Chapter 3

Complete Genome Sequence of the Virginia Avirulent Strain of Turkey Hemorrhagic Enteritis Virus

3.1 Abstract

Turkey Hemorrhagic Enteritis Virus (THEV) is a *Siadenovirus* that causes disease in turkey poults characterized by splenomegaly, depression, bloody diarrhea, and death. The mechanism responsible for intestinal lesion formation and mortality is not known, though there is strong evidence that it is immune mediated. The Virginia Avirulent Strain (VAS) is a naturally occurring avirulent strain of THEV that replicates effectively in turkeys but does not cause the intestinal hemorrhage or mortality caused by more virulent strains of THEV. All strains of THEV are serologically indistinguishable, and their structural proteins are roughly the same size. Attempts to genetically differentiate isolates of THEV using restriction endonuclease fingerprinting resulted in apparent confirmation that there were genomic differences between strains isolated from chickens, turkeys, and pheasants. The full length genome of a virulent field isolate from Israel (IVS) was sequenced in 1998. The full length genome of the VAS was sequenced in order to determine any genetic differences that could have an impact on virulence.

Comparison of the two full length DNA sequences revealed 49 point mutations in the approximately 26.6 kb genome. Of those mutations, only 14 are expected to result in an alteration of amino acid sequence within viral proteins. Although any of these putative amino acid changes could be responsible for the observed virulence difference, there are a few that are more likely to have an impact on virulence. Changes in fiber length and rigidity have been attributed to virulence changes in other adenoviruses. Two putative amino acid changes found within the fiber gene may alter the structure of the fiber. Two putative amino acid changes found within the IVa2 gene may result in a decrease in viral transcriptional activation, or a change in virion packaging efficiency. E3 and ORF1 are putative viral genes that encode proteins of unknown structure and function. Two

mutations were detected within the E3 gene that may alter its function in the avirulent strain. The E3 putative gene product is not similar to any other known protein based on sequence comparisons using BLAST, so no speculation can be made about its possible function. The ORF1 gene of THEV potentially encodes a protein with some homology to bacterial sialidases. Interestingly, a frame-shift within ORF1 is predicted to result in the alteration of the N-terminus of the sialidase gene product. It is speculated that the loss of a function attributed to virulence is linked to one of the genetic differences found in the VAS and IVS. Unfortunately, not enough is known about the conformational changes that may result from the predicted amino acid changes.

The differences were confirmed by re-sequencing the regions surrounding each mutation. The same regions were then sequenced in five additional strains: three commercially available tissue culture vaccines, the Virginia Virulent Strain (VVS), and a field strain of THEV-V (S-1). Sequence analysis reveals variation in the fiber, ORF1, and E3 genes of all of the strains tested. Consistent variation in these genes between strains of THEV with different phenotypes strongly indicates these genes are key factors affecting virulence.

3.2 Introduction

Adenoviruses are linear, double-stranded DNA viruses that are known to be involved in infectious diseases of virtually every class of vertebrates. Adenoviruses are divided into four genera based on phylogenetic differences: *Mastadenovirus*, *Aviadenovirus*, *Atadenovirus*, and *Siadenovirus*. Mastadenoviruses infect a wide range of mammalian species including humans, while aviadenoviruses infect a wide range of avian species. The most recently added genera, *Atadenovirus* and *Siadenovirus,* are composed of "orphan" adenoviruses whose genetic structure and evolutionary history are not consistent with the host-defined hierarchy of mast- and aviadenoviruses (Benko and Harrach, 1998; Davison *et al*., 2003). The genus *Siadenovirus* is named after the genus-specific open reading frame (ORF) in the E1 region whose putative gene product has high

sequence similarity with bacterial sialidases. The two recognized members of this genus are THEV and Frog Adenovirus 1 (FrAdV-1)(Davison *et al*., 2000; Davison and Harrach, 2002). It has been hypothesized that siadenoviruses originated in amphibians and then adapted to avian species (Davison *et al*., 2003).

Members of the genus *Siadenovirus* have several genes that share no sequence similarity with any proteins studied in other organisms. It is not known what functional importance these genes may have. Very little is known about the replication of FrAdV-1 *in vivo*, and although THEV has been studied extensively for decades, detailed information about the molecular mechanisms of replication and pathogenesis is still lacking. An understanding of the role of unknown genes in viral replication and pathogenesis is essential in order to determine mechanisms by which siadenoviruses cause disease in their respective hosts.

Hemorrhagic enteritis (HE) is a disease of turkeys caused by THEV resulting in depression, splenomegaly, hemorrhagic diarrhea, immunosuppression, and mortality. Mortality is dependent on the virulence of the particular strain with up to 80% mortality observed in experimentally inoculated turkeys (Pierson and Fitzgerald, 2003). Gastrointestinal lesions in turkeys infected with virulent strains of THEV (THEV-V) consist primarily of mucosal congestion and hemorrhage that is not thought to be directly caused by viral cytolysis. Very few infected cells are found associated with the intestinal epithelium in the presence of lesions, and hemorrhage is believed to be the result of endothelial disruption rather than destruction (Gross, 1967; Opengart *et al*., 1992; Suresh and Sharma, 1996). Experimental immunosuppression by corticosteroid and cyclosporine A administration causes a decrease of lesion severity without affecting viral replication, suggesting a link to histamine and T-lymphocyte mediated inflammatory responses (Suresh and Sharma, 1995; Opengart *et al*., 1992; Fitzgerald *et al*., 1995).

Naturally occurring avirulent strains of THEV (THEV-A) have been isolated that replicate efficiently in turkeys, producing splenomegaly and transient immunosuppression, but not resulting in mortality or intestinal lesions (Domermuth *et al*., 1977, 1979a, b; Silim *et al*., 1978). The VAS has been used as a live-virus vaccine for

decades, with no reported reversion to a virulent phenotype (Fadly *et al*., 1985; Domermuth *et al*., 1977; Thorsen *et al*., 1982).

The inherent stability of a double-stranded DNA genome and the observed stability of the VAS genome suggests that the differences between the THEV-A and THEV-V strains are clearly defined. Attempts to genetically differentiate strains using restriction endonuclease fingerprinting indicated that there are genomic differences between strains isolated from chickens, turkeys, and pheasants (Zhang and Nagaraja, 1989). However, comparison of partial sequence data from the THEV-A and THEV-V shows extremely high sequence homology of major functional ORFs (Suresh *et al*., 1995; Jucker *et al*., 1996; Pitcovski *et al*., 1998; McQuiston *et al.,* 1997, not published: GenBank accession # AF036243). Such sequence identity is consistent with serological identity observed between strains, but does not explain the change in virulent phenotype. Determination of the viral factors affecting virulence is further complicated by a lack of detailed understanding of the molecular mechanism of lesion formation and mortality *in vivo*.

Based on stability of phenotypes, the difference between virulent and avirulent strains of THEV should be explainable by a change in genetic content. The purpose of this study was to determine the genome sequence of the VAS for direct comparison to the DNA sequence of a virulent field isolate from Israel (IVS)(Pitcovski *et al*., 1998). Genetic variations between the strains of THEV with different phenotypes will indicate which genes affect virulence.

3.3 Materials and Methods

Viruses

The VAS was originally isolated in 1977 from enlarged spleens of pheasants suspected of having marble spleen disease (Domermuth *et al*., 1977). It has been maintained as a low passage splenic vaccine strain for THEV-V, with 5-6 live-bird passages since 1985. Six week-old medium white turkey poults provided by British United Turkeys of America (BUTA, Lewisburg, WV) were intravenously inoculated with

100 TID_{50} VAS and euthanized by cervical dislocation 3 days post-inoculation. Spleens were collected from each bird and stored at -20°C until used.

The other strains used in this study were two virulent field strains and three commercially available tissue culture vaccine products. The Virginia Virulent Strain (VVS) was originally isolated from enlarged spleens of turkeys that had died from clinical HE (Domermuth *et al*., 1977). A strain of THEV-V was isolated during the summer of 2005, from a suspected case of clinical HE in the Shenandoah Valley, Virginia. Spleen samples from this case tested positive for THEV by PCR and AGID, and it is referred to as Suspect 1 (S-1). Three commercially available tissue culture vaccine strains were also included (TC-B, C, D). The original source of these strains is thought to be the VAS, but the true source is unknown.

The full length genome of the IVS was sequenced in 1998, and is included in the comparative analysis (Pitcovski *et al*., 1998).

DNA Extraction

All DNA extraction was performed using InstaGene Matrix (BioRad). Briefly, spleens were homogenized 50:50 v/v in sterile phosphate buffered saline (PBS). 250 uL tissue homogenate or vaccine was added to 200 uL InstaGene Matrix, mixed well, and incubated at 56°C in a hot water bath for 30 minutes. Following incubation, samples were vortexed thoroughly and incubated in a 100°C heat block for 8 minutes. Samples were then centrifuged at 15000 rpm for 10 minutes at room temperature in a microcentrifuge (Tomy MTX-150). Supernatants were collected and used in subsequent PCR amplifications. A standard PCR assay using primers specific for the THEV-A hexon gene (NHEVF-5'-gtggttcagcagaaagttctt-3'; NHEVR-5'-cagtagactcataagcaactat-3') was performed to verify presence of THEV DNA.

Complete Genome Sequencing

PCR primers were designed to amplify overlapping segments of the VAS genome that collectively spanned the complete viral genome. High sequence homology

previously observed between THEV strains enabled PCR primers to be designed based on the complete DNA sequence of a virulent field strain from Israel (IVS)(Pitcovski *et al.,* 1998). These segments were then sequenced by "primer walking" on both strands using 82 overlapping primer sets. Sequencing was carried out using an Applied Biosystems 3730 DNA Analyzer and assembly of the full length sequence was performed using Sequencher version 4.1. Amplification and sequencing of PCR fragments was performed by Sequetech Corp. (Mountain View, CA). The complete nucleotide sequence has been submitted to GenBank (Accession Number AY849321).

Partial Genome Sequencing

A total of ten regions of each strain were amplified by standard PCR in overlapping amplicons ranging in size from 120 to 624 bp (Table 1). Sequence homology previously observed between THEV strains enabled PCR primers to be designed based on the complete DNA sequence of the IVS (Pitcovski *et al*., 1998). Amplification products were electrophoresed in a 1% agarose gel and visualized with ethidium bromide. DNA was extracted from bands excised from the gel using the QiaQuick Gel Extraction kit (Qiagen) and re-suspended at 10 ng/uL for sequencing. Sequencing was carried out by the Core Laboratory at the Virginia Bioinformatics Institute using an Applied Biosystems 3730 DNA Analyzer.

Sequence Analysis

Sequence alignment and ORF prediction was carried out by BioEdit version 5.0.9 (Hall, 1999). Amino acid homology comparisons of putative ORFs were performed using TBLASTN version 2.2.9 (Altschul *et al*., 1997).

3.4 Results

The genome length, G+C content, and length of inverted terminal repeats (ITRs) were very similar (Table 2). The mutations identified in the sequence comparisons are

Table 3-1: Sequencing PCR Primers

Amplicon		Sense Primer		Antisense Primer	
Name	Size (bp)	Name	Position[+]	Name	Position[+]
Orf1	485	Orf11-L	40	Orf1-R	525
Orf1B	323	Orf1-B-L	1495	Orf1-B-R	1818
IVa2	338	IVa2-L	2938	IVa2-R	3276
Pol1	208	Pol1-L	3621	Pol1-R	3829
Pol2	149	Pol2-L	5050	Pol2-R	5199
Polptp	624	Polptp-L	6401	Polptp-R	7025
pVII	149	pVII-L	12498	pVII-R	12647
Hex	120	Hex-L	14125	Hex-R	14245
E3	377	E3-L	21597	E3-R	21974
Fib	233	Fib-L	23519	Fib-B-R	23752

[+]Nucleotide position relative to VAS sequence.

*Sequence of each primer is listed in Appendix A.

Table 3-2: *Siadenovirus* Genome Comparison

	THEV-A	THEV-V	FrAdV-1
Genome size (bp)	26266	26263	26163
ITR (bp)	40	39	36
(G+C)%	34.9	34.9	37.9
GenBank Accession	AY849321	AF074946	AF224336
Reference	This study	Pitcovski *et al*.,1998	Davison *et al*., 2000

listed in Table 3. The sequence of the VAS was found to be 99.9% identical to that of the IVS. All major ORFs and other sequence features were found in expected locations and conformation (Figure 1). The hexon and penton base proteins of VAS and IVS were found to have identical putative amino acid sequences. This was consistent with the observations that THEV strains are serologically indistinguishable. There were 49 point mutations found in the entire genome, resulting in only 14 putative amino acid changes within 7 functional ORFs.

A major mutation was found in ORF1, where a point deletion caused a frame-shift that alters the N-terminus of the putative gene product. A possible initiator for THEV ORF1 was located at nucleotide (nt) 282, followed downstream by a start ATG at nt 399. ORF1 in IVS then extended for 1550 bp, possibly encoding a single 517 amino acid (aa) polypeptide. A deletion of a thymine was found in VAS at nt 456, which caused a frame shift resulting in a double stop codon. The VAS may utilize an alternative start ATG at nt 313, restoring the reading frame but altering the N-terminal 20 aa residues. Interestingly, translation starting at the ATG at nt 313 by IVS would lead to a short 51 aa polypeptide identical to the altered N-terminus of VAS ORF1. It is possible that the thymine insertions and deletions were artifacts of sequencing IVS.

The mutations in the VAS leading to aa changes were all confirmed by the re-sequencing of ten key regions. The sequencing of those regions in five other THEV strains revealed five unique mutations in the ORF1, E3, and fiber genes.

3.5 Discussion

All structural components of the viral capsid had identical amino acid sequences, except for two differences found in the fiber knob. This is significant, as it supports the observation that THEV strains are indistinguishable by serology.

HE lesion formation is immune-mediated, so genetic differences in the VAS and the IVS must directly alter interaction of the virus with the host immune system. The sequence comparison of VAS and IVS permitted the generation of two hypotheses

Table 3-3: Sequence Differences

Mutation[a]	Nucleotide Position[b]	Amino Acid Change[a]	Amino Acid Position[b]	Gene Affected	Strain Affected
C → G	41	-	-	-	VAS or IVS[d]
T DELETE	104	-	-	-	IVS
A → C	253	N→ H[c]	-	n/a[c]	VAS
A → G	396	I → M	28	ORF1	S-1
T INSERT	457	Frameshift	-	ORF1	IVS
A → T	870	silent	-	ORF1	VAS or IVS[d]
A → G	1548	silent	-	ORF1	IVS
G → A	1585	V → I	425	ORF1	TC-B,C,D
A → C	1731	Q → H	473	ORF1	VAS
C → T	1755	silent	-	ORF1	VAS
C → T	2131	-	-	-	VAS or IVS[d]
A → G	2898	silent	-	IVa2	VAS or IVS[d]
C → T	3005	D → N	145	IVa2	IVS
G → C	3203	H → D	79	IVa2	IVS
C → T	3672	E → K	1033	AdPOL	VAS
A → T	3765	Y → N	1002	AdPOL	IVS
G → A	4543	silent	-	AdPOL	VAS or IVS[d]
C → G	5129	C → S	547	AdPOL	IVS
A → G	5268	silent	-	AdPOL	VAS or IVS[d]
A → G	5778	silent	-	AdPOL	VAS or IVS[d]
A → G	6010	silent	-	AdPOL	VAS or IVS[d]
C → T	6283	silent	-	AdPOL	VAS or IVS[d]
G → A	6453	H → Y	106	AdPOL	IVS
T → G	6607	silent	-	AdPOL	IVS
G → C	6977	L → V	528	pTP	VAS
G → A	7245	silent	-	pTP	VAS or IVS[d]
G → A	11321	silent	-	penton	VAS or IVS[d]
C → G	12563	Q → E	73	pVII	VAS
T → C	13927	silent	-	hexon	VAS or IVS[d]
T → C	14188	silent	-	hexon	IVS

[a] mutation reported relative to common ancestor

[b] positions reported relative to VAS

[c] located before putative start codon, not expected to be within translated region

[d] Only nucleotides of IVS and VAS sequenced at this position, sequence of common ancestor not known.

Table 3-3 Continued: Sequence Differences

Mutation[a]	Nucleotide Position[b]	Amino Acid Change[a]	Amino Acid Position[b]	Gene Affected	Strain Affected
A → T	14203	silent	-	hexon	IVS
C → T	14518	silent	-	hexon	VAS or IVS[d]
G → A	15241	silent	-	hexon	VAS or IVS[d]
C → T	16039	silent	-	hexon	VAS or IVS[d]
T → C	16222	silent	-	hexon	VAS or IVS[d]
C → T	17741	silent	-	DBP	VAS or IVS[d]
A → G	18208	silent	-	100K	VAS or IVS[d]
T → C	19192	silent	-	100K	VAS or IVS[d]
G → A	19273	silent	-	100K	VAS or IVS[d]
A → C	21650	T → K	146	E3	IVS
C → A	21710	P → H	166	E3	S-1
A → C	21730	T → P	173	E3	TC-B,C,D
G → C	21930	K → N	239	E3	VAS
T → G	23016	silent	-	fiber	VAS or IVS[d]
G → A	23580	M → I	354	fiber	IVS
T → C	23645	M → T	376	fiber	IVS
A → G	23650	N → D	378	fiber	VVS
C → G	23984	-	-	-	VAS or IVS[d]
C → G	23992	-	-	-	VAS or IVS[d]
T INSERT	24275	-	-	-	VAS or IVS[d]
T INSERT	24277	-	-	-	VAS or IVS[d]
A → G	24811	silent	-	ORF7	VAS or IVS[d]
C → T	24940	silent	-	ORF7	VAS or IVS[d]
A DELETE	26037	-	-	-	VAS or IVS[d]

[a] mutation reported relative to common ancestor

[b] positions reported relative to VAS

[d] Only nucleotides of IVS and VAS sequenced at this position, sequence of common ancestor not known.

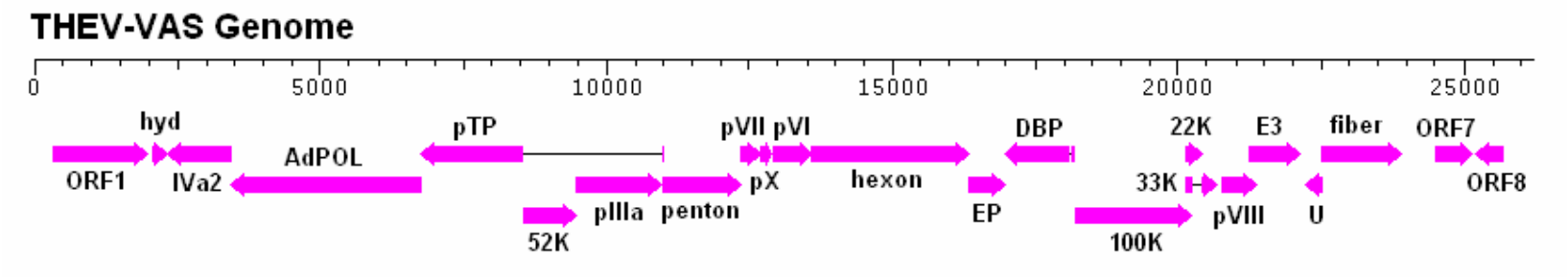

Figure 3-1: Genetic Organization of the Virginia Avirulent Strain of THEV

Putative open reading frames expected to code for actual viral proteins based on homology with the other member of the genus *Siadenovirus*, FrAdV-1. THEV and FrAdV-1 have the fewest putative genes of any adenoviruses sequenced to date.

regarding possible mechanisms for this alteration. First, THEV-V may produce a protein that directly activates T-lymphocytes and initiates lesion formation. Alteration of this protein could remove this functionality. Alternatively, a change resulting in a loss of THEV-A replication efficiency resulting in fewer infected cells and lower viral load in the spleen at peak infection could prevent intestinal lesions from forming.

Studies on mastadenovirus infection suggest individual viral genes can greatly impact pathogenesis. Mastadenoviruses utilize several viral gene products to mediate the host cell cycle and systemic immune response. The mastadenovirus E1A region encodes a transcription factor known to deregulate cellular growth control in order to make a favorable environment for viral replication. Without the interaction of the mastadenovirus E1B gene products, the disruption of well regulated cellular processes caused by E1A often results in apoptosis of the infected cell (Grand, 1987). Mastadenovirus E3 gene products are known to protect viral infected cells from cytolysis by immune-mediated tumor necrosis factor (TNF) and cytotoxic T lymphocytes (CTLs)(Ginsberg *et al*., 1989). The early viral gene products act together to promote favorable conditions for viral replication and at the same time protect infected cells from the host immune response and apoptosis (Shenk, 1996).

Characterization of the THEV genome permitted identification of potential genes that code for proteins that could carry out similar functions. ORF1, hyd, E3, ORF7, and ORF8 may encode the viral proteins needed to antagonize specific cellular pathways in order to increase viral replication and protect infected cells from apoptosis. Based on sequence similarity with FrAdV-1, it is expected that each of these ORFs encode true viral proteins. Unfortunately, due to the lack of clear homology between these genes and other proteins with known function, it was impossible to determine which ones may be orthologous to mastadenoviral E1A, E1B, and E3 gene products. However, putative amino acid differences were discovered in the ORF1 and E3 genes of the VAS which could play a major role in virulence.

ORF1

ORF 1 has been referred to as sialidase because it has a region with a high level of similarity to bacterial sialidase proteins. ORF1 is considered to be genus specific for the siadenoviruses, though its function is not known (Davison *et al.*, 2000; Davison and Harrach, 2002). It is not known whether the conserved sialidase domain is functional within the viral replication cycle. The sequence comparison of the VAS and IVS revealed an altered putative N-terminus. Partial sequencing of the ORF1 gene confirmed that the N-terminal amino acid sequence of the VAS was shared by all of the North American isolates tested. The consequence of this change is unknown. A novel mutation was found in the S-1 strain, also near the N-terminus. Further characterization of this viral gene product is necessary before speculation about its possible role in the mechanism of lesion formation in THEV-V infection.

E3

Only mastadenoviruses and siadenoviruses encode genes in the region between the pVIII and fiber ORFs known as E3 (Davison *et al*., 2003). In mastadenoviruses, the E3 genes encode proteins that aid the virus in the avoidance of the host immune response. The E3 19kDa protein results in decreased MHC-I antigens on the surface of infected cells, which prevents efficient antigen presentation to T-lymphocytes. E3 deletion mutants caused increased lymphocyte and macrophage activation and inflammatory response to infection (Ginsberg *et al*., 1989). However, the single ORF found within the THEV E3 region shared no sequence homology with any of the genes found within the mastadenovirus E3 region. Unlike ORF1, there was no apparent similarity of THEV's E3 ORF to any protein studied in other organisms, preventing prediction of the structure and function of its gene product. Four point mutations leading to amino acid changes were found in the E3 ORF of several THEV strains, but it was impossible to predict the impact of individual amino acid changes on protein function without any knowledge of the protein itself.

Genetic changes altering amino acids in genus-common proteins could result in a general replication deficiency of THEV-A. Although many of the mutations found in these genes are conserved in all of the North American strains, regardless of their virulence, it is possible that other mutations affecting these genes could result in virulence differences. Differences observed in certain proteins could affect viral replication by altering viral infectivity, transcription levels, genome replication, and virion packaging. This deficiency of viral production could directly affect the stimulation of the host immune system, leading to decreased macrophage activation, less severe T-lymphocyte mediated response, and a subsequent decrease in immune-mediated lesion formation.

Fiber

Changes in the knob domain of the fiber protein in the VAS could lead to a decrease in receptor binding efficiency, resulting in a net decrease of viral infectivity. If the virus is not able to infect cells efficiently, the number of infected cells and consequently the number of progeny virus generated would decrease. Two mutations leading to amino acid changes were found in the fiber gene. The adenoviral fiber structure is composed of a trimer of fiber monomers. Each fiber monomer has an N-terminal penton-binding domain and a C-terminal knob domain, linked together by a shaft domain made up of short repeats of a triple beta-spiral motif (van Raaij *et al*., 1999; Pitcovski *et al*., 2005). The C-terminal knob domain is responsible for receptor binding and target cell adhesion, and the predicted amino acid differences in VAS were both found in this part of the protein. All sequence changes found in the VAS fiber gene were identical to those found in a previous sequence of the VAS fiber (McQuiston *et al*., 1997, not published; GenBank accession # AF036243). All of the North American isolates shared the same sequence in the fiber knob, except for a unique mutation found in the VVS. It is not known whether the amino acid differences result in a major conformational shift.

A lower level of viral transcription from the major late promoter (MLP) could result in a net decrease in viral replication efficiency. IVa2 is a transcriptional activator and plays a major role in virion packaging (Zhang *et al*., 2001; Zhang and Imperiale, 2003; Pardo-Mateos and Young, 2004b; Lutz and Kedinger, 1996; Tribouley *et al*., 1994). Amino acid changes found in IVa2 could affect transcriptional activation of viral promoters. It is not clearly understood which domains of the IVa2 gene product are necessary or sufficient for each function, though an N-terminal truncated form of the Human adenovirus 2 IVa2 protein was still capable of binding the MLP and facilitation of virion packaging (Pardo-Mateos and Young, 2004a). Two mutations leading to amino acid changes were found in the IVa2 gene. It is possible that mutations in the IVa2 protein could alter transcription levels from the virus major late promoter. During viral replication, accumulating IVa2 binds to the MLP at specific sites and activates transcription of the viral late genes. A point mutation changing the binding efficiency of IVa2 to these sites could result in a decrease in overall transcription from the MLP (Pardo-Mateos and Young, 2004b).

Genomic packaging efficiency may also be altered, resulting in a net change in replication efficiency. It is not fully understood how IVa2 functions in packaging, but it involves the binding of the genome near the left terminus and interaction with other viral proteins such as 52K and pVII in order to assemble the genome within the viral capsid (Gustin *et al*., 1996; Zhang *et al*., 2001). Viral protein pVII is also known to play a role in packaging of the genome into whole virions. pVII is the major viral histone-like protein, primarily responsible for packaging the viral chromosome (Zhang and Arcos, 2005). Mutations affecting DNA binding of pVII may destabilize the viral genome, resulting in less efficient packaging. The mutation in pVII is unique to the VAS, and appears to be a relatively recent change. Two previous partial sequences of the region including the pVII ORF in the VAS do not contain this mutation (Jucker *et al*., 1996; Suresh *et al*. 1995).

DNA replication could be affected by changes in the viral DNA polymerase and terminal proteins. The terminal protein binds the genome terminus, facilitating entry of the genome into the host nucleus and acting as a protein primer for DNA replication (Liu

et al., 2003; Freimuth and Ginsberg, 1986; Greber *et al*., 1997). The two proteins work together to bind the genome terminus and replicate the viral genome. Amino acid changes altering their interactions could decrease the number of complete genomes available for incorporation into progeny virions. Four mutations leading to amino acid changes were found in the viral DNA polymerase gene, but none within well-conserved domains of polymerase function (Liu *et al*., 2003).

A general replication deficiency may cause a difference in viral load achieved within the spleens of THEV infected turkeys. Splenic lesions have been reported to vary based on the amount of viral inoculum used (Ossa *et al*., 1983). Incidence and severity of intestinal lesions associated with THEV-V have been found to appear in a dose dependent manner (Nazerian and Fadly, 1986; Gross and Moore, 1967). It is not known whether a change in viral load in the spleen of infected birds could alone be responsible for the differences seen in virulence.

The comparison of the complete sequences of two strains of THEV allowed for speculation about the molecular pathways involved in the pathogenesis of THEV-V *in vivo*. However, it did not allow for definitive conclusions to be drawn about the role of unknown viral gene products during infection. Partial sequencing has revealed that the ORF1, E3, and fiber genes are regions of high variation between THEV strains. Complete sequencing of these ORFs will be performed in several strains with different levels of virulence to further explore the roles each might play in pathogenesis. Further characterization of the transcription patterns and mechanisms of pathogenesis must be performed before the genetic change responsible for variation of virulence of THEV-V and THEV-A can be confirmed.

3.6 References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nuc Acids Res* 25, 3389-3402.

Benko, M., & Harrach, B. (1998). A proposal for a new (third) genus within the family adenoviridae. *Arch Virol* 143, 829-837.

Davison, A.J., Wright, K.M., and Harrach, B. (2000). DNA sequence of frog adenovirus 1. *J Gen Virol* 81, 2431-2439.

Davison, A.J., and Harrach, B. (2002). Siadenovirus. In *The Springer Index of Viruses,* pp.29-33. Edited by C.A. Tidona and G. Darai. New York: Springer-Verlag.

Davison, A.J., Benko, M., and Harrach, B. (2003). Genetic content and evolution of adenoviruses. *J Gen Virol* 84, 2895-2908.

Domermuth, C.H., Gross, W.B., Douglass, C.S., Dubose, R.T., Harris, J.R., and Davis, R.B. (1977). Vaccination for hemorrhagic enteritis of turkeys. *Avian Dis* 21, 557-565.

Domermuth, C.H., Gross, W.B., Schwartz, L.D., Mallinson, E.T., and Britt, R. (1979a). Vaccination of ring-necked pheasant for marble spleen disease. *Avian Dis* 23, 30-38.

Domermuth, C.H., Harris, J.R., Gross, W.B., and Dubose, R.T. (1979b). A naturally occurring infection of chickens with a hemorrhagic enteritis/marble spleen disease type of virus. *Avian Dis* 23, 479-484.

Fadly, A.M., Nazerian, K., Nagaraja, K., and Below, G. (1985). Field vaccination against hemorrhagic enteritis of turkeys by a cell-culture live-virus vaccine. *Avian Dis* 29, 768-777.

Fitzgerald SD, Reed WM, Furukawa AM, Zimels E, Fung L. (1995). Effect of T-lymphocyte depletion on the pathogenesis of marble spleen disease virus infection in ring-necked pheasants. *Avian Dis* 39, 68-73.

Freimuth, PI, Ginsberg, H.S. (1986). Codon insertion mutants of the adenovirus terminal protein. *Proc Natl Acad Sci* 83, 7816-20.

Ginsberg, H.S., Lundholm-Beauchamp, U., Horswood, R.L., Pernis, B., Wold, W.S.M., Chanock, R.M., and Prince, G.A. (1989). Role of early region 3 (E3) in pathogenesis of adenovirus disease. *Proc Natl Acad Sci* 86, 3823-3827.

Grand, R.J.A. (1987). The structure and function of the adenovirus early region 1 proteins. *Biochem J* 241, 25-38.

Greber, U.F., Suomalainen, M., Stidwell, R.P., Boucke, K., Ebersold, M.W., and Helenius, A. (1997). The role of the nuclear pore complex in adenovirus DNA entry. *EMBO J* 16, 5998-6007.

Gross, W.B. (1967). Lesions of hemorrhagic enteritis. *Avian Dis* 11, 684-693.

Gross, W.B., and Moore, W.E.C. (1967). Hemorrhagic enteritis of turkeys. *Avian Dis* 11, 296-307.

Gustin, K.E., Lutz, P., and Imperiale, M.J. (1996). Interaction of the adenovirus L1 52/55-kilodalton protein with the IVa2 gene product during infection. *J Virol* 70, 6463-7.

Hall, T.A. (1999). Bioedit: a user-friendly biological sequence alignment editor and analysis (program for Windows 95/98/NT). *Nucl Acids Symp Ser* 41, 95-98.

Jucker, M.T., McQuiston, J.R., van den Hurk, J.V., Boyle, S.M., and Pierson, F.W. (1996). Characterization of the haemmorhagic enteritis virus genome and the sequence of the putative penton base and core protein genes. *J Gen Virol* 77, 469-479.

Liu, H., Naismith, J.H., Hay, R.T. (2003). Adenovirus DNA replication. *Curr Top Microbiol Immunol* 272, 131-164.

Lutz, P., Kedinger, C. (1996). Properties of the adenovirus IVa2 gene product, an effector of late-phase-dependent activation of the major late promoter. *J Virol* 70, 1396-1405.

Nazerian, K., and Fadly, A.M. (1986). Further studies on *in vitro* and *in vivo* assays of hemorrhagic enteritis virus (hev). *Avian Dis* 31, 234-240.

Opengart, K. N., Eyre, P., and Domermuth, C.H. (1992). Increased numbers of duodenal mast cells in turkeys inoculated with hemorrhagic enteritis virus. *Am J Vet Res* 53, 814-819.

Ossa, J.E., Bates, R.C., and Schurig, G.G. (1983). Hemorrhagic enteritis in turkeys: purification and quantification of the virus. *Avian Dis* 27, 235-245.

Pardo-Mateos, A., and Young, C.S.H. (2004a). A 40kda isoform of the type 5 adenovirus IVa2 protein is sufficient for virus viability. *Virol* 324, 151-164.

Pardo-Mateos, A., and Young, C.S.H. (2004b). Adenovirus IVa2 protein plays an important role in transcription from the major late promoter *in vivo*. *Virol* 327, 50-59.

Pierson, F.W., and Fitzgerald, S.D. (2003). Hemorrhagic enteritis and related infections. In *Diseases of Poultry,* 11th ed., pp. 237-247. Edited by Y.M. Saif. , H.J. Barnes, J.R. Glisson, A.M. Fadly. L.R. McDougald, and D.E. Swayne. Ames: Iowa State University Press.

Pitcovski, J., Mualem, M., Rei-Koren, Z., Krispel, S., Shmueli, E., Peretz, Y., Gutter, B., Gallili, G.E., Michael, A., and Goldberg, D. (1998). The complete DNA sequence and genome organization of the avian adenovirus, hemorrhagic enteritis virus. *Virology* 249, 307-315.

Pitcovski, J., Fingerut, E., Gallili, G., Eliahu, D., Finger, A., Gutter, B. (2005). A subunit vaccine against hemorrhagic enteritis adenovirus. *Vaccine* 23, 4697-4702.

Shenk, T. (1996). Adenoviridae: the viruses and their replication. In *Fields Virology,* 3rd ed., pp. 2111-2148. Edited by Fields, B.N., Knipe, D.M., and Howley, P.M. Philadelphia: Lippincott – Raven.

Silim, A., Thorsen, J., and Carlson, H. (1978). Experimental infection of chickens with hemorrhagic enteritis virus. *Avian Dis* 22, 106-114.

Suresh, M. And Sharma, J.M. (1995). Hemorrhagic enteritis virus induced changes in the lymphocyte sub-populations in turkeys and the effect of experimental immunodeficiency on viral pathogenesis. *Vet Immunol Immunopathol* 45, 139-150.

Suresh, M., Cyr, S.St., and Sharma, J.M. (1995). Molecular cloning and sequence analysis of the penton base genes of type II avian adenoviruses. *Virus Res* 39, 289-297.

Suresh, M., and Sharma, J. M. (1996). Pathogenesis of type II avian adenovirus infection in turkeys: *in vivo* immune cell tropism and tissue distribution of the virus. *J Virol* 70, 30-36.

Thorsen, J., Weninger, N., Weber, L., Van Dijk, C. (1982). Field trials of an immunization procedure against hemorrhagic enteritis virus of turkeys. *Avian Dis* 26, 473-477.

Tribouley, C., Lutz, P., Staub, A., Kedinger, C. (1994). The product of the adenovirus intermediate gene IVa2 is a transcriptional activator of the major late promoter. *J Virol* 68, 4450-4457.

van Raaij, M.J., Mitraki, A., Lavigne, G., Cusack, S. (1999). A triple B-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. *Nature* 401, 935-938.

Zhang, C., and Nagaraja, K.V. (1989). Differentiation of avian adenovirus type-II strains by restriction endonuclease fingerprinting. *Am J Vet Res* 50, 1466-1470.

Zhang, W., Low, J.A., Christensen, J.B., and Imperiale, M.J. (2001). Role for the adenovirus IVa2 protein in packaging of viral DNA. *J Virol* 75, 10454-10466.

Zhang, W., and Imperiale, M.J. (2003). Requirement of the adenovirus IVa2 protein for virus assembly. *J Virol* 77, 3586-3594.

Zhang, W., Arcos, R. (2005). Interaction of the adenovirus major core protein precursor, pVII, with the viral DNA packaging machinery. *Virology* 334, 194-202.

Chapter 4

Sequence Comparison of ORF1, E3 and Fiber Genes from Twelve Different Isolates of Turkey Hemorrhagic Enteritis Virus

4.1 Abstract

Turkey Hemorrhagic Enteritis Virus (THEV) is a *Siadenovirus* that causes disease in turkey poults characterized by splenomegaly, depression, bloody diarrhea, and mortality. The purpose of this study is to determine which viral genes are involved in virulence. A previous study comparing the full-length genome sequences of a virulent field isolate and an avirulent vaccine strain revealed point mutations resulting in putative amino acid changes in seven viral gene products. Further sequencing allowed narrowing of the focus from seven genes to three: ORF1, E3, and fiber. ORF1 and E3 encode non-structural proteins of unknown function. The Fiber is the structural protein responsible for target cell adhesion. The ORF1, E3, and fiber genes were sequenced in eleven different strains of THEV: four commercial tissue culture vaccine strains, four virulent field isolates, the Virginia Avirulent Strain (VAS), the Virginia Virulent Strain (VVS), and a pheasant vaccine strain of turkey origin (MSV). PCR primers amplifying overlapping gene segments were designed based on the sequence of the VAS. After amplification, both strands of each PCR product were sequenced. Although the data revealed a high level of homology (99.9%) between all of the strains, there were several mutations found in each strain. Virulent strains could be differentiated from the vaccine strains based on mutations in the ORF1 and E3 genes. The virulent field isolates differed substantially from the VVS in all three genes. It was impossible to predict which of the three candidate genes is responsible for virulence, or if it is a combination of factors. The mechanism responsible for intestinal lesion formation and mortality is not known, though there is strong evidence that it is immune mediated. Several mutations were found in the fiber knob, but none that clearly differentiate the virulent strains from the vaccine strains. It is possible that ORF1 and E3 are responsible for interference with host cell cycle

regulation much like E1A in mammalian adenoviruses. Future studies will attempt to clarify the roles of ORF1 and E3 during infection.

4.2 Introduction

Hemorrhagic enteritis (HE) is a disease of turkeys caused by THEV resulting in depression, splenomegaly, hemorrhagic diarrhea, immunosuppression, and mortality. Mortality is dependent on the virulence of the particular strain with up to 80% mortality observed in experimentally inoculated turkeys. Strains of THEV have been isolated based on their ability to cause disease in other avian species such as chickens and pheasants, with variation in mortality and lesion formation. These include Marble Spleen Disease Virus (MSDV) of pheasants and *Aviadenovirus* Type II-Associated Splenomegaly Virus (AASV) of chickens. All of the subtypes are serologically identical regardless of the host from which they were isolated, and most experimental evidence indicates the three viruses are actually a single species (Pierson and Fitzgerald, 2003).

Avirulent strains of THEV (THEV-A) have been successfully used as live virus vaccines for decades. These strains actively replicate in the spleens of infected birds, causing splenomegaly and some degree of immunosuppression, without intestinal hemorrhage or mortality. A naturally occurring avirulent strain, the Virginia Avirulent Strain (VAS) was isolated from pheasants in the 1970s (Domermuth *et al*., 1977). The use of MDTC-RP19 lymphoblastoid cells to propagate virus *in vitro* has allowed for the creation of tissue culture attenuated strains, which are now commercially available.

Although the phenotypic differences between avirulent and virulent strains of THEV are significant, the actual cause for the virulence shift is unknown. The full length genome of a virulent field isolate from Israel (IVS) was sequenced in 1998 (Pitcovski *et al*., 1998). Comparison of sequence data from the VAS revealed extremely high sequence homology of major structural ORFs (Suresh *et al*., 1995; Jucker *et al*., 1996; Pitcovski *et al*., 1998; McQuiston *et al*., 1997, not published; GenBank accession # AF036243). The

full length genome of the VAS was sequenced in 2004 in order to determine the location of these key genetic differences (Chapter 3).

It is hypothesized that in order for two viral strains to have a major phenotypic difference, they must have key genetic differences. Comparison of the VAS to the IVS resulted in the identification of 14 point mutations, leading to putative amino acid changes in seven different genes. The regions surrounding these changes were then sequenced in five additional isolates with varying virulence. Six of the 14 changes were determined to be possessed solely by the VAS, and the rest of the changes were possessed solely by the IVS. Although the 5 isolates did not possess any of the specific point changes found in the IVS and VAS, novel changes were discovered within some of the regions sequenced (Chapter 3). The presence of novel changes only in ORF1, E3, and the fiber knob domain permitted a shift of focus to those three genes. The purpose of this study was to determine the complete sequence of the ORF1, E3, and fiber knob genes in 11 distinct North American isolates of THEV with varying degrees of virulence. Genetic changes that are consistently seen in the strains may be responsible for shifts in virulence.

4.3 Materials and Methods

Viruses

The viruses used in this study were isolated from commercially available vaccines and various virulent field isolates (Table 1). The Virginia Avirulent Strain (VAS) was originally isolated in 1977 from enlarged spleens of pheasants suspected of having marble spleen disease. It has been maintained as a low passage splenic vaccine strain for THEV-V, with 5-6 live-bird passages since 1985. The Virginia Virulent Strain (VVS) was originally isolated from enlarged spleens of turkeys that had died from clinical HE. The Marble Spleen Vaccine strain (MSV) is thought to be the Avirulent II strain isolated from the litter of a turkey farm whose poults displayed natural resistance to experimental THEV-V infection (Domermuth *et al*., 1977). Four strains of THEV-V were isolated during the summer of 2005, from four cases of suspected clinical HE in the Shenandoah

Table 4-1: Viruses Sequenced

Isolate Designation	Source	Year	GenBank Accession numbers
Virginia Avirulent Strain (VAS)	Pheasant spleen	2004	AY849321
Virginia Virulent Strain (VVS)	Turkey spleen	1996	DQ868929
Israel Virulent Strain (IVS)	Israel field isolate	1998	AF074946
Marble Spleen Vaccine (MSV)	Turkey litter	1987	DQ868930
Tissue Culture Vaccine A (TC-A)	Turkey spleen	2005	DQ868935
Tissue Culture Vaccine B (TC-B)	Commercial	1991	DQ868936
Tissue Culture Vaccine C (TC-C)	Commercial	1992	DQ868937
Tissue Culture Vaccine D (TC-D)	Commercial	1992	DQ868938
Virulent Field Isolate 1 (S-1)	Virginia Field Isolate	2005	DQ868931
Virulent Field Isolate 2 (S-2)	Virginia Field Isolate	2005	DQ868932
Virulent Field Isolate 3 (S-3)	Virginia Field Isolate	2005	DQ868933
Virulent Field Isolate 4 (S-4)	Virginia Field Isolate	2005	DQ868934

Valley, Virginia. Spleen and gut samples from these cases tested positive for THEV by PCR and AGID, and are referred to as Suspect one through four (S1-4). Four commercially available tissue culture vaccine strains were also included. The original source of these strains is thought to be the VAS, but the true source is unknown. The full length genome of a virulent field isolate from Israel (IVS) was sequenced in 1998, and is included in the comparative analysis (Pitcovski *et al*., 1998).

DNA Extraction

Viral DNA was extracted from all isolates using InstaGene Matrix (BioRad). Briefly, spleen and intestinal samples were homogenized 50:50 v/v in sterile phosphate buffered saline (PBS). Lyophilized tissue culture vaccines were re-suspended in sterile PBS at a final concentration of 100-1000 doses/mL. 250 uL tissue homogenate or vaccine was added to 200 uL InstaGene Matrix, mixed well, and incubated at 56°C in a hot water bath for 30 minutes. Following incubation, samples were vortexed thoroughly and incubated in a 100°C heat block for 8 minutes. Samples were then centrifuged at 15,000 rpm for 10 minutes at room temperature in a microcentrifuge (Tomy MTX-150). Supernatants were collected and used in subsequent PCR amplifications.

PCR Amplification

The complete ORF1, E3, and fiber knob domain were amplified by standard PCR using overlapping amplicons ranging in size from 298 bp to 751 bp (Table 2). Sequence homology previously observed between THEV strains enabled PCR primers to be designed based on the complete DNA sequence of the IVS (Pitcovski *et al*., 1998). Agar gel electrophoresis was performed on amplification products at 100 volts using a 1% agarose gel and visualized with ethidium bromide. DNA was extracted from bands excised from the gel using the QiaQuick Gel Extraction kit (Qiagen) and re-suspended at 10 ng/uL for sequencing.

Table 4-2: Sequencing PCR Primers

Amplicon		Sense Primer		Antisense Primer	
Name	Size (bp)	Name	Position[+]	Name	Position[+]
ORF1A	485	ORF1-L	40	ORF1-R	525
ORF1B	636	0.1	329	WWW	965
ORF1C	751	ORF1-C-L	831	UUU	1582
ORF1D	323	ORF1-B-L	1495	ORF1-B-R	1818
ORF1E	489	0.5	1639	SSS	2128
E3A	473	0.63	21033	E3-C-R	21506
E3B	458	E3-D-L	21391	K	21849
E3C	377	E3-L	21597	E3-R	21974
E3D	520	E3-B-L	21810	E3-B-R	22330
FIBA	298	FIB-B-L	23350	FIB-R	23648
FIBB	396	FIB-L	23519	FIB-OLD	23915

[+]Nucleotide position relative to the VAS sequence.

*Sequence of each primer is listed in Appendix A.

Sequencing and Analysis

Sequencing was carried out by the Core Laboratory at the Virginia Bioinformatics Institute using an Applied Biosystems 3730 DNA Analyzer. Overlapping sequences were aligned and compiled using BioEdit version 5.0.9 (Hall, 1999).

4.4 Results

The mutations discovered by the sequence comparison are listed in Table 3. Each isolate was sequenced in three separate regions of the viral genome: ORF1 region (nt 70-2200); E3 region (nt 21070-22300); and fiber knob region (nt 23386-23880). The total length sequenced from each isolate was approximately 3850 bp, of a total 26266 bp, which is 14.7% genome coverage. Every isolate was nearly identical, with the greatest genetic change being individual point mutations. Of the twelve isolates whose sequences were compared, there were only eight distinct strains. TC-A, TC-B, and TC-C all had identical DNA sequences. S-2, S-3, and S-4 also had identical sequences to one another. A total of 28 point mutations were discovered in the sequenced regions of the 12 isolates. Some of these were shared among multiple isolates, but most were identified in only one. Of the 28 mutations, nine were silent or did not appear within a putative ORF. The remaining 19 changes are expected to alter the amino acid sequence of the translated viral proteins.

ORF1 Region

Sequence comparison reveals the presence of several unique and common point mutations within ORF1 (Figure 1). Sixteen point mutations were discovered in the 2130 bp region sequenced, eight of which were silent or not within the expected ORF1 translated region.

Importantly, none of the isolates from North America showed the T deletion/insertion found in the IVS during comparison to the full length sequence of the VAS. The putative N-terminus of ORF1 is therefore conserved among all the North

Table 4-3: Sequence Differences

Mutation	Nucleotide Position[+]	Amino Acid Change	Amino Acid Position[+]	Gene affected	Strain Affected
T Deletion	104	N/A	N/A	N/A	IVS
A → C	253	N → H*	N/A	N/A*	VAS
A → G	396	I → M	28	ORF1	MSV, S-1,2,3,4
T Insert	458	Frameshift	N/A	ORF1	IVS
C → G	557	A → G	82	ORF1	VVS
T → C	793	Silent	N/A	ORF1	VVS
T → C	836	V → A	175	ORF1	TC-D
A → T	870	Silent	N/A	ORF1	IVS
A → G	1548	Silent	N/A	ORF1	IVS
G → A	1585	V → I	425	ORF1	TC-A,B,C,D
G → T	1623	Silent	N/A	ORF1	S-2,3,4
A → C	1731	Q → H	473	ORF1	VAS
C → T	1755	Silent	N/A	ORF1	VAS
A → G	1796	Q → R	495	ORF1	TC-A,B,C,D
G → T	1845	L → F	511	ORF1	VVS
C → T	2131	N/A	N/A	N/A	IVS
A → C	21650	K → T	146	E3	IVS
C → A	21710	P → H	166	E3	S-1,2,3,4
A → C	21730	T → P	173	E3	TC-A,B,C,D
G → C	21930	K → N	239	E3	VAS
T → A	22056	Silent	N/A	E3	VVS
C → T	22191	N/A	N/A	N/A	VVS, MSV, S-1,2,3,4
G → A	23468	G → D	317	Fiber knob	S-1,2,3,4
G → A	23580	M → I	354	Fiber knob	IVS
T → C	23645	M → T	376	Fiber knob	IVS
A → G	23650	N → D	378	Fiber knob	VVS
C → A	23816	P → H	433	Fiber knob	S-2,3,4
C → T	23821	P → S	435	Fiber knob	TC-D

[+] positions reported relative to VAS

* located before putative start codon, not expected to be within translated region

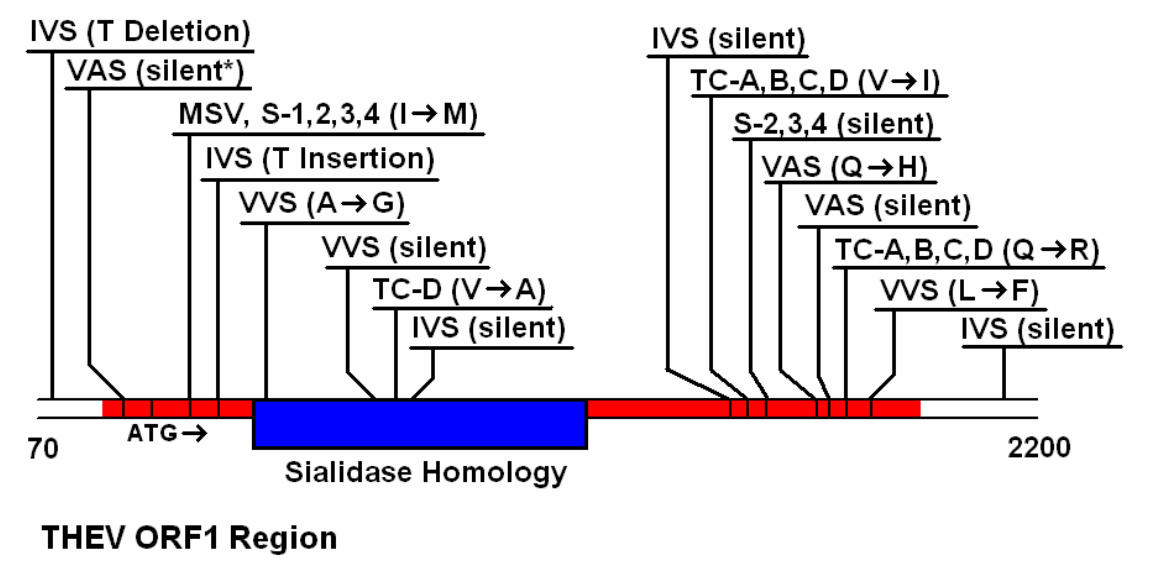

Figure 4-1: Mutations in the Putative ORF1 Gene

ORF1 is also known as sialidase, due to its putative amino acid homology with bacterial sialidase proteins. Sixteen point mutations were discovered within the ORF1 region.

American THEV isolates tested. A point mutation at nucleotide (nt) 396 of MSV, S1, S2, S3, and S4 was found, which resulted in an isoleucine to methionine change near the N-terminus of ORF1. Another mutation leading to amino acid change near the N-terminus of the ORF was found in the VVS at nt 557. This alanine to glycine change occurred within the region of ORF1 that has high homology with bacterial sialidases. The only other non-silent mutation within the sialidase-like region was found in TC-D, which occurred near silent mutations found in the VVS and IVS. Two mutations were found in all four tissue culture isolates, at nt 1585 and nt 1796. The valine to isoleucine change in particular occurred in a region of very high amino acid similarity when compared to the ORF1 of FrAdV-1. One mutation was found in the VAS at nt 1731 resulting in a glutamine to histidine change. A non-silent mutation was found near the C-terminus in the VVS at nt 1845.

E3 Region

Six point mutations were discovered in the 1230 bp E3 region, two of which were silent (Figure 2). Three non-silent mutations were found within a 27 aa region in the middle of the ORF (aa 146-173). The changes found in the suspect and tissue culture strains occur in a region of extreme variability compared to FrAdV-1. An additional non-silent mutation was found in the VAS at nt 21930. Interestingly, the only single point mutation conserved in all 6 of the North American virulent isolates was found at nt 22191, just downstream of the E3 ORF stop codon.

Fiber-knob Domain

Six point mutations were found in the fiber knob domain, and none of them were silent (Figure 3). All of the mutations occurred in virulent strains, with the exception of TC-D. The mutations were evenly distributed throughout the domain, with some apparent clustering. The amino acids involved in the changes were similar: a gain of aspartic acid in two mutations; the loss of methionine in two mutations, and the loss of proline in two mutations.

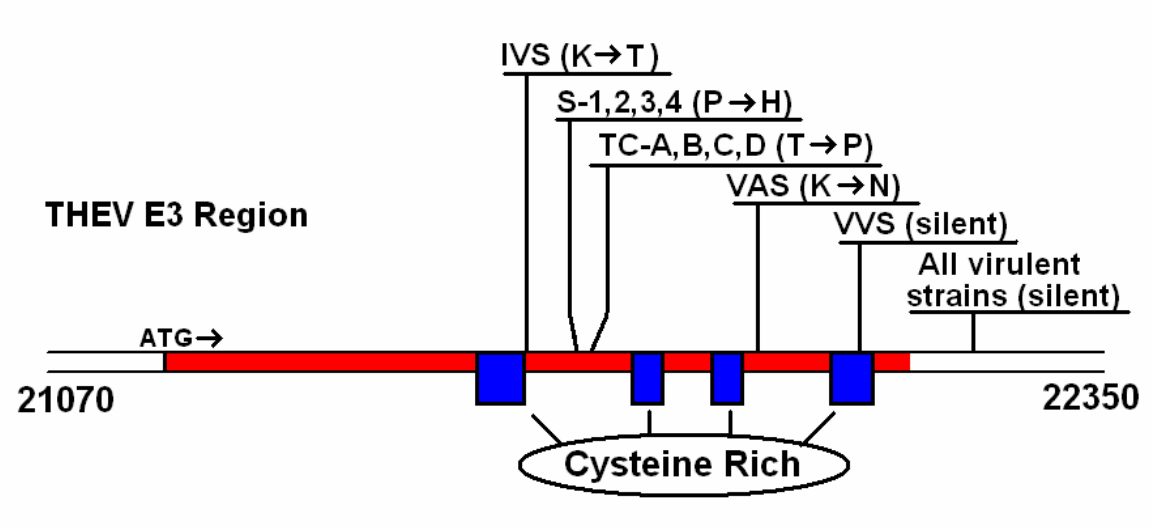

Figure 4-2: Mutations in the Putative E3 Gene

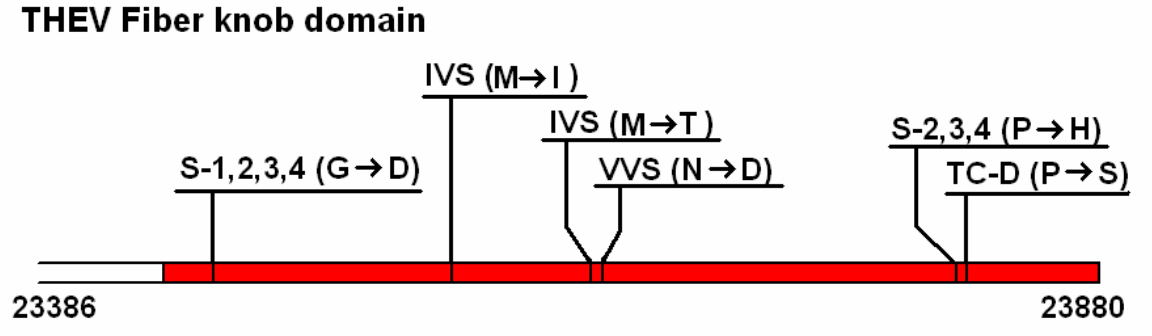

Figure 4-3: Mutations in the Fiber-knob Domain

Using phylogenetic analysis, the sub-types of THEV were grouped into distinct lineages (Figure 4). All four tissue culture vaccine strains were found to possess at least three mutations relative to the common ancestor. The VAS and the IVS were found to be different from all of the other strains. The North American virulent isolates grouped together, with the suspect field isolates most closely related to the MSV strain.

4.5 Discussion

Viruses are selected based on their ability to avoid the host immune system. In addition, viruses that kill their hosts during infection are at a disadvantage because a dead host cannot produce more viral progeny. Therefore, genetic mutations should generally progress toward a less virulent phenotype, and at the same time provide the virus some competitive advantage (Tizard, 2000a). All strains of THEV are very similar, replicating efficiently in the spleen causing immunosuppression and splenomegaly. Only certain strains cause the disease hemorrhagic enteritis, characterized by duodenal hemorrhage and death. The mechanism of intestinal lesion formation during THEV infection is unknown, but it is linked to a T-lymphocyte induced systemic shock in response to viral infection (Pierson and Fitzgerald, 2003). It is hypothesized that some difference in one or more viral proteins in avirulent strains results in a decreased T-lymphocyte activation, and lesions associated with virulence are not created. It is not known whether THEV strains have mutated over time to become more virulent in turkeys, or vice versa. The spontaneous emergence of naturally occurring avirulent strains in the 1970s and the conspicuous absence of reversion to virulence of these strains that have been used in turkey vaccines for decades suggests that THEV may progressively evolve toward a less virulent phenotype. Regardless of whether this is true, it is hypothesized that some of the genetic differences found in the ORF1, E3, and fiber knob provide either the virulent or avirulent viruses with a competitive advantage.

The MSV strain is thought to be the same as the Avirulent II strain isolated in

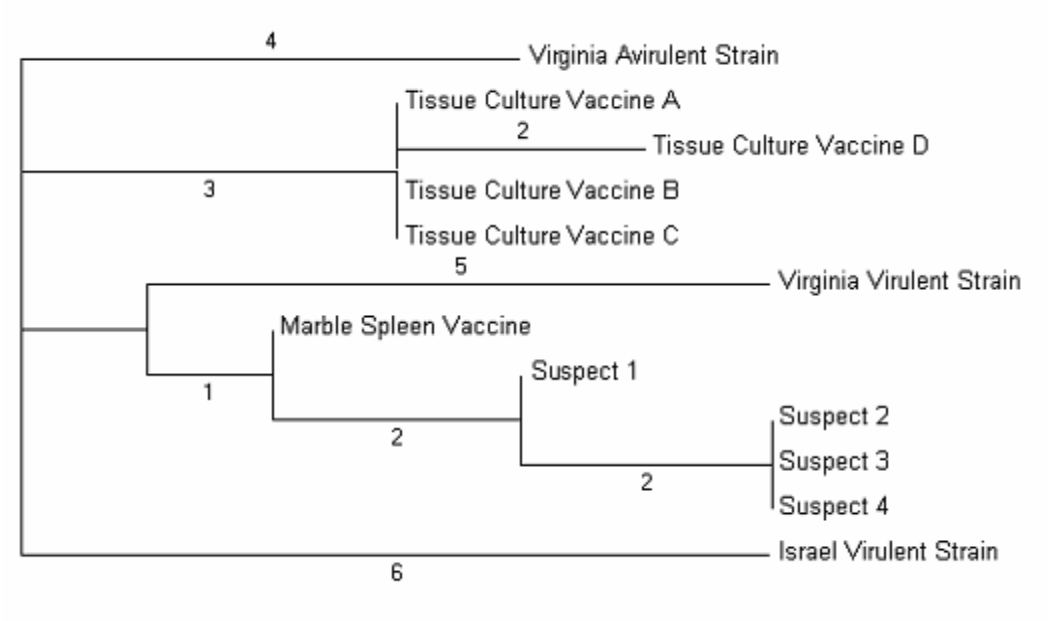

Figure 4-4: Relationship Between Strains Sequenced

Numbers on horizontal bars represent the number of unique individual point mutations discovered within the ORF1, E3, and fiber knob. The VAS and IVS strains shared no common mutations with any of the other isolates sequenced. The tissue culture strains shared three unique mutations, and TC-D had two additional mutations. All of the virulent strains isolated in Virginia shared at least one common mutation.

1977 and tested for use as a turkey vaccine strain. In the original tests, 25% of turkey poults inoculated with this virus had moderate intestinal lesions with no mortality, and a small spike in mortality during field vaccination trials. Due to the HE lesions caused by the virus, it was only used as a vaccine against MSD in pheasants (Domermuth *et al*., 1977). As a vaccine, it didn't cause clinical lesions in pheasants, although clinical MSD is rarely if ever seen in experimentally inoculated pheasants (Domermuth *et al*., 1979a). Certainly, the intestinal lesions the MSV strain caused in turkeys were not nearly as severe as those caused by the VVS. Based on this information, the MSV strain may be seen as an intermediate virulent strain. Other reports of virulent field isolates with varying degree of pathogenicity seem to support the possibility of varying degrees of virulence (Itakura *et al*., 1974). Unfortunately, this possibility only further complicates comparative analyses.

The sequence comparisons did not indicate major changes in ORF1, E3, or the fiber knob. There were no deletions or insertions discovered, only single point mutations. No single point mutation was conserved among all of the avirulent isolates. Several point changes were conserved among all tissue culture isolates, but did not share any common mutations with the VAS. No single point mutation was conserved among all of the virulent isolates, either. Of the seven virulent strains studied, six were isolated in Virginia. The VVS has been maintained in a laboratory setting, and had several unique mutations. The suspect strains were recent field isolates, and shared several common mutations. The lack of common mutations in the IVS suggests a more distant ancestry.

It cannot be assumed that all non-silent mutations have an effect on viral phenotype. An undefined proportion of the mutations discovered have no impact on the structure or function of the proteins they affect. Many mutations are simply random changes that have occurred over years of replication. In this way, the mutations only serve as phylogenetic markers that enable the tracing of the strains back to a common ancestor.

Based on the theory that MSV is a partially virulent strain, it makes sense that field strains isolated from spontaneous HE in unvaccinated flocks would be somewhat

related. If the four suspect strains are directly descended from the MSV strain, the mode of transition is unknown. It is possible that MSV used in pheasant MSD vaccination protocols was spread into wild pheasants or turkeys. Circulation within wild populations could maintain the virus for long periods of time, and unvaccinated turkeys would be particularly susceptible to infection.

The lack of common mutations prevented a definitive identification of mutations that alter virulence, however the pattern of the mutations indicated regions within certain genes that may affect virulence in some way. It is difficult to predict the effect of single amino acid changes in a well-characterized protein, and impossible when the protein has unknown structure and function. The mutations were compared in each gene based on the amino acid changes that result, although structural information for the protein is unknown. The similarity was evidence that these are key changes that affect viral phenotype.

Fiber knob

The fiber knob is responsible for host receptor interaction, so it is possible that any amino acid changes could result in alteration of cell tropisms or receptor affinity (Howitt *et al*., 2003; Arnberg *et al*., 1997). Unfortunately, the fiber knob of THEV is so different from that of the well characterized mastadenovirus fiber that structural predictions are still impossible. TC-D was the only avirulent strain that had a fiber knob mutation. Each of the virulent strains had a unique mutation in the fiber knob, except MSV. The fiber knob of MSV was identical to that of four of the five avirulent strains. If MSV is only partially virulent, the fiber may be one of several factors contributing to virulence.

The fiber is the only adenoviral protein that is glycosylated. Mutations proximal to or resulting in alteration of N-linked glycosylation sequences in the fiber knob have been observed to result in different hemagglutination profiles (Arnberg *et al*., 1997). Putative N-linked glycosylation sites (-N-X-T/S-) were located at amino acid (aa) 356 and 378 in the THEV strains sequenced. Both of the IVS mutations were exactly two

amino acids before these glycosylation sequences. The VVS asparagine to aspartate mutation resulted in complete removal of the glycosylation sequence at aa 378. Similar differences in amino acid sequences observed in Human adenovirus 19 resulted in a shift in hemagglutination efficiency (Arnberg *et al*., 1997). The consequences of such a change in THEV are unknown, and alteration in hemagglutination efficiency was not experimentally confirmed.

The suspect strain mutations and the TC-D mutation at aa 433 and aa 435 both resulted in the loss of a proline. Prolines tend to be located in regions of rigid protein structure, due to their unique chemical structure. Loss of rigid amino acids may increase the flexibility of the knob, and result in conformational change. In the same way, the gain of a charged residue in the suspect strains at aa 317 may increase polypeptide chain flexibility.

The actual role of the ORF1 and E3 genes in the formation of clinical HE in turkeys is unknown. The expression of early genes by replication-deficient mastadenoviruses has been shown to cause clinical lesions in infected mice (Ginsberg *et al*., 1990, 1991). In this way, early genes in THEV could be solely responsible for generation of clinical HE lesions in turkeys.

E3

Only mastadenoviruses and siadenoviruses encode genes in the region between the pVIII and fiber ORFs, known as E3. In mastadenoviruses, the E3 genes encode proteins that aid the virus in the avoidance of the host immune response. The E3 19kDa protein results in decreased MHC-I antigens on the surface of infected cells, which prevents efficient antigen presentation to T-lymphocytes. E3 deletion mutants caused increased lymphocyte and macrophage activation and inflammatory response to infection (Ginsberg *et al*., 1989). However, the single ORF found within the THEV E3 region shared no sequence homology with any of the genes found within the mastadenovirus E3 region. Unlike ORF1, there was no apparent similarity of THEV's E3 ORF to any protein

studied in other organisms, preventing prediction of the structure and function of its gene product. One major feature of the putative E3 gene product was a high number of cysteine residues that are conserved in both FrAdV-1 and THEV. A conserved cysteine-rich motif is repeated three times in the C-terminal half of siadenoviral E3 (CX_2CX_6CXC). This motif is very similar to one found in the GAL4 protein of yeast. GAL4 is a DNA binding transcriptional activator that has a very similar set of conserved cysteines (Marmorstein *et al*., 1992). The E3 cysteine rich regions also resembled the zinc-binding CR3 domain of the mastadenovirus E1A 13S protein. E1A 13S binds to the TATA-binding protein and other cellular transcription factors, and activates transcription of viral genes (Bruder and Hearing, 1989; Culp *et al*., 1988; Geisberg *et al*., 1994, 1995; Parker *et al*., 1997). E1A 13S activates the TNF gene in inflammatory cells, and this activation is dependent on this CR3 domain. An alternatively spliced form of E1A lacking the CR3 domain does not stimulate TNF production (Metcalf, 1996).

The MSV strain was identical to the VVS within the E3 gene, as neither had unique mutations relative to the other strains. As both are virulent strains, the changes in E3 of the avirulent strains may cause the virus to be less virulent. In E3, all suspect strains lost a proline and all of the tissue culture strains gained a proline in the central part of the gene. As in the fiber, it is expected that changes in prolines result in an overall change in polypeptide flexibility, with unknown structural consequences. The change in prolines occurred in a unique region of the E3 gene that has no similarity with the E3 gene of FrAdV-1. This may be an indication of a type-specific region of importance, or rather a non-essential region linking two domains of functional importance.

ORF1

ORF 1 has been called the sialidase protein because it has a region with a high level of similarity to bacterial sialidase proteins. ORF1 is considered to be genus specific for the siadenoviruses, though its function is not known (Davison *et al*., 2000; Davison and Harrach, 2002). It is not known whether the conserved sialidase domain is of functional importance. The ORF1 gene is large, and the number of unique mutations

complicated predictions. Mutations leading to putative amino acid changes in the N-terminal 20% of ORF1 occurred only in isolates expected to be virulent in turkeys. With the exception of the mutation found in TC-D at aa 175, every non-silent mutation in avirulent isolates was found in the C-terminal 20% of ORF1. The clustering of mutations in the terminal portions of ORF1 indicates the presence of functionally important domains.

Several of the putative changes do not likely result in any conformational change, due to the very close similarity of the amino acids involved. These include the valine to isoleucine change in the tissue culture strains at aa 425, the valine to alanine change in TC-D at aa 175. The isoleucine to methionine change at aa 28 is interesting in that it could serve as an alternative start ATG during translation of the ORF1 mRNA. Structurally speaking, isoleucine is almost identical to methionine, and is not expected to alter conformation.

All avirulent strains had a mutation resulting in the loss of a glutamine residue near the C-terminus. The glutamine lost in the VAS is conserved in the FrAdV-1 ORF1 protein. Substitution of glutamine with charged amino acids could alter the flexibility of the polypeptide chains and result in different conformation.

The two mutations found in the VVS could lead to significant conformational changes. The alanine to glycine change in the VVS at aa 82 could result in increased rotational freedom of the polypeptide chain. The size difference in phenylalanine and leucine at aa 511 is significant, and could cause a difference in hydrophobic packing.

Gastrointestinal lesions in turkeys infected with virulent strains of THEV consist primarily of mucosal congestion and hemorrhage that is not thought to be directly caused by viral cytolysis. Very few infected cells are found associated with the intestinal epithelium in the presence of lesions, and hemorrhage is believed to be the result of increased vascular permeability rather than endothelial destruction (Opengart *et al*., 1992; Suresh and Sharma, 1996). Such lesions are characteristic of a type I hypersensitivity reaction. The fact that duodenal mast cell numbers increase prior to lesion formation

indicates an inflammatory reaction caused by release of heparin, histamine, and cytokines such as TNF-alpha (Aggarwal, 2003). The T-lymphocyte population that is expected to mediate this mast cell response is the T-helper type 2 (T_{H2}). T_{H2} cell activation could lead to release of cytokines responsible for increasing mast cell proliferation (Tizard, 2000b; Prussin and Metcalfe, 2003). The mechanism of IgE dependent mast cell degranulation is not well characterized in avian species, but is assumed to follow a similar signaling pathway. It appears that THEV-V and THEV-A may be differentiated based on their ability to activate T_{H2} cells. THEV-V could directly stimulate T_{H2} cells via up-regulation of necessary cytokines, or by production of an interleukin-analog.

A major consequence of THEV infection is the depletion of IgM bearing cells in the spleen during peak viral replication (Suresh and Sharma, 1995, 1996; Saunders *et al*., 1993). THEV-V induces apoptosis in spleen cells of infected birds and primes splenocytes to release tumor necrosis factor (TNF) and interleukin 6 (IL-6) upon stimulation with Concanavalin-A (Con-A). Apoptosis in the spleen is observed concurrently with increasing viral replication, though it is not known whether THEV-V encodes a gene product that is capable of directly inducing apoptosis of target cells (Rautenschlein *et al*., 2000). It may be that E3 acts in a similar way to E1A 13S, up-regulating TNF production and causing apoptosis. TNF-alpha treatment has been shown to cause hemorrhagic intestinal lesions similar to HE in experimental animals. Treatment with thalidomide, a specific antagonist of TNF-alpha, has been shown to prevent clinical HE in turkeys infected with THEV-V (Suresh and Sharma, 1996; Rautenschlein and Sharma, 2000).

It is unusual for a wild-type adenovirus to cause apoptosis. The mastadenovirus E1A 13S protein is known to cause apoptosis, but only in mutant viruses lacking E1B and E3 gene functions. The inability of THEV-V to inhibit apoptosis in infected cell populations may result in uncontrolled cytolysis in the spleen. Such cytolysis in the spleen could produce a massive cytokine release, triggering an inflammatory shock response consistent with the intestinal lesions and mortality associated with THEV-V infection. If this is the case, it may be that the changes found in the THEV-A E3 protein

change its activity and prevent apoptosis. Alternatively, a loss of anti-apoptosis function in THEV-V or a gain of anti-apoptosis function in THEV-A ORF1 would result in a similar phenotype. It has not been determined whether THEV-A causes apoptosis, so it is unclear if apoptosis could be solely responsible for clinical lesion formation.

The lack of clear sequence homology between E3, ORF1, and mastadenoviral early gene products inhibited predictions about functional similarities. While the sequences determined in this study have allowed us to group the isolates based on common mutations, it is impossible to definitively identify the genes responsible for virulence until there is more knowledge of the molecular basis of virulence.

4.6 References

Aggarwal, B.B. (2003). Signalling pathways of the TNF superfamily: a double-edged sword. *Nature Reviews* 3, 745-756.

Arnberg, N., Mei, Y., & Wadell, G. (1997). Fiber genes of adenoviruses with tropism for the eye and genital tract. *Virology* 227, 239-244.

Bruder, J.T., and Hearing, P. (1989). Nuclear factor ef-1a binds to the adenovirus E1A core enhancer element and to other transcriptional control regions. *Mol Cell Biol* 9, 5143-5153.

Culp, J.S., Webster, L.C., Friedman, D.J., Smith, C.L., Huang, W., Wu, F.Y., Rosenberg, M., and Ricciardi, R.P. (1988). The 289-amino acid E1A protein of adenovirus binds zinc in a region that is important for trans-activation. *Proc Natl Acad Sci* 85, 6450-6454.

Davison, A.J., Wright, K.M., and Harrach, B. (2000). DNA sequence of frog adenovirus 1. *J Gen Virol* 81, 2431-2439.

Davison, A.J., and Harrach, B. (2002). Siadenovirus. In *The Springer Index of Viruses,* pp.29-33. Edited by C.A. Tidona and G. Darai. New York: Springer-Verlag.

Domermuth, C.H., Gross, W.B., Douglass, C.S., Dubose, R.T., Harris, J.R., and Davis, R.B. (1977). Vaccination for hemorrhagic enteritis of turkeys. *Avian Dis* 21, 557-565.

Domermuth, C.H., Gross, W.B., Schwartz, L.D., Mallinson, E.T., and Britt, R. (1979a). Vaccination of ring-necked pheasant for marble spleen disease. *Avian Dis* 23, 30-38.

Geisberg, J.V., Lee, W.S., Berk, A.J., and Ricciardi, R.P. (1994). The zinc finger region of the adenovirus E1A transactivating domain complexes with the TATA box binding protein. *Proc Natl Acad Sci* 91, 2488-2492.

Geisberg, J.V., Chen, J.L., Ricciardi, R.P. (1995). Subregions of the adenovirus E1A transaction domain target multiple components of the TF-IID complex. *Mol And Cell Biol* 15, 6283-6290.

Ginsberg, H.S., Lundholm-Beauchamp, U., Horswood, R.L., Pernis, B., Wold, W.S.M., Chanock, R.M., and Prince, G.A. (1989). Role of early region 3 (E3) in pathogenesis of adenovirus disease. *Proc Natl Acad Sci* 86, 3823-3827.

Ginsberg, H.S., Horswood, R.L., Chanock, R.M., and Prince, G.A. (1990). Role of early genes in pathogenesis of adenovirus pneumonia. *Proc Natl Acad Sci* 87, 6191-6195.

Ginsberg, H.S., Moldawer, L.L., Sehgal, P.B., Redington, M., Kilian, P.L., Chanok, R.M., and Prince, G.A. (1991). A mouse model for investigating the molecular pathogenesis of adenovirus pneumonia. *Proc Natl Acad Sci* 88, 1651-1655.

Hall, T.A. (1999). Bioedit: a user-friendly biological sequence alignment editor and analysis (program for Windows 95/98/NT). *Nucl Acids Symp Ser* 41, 95-98.

Howitt, J., Bewley, M.C., Graziano, V., Flanagan, J.M., and Freimuth, P. (2003). Structural basis for variation in adenovirus affinity for the cellular coxsackievirus and adenovirus receptor. *J Biol Chem* 278, 26208-26215.

Itakura, C., Carlson, H.C., Lang, G.N. (1974). Experimental transmission of haemorrhagic enteritis of turkeys. *Avian Pathol* 3, 279-292.

Jucker, M.T., McQuiston, J.R., van den Hurk, J.V., Boyle, S.M., and Pierson, F.W. (1996). Characterization of the haemmorhagic enteritis virus genome and the sequence of the putative penton base and core protein genes. *J Gen Virol* 77, 469-479.

Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S.C. (1992). DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* 356, 408-414.

Metcalf, J.P. (1996). Adenovirus E1A 13S gene product upregulates tumor necrosis factor gene. *Am J Physiol* 270, L535-L540.

Opengart, K. N., Eyre, P., and Domermuth, C.H. (1992). Increased numbers of duodenal mast cells in turkeys inoculated with hemorrhagic enteritis virus. *Am J Vet Res* **53**, 814-819.

Parker, S.F., Felzien, L.K., Perkins, N.D., Imperiale, M.J., Nabel, G.J. (1997). Distinct domains of adenovirus E1A interact with specific cellular factors to differentially modulate human immunodeficiency virus transcription. *J Virol* 71, 2004-2012.

Pierson, F.W., and Fitzgerald, S.D. (2003). Hemorrhagic enteritis and related infections. In *Diseases of Poultry,* 11th ed., pp. 237-247. Edited by Y.M. Saif. , H.J. Barnes, J.R. Glisson, A.M. Fadly. L.R. McDougald, and D.E. Swayne. Ames: Iowa State University Press.

Pitcovski, J., Mualem, M., Rei-Koren, Z., Krispel, S., Shmueli, E., Peretz, Y., Gutter, B., Gallili, G.E., Michael, A., and Goldberg, D. (1998). The complete DNA sequence and genome organization of the avian adenovirus, hemorrhagic enteritis virus. *Virology* 249, 307-315.

Prussin, C., Metcalfe, D.D. (2003). IgE, mast cells, basophils, and eosinophils. *J Allergy Clin Immunol* 111, 486-494.

Rautenschlein, S., Suresh, M., and Sharma, J.M. (2000). Pathogenic avian adenovirus type II induces apoptosis in turkey spleen cells. *Arch Virol* 145, 1671-1683.

Rautenschlein, S., Sharma, J.M. (2000). Immunopathogenesis of haemorrhagic enteritis virus in turkeys. *Dev Comp Immunol* 24, 237-246.

Saunders, G. K., Pierson, F.W., and van den Hurk, J.V. (1993). Haemorrhagic enteritis virus infection in turkeys: a comparison of virulent and avirulent virus infections, and a proposed pathogenesis. *Avian Pathol* 22, 47-58.

Suresh, M. And Sharma, J.M. (1995). Hemorrhagic enteritis virus induced changes in the lymphocyte sub-populations in turkeys and the effect of experimental immunodeficiency on viral pathogenesis. *Vet Immunol Immunopathol* 45, 139-150.

Suresh, M., Cyr, S.St., and Sharma, J.M. (1995). Molecular cloning and sequence analysis of the penton base genes of type II avian adenoviruses. *Virus Res* 39, 289-297.

Suresh, M., and Sharma, J. M. (1996). Pathogenesis of type II avian adenovirus infection in turkeys: *in vivo* immune cell tropism and tissue distribution of the virus. *J Virol* 70, 30-36.

Tizard, I.R. (2000a). Resistance to viruses. In *Veterinary Immunology: An Introduction*, 6th ed., pp. 267-279. Philadelphia: W.B. Saunders.

Tizard, I.R. (2000b). Type I hypersensitivity. In *Veterinary Immunology: An Introduction*, 6th ed., pp. 307-323. Philadelphia: W.B. Saunders.

Chapter 5

Sequence Comparison of Turkey Hemorrhagic Enteritis Virus and Frog Adenovirus 1: Prediction of Genetic Features of the Genus *Siadenovirus*

5.1 Abstract

The viral family *Adenoviridae* is divided into four genera based on sequence differences: *Mastadenovirus*; *Aviadenovirus*; *Atadenovirus*; and *Siadenovirus*. The genus *Siadenovirus* has two members that infect very different species: Frog Adenovirus 1 (FrAdV-1) and Turkey Hemorrhagic Enteritis Virus (THEV). FrAdV-1 was isolated in 1971 from a frog kidney, but it can only be propagated *in vitro* in a reptilian cell line. THEV is known to infect several avian species, including turkeys, chickens, and pheasants. The purpose of this study is to determine genetic and structural elements common to all members of the genus *Siadenovirus*, and determine differences that may contribute to host specificity. DNA sequences and putative gene products from THEV and FrAdV-1 were analyzed and compared (accession # AY849321; AF224336).

The five open reading frames (ORFs) found only in members of the genus *Siadenovirus* (ORF1, hyd, E3, ORF7, and ORF8) are predicted to code for non-structural proteins with functions that are currently unknown. ORF1 shares a significant region of homology with bacterial sialidases, though it is not known whether this indicates a functional similarity. Putative proteins Hyd and ORF7 are predicted to be real products based on similar position and length, despite only sharing 18-19% amino acid identity. E3 contains a high number of cysteine residues, and a conserved motif of CX_2CX_6CXC is repeated three times in the C-terminal half of the putative protein. Moreover, the putative promoter for ORF7 is relatively well conserved, and ORF8 contains three conserved prolines and two conserved cysteines.

TATA signals (RNA PolII binding sequence), inverted CCAAT boxes, and initiator sequences were predicted for the genus-specific ORFs as well as the major late

promoter (MLP). The siadenoviral MLP was found to be located in the same region as other adenoviruses, between the pol I and pol III domains of the DNA Polymerase gene, on the reverse coding strand.

The fiber protein is responsible for binding to specific host cell receptors, and has three domains: the tail, the shaft, and the knob. A variation of the highly conserved mastadenoviral penton interaction sequence (LNL(A/V)YPF) was found in the tail domain. The shaft of the siadenoviral fiber protein contains 16 repeats of the triple-beta helix motif common to all adenoviral fibers. The amino acid sequence of the knob domain of THEV is very different from that of FrAdV-1, suggesting significant difference in primary receptor structure.

Membership in the genus *Siadenovirus* is well defined based on inverted terminal repeat (ITR) length and sequence, G+C content, and possession of four genus-specific genes. Comparison of the putative gene products of members of the genus *Siadenovirus* reveals several key differences that may be responsible for host specificity. Further characterization of the roles of the genus-specific genes is necessary to fully understand how these genetic differences affect host specificity.

5.2 Introduction

The genus *Siadenovirus* is the most recently recognized addition to the family *Adenoviridae.* Only two members have been identified to date: Turkey Hemorrhagic Enteritis Virus (THEV) and Frog Adenovirus 1 (FrAdV-1)(Davison and Harrach, 2002). Strains of THEV responsible for three distinct clinical diseases have been isolated from chickens (AASV), turkeys (THEV), and pheasants (MSDV). Electron microscopy of intra-nuclear inclusion bodies and cesium chloride-purified virions from several strains of the virus revealed non-enveloped, icosahedral virions with diameters ranging in size from 60-90 nanometers (Pierson and Fitzgerald, 2003).

FrAdV-1 was originally isolated from a granulomatous nodule within the kidney of a Vermont leopard frog. Kidney tissues were used to inoculate TH-1 cells, which is a

turtle heart cell line that is known for its unusual susceptibility to human adenoviruses. Cytopathic effect was seen after 16 days at 23 C. After approximately twenty passages, pale eosinophilic intra-nuclear inclusions were seen in infected cells. Electron microscopy of the inclusions revealed a non-enveloped icosahedral virus approximately 78 nm in diameter. Virions were packed in semi-crystalline arrays within the nuclei of infected cells. The virus was not able to infect amphibian, mammalian, or avian cell lines despite the amphibian origin of the isolate (Clark *et al.,* 1973). The complete genome sequence of FrAdV-1 was published in 2000 (Davison *et al*., 2000).

Based on sequence similarity, THEV and FrAdV-1 were placed in a new adenovirus genus, *Siadenovirus* (Davison *et al.,* 2000; Davison and Harrach, 2002). According to phylogenetic comparison with other adenoviruses, siadenoviruses are among the most primitive. However, there is no evidence to suggest THEV is capable of infecting amphibian hosts. The lack of understanding of the natural host range of FrAdV-1 limits comparison of pathology between the two viruses.

Members of the genus *Siadenovirus* have several genus-specific genes that share no sequence similarity with any proteins studied in other organisms. The name of the genus *Siadenovirus* is derived from the genus-specific open reading frame (ORF) in the E1 region whose putative gene product has high sequence similarity with bacterial sialidases (Davison *et al.,* 2000). It is not known what functional importance these genes may have. Very little is known about the replication of FrAdV-1 *in vivo* and, although THEV has been studied extensively for decades, detailed information about the molecular mechanisms of replication and pathogenesis is still lacking. An understanding of the role of genus-specific genes in viral replication and pathogenesis is essential in order to determine mechanisms by which siadenoviruses cause disease in their respective hosts. The purpose of this study was to determine putative genetic and structural elements common to all members of the genus *Siadenovirus*, and determine differences that may contribute to host specificity.

5.3 Materials and Methods

Sequences

Complete sequences of the Virginia Avirulent Strain (VAS) and FrAdV-1 were compared for the determination of genetic characteristics of the genus *Siadenovirus.* The VAS was chosen as the THEV representative because it most consistently shared the sequence of the other isolates sequenced, and is expected to be more representative of THEV strains present in North America (Chapter 3; GenBank accession # AY849321). FrAdV-1 from the ATCC collection was sequenced in 2000 (Davison *et al*., 2000; GenBank accession # AF224336). Sequence alignment and ORF prediction was carried out by BioEdit version 5.0.9 (Hall, 1999). Similarity of putative ORFs was calculated using BioEdit optimal global pair-wise alignment, with the PAM250 similarity matrix.

5.4 Results/Discussion

The genomic organization of THEV and FrAdV-1 was nearly identical (Table 1). The G+C% of FrAdV-1 (37.9%) was slightly higher than THEV (34.9%), though both were much lower than most other adenoviruses. The inverted terminal repeat regions were also similar in length, and the extreme ends had 90% nucleotide identity (Figure 1). Only 36-40 nt in length, the ITRs of the siadenoviruses are the shortest of any adenovirus yet sequenced. By comparison, atadenoviruses and aviadenoviruses have ITRs of approximately 50 nt, and mastadenoviruses have ITRs over 100 nt in length (Davison *et al.,* 2003). A consensus AdPOL-pTP binding site is located within the extreme 7-18 nt. This binding site is the origin of replication of the viral genome during infection, and similarity in the sequence indicates common structure and function of the siadenoviral AdPOL and pTP proteins with the rest of the adenoviruses (Liu *et al*., 2003).

The reported length of FrAdV-1 is 26163 bp, which includes a direct repeat of a partial ORF7 near the right terminus (Davison *et al*., 2000). The total length of the VAS

Table 5-1: Genetic Features of the Members of the Genus *Siadenovirus*

Feature	Nucleotide position		strand	% Identity*	% Similarity*	Protein Size (aa)	
	THEV-VAS	FrAdV-1				THEV	FrAdV-1
ITR	1-40	1-36		55	-	-	-
CAAT Box	222-226	159-164	-	-	-	-	-
E1 TATA	261-267	207-213	+	-	-	-	-
E1 Initiator	284-290	231-237	+	-	-	-	-
ORF1	313-1953	323-2029	+	32	57	546	568
hyd	2098-2325	2029-2232	+	18	40	75	67
E1 AATAAA	2321-2326	2258-2263	+	-	-	-	-
E2B AATAAA	2330-2335	2250-2255	-	-	-	-	-
IVa2	2334-3437	2270-3373	-	53	77	367	367
AdPOL	3430-6768	3366-6692	-	50	70	1112	1108
CAAT Box	4122-4126	4052-4056	-	-	-	-	-
MLP TATA	4173-4180	4103-4110	+	-	-	-	-
MLP Initiator	4197-4203	4127-4133	+	-	-	-	-
pTP	6765-10995	6692-10941	-	45	67	597	609
52K	8570-9472	8578-9471	+	41	70	300	297
pIIIa	9462-10979	9461-10915	+	42	70	505	484
L1 AATAAA	11495-11500	11286-11291	+	-	-	-	-
Penton	11001-12347	10947-12284	+	59	80	448	445
pVII	12347-12709	12281-12730	+	34	54	120	149
pX	12712-12888	12762-12935	+	52	84	58	57
pVI	12906-13601	12953-13606	+	45	72	231	217
L2 AATAAA	13733-13738	13737-13742	+	-	-	-	-
hexon	13610-16330	13614-16376	+	66	83	906	920
protease	16332-16940	16376-16990	+	55	71	214	204
L3 AATAAA	16901-16906	16989-16994	+	-	-	-	-
L3 AATAAA	16919-16925	17002-17007	+	-	-	-	-
E2A AATAAA	16947-16952	-	-	-	-	-	-
E2A AATAAA	16961-16966	16986-16991	-	-	-	-	-
DBP	16973-18186	17028-18245	-	47	75	380	378
100K	18230-20227	18293-20188	+	40	63	665	631

*Similarity of putative ORFs calculated using BioEdit pair-wise alignment, optimal global alignment, PAM250 similarity matrix.

Table 5-1 Continued: Genetic Features of the Members of the Genus *Siadenovirus*

Feature	Nucleotide position		strand	%	%	Protein Size (aa)	
	THEV-VAS	FrAdV-1		Identity*	Similarity*	THEV	FrAdV-1
33K	20142-20699	20085-20695	+	30	53	120	129
22K	20142-20411	20085-20408	+	35	55	89	107
pVIII	20769-21371	20739-21320	+	37	59	200	193
L4 AATAAA	21511-21516	21969-21974	+	-	-	-	-
CAAT Box	20969-20973	20998-21002	-	-	-	-	-
CAAT Box	20979-20983	21008-21012	-	-	-	-	-
E3 TATA1	21014-21020	21053-21059	+	-	-	-	-
CAAT Box	21032-21036	21068-21072	-	-	-	-	-
E3 Initiator 1	21038-21044	21077-21083	+	-	-	-	-
E3 TATA2	21069-21076	21107-21113	+	-	-	-	-
E3 Initiator 2	21096-21102	21131-21137	+	-	-	-	-
E3	21247-22116	21211-22128	+	29	53	289	305
E3 AATAAA	-	22170-22175	+	-	-	-	-
E3 AATAAA	22239-22244	22190-22195	+	-	-	-	-
U exon	22263-22520	22111-22331	-	48	65	86	74
fiber	22519-23883	22343-23632	+	26	59	454	429
L5 AATAAA	23887-23892	23628-23633	+	-	-	-	-
THEV unique L5 ORF	23899-24099	-	+	-	-	67	-
CAAT Box	24429-24433	-	-	-	-	-	-
CAAT Box	24434-24438	-	-	-	-	-	-
CAAT Box	-	23629-23633	+	-	-	-	-
E5 TATA	24479-24485	23674-23680	+	-	-	-	-
E5 Initiator	24505-24511	23700-23706	+	-	-	-	-
ORF7	24512-25168	23711-24274	+	19	45	218	187
E4 AATAAA	-	24311-24316	-	-	-	-	-
E5 AATAAA	25164-25169	24327-24332	+	-	-	-	-
E4 AATAAA	25200-25205	24360-24365	-	-	-	-	-
ORF8	25204-25701	24365-24919	-	27	49	165	184
E4 Initiator	25778-25785	25014-25021	-	-	-	-	-
E4 TATA	25810-25816	25047-25054	-	-	-	-	-
ITR	26227-26266	26128-26163		55	-	-	-

*Similarity of putative ORFs calculated using BioEdit pair-wise alignment, optimal global alignment, PAM250 similarity matrix.

Figure 5-1: Comparison of Sequences at the Genome Termini

The extreme termini of adenoviruses contain the binding site for the pTP/AdPOL protein hetero-dimer. The consensus sequence was roughly conserved in FrAdV-1 and THEV, and the first 28 bp of their ITRs are 90% identical.

was 26266 bp. Two significant differences in the two viral genomes were found in non-coding regions. An addition of approximately 150 nucleotides was found in THEV, between the E3 and fiber coding regions. An addition of approximately 540 nucleotides was discovered between the fiber coding region and the putative E5 promoter region. There is a single 67 aa ORF located within this region, though it shares no sequence in common with any other proteins to date. If this is a true coding region, it would be the only gene unique to THEV.

The locations of the E1, E3, E4, E5, and major late promoters in siadenoviruses have not been previously described, and were determined based on comparison of THEV and FrAdV-1 sequences (Table 2). However, the relative position and conservation of nucleotides at these locations strongly indicate the presence of a functional importance. Each promoter is composed of three elements: a TATA signal, an initiator site, and an upstream regulatory sequence. The TATA is defined as a T/A rich region at which the TATA-binding protein (TBP) binds to initiate formation of the basal transcription complex (BTC). The BTC is made up of several proteins that interact with cellular RNA polymerase II and cause it to bind to the initiator (INR) sequence (Hahn, 2004). The consensus INR sequence is YYANWYY, in which Y can be thymine or cytosine, N can be any nucleotide, and W can be a thymine or adenine. The upstream regulatory sequence that was searched for was the inverted CCAAT box (Young, 2003). This was found in most of the siadenoviral promoters, though it is not known for certain that this sequence is of any functional value in the avian and amphibian cells the viruses infect.

The MLP is the site of transcriptional regulation of all of the late genes. It was located in the siadenoviruses in the same relative place as all of the other adenoviruses: between the Pol I and Pol III domains of the AdPOL gene on the (+) strand (Young, 2003)(Figure 2). The MLP TATA signal was well conserved in the two viruses. The E1 promoter is expected to regulate transcription of the E1 genes ORF1 and hyd. It is located on the plus strand, near the left terminus. The TATA signal was very similar in both viruses, and the other elements were homologous with the consensus. Two discreet sets

Table 5-2: Comparison of Promoter Elements

Promoter	Element	THEV	FrAdV-1
E1	CCAAT	CCAAT	CCAAAT
E1	TATA	TGTAATT	TGTTATT
E1	Initiator	TGACCTT	ACACTTT
MLP	CCAAT	ACAACT	CCAAT
MLP	TATA	TGTAAATA	TGTATAAG
MLP	Initiator	TGACGTC	CGACATC
E3-1	CCAAT	GCAAT	CCAAT
E3-1	CCAAT	GCAAT	GCAAT
E3-1	TATA	TATAGTT	TTTACAA
E3-1	Initiator	TGACTTG	TGAATCT
E3-2	CCAAT	TCAAT	CCAAG
E3-2	TATA	TCTAATCT	TTTAACT
E3-2	Initiator	GCACTTG	TGACTTT
E5	CCAAT	GCAAT	-
E5	CCAAT	TCAAT	GCAAT*
E5	TATA	TATAAAA	TATAAAT
E5	Initiator	TTACAAT	CGACACT
E4	TATA	TAATACT	TAACGTTT
E4	Initiator	TCATATTT	TCATATTT

*non-inverted GCAAT sequence was only one found

- Conserved nucleotides are underlined

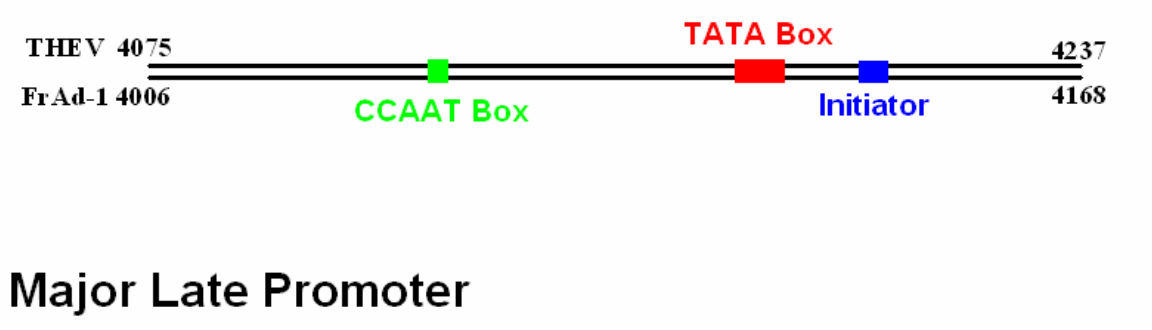

Figure 5-2: Spatial Layout of the Major Late Promoter

of promoter elements were discovered just upstream of the E3 coding region (Figure 3). It is not known if both are functional, or if the redundancy indicates particular functional importance of the E3 protein. The TATA signal and initiator of the second E3 promoter was more highly conserved between the two viruses. This may suggest a host-species determinant binds to the first initiator, and a more common transcription factor binds to the second. The E5 TATA signal was the closest to the consensus TATA of all of the predicted promoters. There did not seem to be conservation of the upstream sequences, which was not surprising considering the 540 nt of extra DNA between the fiber gene and the E5 TATA of THEV. The E4 promoter was located in the region between the right terminal repeat and ORF8 on the minus strand. The initiator sequence was identical in both viruses, though the TATA signal was not highly conserved. There were no sequences that were found in the E4 promoter that resembled CCAAT signals. The differences found in the promoter sites of the two viruses are expected to play a role in host species determination.

Candidate poly-adenylation signals (AATAAA) were also found for each of the coding regions of the siadenoviruses (Edmonds, 2002). Most of the signals were found to be highly conserved and in the same relative position in FrAdV-1 and THEV, with a few exceptions. The conservation of most of the polyA signals further indicates the similarity of transcription patterns of FrAdV-1 and THEV.

In addition to the genetic features of the viruses, the features of the putative viral proteins were compared. The genus-common genes were all assumed to have the same functionality. Any major differences in amino acid sequence were expected to result in conformational changes that directly affect host specificity. In the case of the genus-specific genes ORF1, hyd, E3, ORF7, and ORF8, common amino acid sequences were expected to reflect regions of functional importance within these proteins of unknown function.

Based on electron microscopy, the morphology and characteristic intra-nuclear inclusions of THEV and FrAdV-1 are remarkably similar. This observation was supported by the high level of sequence identity seen in the major structural proteins. The

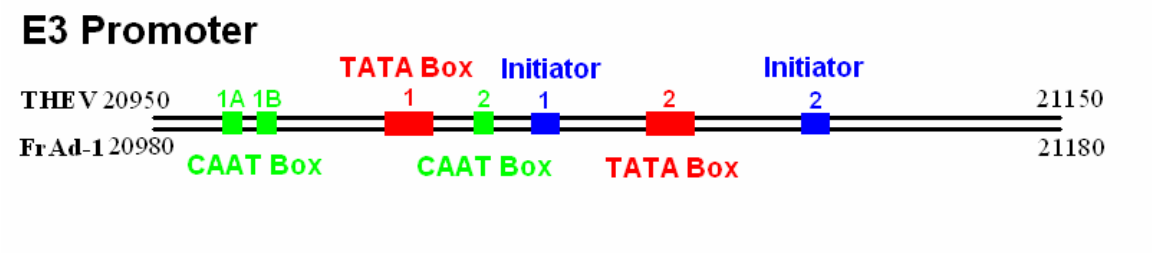

Figure 5-3: Spatial Layout of the E3 Promoter

hexon and penton proteins shared 66% and 59% amino acid identity, respectively. It is not known whether such structural similarity would be enough to result in serologic cross-reactivity. Interestingly, the fiber proteins of the two viruses were extremely different, especially in the knob domain.

Fiber

The adenovirus fiber is responsible for attachment of the virus to host cells via a specific receptor protein. The fiber structure is made up of a trimer of fiber proteins. Each fiber has three distinct domains: the tail, the shaft, and the knob. The N terminal tail is responsible for the non-covalent binding of the fiber to the penton protein situated at each vertex of the icosahedral capsid. This interaction takes place at a well conserved amino acid motif in all adenoviruses (Shenk, 1996).

The predicted structural features of the fiber protein of FrAdV-1 and THEV are shown in Figure 4. Each had a penton interaction domain that was very close to the consensus sequence. In addition, the N-terminus contained a lysine and arginine-rich nuclear localization signal that allows the translated fibers to be transported into the nucleus where viral assembly takes place. The shaft determines the length of the fiber, and is made up of several repeats of a beta-rich motif, each approximately 15 amino acids in length (van Raaij *et al*., 1999). Based on comparative analysis, the shafts of both viruses were predicted to have 16 repeats, and therefore have roughly the same length. Although there was very limited aa identity in the shaft, the residues located at key beta-turns and hydrophobic residues were conserved. The knob is responsible for recognition and binding of the host cell receptor (Magnusson *et al*., 2001; Stevenson *et al*., 1995; Roelvink *et al.,* 1999). THEV and FrAdV-1 shared only 15% aa identity in the fiber knob. This difference was not surprising considering that the two viruses infect two different classes of vertebrates. Despite the apparent differences in the knob sequence, there were conserved N-linked glycosylation signals (-N-X-T/S-) in the shaft and knob of both viruses (Arnberg *et al*., 1997).

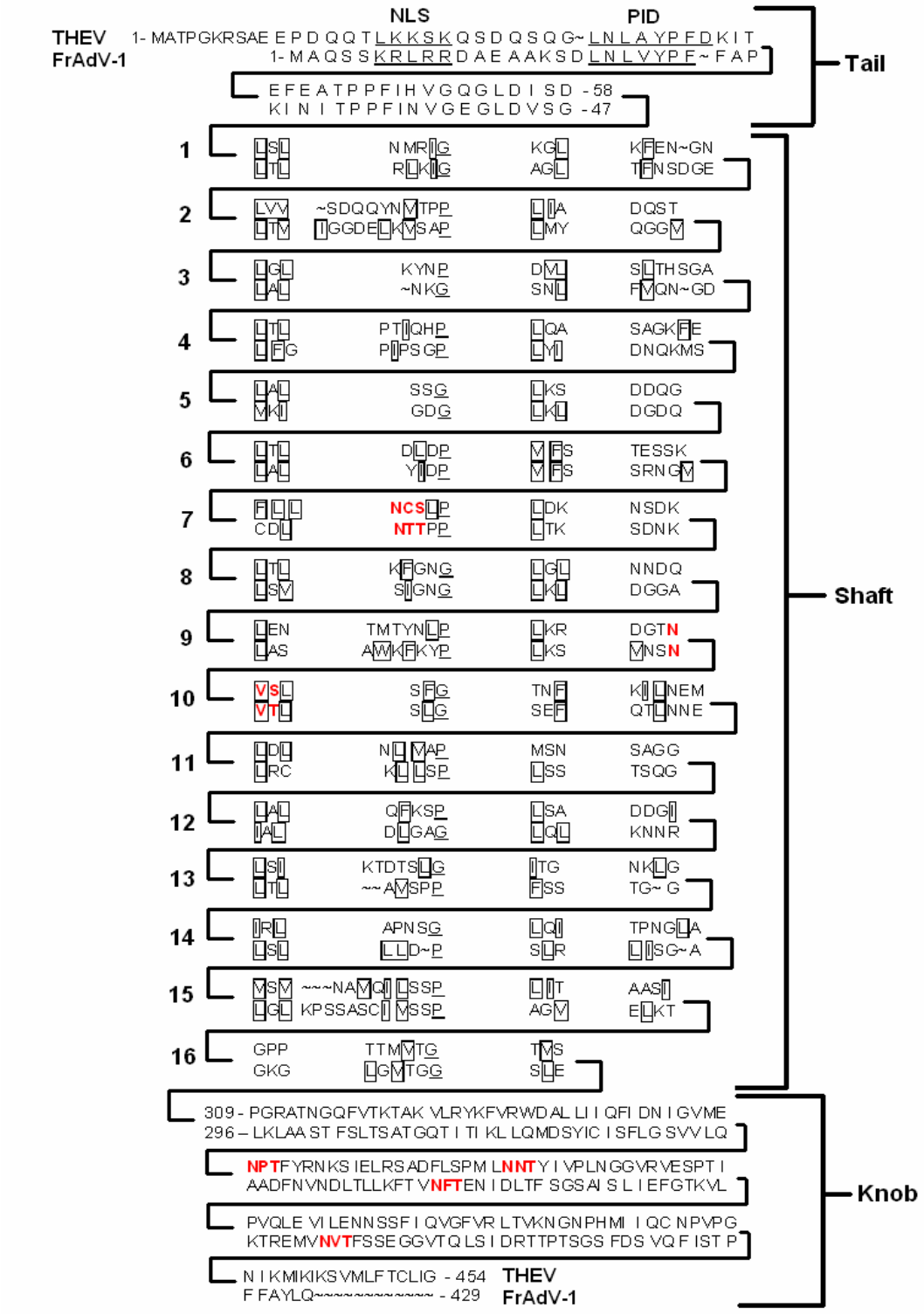

Figure 5-4: Features of the Fiber Protein

RED – putative N-linked glycosylation sites; BOX – hydrophobic amino acids in shaft

P or G – amino acid located at beta-bend of each shaft repeat

E3

Only mastadenoviruses and siadenoviruses encode genes in the region between the pVIII and fiber ORFs, known as E3 (Davison *et al*., 2003). In mastadenoviruses, the E3 genes encode proteins that aid the virus in the avoidance of the host immune response. The E3 19kDa protein results in decreased MHC-I antigens on the surface of infected cells, which prevents efficient antigen presentation to T-lymphocytes. E3 deletion mutants caused increased lymphocyte and macrophage activation and inflammatory response to infection (Ginsberg *et al*., 1989). However, the single ORF found within the E3 region of the siadenoviruses shared no sequence homology with any of the genes found within the mastadenovirus E3 region. There is no apparent similarity of siadenovirus E3 to any protein studied in other organisms, preventing prediction of the structure and function of its gene product. One key feature of the putative E3 gene product is a high number of cysteine residues that were conserved in both FrAdV-1 and THEV (Figure 5). A conserved cysteine-rich motif was repeated three times in the C-terminal half of siadenoviral E3 (CX_2CX_6CXC). This motif is very similar to one found in the GAL4 protein of yeast. GAL4 is a DNA binding transcriptional activator that has a very similar set of conserved cysteines (Marmorstein *et al*., 1992). The E3 cysteine rich regions also resemble the zinc-binding CR3 domain of the mastadenovirus E1A 13S protein. E1A 13S binds to the TATA-binding protein and other cellular transcription factors, and activates transcription of viral genes (Bruder and Hearing, 1989; Culp *et al*., 1988; Geisberg *et al*., 1994, 1995; Parker *et al*., 1997). E1A 13S activates the TNF gene in inflammatory cells, and this activation is dependent on this CR3 domain (Metcalf, 1996).

A central region of the THEV E3 gene was found that had no similarity with the E3 gene of FrAdV-1. In FrAdV-1 there was a glutamate-rich span of 31 amino acids, and in THEV there was a proline-rich span of 15 amino acids. This may be an indication of a type-specific region of importance, or rather a non-essential region linking two domains of functional importance.

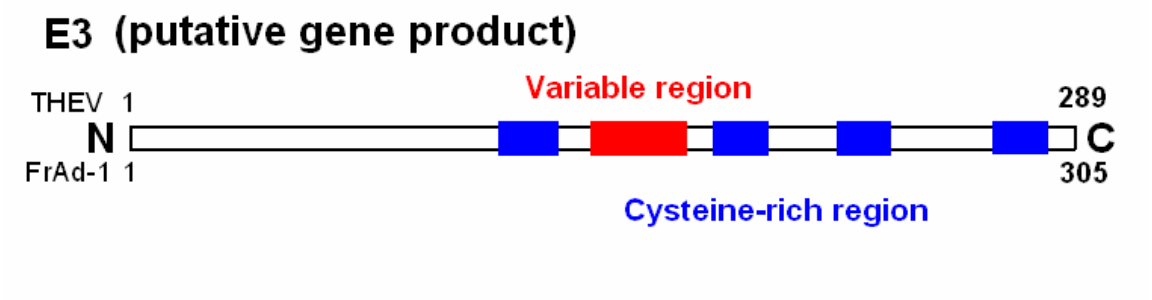

Figure 5-5: Features of the Putative E3 Protein

E3 is a putative *Siadenovirus*-specific protein with unknown function. There were four regions found with several conserved cysteine residues. A highly variable central domain was discovered that may be partially responsible for host species tropism.

ORF1

ORF 1 has been called the sialidase protein because it has a region with a high level of similarity to bacterial sialidase proteins. ORF1 is considered to be genus specific for the siadenoviruses, though its function is not known (Davison *et al*., 2000; Davison and Harrach, 2002). It is not known whether the conserved sialidase domain is of functional importance.

There seemed to be several semi-conserved proline residues in the N terminus of each virus ORF1, which had 66% identity over the first 15 aa (Figure 6). There was an additional domain of similarity between the viruses located in the C terminal half of the protein. There was a span of 10 amino acids in this region with 80% identity between the two viruses.

Hyd, ORF7, and ORF8

The least amount of sequence identity was found in the hyd, ORF7, and ORF8 putative gene products. Especially in the hyd and ORF7 proteins, there were no significant regions of amino acid homology that would indicate specific functionality. The hydrophobic and ORF7 proteins of both viruses have a large number of leucine residues, but the spacing is not consistent enough to indicate a motif. The ORF7 proteins of both viruses have several proline residues whose spacing could indicate structural features. There are two regions of distinct similarity in the ORF8 proteins (Figure 7). In a central region, there was a span of seven aa that was perfectly conserved: KLSEYNF. In a second region of homology near the C terminus, several conserved proline, cysteine, and hydrophobic residues were found in each virus. The functional value of these domains cannot yet be ascertained.

This was the first in-depth comparison of FrAdV-1 and THEV for the prediction of genetic features common to all of the siadenoviruses. The discovery of additional genus members will be of great importance so that comparative analysis can be extended to include them. Experimental confirmation of the predicted promoter sites and gene expression is planned in future studies of THEV.

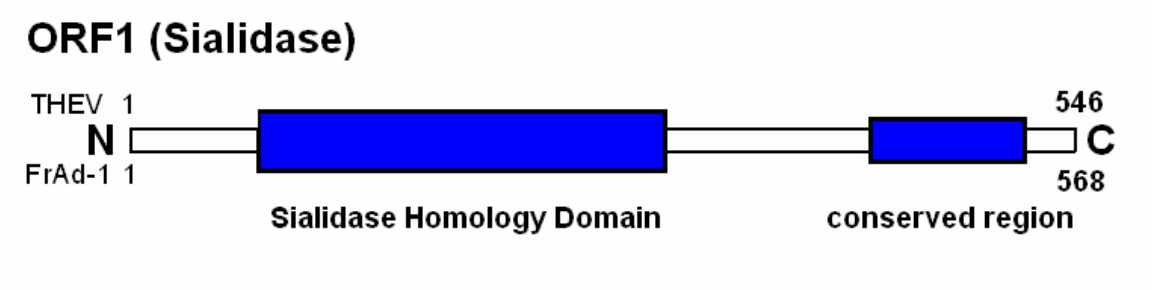

Figure 5-6: Features of the Putative ORF1 Protein

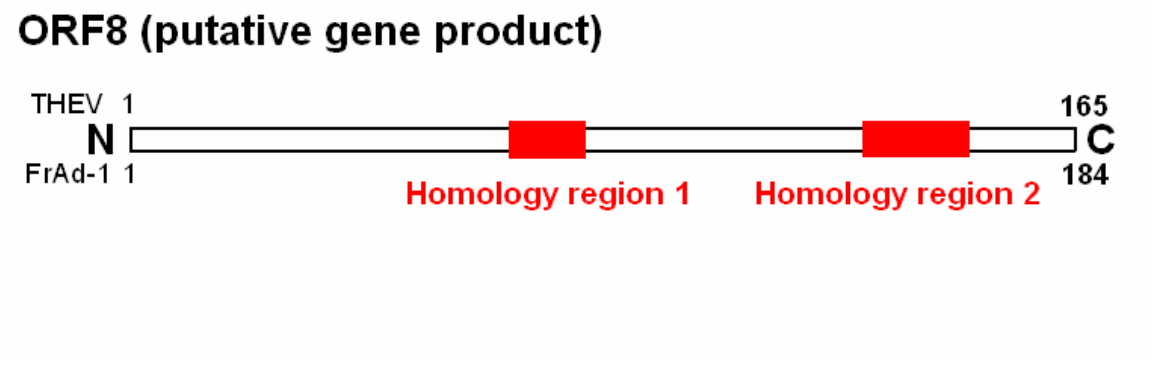

Figure 5-7: Features of the Putative ORF8 Protein

5.5 References

Arnberg, N., Mei, Y., & Wadell, G. (1997). Fiber genes of adenoviruses with tropism for the eye and genital tract. *Virology* 227, 239-244.

Bruder, J.T., and Hearing, P. (1989). Nuclear factor ef-1a binds to the adenovirus E1A core enhancer element and to other transcriptional control regions. *Mol Cell Biol* 9, 5143-5153.

Clark, H.F., Michalski, F., Tweedell, K.S., Yohn, D., and Ziegel, R.F. (1973). An adenovirus, fav-1, isolated from the kidney of a frog (rana pipiens). *Virology* 51, 392-400.

Culp, J.S., Webster, L.C., Friedman, D.J., Cmith, C.L., Huang, W., Wu, F.Y., Rosenberg, M., and Ricciardi, R.P. (1988). The 289-amino acid E1A protein of adenovirus binds zinc in a region that is important for trans-activation. *Proc Natl Acad Sci* 85, 6450-6454.

Davison, A.J., Wright, K.M., and Harrach, B. (2000). DNA sequence of frog adenovirus 1. *J Gen Virol* 81, 2431-2439.

Davison, A.J., and Harrach, B. (2002). Siadenovirus. In *The Springer Index of Viruses,* pp.29-33. Edited by C.A. Tidona and G. Darai. New York: Springer-Verlag.

Davison, A.J., Benko, M., and Harrach, B. (2003). Genetic content and evolution of adenoviruses. *J Gen Virol* 84, 2895-2908.

Edmonds, M. (2002). A history of poly-A sequences: from formation to factors to function. *Prog Nucl Acid Res Mol Biol* 71, 285-389.

Geisberg, J.V., Lee, W.S., Berk, A.J., and Ricciardi, R.P. (1994). The zinc finger region of the adenovirus E1A transactivating domain complexes with the TATA box binding protein. *Proc Natl Acad Sci* 91, 2488-2492.

Geisberg, J.V., Chen, J.L., Ricciardi, R.P. (1995). Subregions of the adenovirus E1A transaction domain target multiple components of the TF-IID complex. *Mol And Cell Biol* 15, 6283-6290.

Ginsberg, H.S., Lundholm-Beauchamp, U., Horswood, R.L., Pernis, B., Wold, W.S.M., Chanock, R.M., and Prince, G.A. (1989). Role of early region 3 (E3) in pathogenesis of adenovirus disease. *Proc Natl Acad Sci* 86, 3823-3827.

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nature* 11, 394-403.

Hall, T.A. (1999). Bioedit: a user-friendly biological sequence alignment editor and analysis (program for Windows 95/98/NT). *Nucl Acids Symp Ser* 41, 95-98.

Liu, H., Naismith, J.H., Hay, R.T. (2003). Adenovirus DNA replication. *Curr Top Microbiol Immunol* 272, 131-164.

Magnusson, M.K., Hong, S.S., Boulanger, P., and Lindholm, L. (2001). Genetic retargeting of adenovirus: novel strategy employing "deknobbing" of the fiber. *J Virol* 75, 7280-7289.

Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S.C. (1992). DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* 356, 408-414.

Metcalf, J.P. (1996). Adenovirus E1A 13S gene product upregulates tumor necrosis factor gene. *Am J Physiol* 270, L535-L540.

Parker, S.F., Felzien, L.K., Perkins, N.D., Imperiale, M.J., Nabel, G.J. (1997). Distinct domains of adenovirus E1A interact with specific cellular factors to differentially modulate human immunodeficiency virus transcription. *J Virol* 71, 2004-2012.

Pierson, F.W., and Fitzgerald, S.D. (2003). Hemorrhagic enteritis and related infections. In *Diseases of Poultry,* 11th edn, pp. 237-247. Edited by Y.M. Saif. , H.J. Barnes, J.R. Glisson, A.M. Fadly. L.R. McDougald, and D.E. Swayne. Ames: Iowa State University Press.

Roelvink, P.W., Lee, G.M., Einfeld, D.A., Kovesdi, I., and Wickham, T.J. (1999). Identification of a conserved receptor-binding site on the fiber proteins of car-recognizing adenoviridae. *Science* 286, 1568-1571.

Shenk, T. (1996). Adenoviridae: the viruses and their replication. In *Fields Virology,* 3rd edn, pp. 2111-2148. Edited by Fields, B.N., Knipe, D.M., and Howley, P.M. Philadelphia: Lippincott – Raven.

Stevenson, S.C., Rollence, M., White, B., Weaver, L., and McClelland A. (1995). Human adenovirus serotypes 3 and 5 bind to two different cellular receptors via the fiber head domain. *J Virol* 69, 2850-2857.

van Raaij, M.J., Mitraki, A., Lavigne, G., Cusack, S. (1999). A triple B-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. *Nature* 401, 935-938.

Young, C.S.H. (2003). Structure and function of the adenovirus major late promoter. *Curr Top Microbiol Immunol* 272, 213-249.