

Background

Epidemiology

Astroviruses are a major cause of diarrhea in children and immunocompromised populations. Estimates range from 7 - 20% of diarrhea in these groups.¹ Infections are seasonal in nature, following the rainy season in tropical areas, and the winter season in temperate areas.² The different serotypes have varying clinical presentations, with diarrhea, fever, and vomiting being common symptoms.³

Astroviruses are transmitted by the oral-fecal route, but basic hygiene can be preventative. People in institutionalized care settings (such as nursing homes and day care centers) are at greater risk for infection.⁴

Basic Biology

Astroviruses are non-enveloped, single-stranded, positive sense RNA viruses, with a genome of about 7kb. The genome has 3 open reading frames, ORF 1a, 1b, and 2. ORF 2 encodes the structural capsid protein.⁵ Within the *Astroviridae* family, there are two genera: mammalian (*mamastroviruses*) and avian (*avastroviruses*). There are 8 major serotypes of human astroviruses (HAstV), including HAstV-1, the most common serotype, and the one considered in this study.⁶

The HAstV capsid protein contains 5 distinct structural domains (Fig. 1). The N-terminus contains a positively charged domain that binds the RNA genome, followed by the inner and outer core domains, which form the main capsid shell encapsidating the RNA genome. The spike and acidic domains project outside of the main capsid shell, and their sequences are variable between strains. The spike domain is thought to be the receptor-binding domain because antibodies that bind this domain block virus

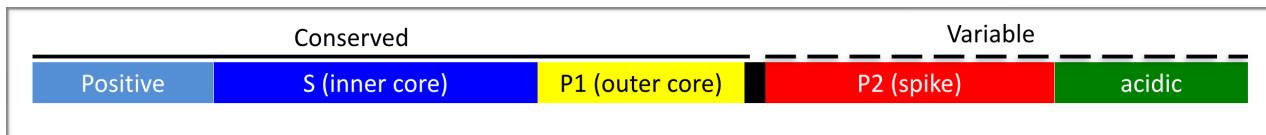


Figure 1: Capsid protein domains

attachment and are neutralizing⁷ however the host cell receptor is unknown. The crystal structures of the spike domains from HAstV-1 and turkey astrovirus -2 (TAstV-2) have been solved.^{8 9}

Little is known about the structure or function of the astrovirus acidic domain, but it is thought to be involved in the maturation process, and may be cleaved by caspases.¹⁰ ORF 2 produces a 90 kDa protein that appears to go through at least two processing and maturation steps. First, the 90kD protein becomes cleaved to form a 70 kDa protein.¹¹ The 70 kDa assembles with viral RNA genome into immature virus particles, which are secreted extracellularly and undergo a second proteolysis maturation step¹⁰. *In vitro*, this second proteolysis step is often performed using trypsin and is required to obtain infectious virions.

The difference between the 90 kDa and 70 kDa capsid proteins is the C-terminal acidic domain, which may be cleaved by a member of the caspase family, a group of proteins involved in apoptosis. Evidence exists that Casp-3 and/or Casp-9 play a role in cleavage intracellularly, at least for the HAstV-8 serotype¹². Putative caspase cleavage sites have been described in the HAstV-1 serotype as well (DEDDEADRFIDIIDTSDEED and IETD)¹⁰. The cleavage of the 90 kDa to 70 kDa form is important for virus particles to be secreted from the cell.

The 90 kDa and 70 kDa capsid proteins also have different cellular localization patterns. The 90 kDa version is found co-localized with viral RNA, and associated with cell membranes. This indicates that this version is associated with the possibly membrane-bound immature virus. The 70 kDa version (lacking the acidic domain) is found extracellularly, associated with the production of highly infectious viral particles following trypsin digestion.¹³

This evidence points toward a potential role for the acidic domain in localizing the immature viral particles at the cell membrane, and facilitating particle assembly and release from the cell. Since cleavage of the acidic domain is thought to be mediated by caspases, control of the apoptotic pathway may be critical for maturation and release of the virus from host cells¹⁰.

Goals

The goal of this project is to understand the role of the astrovirus capsid acidic domain in the virus life cycle. Structural and biochemical experiments will be used to elucidate its function. Specifically, bioinformatics methods will be used to predict the sequence

boundaries and predicted structural features of the astrovirus capsid acidic domain. Molecular cloning techniques will be used to introduce the acidic domain DNA sequence into an expression plasmid, and the recombinant protein will be expressed in *Escherichia coli*. Affinity, ion exchange, and size-exclusion chromatography techniques will be used to isolate highly-pure astrovirus acidic domain protein. Purified acidic domain will be subjected to structural studies using circular dichroism, nuclear magnetic resonance (NMR) spectroscopy, and X-ray crystallography. In addition to solving the structure, various biochemical and functional assays will be developed to help understand the role of the acidic domain in the viral life cycle. These might include fluorescence imaging to localize the protein within cells, as well as tests to see if the acidic domain alone can affect the opening of tight junctions, as has been observed for full-length capsid protein¹⁴.

Methods

The HHpred server was used to predict the secondary structure of the HAstV-1 capsid from its primary amino acid sequence and find possible structural homologs in the Protein Data Bank (PDB).¹⁵ No close matches were found for the acidic domain,

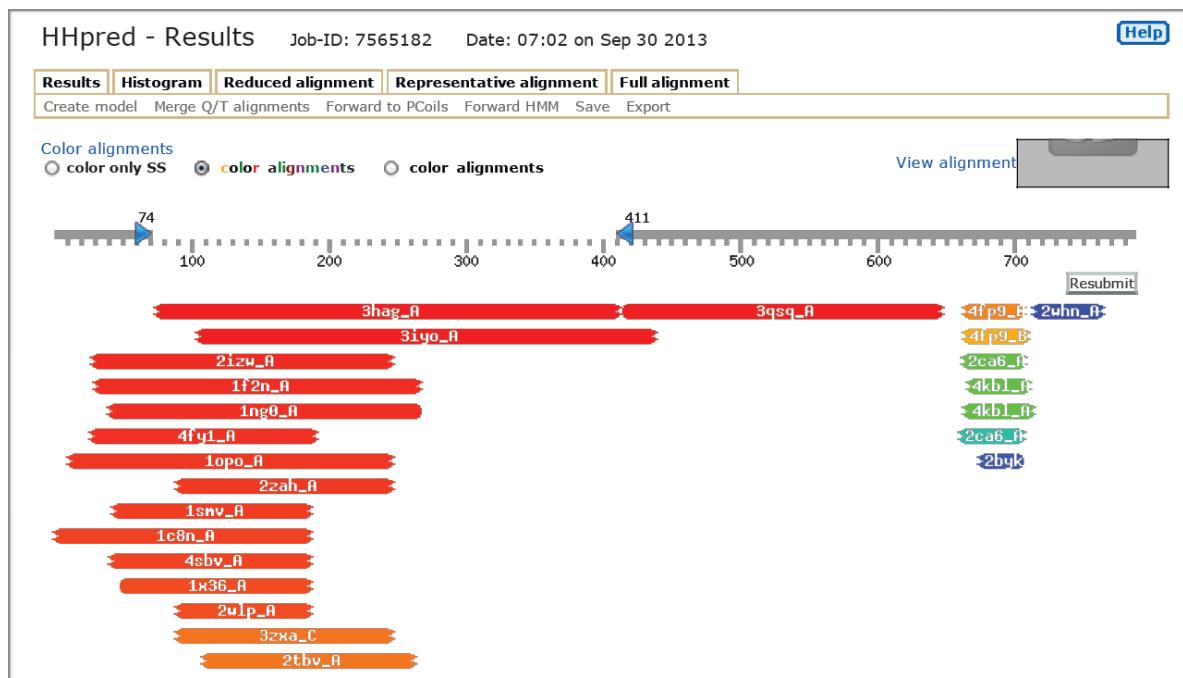


Figure 2: HHpred results, showing no close homologs for the acidic domain (from ~630 to the end)

however it was predicted to contain helices. Structural homologs were predicted for the core and spike domains (shown in red in Fig. 2).

Using this information, as well as the amino acid composition, the final 141 amino acids were used as the target sequence for the recombinant capsid protein acidic domain.

This led to an overall DNA length of 423 nucleotides. Primers were designed and ordered with Nco1 and Not1 restriction endonuclease recognition sites, for a PCR reaction using the cDNA sequence of the HAstV-1 capsid (derived from isolated virus, GenBank: GQ925363.1) as the template (with the intention of ligating the insert into the pET52b plasmid cloning site). No stop codon was included in the primer, leaving the final construct with a C-terminal His-tag.

```
srasgygyesdnteyleadapsadqfnediettdiesteddeddeadrfdiidtsdeede setdrvtilstlvnqgmtitratkiarrafptlsdrirkrgvymdlasgaspgnawshaceearkaageinpctsgsrghae  
ctcggcgccgcggctccagatgtcagggattggttcccccactgcttgcgtgcctctcacacgcatacgtgaccatg  
catggcctggctaccccgagacaaggcaggtccatgtatactccacgcgtgatcctatcgaaagcgtgggaatgcgc  
gccgtgcataccgtgtggcacgcgtattgtcattccgtattacgagagttacacgggttacacggctgtctcatttcca  
tcttcattcatcagaagtgtctatgatatcaaacctgtccgctcgtcgtcctctgtactctcaatgtctgtctgtctatgt  
cttcattaaactggtcagcagactggagcatccaagtattcgatgttgactctcataaccatgaccagatgccctgga
```

Figure 3: Amino acid and nucleotide sequences, with acidic residues in pink

In addition to HHpred, the Robetta structure prediction server was also used.¹⁶ This tool combines the Rosetta@home distributed computing cluster with secondary structure prediction methods. (Rosetta is a *de novo* structure prediction tool based on molecular dynamics calculations combined with Monte Carlo fold optimization.)

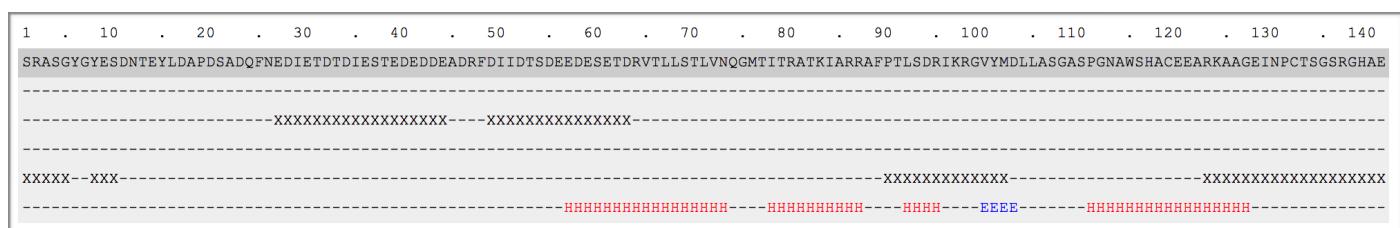


Figure 4: Secondary structure prediction, with red H's indicating helical regions.

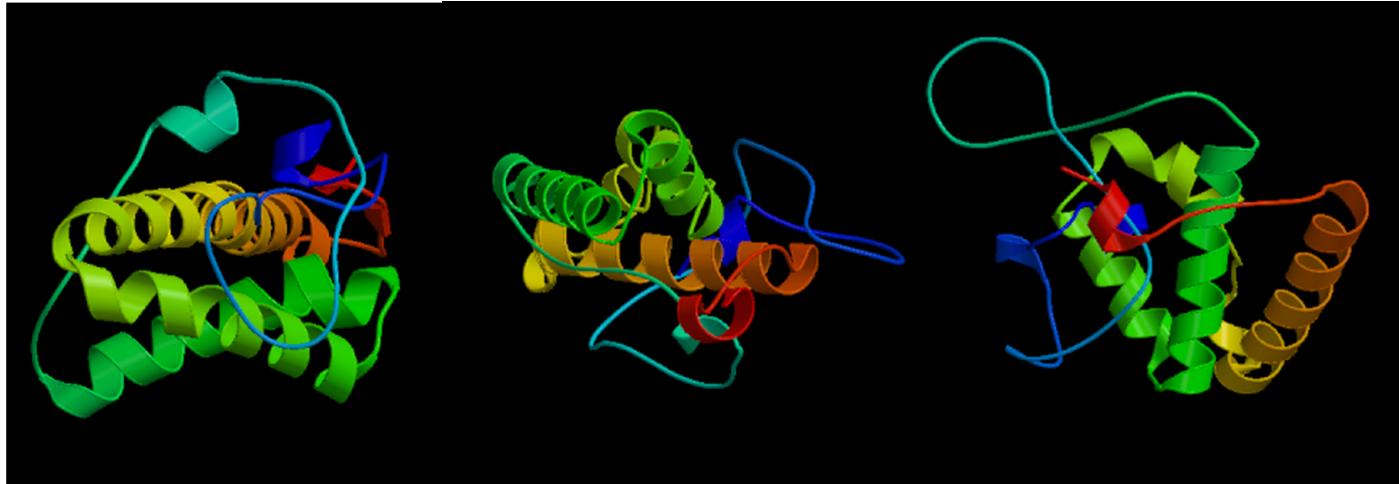


Figure 5: Tertiary predictions 1-3 (of 5) presented by the Robetta structure prediction server

The Robetta results indicated that the secondary structure of the acidic domain would be approximately 40% helical (concentrated toward the C-terminus) with the rest being disordered.

Tertiary structure predictions were also produced, showing bundles of 4 helices, combined with a long unstructured region. The predictions differ in where the helices break, as well as in their spatial orientation. They all agree that a helical bundle is present (as opposed to a string of “floppy” unbundled helices).

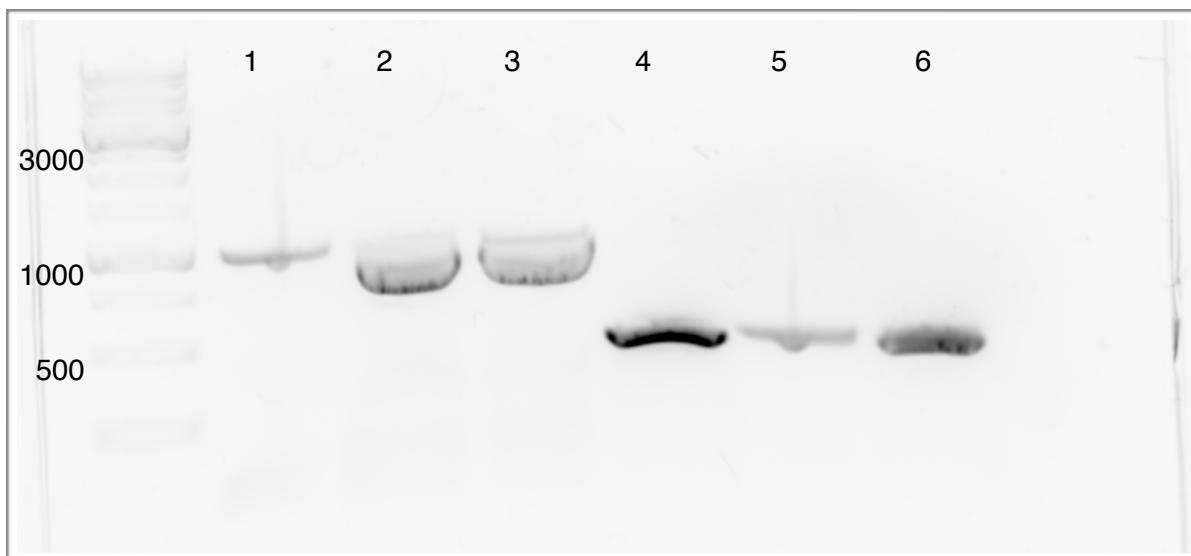


Figure 6: PCR Results (wells 4-6), showing target at ~500bp

PCR was performed with the primers ordered, Phusion polymerase and buffer, and the cDNA template described above (30 cycles of denaturation: 95°C x 10 sec., annealing: 58°C x 30 sec., extension: 72°C x 30). Gel electrophoresis was performed in a 1% agarose gel in TBE buffer with Sybr Safe stain (Fig. 6), and the desired fragment was extracted using the Promega gel clean-up kit protocol.

A restriction digest was performed on both the purified insert and the stock pET52b plasmid (with NOT1, NCO1, and CutSmart buffer) overnight, and then purified by gel electrophoresis and gel extraction. Unfortunately, the gel clean-up kits were only giving a yield of about 10%. Repeated rounds of clean-up led to concentrations so low that the spectrophotometer was unable to confirm the presence of DNA.

Despite the low yields, I attempted to concentrate the solutions of insert and plasmid using a Speed-Vac. These samples were then ligated (50 ng plasmid with 15 ng insert, using the Quick Ligation protocol) and transformed into DH5-a chemically competent *E. coli* cells. These cells were then heat-shocked, plated, and incubated at 37 °C overnight. No colonies were present the following day.

Repeated attempts at concentration, ligation, and transformation into DH5-a cells were unsuccessful, so a sub-cloning step into a blunt-ended TOPO plasmid was attempted. This construct was then transformed into TOP10 chemically competent *E. coli* cells. This resulted in a number of colonies, which were then cultured overnight at 37 °C in 5 ml of LB/Kanamycin. The plasmids were extracted from the 5 ml cultures using the Qiagen Mini-Prep kit, and DNA sequencing of plasmids with M13 primers was used to ensure they contained the correct insert (Fig. 7).

While this step confirmed that the primers were correct and that the insert was present with no mutations, it was unfortunately not possible to recover the insert through restriction digest due to the complication that the Ncol- and Notl-digested TOPO plasmid generated another fragment of approximately the same size as the insert.

To avoid the earlier restriction digest and ligation problems, a new forward primer was designed with 3 additional 5' nucleotides, which should improve the efficiency of the Nco1 enzyme. Ligations and restriction digests were also performed overnight, rather than for the times listed on the manufacturer's protocols. To avoid the ongoing issues of low yield after gel extraction, PCR clean-up was used whenever possible, and the largest practical amount of insert DNA was used for ligation. Transformed cells (DH5-a) were spun down and resuspended in a smaller volume (100 µL) prior to plating to increase concentration.

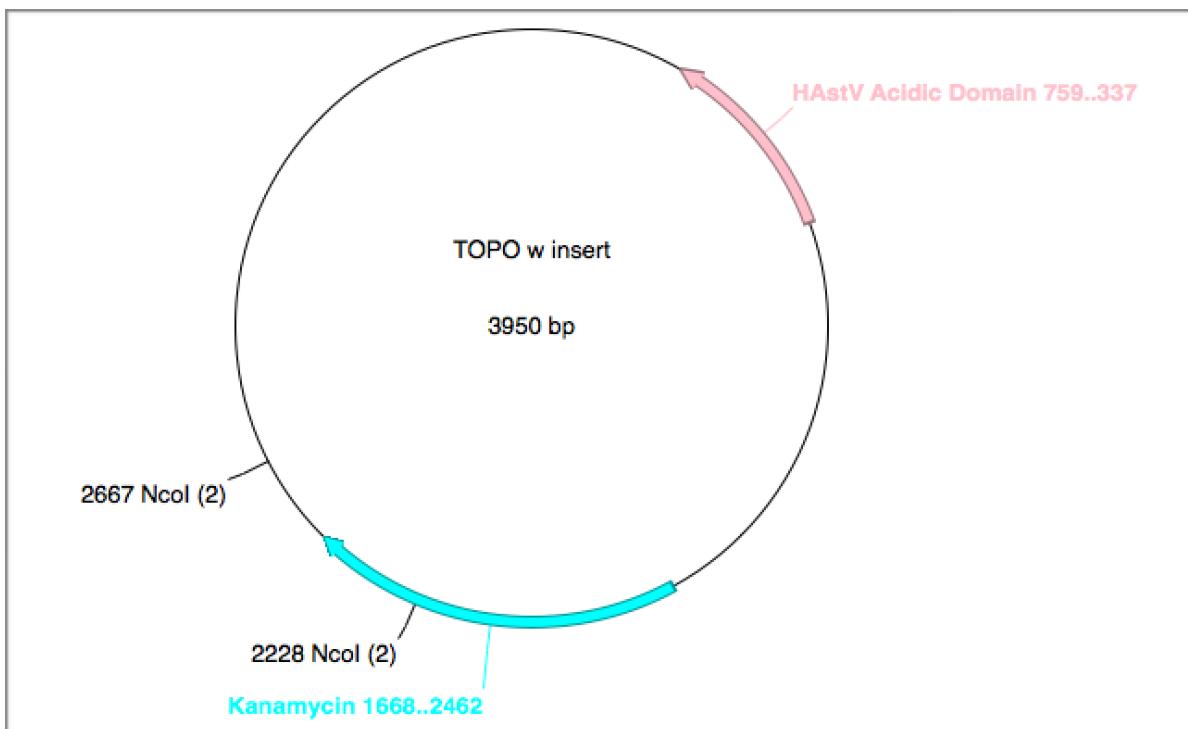


Figure 7: TOPO plasmid diagram showing additional fragment created during digest.

This process resulted in liberal colonies after overnight incubation, and colony PCR and DNA sequencing were performed to confirm the presence of the desired insert (Fig. 8). A 25 mL culture was prepared from a positive colony, incubated overnight, and purified using a Qiagen Midi-Prep kit.

These plasmids were then transformed into 4 cell lines for expression tests (BL21, BL21 pLysS, Rosetta 2 pLysS, and C43 pLysS), plated, and incubated overnight. The BL21, BL21 pLysS, and C43 pLysS lines produced colonies, and one of each was chosen for 5 mL overnight cultures. These 5mL cultures were then used to inoculate 50 mL cultures. Each 50 mL culture was incubated at 37 °C until the 600nm optical density (O.D. 600) reached 0.8, and then induced with 1mM IPTG. After induction, each culture was allowed to express overnight at 18 °C.

Two protein gels (sodium dodecyl sulfate polyacrylamide gel electrophoresis - SDS PAGE) were prepared with pre- and post-IPTG samples of each culture to test for expression. One was dyed with Coomassie stain, and the other was used to perform a western blot (using a horse-radish peroxidase-conjugated anti-His-tag antibody, a semi-dry transfer cell, and the Snap ID 2.0 washing and blocking protocol). The results of the

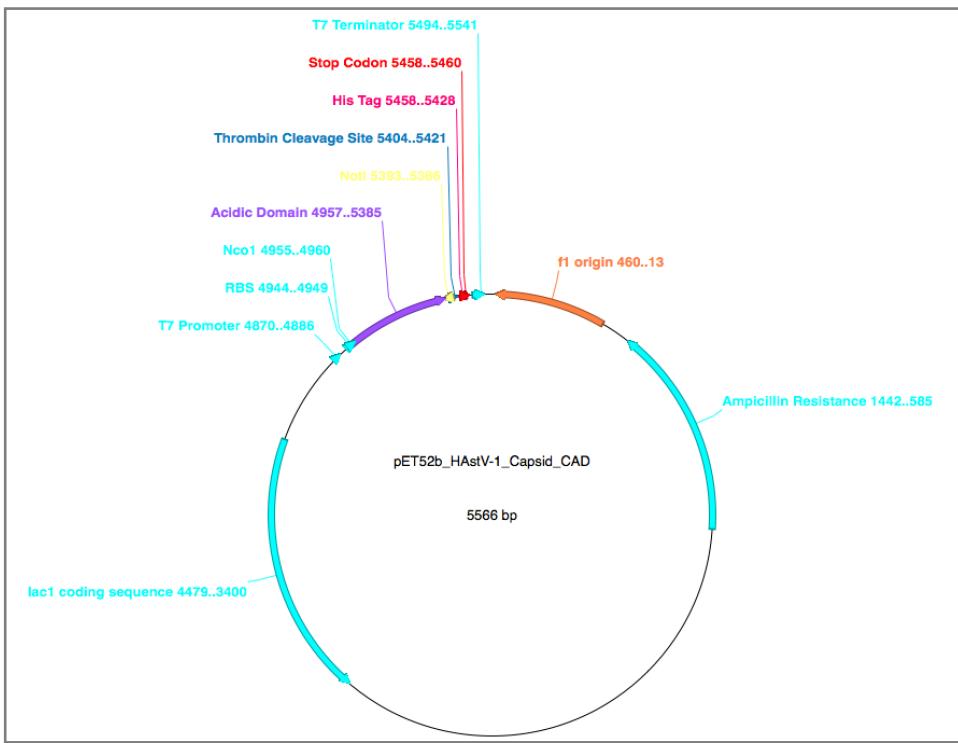


Figure 8: pET52b plasmid diagram showing insertion of the capsid acidic domain.

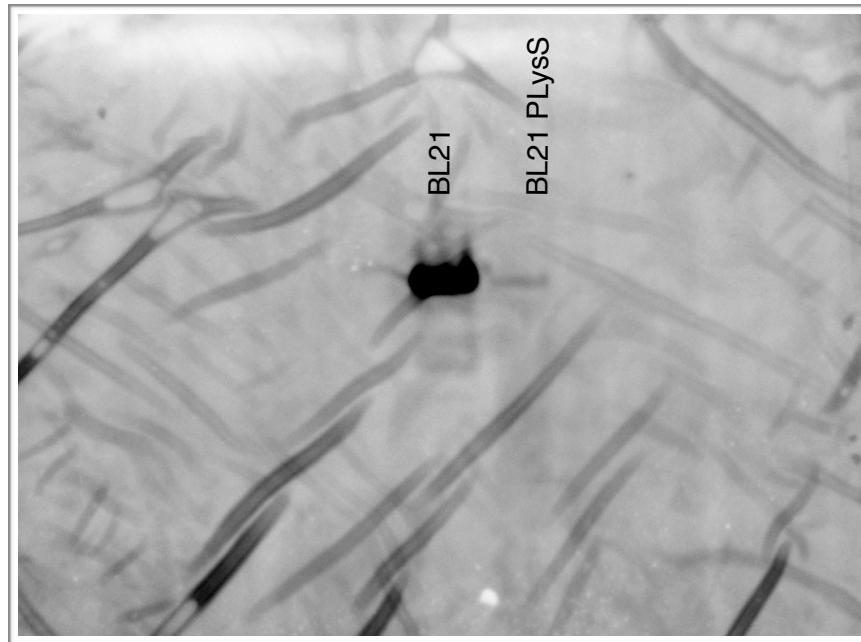


Figure 9: Western blot showing protein expression in the BL21 (dark, left side) and BL21 PLysS (light, right side) cell lines.

gel were somewhat ambiguous (due to the large number of *E. coli* proteins present) (data not shown), but the western blot confirmed that the BL21 strain was expressing large amounts of the target protein (the BL21 pLysS strain was also showing a lesser degree of expression) (Fig. 9).

Moving forward with the sample from the BL21 strain, a small scale purification test was performed to test the solubility of the protein and it's ability to bind to cobalt beads (which bind the His-tag, as a test prior to using a His-trap purification column). The acidic domain was present in the soluble fraction, as well as in the eluent fraction (data not shown).

Given these results, a 7.5 L large-scale culture was grown to obtain large quantities of the recombinant acidic domain protein (same growth and expression procedure as outlined above). When overnight expression was completed, the cells were pelleted, resuspended (in a buffer containing 10 mM Tris @ pH 8.0, 500 mM NaCl, and 30 mM Imidazole) and lysed by sonication in the presence of Benzonase (to remove DNA and RNA) and MgCl₂. The lysed cells were then subjected to centrifugation for 30 minutes (at 3200 x g), and the supernatant was filtered through a 0.22 micron filter in preparation for FPLC purification.

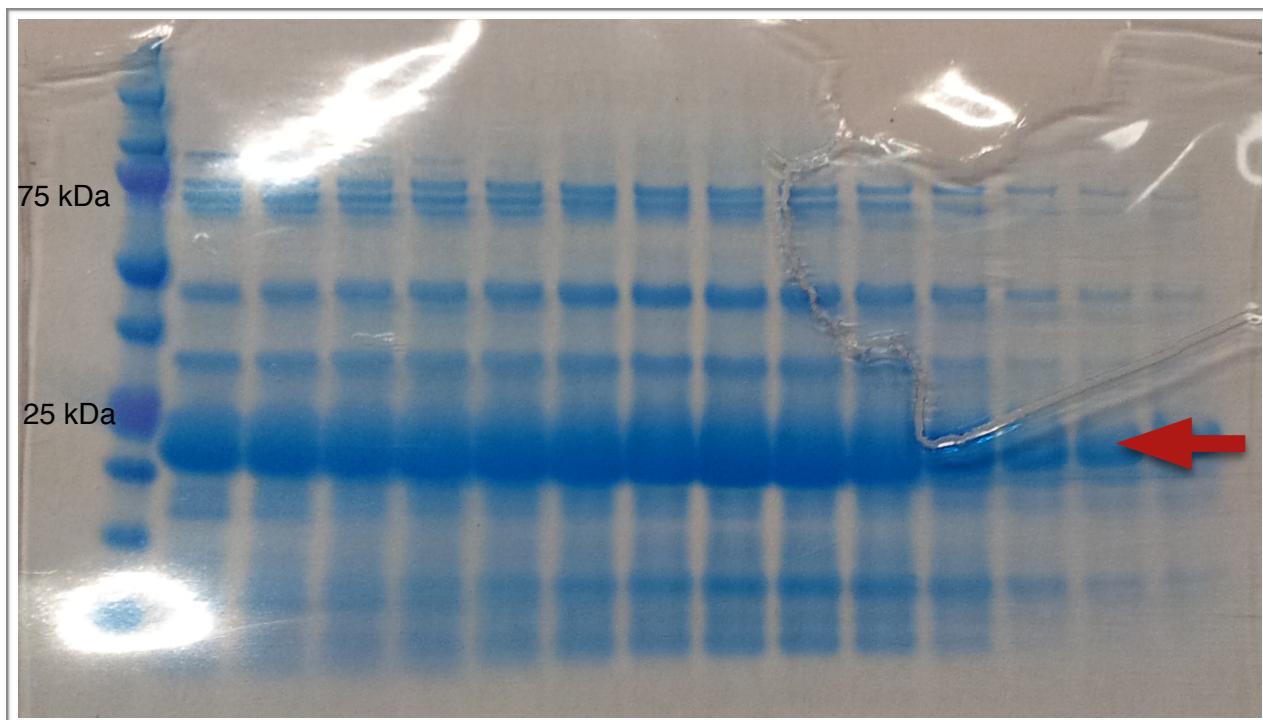


Figure 10: Post-His-trap SDS PAGE showing significant *E. coli* protein contamination (red arrow is HAstV 1 acidic domain)

The acidic domain was purified on a His-Trap affinity column with an imidazole gradient elution, and the results were analyzed by SDS PAGE (Fig. 10). Interestingly, the protein appeared to run approximately 7 kDa larger than its predicted size (16.2 kDa).

In an attempt to refine the purification process, a sample was dialyzed into a low salt buffer (30 mM NaCl, 10 mM Tris @ pH 8.0), and purified by anion exchange chromatography. Since the pI of the acidic domain is approximately 4.2, it was highly negatively charged under these conditions. Elution was performed with a gradient into 10mM Tris pH8 buffer containing 1M NaCl, and fractions were analyzed by SDS PAGE (Fig. 11). The gel showed that little purification was achieved using this anion exchange protocol.

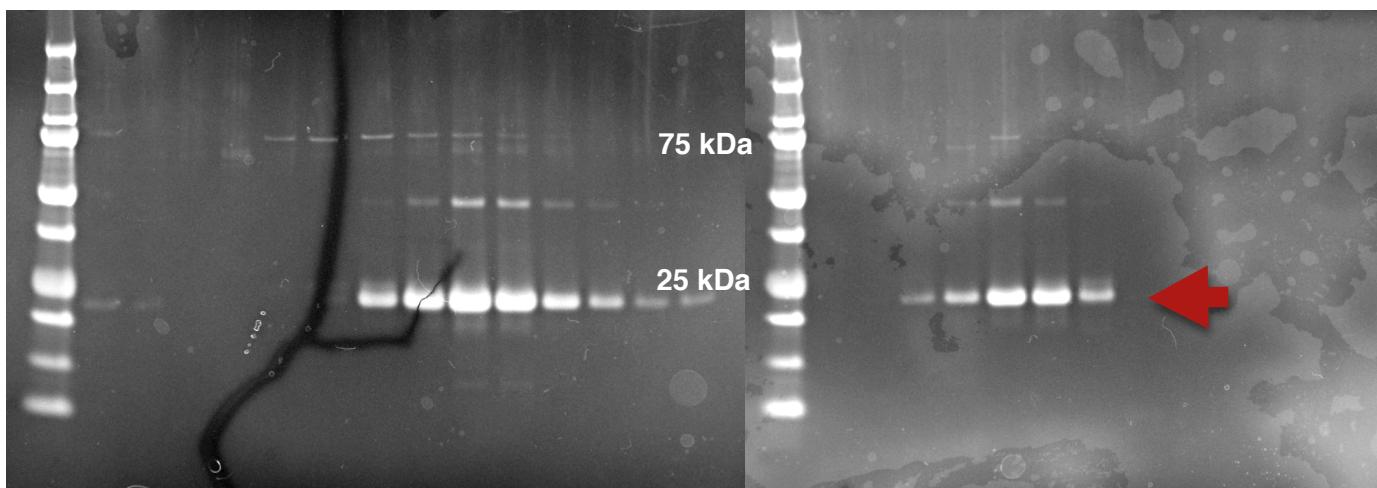


Figure 11: Anion exchange (left) and size exclusion (right) chromatography results (SDS PAGE of collected fractions). Red arrow indicates acidic domain band.

Post-His-Trap samples were also dialyzed into a buffer with 150 mM NaCl and 10 mM Tris, and purified on a Superdex 200 10/300 size exclusion column. Fractions were analyzed by SDS PAGE and revealed some improvement in sample purity (Fig. 11).

Size exclusion chromatography was also repeated with 1 mM DTT (to reduce any disulfide bonds). Both with and without the reducing agent, the acidic domain ran on the sizing column as if it were a much larger protein (in the range of 150-200 kDa). (Fig. 12)

This behavior suggests that the protein might be largely unstructured, since size exclusion relies on the ability of the protein to become temporarily lodged in the pores of

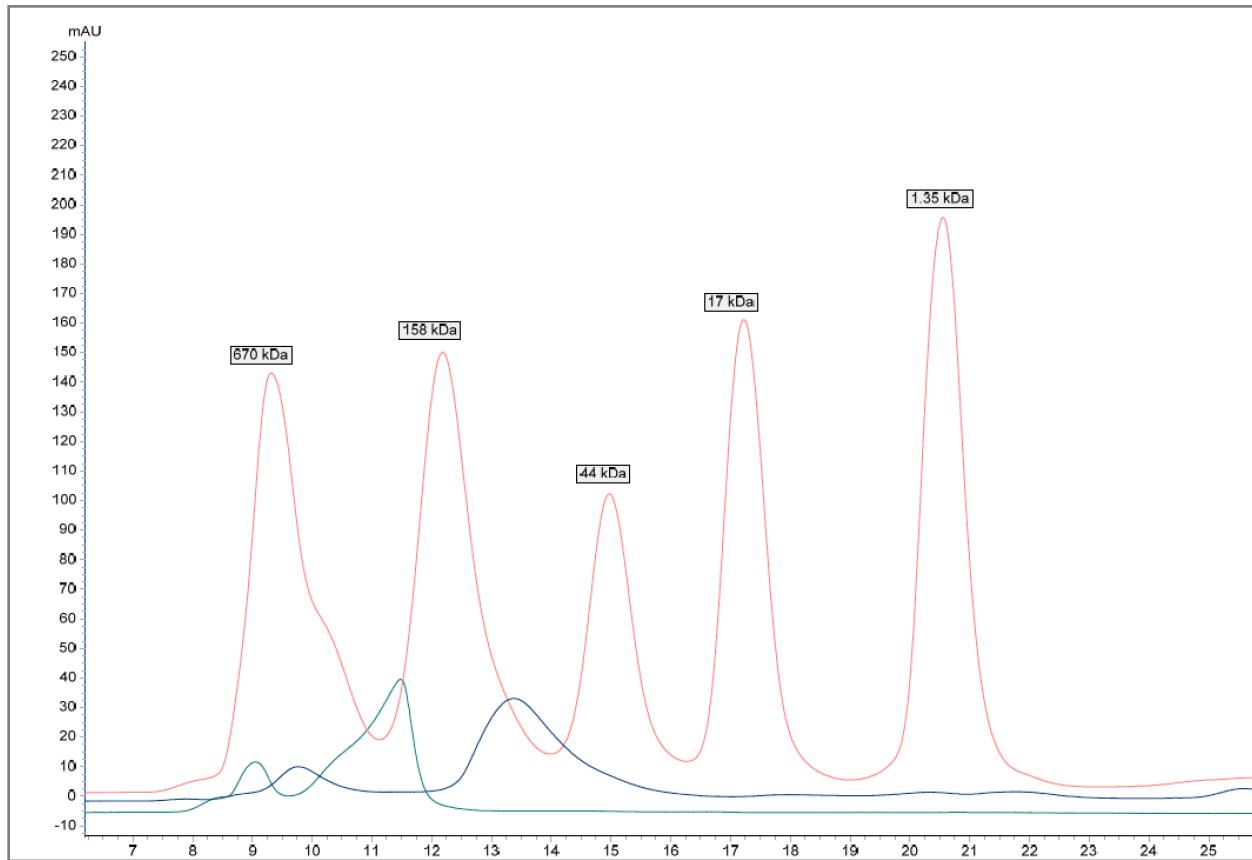


Figure 12: Size exclusion chromatography results with (left) and without (right) reducing agent (DTT).

the media, delaying elution. If it is “floppy” and disordered, it would be unable to fit in these spaces, and would therefore flow quickly through the column like a much larger molecule.

A sample was also taken from the first His-Trap run to use in a trial thrombin protease digest to test the feasibility of removing the His-tag prior to crystallization trials. Unfortunately the thrombin also appeared to cut at an additional site, yielding two additional fragments (Fig. 13). The size of the two resulting fragments suggested that cleavage site may be after the arginine which lies between two predicted helices approximately 40 amino acids from the C-terminus. Thus, these data reveal that cleavage of the His-tag using thrombin protease is impractical.

The His-tag was retained, and all remaining protein was combined into a single container. This sample was concentrated to approximately 5 mL using a 10 kDa MWCO spin concentrator, and the entire volume was purified on the Superdex 200 10/300 size

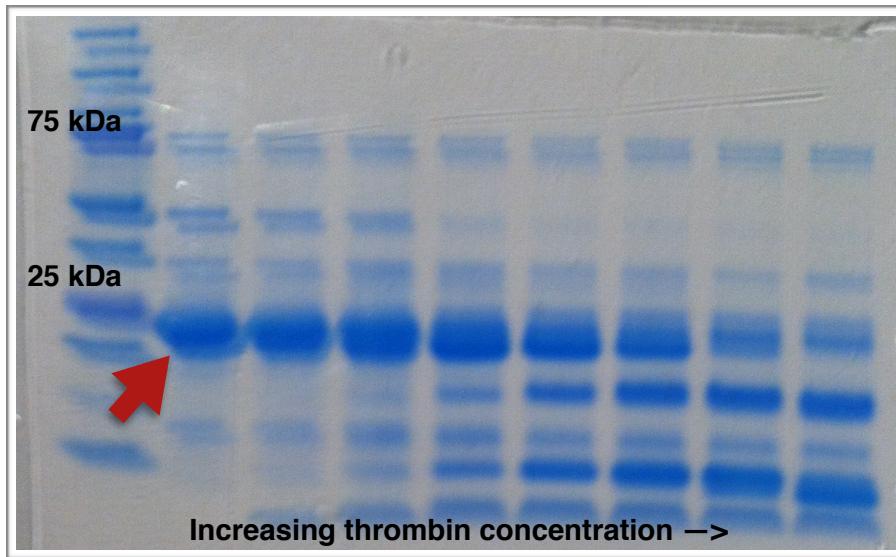


Figure 13: Thrombin digest SDS PAGE, showing the original undigested protein (red arrow), protein with cleaved His-tag (slightly shifted down), and several additional fragments (especially at about 5 and 10 kDa, corresponding to the additional cleavage site).

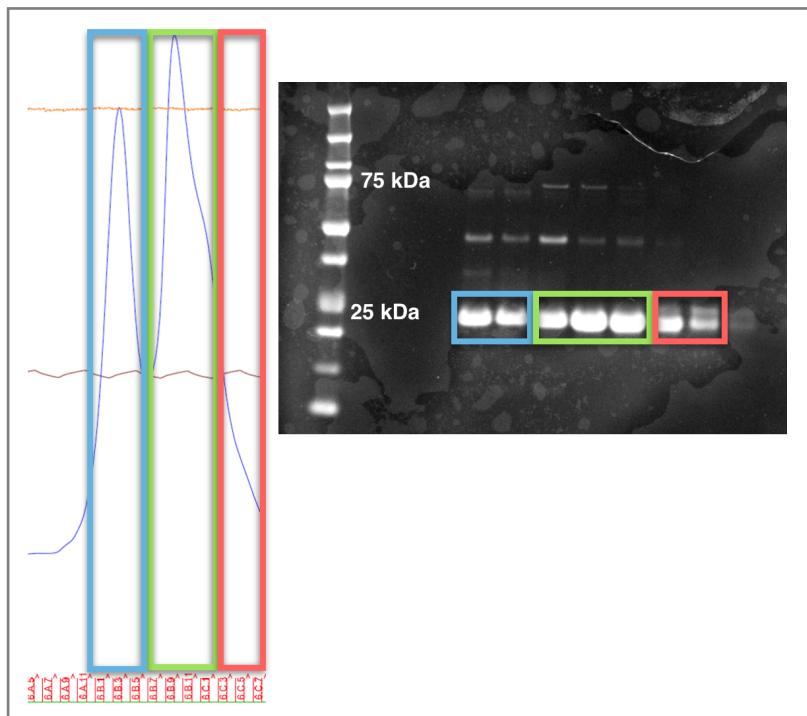


Figure 14: Large scale sizing results, with separations into 3 fractions (blue = fraction 1, green = fraction 2, red = fraction 3) based on purity and concentration.

exclusion column (with 5 mM DTT) (Fig. 14). SDS PAGE was used to analyze samples of each fraction of interest on the chromatogram, and the protein was pooled into 3 fractions according to purity and concentration.

In an attempt to get more information about the secondary structure of this protein, a 1-D (proton) nuclear magnetic resonance (NMR) spectrum was obtained. Unlike small molecules, only very general structural information is available from 1-D protein NMR, but it is enough to determine whether or not the protein is folded, or is largely disordered.

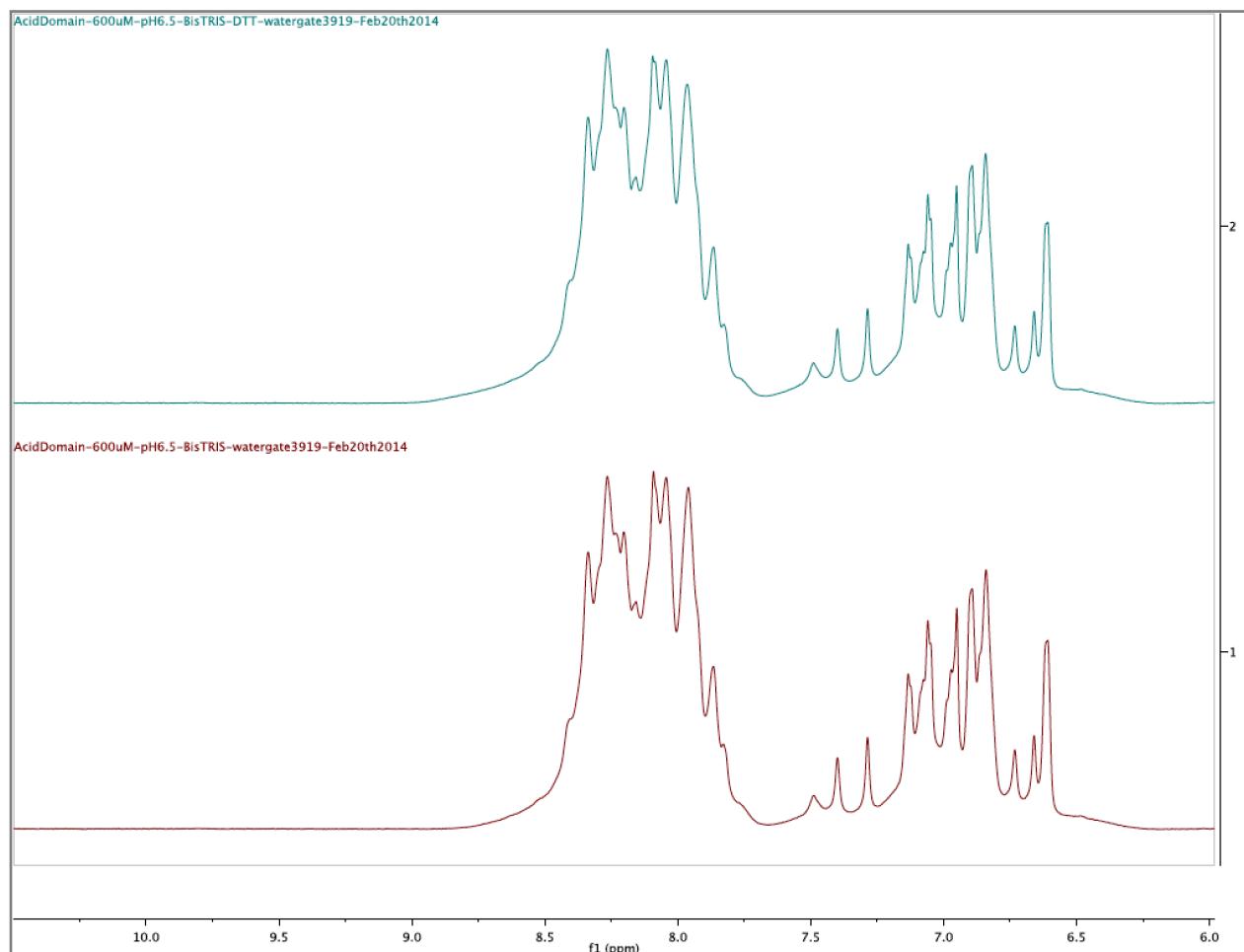


Figure 15: 1-D proton NMR Spectra of the HAstV-1 Acidic Domain, with (top) and without (bottom) DTT, showing the region relevant to secondary structure determination

The spectra were collected using a Varian 600 MHz NMR, equipped with a cryo-probe. The proton spectrum was acquired using the water pulse sequence from Biopack at 25°C. The samples were prepared using 600 uM of acidic domain in 10 mM bis-tris buffer at pH 6.5. 10mM DTT was added to one of the samples.

The NMR spectrum of a protein can get very complicated due to the number of resonances involved (leading to heavily overlapped peaks), but it can provide useful information about protein folding, functions, and the oligomeric state even without the knowledge of peak assignment. For a folded protein, all of the atoms are constrained in a well-ordered conformation. The backbone amide protons that are packed inside the hydrophobic core would have chemical shifts higher than 9.5 ppm and the methyl groups inside the hydrophobic core would have chemical shifts in the 0 to -1 ppm region.

The chemical shifts of the backbone amide protons of the 600 uM acidic domain cluster mainly between 6.5 to 8.8 ppm with sharp peaks (Fig. 15), which is an indication that the protein does not possess ordered secondary or tertiary structures. Proteins primarily composed of helices tend to have less packed hydrophobic regions than proteins primarily composed of beta strands. The chemical shift dispersion for the amide proton region of a helical protein would be smaller (6 to 9 ppm) than that of a beta-stranded protein (6 to 11 ppm).

The predicted alpha-helical region near the C-terminal is likely to be present, but the spectrum is dominated by the unstructured region. The NMR spectrum collected in the presence of 10 mM DTT shows no changes, suggesting that the conformation of the protein is not affected by the reducing environment.

Circular dichroism (CD) is another method for getting general information about secondary structure. CD involves measuring the differential absorption of left and right circularly polarized light at various wavelengths, and correlating that to the ratio of different chiral structures present in a given molecule. In this case, the technique was used to determine the approximate ratio of α -helices to disordered (coil) regions.

The results shown in Fig. 16 demonstrate that the most likely configuration is something similar to 40% helical and 60% unstructured, but is inconclusive due to lack of data below 200 nm.

To see if the disordered region could be trimmed off by limited proteolysis, while leaving the helical bundle intact (which would also provide additional confirmatory evidence of a

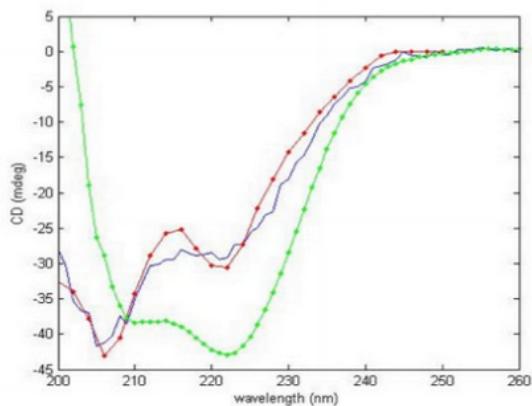


Figure 16: CD Spectrum of the HAstV-1 Acidic Domain (blue) with a linear combination of 40% helix and 60% unstructured regions (red). For contrast, a 77% helical structure is shown in green.

structured helical bundle region), a sample of the acidic domain was subjected to a trypsin protease digest at a variety of concentrations (all incubated at 37°C for 2 hours).

The results indicate that the bundle does exist (a stable band is formed as the concentration of trypsin is increased - Fig. 17), and that it may be possible to separate it from the disordered region. This bundle could provide a more suitable and compact sample for crystallization, in case the disordered region prevents crystallization of the full construct.

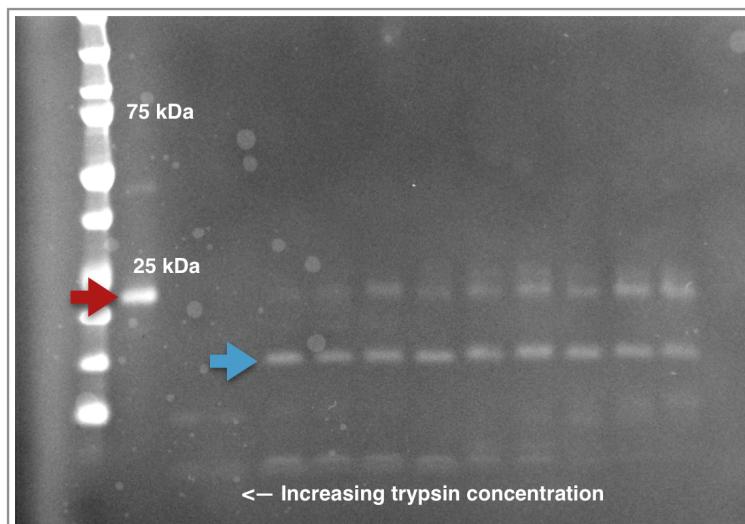


Figure 17: Trypsin digest concentration test ranging from 10:1 to 2500:1 acidic domain:trypsin ratio (by mass). The red arrow corresponds to the full length construct, and the blue arrow corresponds to the stable band

Armed with this information, approximately half of the remaining protein was digested with trypsin, purified by size exclusion, and the stable helical domain was isolated. Unlike previous size exclusion purification steps, the fragment eluted at the volume commensurate with its size (Fig. 18). These fractions were then concentrated to approximately 2 mg/mL using a 10 kDa MWCO spin concentrator.

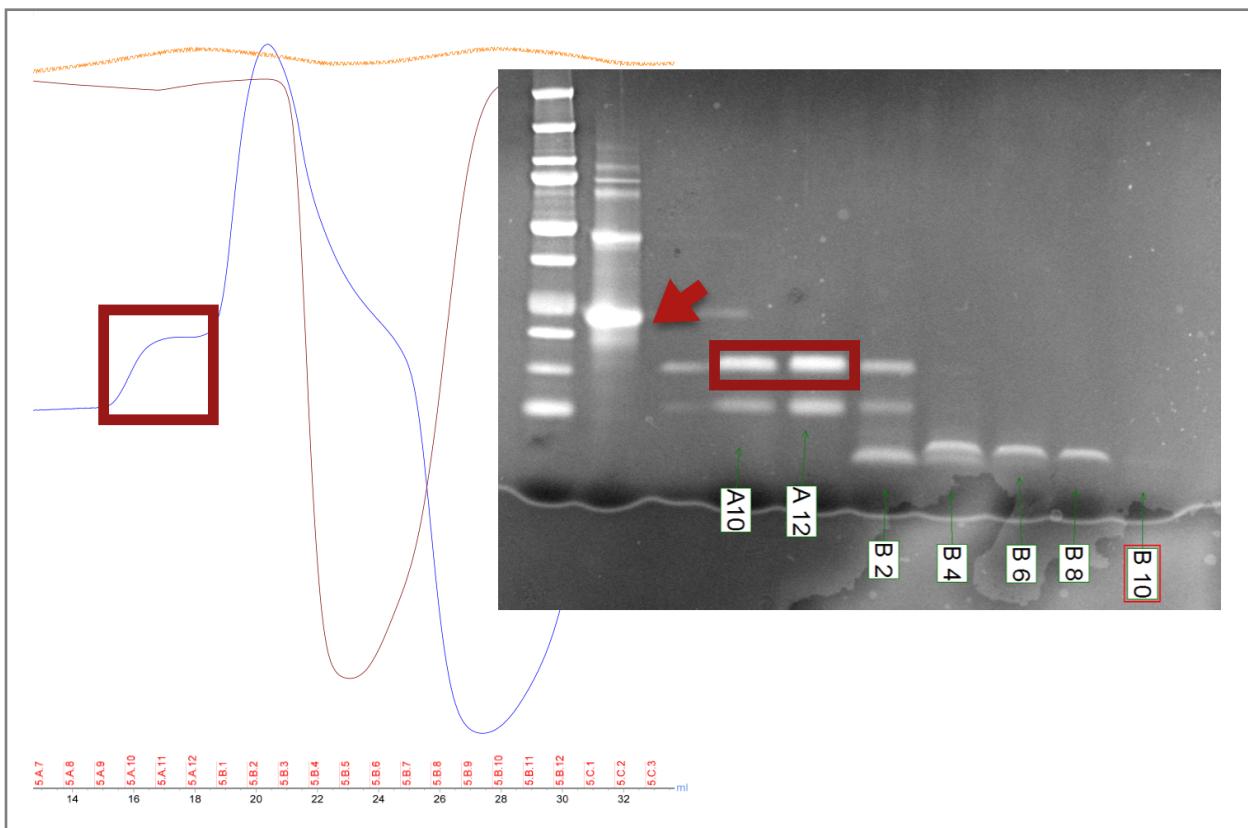


Figure 18: Post-trypsin size exclusion run showing the stable band and the corresponding peak in the chromatograph. The red arrow corresponds to the full-length construct.

All remaining samples of the full length construct were collected into a single container and concentrated, using the same type of concentrator, to about 7 mg/mL. When all samples had reached the desired concentration, they were dialyzed overnight into a fresh buffer containing 10 mM Tris, 150 mM NaCl, and 5 mM DTT. These samples were then set aside to perform crystal trials.

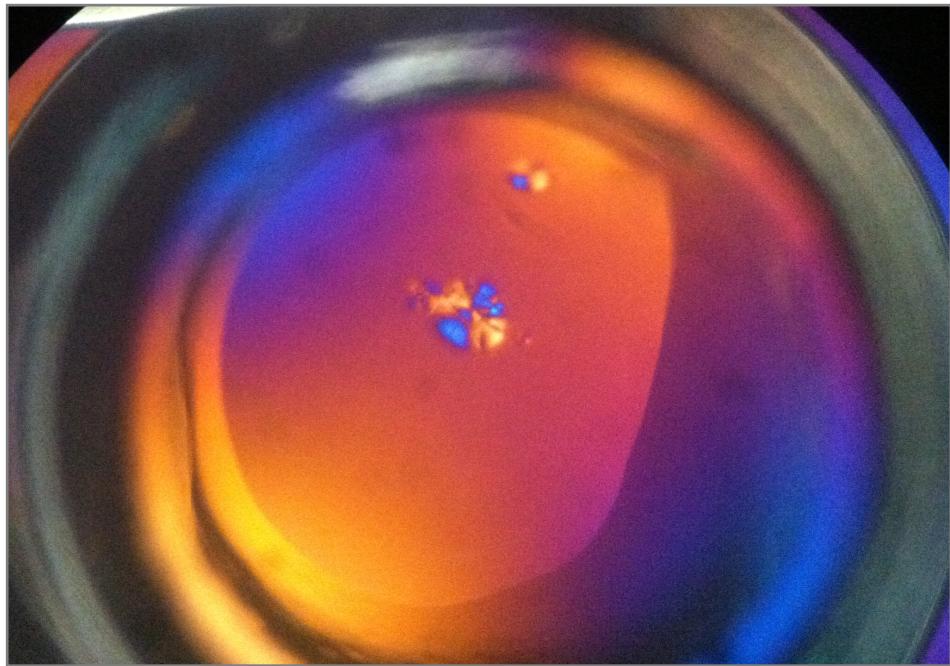


Figure 19: Sample of crystals observed for trypsinized fragment

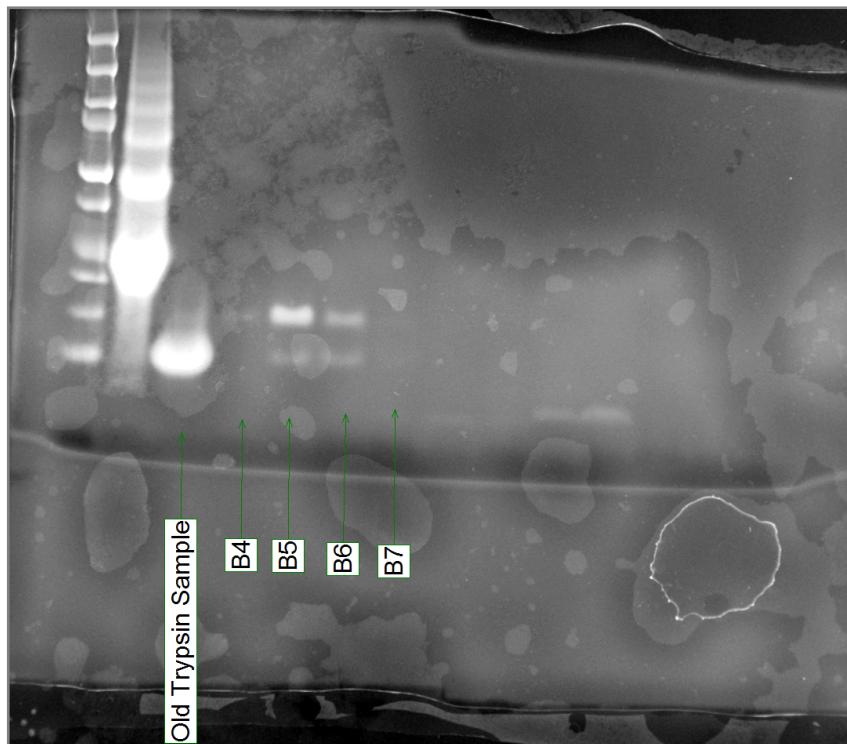


Figure 20: SDS-PAGE with the original trypsinized fragment and post-digest size exclusion results for a second digest (B4-B7)

Crystallization trials were performed with both full length acidic domain and the trypsinized fragment (full length at 3 mg/mL and 5 mg/mL concentrations, fragment at 2 mg/mL) using Hampton HR2-130, HR2-134, HR2-139, and HR2-248 screening kits.

After 10 days, approximately 20 wells contained crystals for the trypsinized fragment. No hits were observed for the full length samples.

Based on the hits found with the screening kits, buffers were formulated for refinement containing sodium citrate at a pH range from 3.0 to 6.0, and with varying concentrations of PEG 3350 (10-20%) and PEG 400 (15-35%) as a precipitant. Each of these solutions produced the same types of crystals as the original screens, each with about the same number of crystals per drop.

Unfortunately, while using the fragment as a control during a second digestion, it was discovered that residual trypsin had caused the complete degradation of the sample. (Fig. 20)

To determine the composition of the observed crystals, a side by side comparison was performed with the old and new trypsin digested fragments. All drops containing

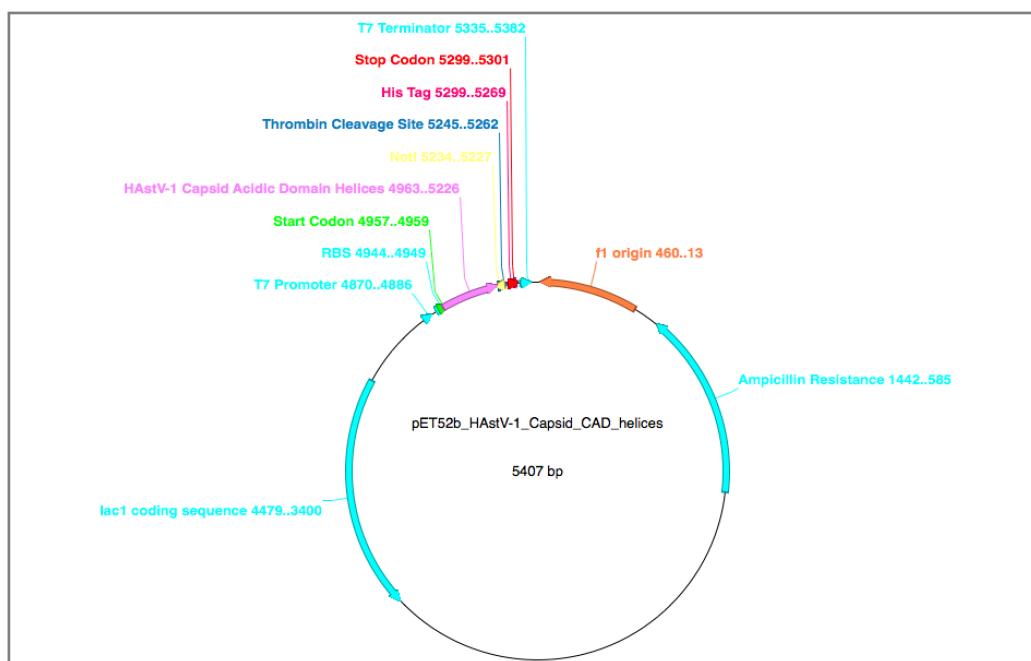


Figure 21: pET52b plasmid diagram showing insertion of the helical region of the capsid acidic domain.

crystals were those containing only the degraded sample, indicating that the crystals were likely composed of these degraded fragments.

Given these difficulties with trypsin digestion, primers were designed to target the helical bundle (as an alternative to digesting the full length acidic domain). Cloning was performed (using the procedures described above), and the presence of the desired insert was confirmed by sequencing in the pET52b expression plasmid (Fig. 21).

Conclusions and Future Directions

While the tertiary structure of the acidic domain is still unknown, protocols for successfully producing and purifying it have been devised. Several lines of evidence (including tryptic digest pattern, circular dichroism, and sizing behavior) have helped confirm the predicted secondary structure of approximately 40% helix and 60% unstructured. Attempts to crystallize both the entire domain and the putative helices isolated by tryptic digest were unsuccessful.

Expression tests (similar to those described above for the full acidic domain) will be performed to determine the best cell line for the production of the helical domain. If these efforts are successful, this protein will then be purified and subjected to crystallization trials. If crystallization proves difficult, a structure may be pursued via NMR.

Functional assays are currently in development, which aim to localize the protein within the cell and determine what effect (if any) this domain has on mammalian cells without the presence of the full virus. Additionally, the potential for association with the host membrane will be tested using a lipid binding assay. Ultimately, the goal remains to understand the role of this domain in the overall lifecycle and epidemiology of the human astrovirus.

Appendices

I.

DNA and amino acid sequences for the full acidic domain:

tccaggcatctggcatggatgagagtgacaacatcgaaatactggatgctccagactctgctgaccagttaaagaag
acatagagacacagacatggagagtacagaggacgaagacgcgaagcggacagggttatcatagacact
tctgtatgatgatgaaatgagacagacggcgttaaccctcctcaactctgttaatcaaggatgacaatgacgc
gtgccacaaggatagcacggcgcgcattccccacgcttccgataggatcaagcgtggagtatacatggacctgtgtct
cggggtaagcccaggcaatgcatggtctcatgcgtgtgaagaggcacfcaagcagttagggaaaccaatccctgc
acatctggaaagccgcggccacgcccag

srasygyesdnteyleapdsadqfnediettdiesteddeddeadrfdiidtsdeedesetdrvllstlvnqgmtitratki
arrafpTLSdrikrgvymdlasgaspgnawshaceearkageinpctsgsrgae

I.

DNA and amino acid sequences for the helical region of the acidic domain:

tctgtatgatgatgaaatgagacagacggcgttaaccctcctcaactctgttaatcaaggatgacaatgacgc
gtgccacaaggatagcacggcgcgcattccccacgcttccgataggatcaagcgtggagtatacatggacctgtgtct
cggggtaagcccaggcaatgcatggtctcatgcgtgtgaagaggcacfcaagcagttagggaaaccaatccctgc
acatctggaaagccgcggccacgcccag

sdeedgnetdrvllstlvnqgmtmtratriarrarfTLSdrikrgvymdlvsgvspgnawshaceearkavgetnpct
sgsrgae

II.

Primers:

HAstV1 AD Forward: (with start codon for Nco1 into pET52B)
5' - [ccatggatccaggcatctggatggatgagagtgacaac](#)

HAstV1 AD Reverse: (for Not1 into pET52B)

5' - gcggccgcctcgccgtggccgcggc

NcoI HAstV1 AD Forward: (start codon for Nco1 into pET52B –extra bps for Nco1, 45bp)

5' - tatccATGgatccaggcatctggatggttatgagagtgaca

KpnI HAstV1 AD Forward: (KpnI to be in frame with N-term strep-tag in pET52B, 45bp)

5' - tatGGTACCGctccaggcatctggatggttatgagagtgaca

NotI HAstV1 AD Reverse: (reverse primer for Not1 into pET52B with stop codon)

5' - atagcggccgcTTAactggcgatggaaatgagacagaccgtg

HAstV1 AD Helices Forward: (to be used with original reverse)

5' - TATccatggatctgatgaagaagatggaaatgagacagaccgtg

References

- ¹ Walter, J. E.; Mitchell, D. K. Astrovirus infection in children. *Curr. Opin. Infect. Dis.* **2003**, 16, 247-253.
- ² Glass, R. I.; et al. The changing epidemiology of astrovirus-associated gastroenteritis: a review. *Arch. Virol. Suppl.* **1996**, 12, 287-300
- ³ Mitchell, D. K.; et al. Molecular Epidemiology of Childhood Astrovirus Infection in Childcare Centers. *J. Infec. Dis.* **1999**, 180, 514-517.
- ⁴ Mitchell, D. K.; et al. Outbreaks of astrovirus gastroenteritis in day care centers. *J. Pediatr.* **1993**, 123, 725-732.
- ⁵ Willcocks, M. M.; Brown, T. D. K.; Madeley, C. R.; Carter, M. J. The complete sequence of a human astrovirus. *J. Gen. Virol.* **1994**, 75, 1785-1788.
- ⁶ Lukashov, V. V.; Goudsmit, J. Evolutionary relationships among Astroviridae. *J. Gen. Virol.* **2002**, 83, 1397-1405
- ⁷ Bass, D., Upadhyayula, U. Characterization of Human Serotype 1 Astrovirus-Neutralizing Epitopes. *J. Virol.* **1997**, 71, 8666-8671.
- ⁸ DuBois, R. M.; et al. Crystal Structure of the Avian Astrovirus Capsid Spike. *J. Virol.* **2013**, 87, 7853.
- ⁹ Dong, J.; Dong, L.; Mendez, E.; Tao, Y. Crystal Structure of the human astrovirus capsid spike. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, 10, 1073.
- ¹⁰ Mendez, E.; et al. Caspases Mediate Processing of the Capsid Precursor and Cell Release of Human Astroviruses. *J. Virol.* **2004**, 78, 16
- ¹¹ Mendez, E.; et al. Proteolytic processing of a serotype 8 human astrovirus ORF2 polyprotein. *J. Virol.* **2002**, 76, 7996-8002.
- ¹² Banos-Lara, M., Mendez, E. Role of individual caspases induced by astrovirus on the processing of its structural protein and its release from the cell through a non-lytic mechanism. *Virology*. **2010**, 401, 322-332.
- ¹³ Mendez, E.; et al. Association of the Astrovirus Structural Protein VP90 with Membranes Plays a Role in Virus Morphogenesis. *J. Virol.* **2007**, 81, 10649-10658.
- ¹⁴ Moser, L., Carter, M., Schultz-Cherry, S. Astrovirus increases epithelial barrier permeability independently of viral replication. *J Virol.* **2007**, 81, 11937-11945

¹⁵Soding, J; Biegert, A; Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nuc. Acids Res.* **2005**, 33, w244-w248.

¹⁶ Kim, D.E.; Chivian, D.; Baker, D. Protein structure prediction and analysis using the Robetta server. *Nuc. Acids Res.* **2004**, 32, w526-w531.