# Design of Recombinant HIV Vaccine by Orthoreovirus 5' Duplication

A Major Qualifying Project:
Submitted to the Faculty of the
WORCESTER POLYTEHCNIC INSTITUTE
in partial fulfillment of the requirements for the
Degree of Bachelor of Science in Biochemistry
and
Degree of Bachelor of Science in Biology
by:

*Cameron W. Habib*

Thursday, April 29, 2010

Approved by:

*Professor Destin W. Heilman, Advisor*
*Department of Chemistry and Biochemistry*
*Worcester Polytechnic Institute*

*Dr. Max L. Nibert, Sponsor*
*Microbiology and Molecular Genetics*
*Harvard Medical School*

## 0.1 – Abstract

Human Immunodeficiency Virus (HIV) is a deadly and incurable illness impacting millions annually. Targeting the T-helper CD4+ cells of the immune system, HIV is a complex enveloped pseudo-diploid RNA virus. Of the nineteen proteins encoded, one of the most readily produced and unique is the p24 inner capsid protein. Current research suggests that gut-associated lymphoid tissue (GALT) may be one of the primary HIV reservoirs in the body. Due to the immense difficulty associated with a cure for the disease, a preventative vaccine remains the best possible defense. Orthoreovirus, a virus causing mild to no pathological conditions, is known to target the GI system. Recent advances in the field of Orthoreovirus reverse genetics allow for the possibility of this virus to be used in the creation of an HIV vaccine. We provide support that based on the segmented nature of the Orthoreovirus genome, 5' duplication of the segments ORF is the best possibility for insertion of the p24 coding sequence for use as a vaccine antigen. We show there exists a maximum practical length for the length of code which can be added via the 5' duplication method, over which, recombination to original wild type occurs. This limit is an important consideration for future work and design with the Orthoreovirus.

## 0.2 – Acknowledgments

I would like to thank Dr. Max Nibert of Harvard Medical School for the gracious use of his lab, the invaluable experience I have received, and all that I have learned throughout my time there. Dr. Oleksandr Demidenko of the Nibert lab, for his patience, endless help, and guidance throughout the project. My advisor, Dr. Destin Heilman of WPI, for his dedication and commitment to the success of the project and for his understanding and assistance throughout. Finally, I would like to thank all those who have helped in any capacity, lending their time and their effort for simple thanks. It has been an honor and a privilege for which I am deeply thankful and humbled to have been a part of.

## 0.3 - Table of Contents

# 1.0 - Introduction

## 1.1 – Human Immunodeficiency Virus

Human Immunodeficiency Virus, otherwise known as HIV, is the virus responsible for Acquired Immunodeficiency Syndrome, or AIDS[1]. The cause of millions of deaths each year[2], AIDS is a potent and deadly disease that currently has no known preventative treatment or cure. HIV is a single stranded positive-sense pseudo-diploid RNA virus[3], placing it within Class VI of the Baltimore Classification System. Due to the enveloped structure and long incubation periods, it is further classified within Class VI in the Retrovirus family and Lentivirus genus.

There are currently two known species of HIV, classified as HIV-1 and HIV-2[4]. The virus is believed to be the result of at least 11 zoonotic infections, arising from the transfer of simian immunodeficiency virus (SIV) in chimpanzees to humans[5]. A study of genomic markers places the first SIV infection in chimpanzees sometime around the year 1492 and the migration to humans by 1963[6]. While HIV-2 is highly genetically similar to SIV, HIV-1 is both more virulent as well as more readily transferred. HIV-1 is classified into three groups, namely group M (main), O (other), and N (non-M/non-O), varying in slight genetic modifications, mainly in the capsid proteins. This variation is believed to have been, and continues to be, the result of viral escape mechanisms[7]. Group M is further broken down into 9 subgroups, again with slight variations. Further groups continue to arise as the result of genetic combination of two or more varying strains.

### 1.1.1 – Genome and Structure

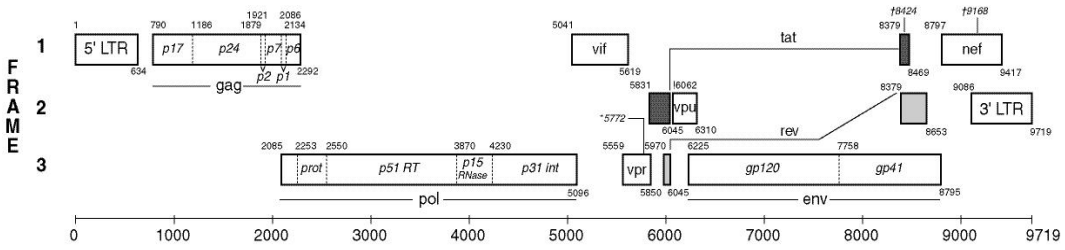


**Figure: 1 – HIV HXB2 Genome. Depiction of genome, displaying three separate reading frames. Frame 1 containing *gag*, encoding for p17 and p14. Frame 3 containing *tat*. Frame 3 containing *pol* and *env*. Source: Los Alamos National Laboratory.**

Several examples of complete HIV-1 genomes have been sequenced, typically around 9700bp in length (Fig. 1)[8]. The length varies slightly between strains due to mutations encountered during replication. Contained within the genome across three various reading frames are nine genes encoding for a total of nineteen proteins. Each gene codes for at least one protein, some coding for up to four, related in function, form or location. The genes referred to as group specific antigen, polymerase, envelope, transactivators, and regulators.

*Group specific antigen* (*gag*) is a single gene responsible for the coding of four proteins, p6, p7, p17, and p24, all related to internal structural or core functions to the virus. Both p6 and p7 are nucleocapsids used to protect the RNA from degradation by way of nucleases, structured in a ratio of one molecule RNA per hexemer of protein[9]. p17 and p24 are both structural proteins, p24 being the main component of the inner capsid, and p17 used as the anchor to attach the inner capsid to the outer envelope as well as the proteins found therein.

*Polymerase* (*pol*) gene encodes three main proteins responsible for the reverse transcriptase, protease, and integrases enzymes. Reverse transcriptase is the enzyme by which RNA is transcribed into DNA. In the case of Retrovirus, this is used in order to produce the material to insert the viral RNA genome into the host cell DNA genome. The enzyme that performs this task

is integrase. Protease is the enzyme responsible for the cleavage of the fusion protein at specific predetermined locations. Without the cleavage performed by protease, the proteins are never separated, thus creating an inactive virion.

*Envelope* (*env*) encodes for the exterior viral proteins. Making use of a host cell enzyme Furin, the fusion protein gp160 coded for my *env*, is cleaved into two smaller subunits, gp120 and gp41[10]. gp41 and gp120 present on the surface of the matured virion in the form of a 3:3 complex. A gp41 trimer is directly attached to the exterior of the virion with the trimer gp120 affixed onto the end of it. The role of gp41 is that of mechanical cellular entry into CD4+ cells[11], as recognized by the gp120 complex[12].

The purpose of a transactivator is to selectively increase or decrease the rate of gene expression. The three produced by HIV are transactivator of transcription (*tat*), regulator of virion (*rev*), and viral protein-r (*vpr*). *Tat* works by enhancing the phosphorylation of cellular factors, thereby increasing their expression[13]. This allows for an explosive increase in copy count, proving difficult for the body to overcome. *Rev* is believed to assist with the exportation of RNA from the cell nucleus to the ctytoplasm, allowing for the increased production of RNA. Contrary to *rev*, *vpr* may potentially work by accelerating the import of viral precursor into the nucleus, to aid in RNA production.

Three other miscellaneous regulatory factors are also present within HIV, viral infectivity factor (*vif*), negative regulatory factor (*nef*), and viral protein-u (*vpu*). *Vif* works to interrupt a main human antiviral protein, APOBEC, by targeting it for degradation by the bodies own immune system. *Nef* decreases the expression of MHC surface protein signals of the host cell, making it less likely to be marked for degradation by the immune system, and thereby increasing the

chances of virion survival. *Vpu* is involved in one of the final step of virion production, the budding of the virion from the host cell.

The mature HIV virion is spherical in shape[14] and roughly 120nm in diameter (Fig. 2)[15]. It contains both an outer and inner membrane. The inner membrane, conical in shape, is coded for by the gag protein p24 and contains the bulk of the encoded viral material, including RNA, reverse transcriptase, protease, and other accessory proteins. The RNA, in its mature packaged form, is tightly bound to p6 and p7 proteins[16]. Immediately surrounding the p24 capsid is a membrane of p17 proteins, which are used to help anchor the outer membrane and its respective proteins to the inner capsid. The viral envelope, gained as a result of the virion buding off from the host cell, is largely composed of host cell material. Inserted into the outer envelope is a series of gp120/gp41 complexes, used to bind to CD4 as well as gain entry into the cell.

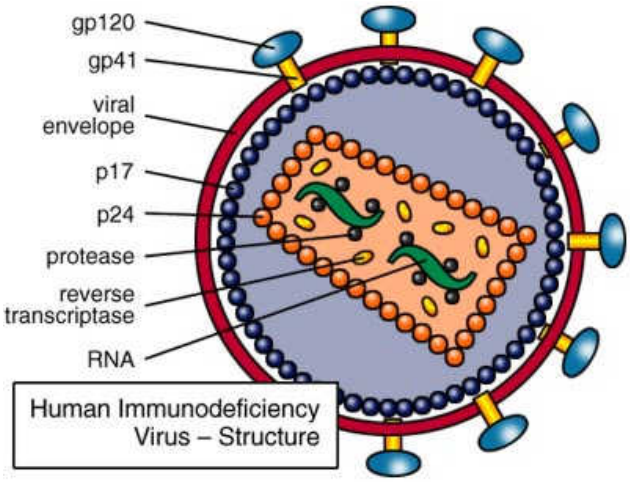


**Figure 2 - HIV Virion Structure. Representation of HIV structure. Gp120/gp41 complex on exterior, p17 matrix protein, p24 capsid protein, and RNS bound with necessary regulatory factors. Source: San Diego State University**

### 1.1.2 – Lifecycle

HIV primarily infects T-cells and macrophages of the human immune system (Fig. 3)[17]. The process is initiated by the infection of a human with a mature HIV virion. Once a mature virion has been transferred, it must come into contact with a T-helper cell. The gp120 protein located on the surface of the virion has a high affinity for the CD4 antigen present on the surface of the target cells, allowing for a strong bond to be formed. This bond is then further stabilized by

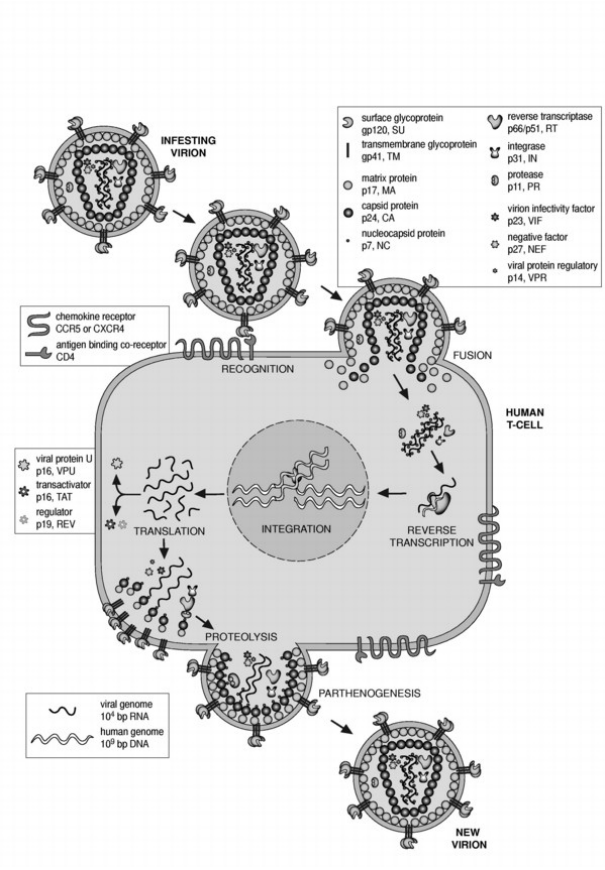


**Figure: 2 - HIV Lifecycle. Figure depicting the lifecycle and replication method of HIV. HIV attachment by mediation of CD4 receptor on T-helper cells, followed by gencome integration, replication, and export of virion. Source: Molmo.**

recruited chemokine receptors on the surface of the host cell, most often CCR5. Once these accessory proteins are moved into the vicinity of both the CD4 and the gp120/gp41 complex, the bond strength and affinity are increased, causing a conformational change of the gp120 to takes place. This change forces the gp41 protein into the membrane of the host causing a fusion of the two membranes, allowing for inner capsid of the virion to be released into the cytoplasm of the host cell.

Due to the fact that the HIV genome is composed of RNA and must be inserted into the DNA host cell genome, once the viral material has entered the host cell, the first step of virion production is reverse transcription. This process, performed by reverse transcriptase, is highly error prone due to the lack of correction as a result of the single stranded nature of the HIV genome. This leads to numerous and frequent mutations within the HIV genome. These mutations, while often times detrimental to the survival of the virus, can at lead to viral escape. Mutations, typically in protein structure, grant a lack of recognition to existing enzymes and pathways designed to recognize and act on the previous protein structure. These select mutations are further amplified and expanded by ways of natural selection; those which are most fit to survive will by probability, be the most likely to survive.

The virus itself lacks many of the pathways and enzymes necessary for replication. As such, once transcription from RNA to DNA takes place, the DNA copy of the viral genome must be inserted into the host cells DNA genome for use of the host cells natural pathways. Integrase facilities this action. Although integration of the genome allows for the possibility of virion production, it does not necessarily indicate it. Several transcription factors, one of the most vital being NF-κB, must be present for replication to occur. The majority of these factors are present only during activation of the host cell in an immune response. This is one of the major leading causes of the varying and prolonged HIV latency periods. Because T-cells can be latent memory cells for long periods of time, the virus genome can remain stored within the cells, dormant, until the cells become activated due to an immune response, initiating the actions of both the host and the viral genome.

Once replication does occur, the initial resulting mRNA is cleaved within the cell nucleus and exported to the cytoplasm, where it is translated into *tat* and *rev*. Once the buildup of these proteins is sufficient, they being to bind to the mRNA particles within the nucleus prior to cleavage. These full-length particles are then exported from the nucleus where they are then used to translate the *gag* and *env* proteins used to build the virion itself. The buildup of *tat* and *rev* proteins continues in a positive feedback look until the maximum production rate of virion is achieved.

The final step of virion production is egression, where it will obtain its viral envelope as well as surface proteins. The process begins as gp160 is produced and cleaved by protease into gp41 and gp120, which are moved to, and subsequently anchored to, the membrane of the infected cell. The adhesion of gp120 and gp41 cause a complex with the viral genome and associated factors

and proteins, including p24, to form. The formation of this complex causes endocytosis and formation of the virion.

### 1.1.3 – Viral Transfer and Infection

HIV is known to be transmitted horizontally by the transfer of blood, semen, vaginal fluid, and breast milk[18]. Infection can also occur vertically, to a child during pregnancy by an HIV positive mother. This risk can be reduced with precautionary measures, most often times the use of antiviral drugs[19]. HIV is not known to be transmitted via airborne methods or by saliva, making unprotected intercourse the primary means of proliferation. Due to lack of awareness and testing, in the 80's and early 90's, blood transfer from HIV positive donors to HIV negative patients was another major cause of proliferation[20]. However, due to advancements in testing, primarily the presence of testing, and population HIV awareness, this risk has been sufficiently eliminated.

Emerging research suggests that CD4+ cells located in the gut-associated lymphoid tissue (GALT) may be a major reservoir of HIV in infected individuals[21-22]. This reservoir is present despite treatment of the individual to reduce blood pathogen level to undetectable amounts, suggesting that current methods of treatments may not be sufficient for the completion eradication of the virus. Levels of CD4+ cell were observed to be approximately 11% in GALT, while normally 40% in HIV positive patients, and 60% in healthy patients. Exchanges between this reservoir and the blood system have been observed, giving a possible cause to latency and prolonged infection periods.

## 1.2 – Orthoreovirus

Reoviridae is a family of non-enveloped double stranded RNA virus in Class III of the Baltimore Classification System. Within the family is a distinction between turreted and nonturreted, with Orthoreovirus being the only genus within the family to infect humans.

The Orthoreovirus genome is separated into a series of ten segments, placed into one of three size dependent groups, equimolar in quantity[23]. Each group encodes for a set of proteins, with small (S) encoding σ proteins, medium (M) encoding μ proteins, and long [L] encoding λ proteins. Segments are transcribed into full length mRNA strands for use in replication within the host cell and can undergo reassortment between virion and strands. The 5'-end of the genome is features a methylated nucleotide cap, important to the replication and proliferation of the virus. 3' ends also lacks polyadenylation. Replication takes place fully in the cytoplasm, indicating that the host cells replication enzymes are not necessary for viral replication.

Spherical in shape, virions are approximately 80nm in diameter, and contain both an inner and outer capsid. The inner capsid is composed of four structural proteins (λ1, λ3, μ2, and σ2) and is approximately 60nm is diameter. The outer capsid, also encoded for by 4 structural proteins (λ2, μ1, σ1, and σ3), is icosahedral in shape with a triangulation number of 13.

Orthoreovirus is the only geneses of Reovirus known to affect humans, causing gastrointestinal and upper respiratory tract infections in infant and young children[24], ranging from mild to asymptomatic[25]. This feature, as well as the common location of infection Orthoreovirus shares with HIV, allow for a wide range of possibilities in the use of a vaccine vector.

## 1.3 – Immunology and Vaccine Construction

Due to the deadly and debilitating nature of HIV and the lack of any known cure, the need arises for a vaccine to create a preventative measure. A vaccine is a step taken to prime the immune system, making it aware of a particular foreign substance. By increasing awareness for a particular toxin, the response time is reduced, and the toxin is eliminated before it is able to replicate and cause damage to the host.

### 1.3.1 – Immune Responses

Mammalian immune systems are composed of two parts, the innate and the adaptive system. The innate system is the omnipresent and continuously active network of precautions put in place by the body. These barriers include those created by cell walls and organs such as the skin, as well as basic responses such as inflammation. These responses do not have memory of previous toxins and cannot adapt in any way to newly presented toxins. The adaptive system, in contrast, does have the ability to learn and adjust to varying toxins, and is the main focus of vaccine delivery.

One of the main functional unit of the adaptive system is the antigen[26]. An antigen is a marker, typically the result of an enzyme digestion of a larger particle, that present on the surface of cells or as a free complex in tissue, to indicate self from non-self. Antigens present by way of a self-enzyme, the Major histocompatability complex (MHC), a multi-subunit integral protein located on the cellular membrane. This complex, upon recognition of a non-self or incorrect self antigen, signals T-cells to destroy the antigen by way of degradation. The MHC complex is distinguished into two classes: Class I that is typically used in response to an endogenous antigen produced interior to the cell, and Class II that is typically in response to an exogenous antigen located exterior to the cell. Class I cells signal CD8+ cytotoxic T-cells while Class II signal CD4+ T-

helper cells. T-helper cells are unique within the immune system in that they do not posses any ability to dispose of a toxin without the use of other immune cells; because of this, an MHC Class II response is generally considered weaker and less effective then a Class I response as it must yet propagate a further immune response. Cells which are CD8+ and CD4- are largely unaffected by HIV, and therefore, are targets of use in the fight against HIV.

### 1.3.2 – Vaccine Forms

Depending on the virulent in question, the chemical composition, structural form, the method by which it enters and replicates, as well as several other mitigating factors, one vaccine form may be chosen over another. The most commonly used methods include attenuated, inactive, subunit, and recombinant viral vaccines[27].

Attenuated vaccines are those that use a live but altered form of the organism. The alterations are selected or performed in such a way as to disable any infectious or pathogenic properties. This allows for the body to mount a response closest to that of an actual infection, producing the strongest and most reliable defense. Because this form of vaccine relies on live particles, there are certain risk involved. It is possible in select cases for the virulent to revert back to infectious form, causing the disease that it was intended to prevent. The vaccine must also be stored and produced in such a way as to insure the virions contained within remain alive. The death of the virions could potentially lead to a lack of immune response from the host, causing a lack of immunity.

Inactive vaccines use a once live organism that has been killed to the point of being unable to infect or reproduce. Even though the virulent cannot infect and reproduce, they are still identified as non-self and an immune response is mounted. As the case with attenuated, this form of

vaccine carries risks with it. If the virulent is not properly killed to a great enough extent, it is possible for the virulent to cause the disease instead of producing immunity to it.

Subunit vaccines are those which use a small part of the whole virulent in order to create an immune response. Because some virulent produce proteins unique to that virulent, it is possible to create an immunity to the virulent as whole by creating an immunity to a unique part of it. Doing so avoids the use of the virulent in its full virion form, which in turns reduces the number of risks, such as infection and disease. The subunit can also be chosen, to select for specific properties or functions. However, because an immunity is only being created to a specific part of a whole, that specific part must be present for the response to take place. This in turn creates an immunity less robust and reliable. There are also a large number of proteins which cannot survive transportation into the body, either by injection or through the gastroenterological system, making the discovery and selection of such a protein difficult.

Recombinant viral vaccine utilizes technologies from several different forms of vaccines, including attenuated and subunit vaccines. In a typical recombinant viral vaccine, a normally non-pathogenic virus, or one that has been altered to become such, is inserted with genetic code to produce a specific part of a virulent for which immunity is to be gained from. The host is infected with the recombinant virus which in turn produces the virulent protein, initiating an immune response, causing an immunity. Because the protein is produced in vivo and not in vitro, the protein does not have to be as robust or posses the ability to survive the transportation into the body. A recombinant virus can also be selected to cause an immune response or infection in a particular region of the host, allowing for selectivity in not only protein produced, but the location in which it is produced.

### 1.3.3 – HIV Vaccine

Due to their ease of creation and effectiveness, two of the most commonly used forms of vaccine include attenuated and inactivated. However, because HIV's viral envelope is, in essence, host material, producing a vaccine against this would cause an autoimmune response, whereby the body initiates an autoimmune response. Combined with the extremely virulent and robust nature of HIV, and the danger that the virus poses, the risks of creating an attenuated or inactivated vaccine for this virus are too high to be considered possible choices.

HIV does present two unique proteins on the surface of its viral envelope, in specific gp120 and gp41, making subunit vaccine a possibility. These proteins, however, provide a poor target for vaccine production[28]. Because both gp120 and gp41 are glycoprotein, the majority of their structures lack the rigid and defined structural characteristics necessary to produce an effective antibody against them. The parts that do contain the necessary structural integrity, on a mature virion, are highly restricted, disallowing to a great degree any possibility of detection and response by the immune system. Because of this, an internal protein must be selected for vaccine investigation.

One of the most readily produced and consistent proteins of HIV is the p24 protein encoded by *gag*. This protein has long been the target of HIV infection tests[29], due in part to the proteins early production and numerous count. p24 is an also a protein exclusive to HIV and absent from uninfected individuals; creating an immunity to this will not cause an autoimmunity condition in any way. Because this protein does not code for any viral properties of the virus, and because it is not currently known to cause any pathology when separated from the viral particle, few if any side reactions will occur by use of this protein in vaccine. An analogue of the HIV p24 protein

also exists in the form of the SIV p24 protein. The existence of an analogue aids in vaccine construction and testing.

Due to the harsh conditions of the digestive system, and the lack of ability for uptake of a protein in a fashion that would invoke an immunological response, a pill form containing such a protein would not be feasible. Injection into the blood stream would, likewise, mount a weak immunological response, leading to minor resistance to subsequent infection. Because of this, the protein must be produced in vivo. Production in vivo would allow for the highest possible chances of an MHC Class I response, allowing a much stronger viral infection resistance. The primary means for such a production is by recombinant viral vaccine. Such a vector would infect the host and produce the specified antigen, allowing the body to gain immunity to it.

As previously discussed, HIV is thought to resonate within GALT. As such, a vector that targets GALT, rather then an alternate tissue, would prove to be the most effect vector. Orthoreovirus is one such possible viral vector that is known to mainly localize to the gut region. This, combined with its relatively low pathogenic characteristics, makes it a possible candidate for vaccine use.

# 2 – Materials and Methods

## 2.1 – S3 Plasmid Construction

An S3 plasmid stock from Type 3 (Dearing) strain of mammalian Ortheovirus was obtained (courtesy of Takeshi Kobayashi). A SIV p24 protein coding sequence was also obtained via consortium resources.

### 2.1.1 – PCR

PCR method for design of construct was established using three part PCR (Table 1). First PCR, designated [L], had primer sequences (L1) annealing at location 3074, and (L2) annealing at location 678. Second PCR, designated [R], had primer sequences (R1) annealing at location 3052, and (R2) annealing at location 469. PCR [L] to amplify OOR, and initial 200bp of T3D ORF for duplication. PCR [R] to amplify the addition of 2A motif and ORF in entirety. Third PCR, designated [C], had primer sequence (C1) annealing to initial 5' end p24, and (C2) to amplify region of SIV-p24 initial 5' end 2A for inclusion into PCR [R]. PCR performed in 25µL, using PFU Turbo (Stratagen cat. #600250) and associated buffer according to manufactures protocols. Products run on 1% agarose gel with Invitrogen TrackIt 1KB DNA ladder in 1x TBE.

### 2.1.2 – Ligation

0.5µL DpnI and associated buffer added to 20 µL PCR [R] and [L] and allowed sit for 1 hour at 37°C for elimination of methylated parent strand. PCR product purified using PureLink PCR Purification Kit (Invitrogen cat. #K3100-01) according to manufactures protocols. PNK with associated buffer added and allowed sit 1 hour at 37°C for 5' phosphate addition due to lack on primer design. 2 µL T4 DNA Ligase and associated buffer added and let sit overnight at 16°C for ligation of PCR products.

## 2.2 – Bacterial Transfection

Previously prepared competent DH5α cell stock in DMSO obtained. 10µL ligated product mixed with 200µL cell aliquot. Allowed incubate 1 hour on ice. Heat shocked 1.5 minute at 42°C. Cells plated to Amp+ plate. Allowed incubate overnight at 37°C. Individual colony selected, transferred to 3mL LB media. Allowed incubate overnight at 37°C. 3mL incubation transferred spun down and plasmid purified using PureLink Quick Plasmid Miniprep Kit (Invitrogen cat. #K2100-11) according to manufactures protocols. Plasmid sent for sequencing. Plasmid added to 200 µL aliquot of competent DH5α cells. Allowed incubate 1 hour on ice. Heat shocked 1.5 minute at 47°C. Added to 100mL LB media for growth. Allowed incubate overnight at 37°C. 100mL LB media incubation growth spun down. Aliquot purified using Endofree Plasmid Maxi Kit (Qiagen cat. #12362) according to manufactures protocols. Plasmid resuspended 500µL TE.

## 2.3 – Reovirus Reverse Genetics

Method for the reverse genetic engineering of reovirus is adapted from that of Takeshi Kobayashi[30].

### 2.3.1 – Infection with Vaccinia Virus

Sonicated Vaccinia Virus (VV) rDIs-T7pol by cup sonication. L cells infected with 0.5-1mL Opti-MEM I (Invitrogen cat. # 31985) containing rDIs-T7pol at MOI 0.5-2.0 $TCID_{50}$. Incubated 1 hour at 37°C with mild agitation at 15 minute intervals.

### 2.3.2 – cDNA Transfection

53.25µL TransIT-LT1 (Mirus cat. #MIR 230x) added into 0.75mL Opti-MEM I. Vortex 1 second to mix. Incubated 15 minute at RT. Added plasmid DNA mixture (Table 2). Mix by gentle pipetting. Incubated 20 minutes at RT. Removed viral adsorption. Washed infected L cells

with growth medium. Added 5mL complete medium to culture dish. Added TransIT-LT1/DNA complex mixture to culture dish. Incubated transfected L cells 5 days at 37°C.

## 2.4 – Viral Plaque Assay

Added 1.2 x $10^6$ cells per well in 2 ml cMEM to Costar 6-well plates. Dispersed cells evenly within well. Incubated overnight at 37°C to form monolayer. Diluted recombinant virus stock into 10 fold delusions using PBS (pH 7.5) + 2 mM $MgCl_2$. Removed existing media from well. Added 0.1mL sample and distributed evenly through well. Incubated at RT for 1 hour with gentle agitation every 15 minutes. Added 2mL prepared serum-free 2x199 with 1% agar and 10 mg/ml trypsin to each well and distributed evenly. Allowed incubate at 37°C until solidified. Two days post inoculation, proceeded with staining. Added 2mL neutral red per 100mL 2x199 mixture. Added 2mL to well, allow incubate 37°C.

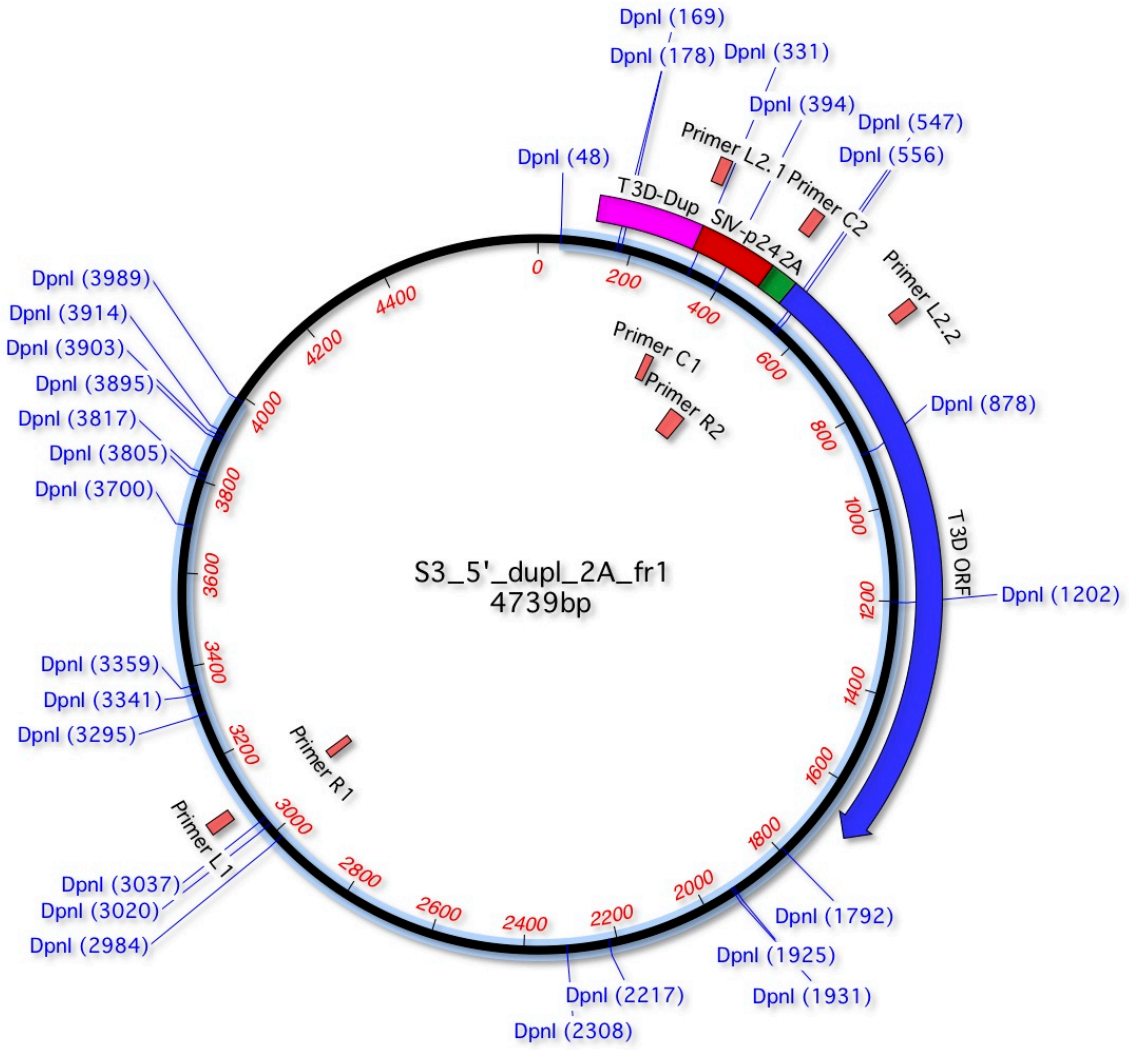


**Figure: 5 – DNA Construct Map. Map showing complete construct of S3 from T3D strain Orthoreovirus. Primers and duplication method indicated.**

| Name | Sequence (5' →3') |
|---|---|
| (L1) | TGA TGA GCA CTT TTA AAG TTC T |
| (L2) | CCT TAG CAT AGC TAA GCA C |
| (R1) | TTG GAA AAC GTT CTT CGG GGC G |
| (R2) | GAG GGC AGG GGC AGC CTG CTG ACC TGC GGC GAC GTG GAG GAG AAC CCC GGC CCC |
| (C1) | CCC GTG CAG CAG ATC GGC GGC AA |
| (C2) | AGC AGG CTG CCC CTG CCC TCG TCG TAG GGT GTA CAG CCC |

**Table: 1 - Primer Sequence. Table displaying the sequence of primers used in the plasmid construction to express p24 protein for use in HIV vaccine.**

| Plasmid | Amount (mg) |
|---|---|
| pT7-L1 | 2.0 |
| pT7-L2 | 2.0 |
| pT7-L3 | 2.0 |
| pT7-M1 | 1.75 |
| pT7-M2 | 1.75 |
| pT7-M3 | 1.75 |
| pT7-S1 | 2.0 |
| pT7-S2 | 1.5 |
| pT7-S3 | 1.5 |
| pT7-S4 | 1.5 |

**Table 1: - DNA Content. Table showing concentrations of DNA added to the viral stock for creation of recombinant Orthoreovirus.**

# 3 – Results

## 3.1 – Construct Design

The recombinant virus, after the alterations necessary for the vaccine protein production, must still properly infect and propagate. Altering the ORF in any way would cause the proteins encoded for by that frame to become altered, possibly losing native functionality. Therefore, a method must be devised in which the ORF remains untouched and code is added into the genome in such a manor as to produce both the vaccine protein and the native coding proteins.

The problem arises as to the location of the inserted code. The initial coding region of the segment reading frame contain vital sequences to the initiation and regulation of transcription, including a 5' region unique to the particular segmented genome. Placement prior to this location would alter this initial 5' region, resulting in a lack of transcription of the code, thereby leading to a lack of production of the protein. Similarly, the end of the reading frame contains coding sequence resulting in the termination and finalization of transcription. Placement of code after this point would likewise result in a lack of transcription of the protein.

The most viable solution to this issue is the replication of the reading frame. By replicating the reading frame, either in its entirety or in portion, a copy is created, containing all necessary transcriptional features, whereby alterations can be made leading to the successful transcription of code. Because only the initial sequence of the reading frame is necessary for protein production, with the central and terminal regions containing protein code and termination sequences, duplication is only necessary of the initial sequence. Depending on the amount of duplication, varying forms of fusion proteins can be created.

Placement of the duplicated reading frame adjacent to the native must be accompanied with a method by which the two regions are separated. Failure to do so could create, depending on the transcriptional factors present in the

The 2A motif is a 54bp oligopeptide sequence that, when present within an ORF, causes a lack of the glycyl-prolyl peptide bond at the C-terminus of the sequence[31]. This allows for the separation of proteins, by way of a lack of bond, without the need for the numerous regulatory factors or physical separation of code. Because of the compact form for the motif, critical in keeping the length of added sequence to a minimum, it was chosen as the method for protein separation. The choice of such a motif greatly reduces the number of regulatory and transcription factors necessary for the code, and eliminates the need for physical separations of the code itself.

Due to the manor in which the 2A motif operates, creating a lack in linkage between the motif and following sequence at the 3' end, the protein at the 5' end will, as a result, contain the 2A motif itself. Because the protein encoded for by the native reading frame is necessary to the replication of the virion, any changes to the sequence could result in a change to the protein function, by way of a structural change. Because of this, the 3' end of the motif must be affixed to the native reading frame, as to insure no superfluous code is transferred to the final resulting protein. This then forces the 5' end to be affixed to the added sequence, forcing duplication and insertion to take place prior to the reading frame, at the 5' end. The resulting construct design is pictured (Fig. 4).

### 3.2 – S3 Plasmid Fabrication

In order to construct the plasmid, a series of PCR's were performed to selectively amplify various regions (Fig.5). Two initial PCR's were performed in order to create the duplication

necessary. Primers for PCR [L] were designed in such as manor as to amplify the 200bp duplication necessary for addition of the p24 sequence, in combination with the later half of the original segment. PCR [R] primer (R1) contained the 2A motif in its entity with overlap of the 5' end, thereby adding the sequence to the 5' end of the native reading frame, while amplifying the initial half of the segment. Following the PCR, an agarose gel was run to insure proper amplification as determined by size (Fig. 6). The gel showed an approximate size of 2600bp for PCR [R] and 2000bp for PCR [L], corresponding to the expected values of 2605bp and 1984bp respectively. A third PCR (PCR [C]) was performed for the amplification of the p24 sequence into the PCR[R] product. Following PCR, all products were ligated and a gel run to insure proper ligation as determined by size (Fig. 7). Gel indicated product of approximately 4700bp, corresponding to expected size of 4739bp. This is a good indication that the PCR's performed as expected.

### 3.3 – Plasmid Selection and Amplification

Following successful ligation, product was transformed into competent bacterial cells for amplification. Plated cells were selected for successful transformation via antibiotic screen and grown via 3mL culture. Plasmids from cultured cells were harvested and sequenced.

Sequencing results yielded 51 unaligned bases prior to the start of the known sequence, as well as numerous frameshifts and basechanges after the 1000bp mark. Both issues are known results of sequencing method, and are in no way indicative of error. The sequence was otherwise clean and in line with expected results, indicating plasmid is as designed (Fig. 8).

## 3.4 – Viral Plaque

Following the confirmation of a correct sequence, a recombinant virus was constructed using the method outlined by Takeshi Kobayashi (§ 2.3). The stock created was used to perform a viral plaque assay.

The viral plaque assay was performed under varying dilutions to yield a single selectable viral line (Fig. 9). A single plaque selected for sequencing. Sequencing yielded a wild type plasmid, as originates from T3D strain (Fig. 10). All modifications to sequence had been eliminated, yielding only the T3D reading frame. No evidence of T3D duplication, p24 sequence, of 2A motif were present. No recombinant plaques were discovered through repeated assays and sequencing.

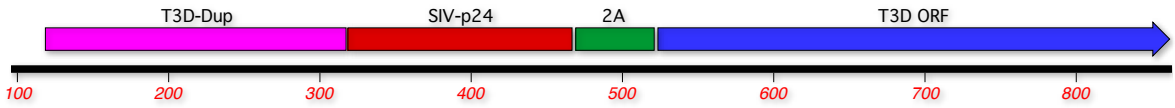


**Figure: 4 – Construct Design. A representation of the construct design. Initial 200bp of T3D strain reading frame replicated, adjacent to region of p24 protein to be expressed. Autocleavage from native reading frame by way of 2A motif. Sequence only partial, OOR and remaining segment not shown.**

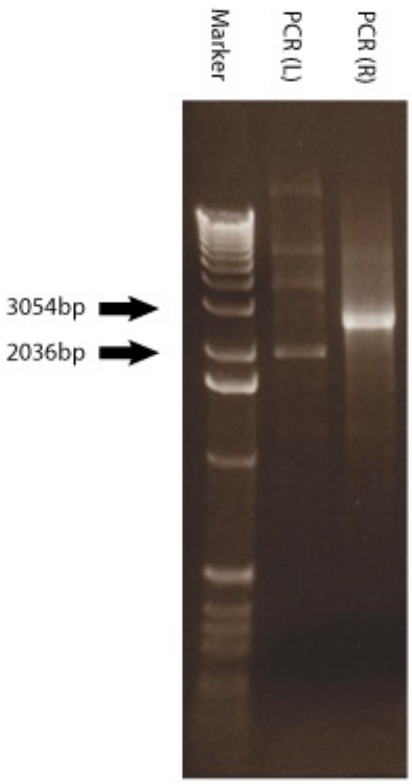


Figure: 6 - PCR [R] [L] Agarose Gel. Gel showing correct amplification of PCR [R] and [L]. Arrows indicate size on DNA marker used. Approximate size of band in PCR[R] determined to be 2600bp and 2000bp in PCR[L], corresponding to expected values of 2605bp and 1984bp respectively.

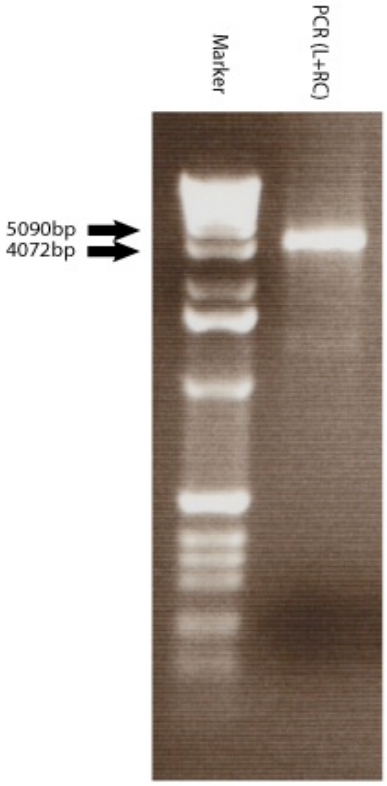


Figure: 7 - PCR [RC+L] Agarose Gel. Gel showing correct ligation of all parts to form plasmid. Arrows and associated size indicate size of band on DNA marker used. Approximate size of plasmid determined to be 4700bp, corresponding to expected value of 4739bp.

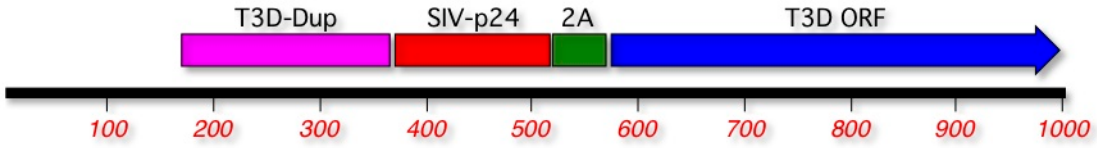


Figure: 8 - S3 Plasmid Sequencing. Sequencing results following the amplification of the construct as performed by competent bacterial cells. Showing correct sequence of 5' duplication, p24 capsid protein, 2A autocleavage motif, and native ORF. No erroneous or unexpected mutations in the first 1000bp as expected.

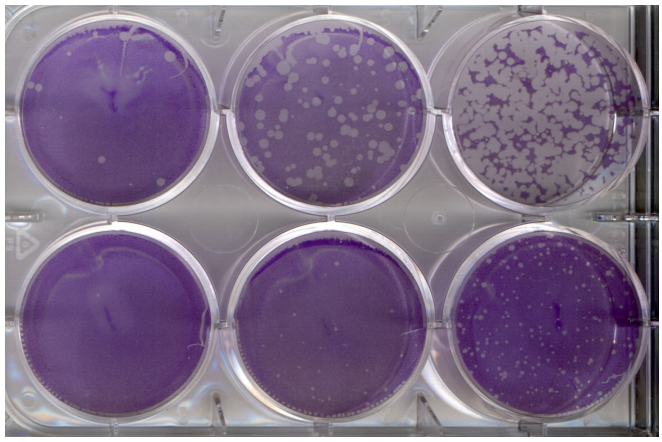

**Figure: 9 – Viral Plaque Assay. Growth and selection of recombinant virions in varying dilutions for sequencing. Wells plated in 10-fold dilution (increasing dilution right → left) series and allowed grow >1 month for the expression of plaques used to sequence obtained Orthoreovirus S3 segment.**

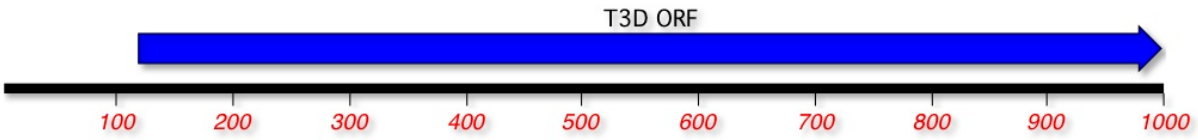


**Figure: 10 – Viral Plaque Sequencing Results. Results from sequencing following viral plaque assaying showing reversion back to wild type S3 segment. 5' duplicated region, p24 capsid protein, and 2A motif all unaccounted for.**

# 4 – Discussion

The deadly and, at this time, incurable nature, of HIV requires the development of a course of treatment against the virus. While a cure would be an optimal solution for those already infected, a vaccine would best for the prevention of the spread of the virus, which will eventually lead to the eradication of the virus from the population. The goal of the 5' duplication method of Orthoreovirus is to create a vaccine by which the p24 HIV capsid protein is expressed, allowing immunity from HIV to be granted to the individual.

## 4.1 – S3 Segment Recombination

The plaque assay results were of particular interest. Virion levels obtained from all assay concentrations were extremely low in nature. In most cases, multiple infections were necessary to obtain a level of virions necessary for sequencing. Since recombination somehow took place during infection or replication, this low production of virions may be the result of the recombination pathway taking place. The extended time observed could also have been the result of selectivity of the pathway, with only a small percentage of recombinant S3 segments being taken up, which over multiple infections, greatly reduced the percentage of recombinant to wild type segments in the stock.

The exact mechanism whereby the recombinant form of the S3 Orthoreovirus segment reverts back into the wild type segment remains unknown. Due to the likeness in sequence of the ORF and the duplicated region, it is feasible that a hairpin created during viral replication causing the excision of the recombinant code. This is unlikely however, due to the presence of the recombinant code after PCR amplification, as well as bacterial transfection. While contamination of the plasmid stock with wild type is possible, the rate of selectivity for wild type over

recombinant type necessary to produce a clean wild type sequence as a result of the plaque assay would indicate otherwise. Should the selectivity of wild type have taken place, a trace amount of recombinant type would still remain present due to the overwhelming amount present, causing the sequencing to produce several errors.

The cause of recombination is most likely due to an unknown pathway within the virion. Recognition of duplication, length of the remaining unmodified segments, or foreign code could have imparted the virion with the ability to select for the recombinant code and remove it to produce the wild type sequence. Preliminary results from further testing indicate possible success with a truncated version of SIV-p24 code inserted (data not shown). Success with a shortened sequence would indicate the possibility of success with this method, assuming further investigation into the method of recombination. This apparent restriction in size is an important consideration that must be taken into account with future use and design involving Orthoreovirus S3 segment.

## 4.2 – S3 5’ Duplication

Despite the issues encountered, the 5’ duplication of the S3 segment remains the best possible approach for the creation of an HIV vaccine by use of Orthoreovirus. The use of a 3’ duplication, should a suitable mechanism for separation even be found, would lead to insufficient production of the recombinant protein, as a result of incomplete translation due to the finished production of the necessary native protein. Insertion mid-coding frame is also not a feasible option, as this would ruin the infectivity of the virus. Therefore, duplication of the 5’ region, conserving the regulatory and structural components of the genome necessary for translation, is the best choice. It is because of this, and the benefits of using Orthoreovirus, including the possibility of inserting multiple foreign codes on multiple segments, as well as the similarities of infection in GALT

cells between Orthoreovirus and HIV, make it currently one of the best possibilities in the fight against HIV. Continued research with this virus may not only lead to results for HIV vaccine, but many other pathogens as well, using a similar method of approach.

### 4.3 – Continued Research

Due to the compact size and native coding sequence, the S3 segment remains one of the best possible segments of the Orthoreovirus genome for use in recombinant vaccines. Investigations into the maximum possible length of foreign code that can be added to this segment would be imperative to determine for its use as a vaccine vector. Determination of this would also most likely be of assistance in studies on subsequent segments. The method by which recombination to wild type takes place is also of great interest, not only in the study of vaccine, but as a study of the native function and pathways of the virus. The possibility exists that recombination takes place by an otherwise unknown pathway, present not only in Orthoreovirus, which may aid in the understanding and development of cures for other pathogens.

While the S3 segment is, in theory, an excellent choice for a vector, other possibilities exist. The L segments, M segments, and remaining S segments are all candidates for the insertion of code, and may prove to lack the problems which has thus far been a factor in the use of S3. Investigation into the feasibility of these segments is an important consideration, not only for the design of an HIV vaccine, but possibly for use in a vaccine that may require a much larger coding region inserted. The degree to which these segments can be altered may vary greatly from S3, possibly allowing a much larger region of code to be inserted, or for a more stable construct to be created.

Various other separation motifs may also be investigated. 2A was used due to its compact size and clean separation. However, should other motifs or fusion protein separation techniques be found to be equally effective, their use into the effect on recombinant code expression in Orthoreovirus should also be investigated. Should work with Orthoreovirus provide no possibility of results, other viruses should be approached, with the in vivo production of p24 targeting GALT CD4+ cells as the goal. This method is most likely to provide the strongest immunity and best possible results in the fight against HIV infection.

## 4.4 – Implications of HIV Vaccine

In many third world and developing countries, such as those found in Africa, the cure of all HIV infected individuals, should a cure be found, would prove to be a highly resource and financially difficult task. Such an undertaking would involve the testing, and subsequent treatment, perhaps with multiple courses of medication over a span of time, the logistics of which alone would prove to be nearly impossible. If eradication of HIV is to be achieved, the prevention of infection is a much more realistic goal. By granting immunity to those not already infected, the virus is unable to spread to a larger portion of the population, and will eventually be removed from that population. Immunity to the virus would also reduce the risk greatly to those in high risk for

## 4.5 – Final Considerations

HIV is perhaps one of the most deadly and feared diseases of our time. With no known cure and almost certain unfavorable outcome, the need for a treatment is massive. Despite the issue of recombination encountered, the 5’ duplication method remains a feasible contender for this treatment. All research and design indicates that success with this method should be obtainable. Should the S3 segment prove to be too problematic for use, not only should the method of recombination be investigated, but other possibilities within the Orthoreovirus genome, and other

viruses, be investigated as well. While the mass production and distribution of a successful HIV medication is almost certainly years away, the technology and tools provided here, as well as in many other avenues, are available now, and should be taken advantage of to the fullest.

# 5 – References


[1] Weiss, R. A. (1993). How does HIV cause AIDS? *Science (New York, N.Y.), 260*(5112), 1273-1279.

[2] Joint United Nations Programme on HIV/AIDS (2006). "Overview of the global AIDS epidemic" (PDF). *2006 Report on the global AIDS epidemic.*

[3] "ICTVdB Virus Description - 00.061.1.06. Lentivirus." *National Center for Biotechnology Information*. Web. 29 Mar. 2010. <http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/00.061.1.06.htm>.

[4] Gilbert, P. B., McKeague, I. W., Eisen, G., Mullins, C., Gueye-NDiaye, A., Mboup, S., et al. (2003). Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in senegal *Statistics in Medicine, 22*(4), 573-593. doi:10.1002/sim.1342

[5] Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, et al. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. Nature. 1999;397:436–441.

[6] Wertheim, J. O., & Worobey, M. (2009). Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2 *PLoS Computational Biology, 5*(5), e1000377. doi:10.1371/journal.pcbi.1000377

[7] Takiguchi, M. (2009). Selection of HIV escape mutants by CTL and adaptation of HIV to HLA class I *Uirusu.Journal of Virology, 59*(2), 145-153.

[8] "Landmarks of the HIV Genome." *HIV Databases*. National Institutes of Health. Web. 28 Mar. 2010. <http://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html>.

[9] Krishna, V., Ayton, G. S., & Voth, G. A. (2010). Role of protein interactions in defining HIV-1 viral capsid shape and stability: A coarse-grained analysis *Biophysical Journal, 98*(1), 18-26. doi:10.1016/j.bpj.2009.09.049

[10] Chan, D. C., Fass, D., Berger, J. M., & Kim, P. S. (1997). Core structure of gp41 from the HIV envelope glycoprotein *Cell, 89*(2), 263-273.

[11] Pancera, M., Majeed, S., Ban, Y. E., Chen, L., Huang, C. C., Kong, L., et al. (2010). Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility *Proceedings of the National Academy of Sciences of the United States of America, 107*(3), 1166-1171. doi:10.1073/pnas.0911004107

[12] Kwong PD, Wyatt R, Robinson J et al. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature 1998; 393:648-659.

[13] Rice, A. P. (2010). The HIV-1 tat team gets bigger *Cell Host & Microbe, 7*(3), 179-181. doi:10.1016/j.chom.2010.03.001

[14] McGovern, S. L., Caselli, E., Grigorieff, N., & Shoichet, B. K. (2002). A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening *Journal of Medicinal Chemistry, 45*(8), 1712-1722.

[15] Harvey, Richard A., Pamela C. Champe, Bruce D. Fisher, and William A. Strohl. *Microbiology*. Philadelphia: Lippincott Williams & Wilkins, 2007. Print.

[16] Jouvenet, N., Simon, S. M., & Bieniasz, P. D. (2009). Imaging the interaction of HIV-1 genomes and gag during assembly of individual viral particles *Proceedings of the National Academy of Sciences of the United States of America, 106*(45), 19114-19119. doi:10.1073/pnas.0907364106

[17] Freed, E. O. (2001). HIV-1 replication *Somatic Cell and Molecular Genetics, 26*(1-6), 13-33.

[18] Ikechebelu, J., Mbamara, S. U., Joe-Ikechebebelu, N. N., & Ezenwabachili, A. O. (2009). Sexual practices of people living with HIV in south eastern nigeria *Nigerian Journal of Clinical Practice, 12*(4), 416-420.
[19] Gumbo, F. Z., Duri, K., Kandawasvika, G. Q., Kurewa, N. E., Mapingure, M. P., Munjoma, M. W., et al. (2010). Risk factors of HIV vertical transmission in a cohort of women under a PMTCT program at three peri-urban clinics in a resource-poor setting *Journal of Perinatology : Official Journal of the California Perinatal Association,* doi:10.1038/jp.2010.31
[20] Takei, T., Amin, N. A., Schmid, G., Dhingra-Kumar, N., & Rugg, D. (2009). Progress in global blood safety for HIV *Journal of Acquired Immune Deficiency Syndromes (1999), 52 Suppl 2*, S127-31. doi:10.1097/QAI.0b013e3181baf0ac
[21] Chun, T. W., Nickle, D. C., Justement, J. S., Meyers, J. H., Roby, G., Hallahan, C. W., et al. (2008). Persistence of HIV in gut-associated lymphoid tissue despite long-term antiretroviral therapy *The Journal of Infectious Diseases, 197*(5), 714-720. doi:10.1086/527324
[22] Yuki S and Wong JK. *Blood and guts and HIV: preferential HIV persistence in GI mucosa.* J Infect Dis 197 (online edition), 2008.
[23] Fields, Bernard N., David M. Knipe, and Peter M. Howley. *Fields Virology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2007. Print.
[24] Tyler, K. L., Barton, E. S., Ibach, M. L., Robinson, C., Campbell, J. A., O'Donnell, S. M., et al. (2004). Isolation and molecular characterization of a novel type 3 reovirus from a child with meningitis *The Journal of Infectious Diseases, 189*(9), 1664-1675. doi:10.1086/383129
[25] Tyler KL. Mammalian reoviruses. In: Knipe DM, Howley PM, editors. *Fields Virology*. Lippincott Williams & Wilkins; Philadelphia: 2001. pp. 1729–1745.
[26] Pancer, Z., & Cooper, M. D. (2006). The evolution of adaptive immunity *Annual Review of Immunology, 24*, 497-518. doi:10.1146/annurev.immunol.24.021605.090542
[27] "Vaccine Types." *National Institute of Allergy and Infectious Diseases*. Web. 29 Mar. 2010. <http://www3.niaid.nih.gov/topics/vaccines/understanding/typesVaccines.htm>.
[28] Wilson, N. A., & Watkins, D. I. (2009). Is an HIV vaccine possible? *The Brazilian Journal of Infectious Diseases : An Official Publication of the Brazilian Society of Infectious Diseases, 13*(4), 304-310.
[29] Wittawatmongkol, O., Vanprapar, N., Chearskul, P., Phongsamart, W., Prasitsuebsai, W., Sutthent, R., et al. (2010). Boosted p24 antigen assay for early diagnosis of perinatal HIV infection *Journal of the Medical Association of Thailand = Chotmaihet Thangphaet, 93*(2), 187-190.
[30] Kobayashi, T., Ooms, L. S., Ikizler, M., Chappell, J. D., & Dermody, T. S. (2010). An improved reverse genetics system for mammalian orthoreoviruses *Virology, 398*(2), 194-200. doi:10.1016/j.virol.2009.11.037
[31] de Felipe, P., Luke, G. A., Brown, J. D., & Ryan, M. D. (2010). Inhibition of 2A-mediated 'cleavage' of certain artificial polyproteins bearing N-terminal signal sequences *Biotechnology Journal, 5*(2), 213-223. doi:10.1002/biot.200900134