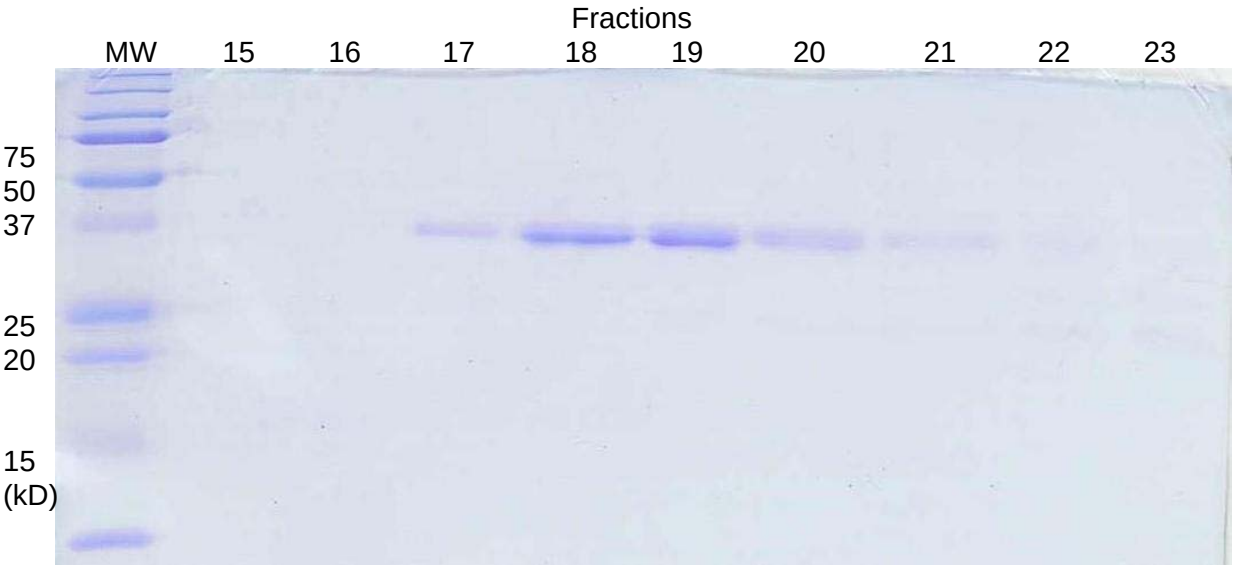

Figure 20: SDS-PAGE of weak anion exchange FPLC fractions (Figure 19). The first lane (MW) contains the Precision Plus Protein® molecular weight marker standard. Fractions 15 – 23 come from Figure 19. SDS-PAGE reveals a protein band around the weight of 35 kD in each of the fractions 17 – 21. Fractions 20 and 21 also contain a faint contaminant band at about 25 kD.

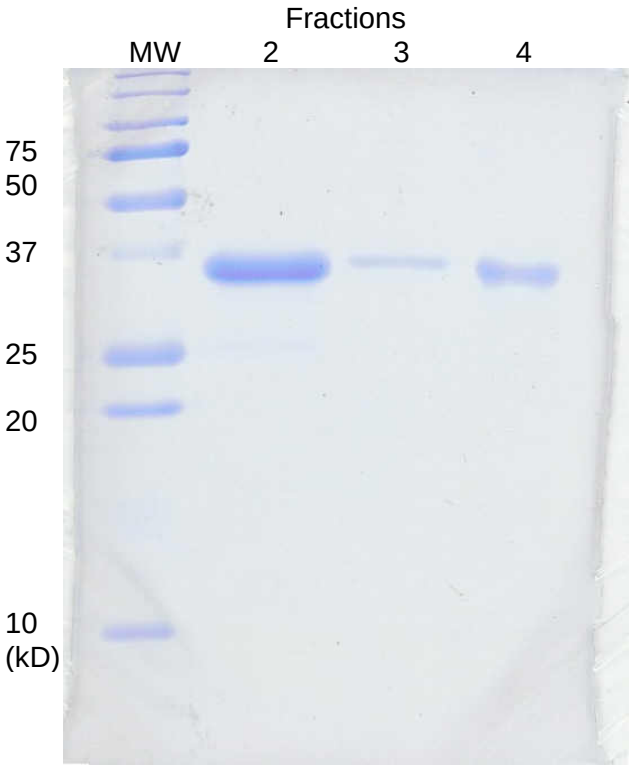

Figure 21: SDS-PAGE portion of a Western Blot of SiaA M79A and SiaA WT solutions. The M79A solution consisted of fractions 17 – 21 from Figure 20. The SiaA WT solutions were provided by Brian Sook. The first lane (MW) contains the Precision Plus Protein® molecular weight marker standard. Lane 2 contained 10 µL of the SiaA M79A solution. Lanes 3 and 4 originally contained 10 µL of a dilute SiaA WT solution and 0.3 µL of a concentrated SiaA WT solution, respectively. SiaA is visible in lane 2 as a dominant band around 35 kD, and a faint contaminant band is observable in lane two at 25 kD.

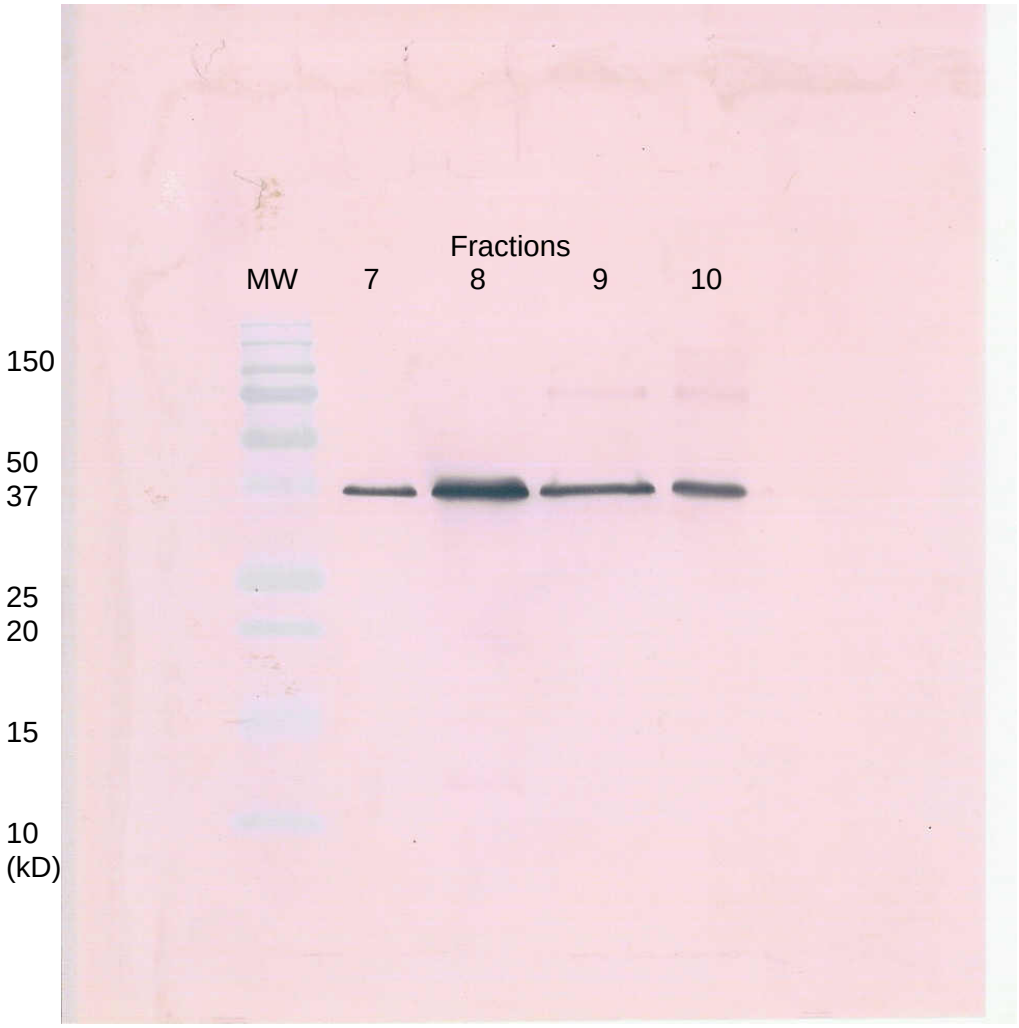

Figure 22: Western blot of SiaA M79A and SiaA WT solutions. The M79A solution consisted of fractions 17 – 21 from Figure 20. The SiaA WT solutions were provided by Brian Sook. Lanes 7 and 8 originally contained 1 µL and 5 µL of the SiaA M79A solution, respectively. Lanes 9 and 10 originally contained 10 µL of a dilute SiaA WT solution and 0.3 µL of a concentrated SiaA WT solution, respectively. No trace of the 25 kD contaminant from lane 2 of Figure 21 is observed in lane 7 or 8. Therefore, the contaminant is not likely a fragment of SiaA.

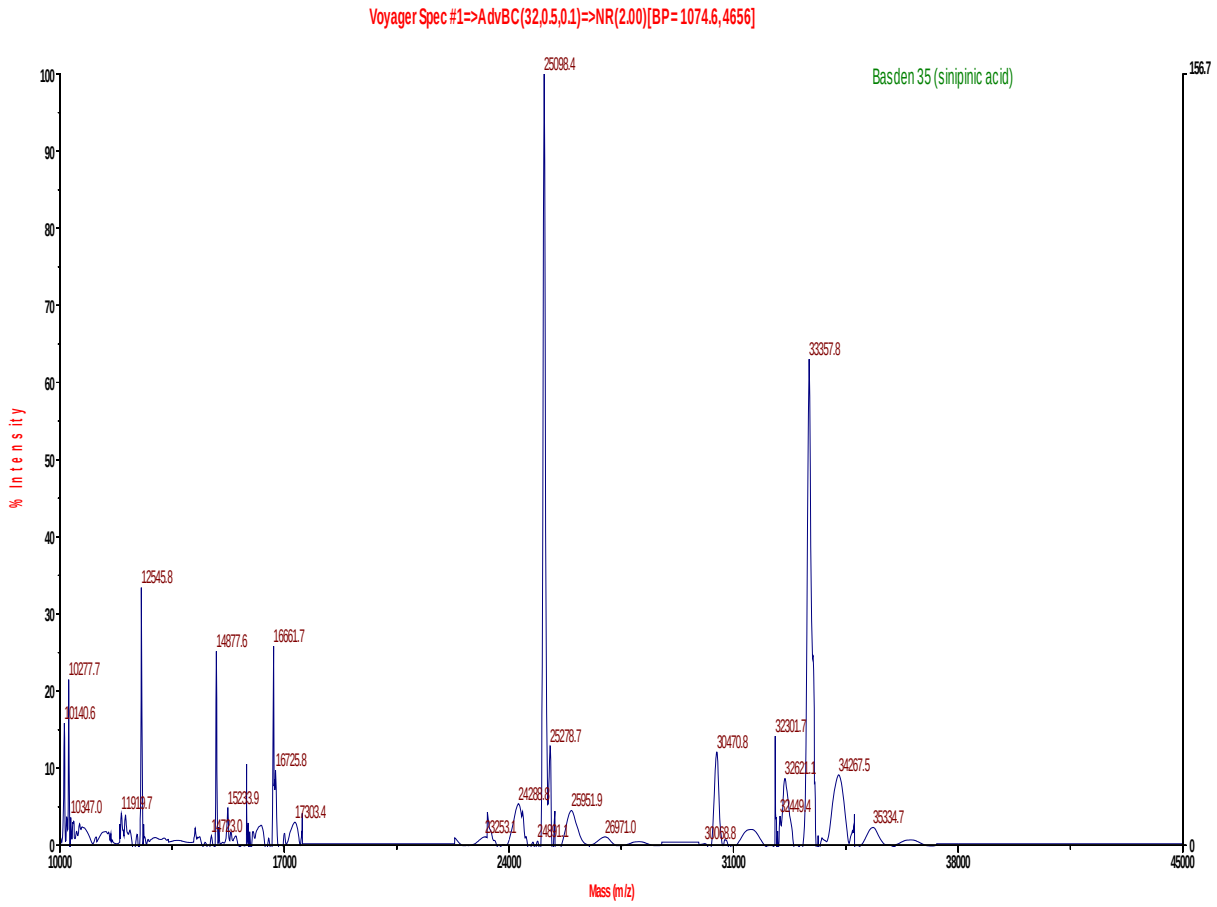

Figure 23: Fractions 17 – 21 were combined from Figure 20. MALDI mass spectrometry (sinipinic acid) reveals a protein solution with two dominant molecular ion peaks: 25098 and 33358 Da, consistent with the SDS-PAGE of Figure 21.

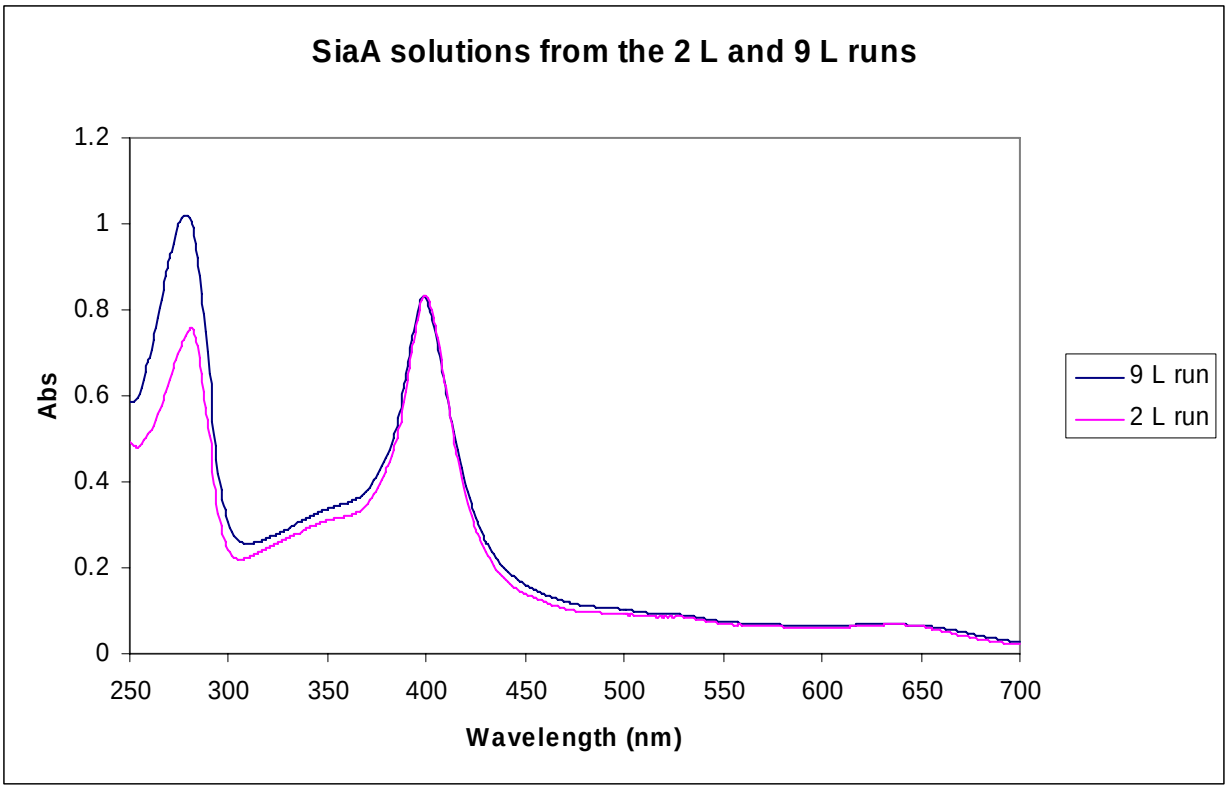

Figure 24: UV-Vis spectral overlay of the SiaA M79A isolated via French press (9 L run) and using rigorous sonication (2 L run). The two spectra are normalized at the Soret. Although the protein from the 9 L run SiaA is larger by about 3000 Da, the spectra appear the same.

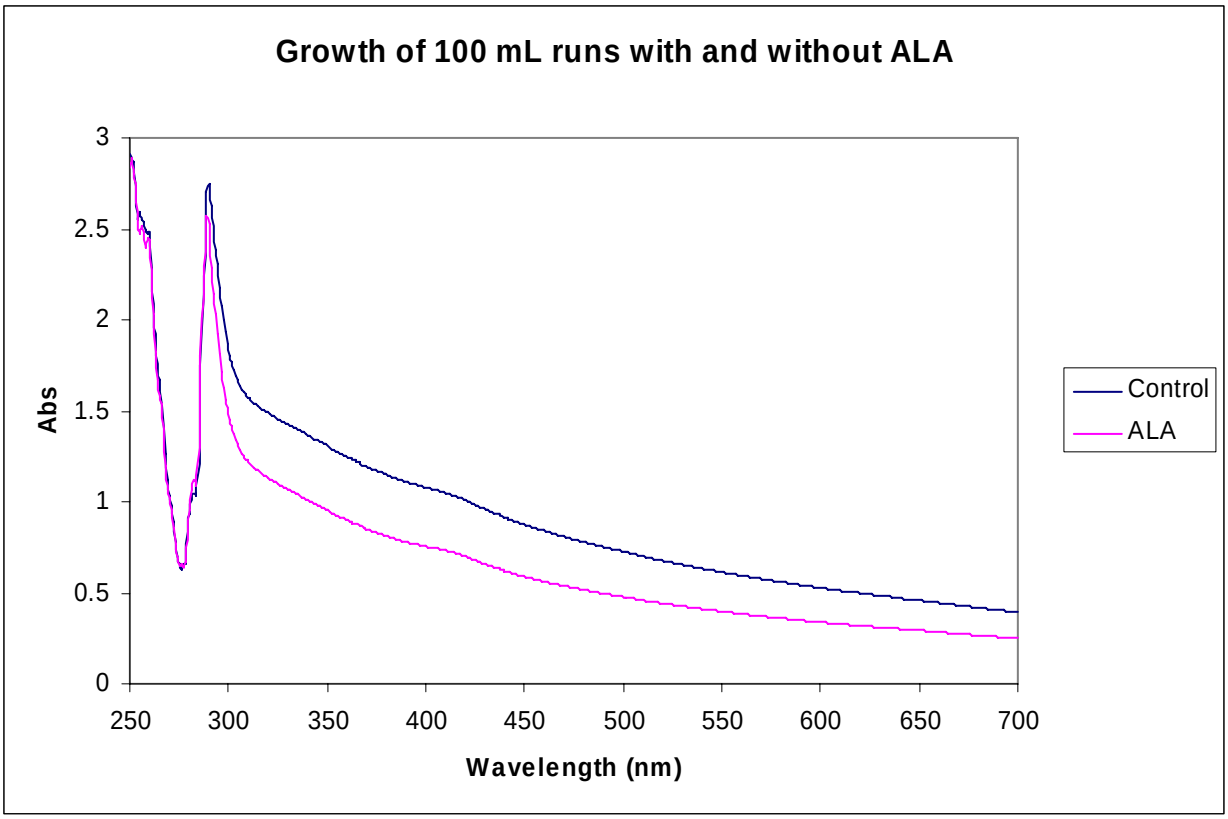

Figure 25: UV spectral overlay of solutions of growth with (pink) and without (purple) ALA. The slight heme peak at approximately 410 nm does not increase when ALA is used in growth conditions.

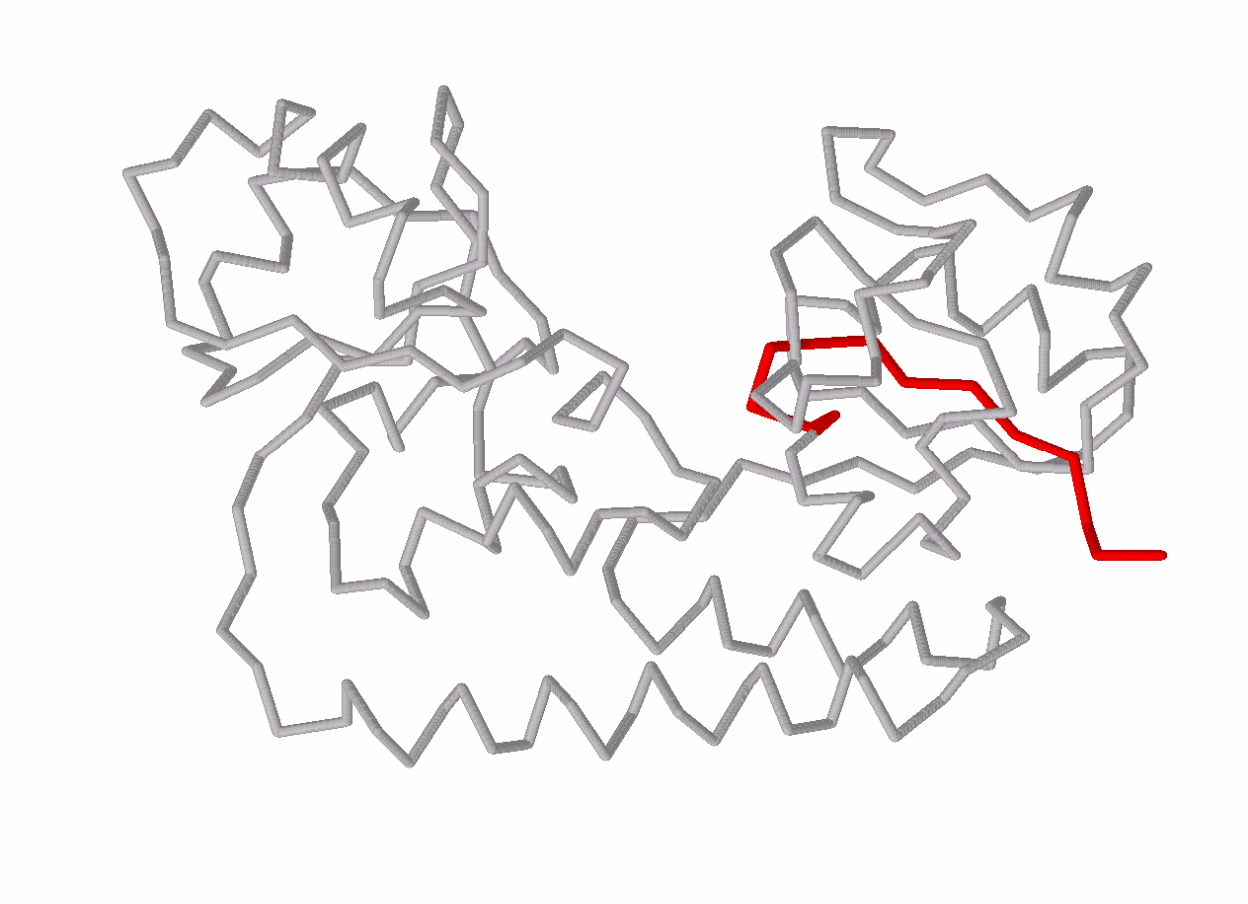

Figure 26: A 3D-JIGSAW homology model of SiaA with BtuF (1N4D), highlighting residues 32 – 44 in red.

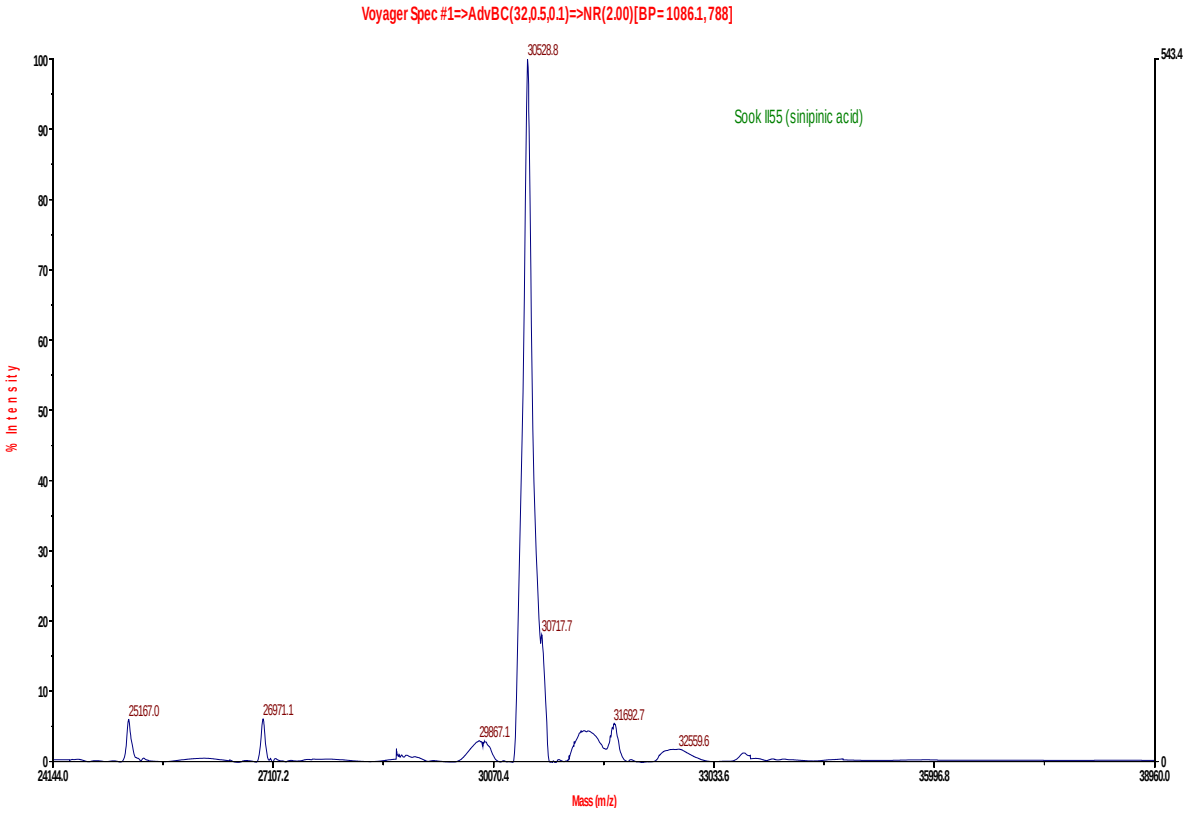

Figure 27: MALDI mass spectrometry (sinipinic acid) of WT SiaA protein solution (Brian Sook) that was sonicated with the rigorous 5 cycles of 2 min each method. The main molecular ion peak is at 30529 Da.

SiaA M79A Cleavage Possibilities

Observed	Predicted	Dev	Theoretical Cleaved Sequence	Sequence #
30464.380	30460.438	-3.942	(H)HGMASMTGGQQMGRDLYDDD DKDRWGSELEVNQHPKTAKE TEQQRIVATSVAVVDICDRL NLDLVGVCDSKLYTLPKRYD AVKRVGLPANPDIELIASLK PTWILSPNSLQEDLEPKYQK LDTEYGFLNLRSVEGMYQSI DDLGNLFQRQQEAKELRQQY QDYYRAFQAKRKGKKKPKVL ILMGLPGSYLVATNQSYVGN LLDLAGGENVYQSDEKEFLS ANPEDMLAKEPDLILRTAHA IPDKVKVMFDKEFAENDIWK HFTAVKEG(K)	10-277
30464.380	30464.545	0.164	(M)TGGQQMGRDLYDDDDKDRWG SELEVNQHPKTAKETEQQRI VATSVAVVDICDRLNLDLVG VCDSKLYTLPKRYDAVKRVG LPANPDIELIASLKPTWILS PNSLQEDLEPKYQKLDTEYG FLNLRSVEGMYQSIDDLGNL FQRQQEAKELRQQYQDYYRA FQAKRKGKKKPKVLILMGLP GSYLVATNQSYVGNLLDLAG GENVYQSDEKEFLSANPEDM LAKEPDLILRTAHAIPDKVK VMFDKEFAENDIWKHFTAVK EGKVYDL(D)	16-282
30464.380	30464.658	0.278	(D)DKDRWGSELEVNQHPKTAKE TEQQRIVATSVAVVDICDRL NLDLVGVCDSKLYTLPKRYD AVKRVGLPANPDIELIASLK PTWILSPNSLQEDLEPKYQK LDTEYGFLNLRSVEGMYQSI DDLGNLFQRQQEAKELRQQY QDYYRAFQAKRKGKKKPKVL ILMGLPGSYLVATNQSYVGN LLDLAGGENVYQSDEKEFLS ANPEDMLAKEPDLILRTAHA IPDKVKVMFDKEFAENDIWK HFTAVKEGKVYDLDNTLFGM SAKLNYP(E)	30-296
30464.380	30465.667	1.287	(P)KTAKETEQQRIVATSVAVVD ICDRLNLDLVGVCDSKLYTL PKRYDAVKRVGLPANPDIEL IASLKPTWILSPNSLQEDLE PKYQKLDTEYGFLNLRSVEG MYQSIDDLGNLFQRQQEAKE LRQQYQDYYRAFQAKRKGKK KPKVLILMGLPGSYLVATNQ SYVGNLLDLAGGENVYQSDE KEFLSANPEDMLAKEPDLIL RTAHAIPDKVKVMFDKEFAE NDIWKHFTAVKEGKVYDLDN TLFGMSAKLNYPEALDTLTQ LFDHVGDH(P)	45-312

Figure 28: Monoisotopic cleavage possibilities of SiaA M79A with His tag construct (2 L run) for a total of 313 residues. Possibilities generated from the ExPASy proteomic tool FindPept using observed weight from Figure 16.

SiaA WT Cleavage Possibilities

Observed	Predicted	Dev	Theoretical Cleaved Sequence	Sequence #
30528.800	30524.548	-4.251	(M)TGGQQMGRDLYDDDDKDRWG SELEVNQHPKTAKETEQQRI VATSVAVVDICDRLNLDLVG VCDSKLYTLPKRYDAVKRVG LPMNPDIELIASLKPTWILS PNSLQEDLEPKYQKLDTEYG FLNLRSVEGMYQSIDDLGNL FQRQQEAKELRQQYQDYYRA FQAKRKGKKKPKVLILMGLP GSYLVATNQSYVGNLLDLAG GENVYQSDEKEFLSANPEDM LAKEPDLILRTAHAIPDKVK VMFDKEFAENDIWKHFTAVK EGKVYDL(D)	16-282
30528.800	30524.662	-4.138	(D)DKDRWGSELEVNQHPKTAKE TEQQRIVATSVAVVDICDRL NLDLVGVCDSKLYTLPKRYD AVKRVGLPMNPDIELIASLK PTWILSPNSLQEDLEPKYQK LDTEYGFLNLRSVEGMYQSI DDLGNLFQRQQEAKELRQQY QDYYRAFQAKRKGKKKPKVL ILMGLPGSYLVATNQSYVGN LLDLAGGENVYQSDEKEFLS ANPEDMLAKEPDLILRTAHA IPDKVKVMFDKEFAENDIWK HFTAVKEGKVYDLDNTLFGM SAKLNYP(E)	30-296
30528.800	30525.671	-3.129	(P)KTAKETEQQRIVATSVAVVD ICDRLNLDLVGVCDSKLYTL PKRYDAVKRVGLPMNPDIEL IASLKPTWILSPNSLQEDLE PKYQKLDTEYGFLNLRSVEG MYQSIDDLGNLFQRQQEAKE LRQQYQDYYRAFQAKRKGKK KPKVLILMGLPGSYLVATNQ SYVGNLLDLAGGENVYQSDE KEFLSANPEDMLAKEPDLIL RTAHAIPDKVKVMFDKEFAE NDIWKHFTAVKEGKVYDLDN TLFGMSAKLNYPEALDTLTQ LFDHVGDH(P)	45-312
30528.800	30530.552	1.751	(D)LYDDDDKDRWGSELEVNQHP KTAKETEQQRIVATSVAVVD ICDRLNLDLVGVCDSKLYTL PKRYDAVKRVGLPMNPDIEL IASLKPTWILSPNSLQEDLE PKYQKLDTEYGFLNLRSVEG MYQSIDDLGNLFQRQQEAKE LRQQYQDYYRAFQAKRKGKK KPKVLILMGLPGSYLVATNQ SYVGNLLDLAGGENVYQSDE KEFLSANPEDMLAKEPDLIL RTAHAIPDKVKVMFDKEFAE NDIWKHFTAVKEGKVYDLDN TLFGMSA(K)	25-291

Figure 29: Monoisotopic cleavage possibilities of SiaA WT with His tag construct for a total of 313 residues. Possibilities generated from the ExPASy proteomic tool FindPept using observed weight from Figure 27.