Stony Brook University

Construction and Characterization of Codon-Pair Bias Customized Poliovirus

A Dissertation Presented

by

John Robert Coleman

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Molecular Genetics and Microbiology

Stony Brook University

August 2008

Stony Brook University

The Graduate School

John Robert Coleman

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree,

hereby recommend acceptance of this dissertation.

Eckard Wimmer, Ph.D. (Dissertation Advisor), Distinguished Professor
Department of Molecular Genetics and Microbiology

Jorge L. Benach, Ph.D. (Thesis Committee Chair), Chairman, Professor
Department of Molecular Genetics and Microbiology

A. Bruce Futcher, Ph.D., Associate Professor
Department of Molecular Genetics and Microbiology

Nancy C. Reich, Ph.D., Professor
Department of Molecular Genetics and Microbiology

Howard B. Fleit, Ph.D., Associate Professor
Department of Pathology

This dissertation is accepted by the Graduate School.

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Construction and Characterization of Codon-Pair Bias Customized Poliovirus

by

John Robert Coleman

Doctor of Philosophy

in

Molecular Genetics and Microbiology

Stony Brook University

2008

Using the new field of synthetic biology, namely *de novo* synthesis of genomic material, large-scale synonymous changes were applied to the P1 region of the poliovirus (PV) genome. Two DNA fragments were synthesized encoding the P1 structural region of PV in which the existing codons were shuffled, such that novel codon pairings were created. The two constructs, PV-Min and PV-Max, were now encoded by codon-pairs that were statistically under-represented (PV-Min) or over-represented (PV-Max) in the human genome. Importantly, each synthetic fragment still had the exact wild type PV amino acid sequence, but now contained an altered context-dependent Codon-Pair Bias (CPB). The synthetic manipulation of CPB aimed to modulate the efficiency of viral growth, with the intention of creating a model for vaccine development.

Upon investigation, chimeric polioviruses that had a portion of their P1s utilizing under-represented codon pairs were significantly attenuated as a result of decreased

translation efficiency. This attenuation was a reduction of specific infectivity *in vitro* as well as neurovirulence in an animal model. The PV-Max poliovirus, with a P1 region encoded by statistically over-represented codon pairs, showed an enhanced rate of translation but no dramatic increase in virulence. The attenuated, chimeric PV-Min viruses were administered sub-lethally to susceptible transgenic mice providing protective immunization, yet these viruses will need additional modification to be suitable vaccines. Since this method targets an elementary function of all viruses, namely protein translation, it is hypothesized that CPB alteration could be generally applicable to designing new anti-viral vaccines for many types of viruses.

Table of Contents

List of Figures

List of Tables

Acknowledgements

The research and work presented in the following chapters could not have been completed without the help and guidance of many individuals. Firstly, my grandfather Dr. John H. Coleman who has had a profound impact on my life. His advise has influenced both the direction of my career as well as my personal life. Without John, I would not be the person I am today. He has shown me not only how to conduct myself as a scientist, but also as a gentlemen. Throughout my life I will always strive to follow his example - scientist, entrepreneur, gentlemen, and father.

I am also extraordinary grateful too my advisor Dr. Eckard Wimmer for his dedication to poliovirus and my research. He was an enthusiastic supporter of my work and it could not have been completed without his guidance and wisdom. I was so fortunate to join his lab and will look back on his mentorship with fondness and admiration. Another member of the Wimmer lab for whom I am thankful was Dr. Steffen Mueller due to his beneficial impact on my work. He was always there to answer a question and bounce ideas off of. His experimental skill and technical ability were a pleasure to witness and to try to emulate. All other members of the Wimmer lab were supportive, including Dr. Aniko Paul, the chief lab scientist, who kept me in-line and organized, as well as all other members who were a pleasure to work along side: Ping, Ying, Shaun, Chungling, Nusrat, Nidia, and JoAnne. Also, a former lab member, Dr. Hidemi Toyoda, was a great influence simply through observing his dedication to science and research. He inspired me to work hard. Dr. Jeronimo Cello's pure ability and genius as a scientist were a pleasure to be around during my dissertation work and his experimental advice was irreplaceable. I am very appreciative for all his help.

The synthetic biology group also had a great influence on my work. The computer scientists in the group, Dr. Steve Skiena and Dr. Dimitris Papamichail, produced the algorithms this dissertation is based upon. Without their skill with math and codon shuffling, this work would not have been possible. The weekly meetings with them was a pleasure and stimulated me to explore new avenues in my research and career. Thank you Steve and Dimitris. The other member of this group, Dr. Bruce Futcher, was also quite influential on my work and I greatly benefited from his skill as a scientist (and writer), so I am grateful for his guidance and influence. I thank the members of my thesis committee, Dr. Jorge Benach, Dr. Nancy Reich, Dr. Bruce Futcher, and Dr. Howard Fleit, for their constructive criticism, enthusiastic support, and knowledgeable guidance. Dr. Benach's guidance as a committee chair was irreplaceable and I will forever be thankful for his influence.

Finally, my family who have provided me with happiness and love. I am eternally indebted to them for their support. My fiancée Dr. Lisa Runco, who has brought joy into my life and who has provided unwavering support. Her presence has forever beneficially impacted my life and I love her dearly. My grandparents, Grammy and Baba, who both taught me the importance of striving to learn everyday. My mother, Lydia, who has always encouraged me to succeed in all my endeavors and supported me through my youth. My grandmother, Ginnie, for her encouragement and support in my life and career. Lastly, to my future in-laws, the Runcos. Maria Runco whose skill as a chef and unwavering support have both been much appreciated. Also, thank you to Diana, Brian, and Michael Martinez for their support and friendship. I feel truly blessed and I thank God everyday for all his blessings.

Chapter 1: Introduction

Synthetic Biology

The emerging field of synthetic biology is a discipline of the biological sciences that has a potential global impact on a variety of fields. These potential impacts include re-designing microorganisms, solving energy production concerns, or solving pressing biomedical issues. The term 'synthetic biology' is met by some with an apprehensive feeling due to its broad definition and the perception of synthetic biologists as manipulators of nature; however, if utilized correctly and for just purposes, synthetic biology could significantly and beneficially influence humanity in the future. Synthetic biology can be defined as the customization, manipulation, and design of biological systems for useful applications [1]. For example, these applications include engineering bacteria to produce bio-fuels (ethanol, bio-diesel) [2], engineering a virus to serve as a vaccine [3,4], or providing a tool to investigate previously experimentally impractical biological phenomena [3].

The important advance in the biotechnology industry that has truly enabled synthetic biology and opened the door to a vast arena of new experiments is the scientific community's ability to synthesize, without a natural template, genetic material in the form of DNA. Currently, the cost of synthesis is approximately 60 cents per base synthesized; however, based on the trajectory of the price decline and the rapid improvement of this technology, the price per base is estimated to decrease throughout the decade. Also, there are now 39 companies that offer complete gene synthesis, a

remarkable number [5]. This ability to synthesize DNA now allows researchers to manipulate gene structure and function in ways that were previously impossible. Examples of ground-breaking experiments using this ability to synthesize novel genetic material included re-encoding genes with synonymous changes to study nucleic acid phenomena [3], re-arranging genes so that their expression is manipulated quantitatively and temporally [6], or re-coding tRNAs so that they now recognize novel anti-codons or these novel tRNAs carry an 'unnatural' amino acid [7].

Another use of DNA synthesis, that has received criticism from certain quarters, has been to use the synthesis of genetic material to synthesize organisms as a whole, from 'scratch'. The first experiment of this nature was successfully conducted at Stony Brook University, and the researchers Cello, Paul, and Wimmer were able to successfully synthesize a replicating organism (poliovirus, or PV) starting from inert genetic material (DNA oligonucleotides) [8]. This work was quickly met with criticism since the researchers re-constructed a human pathogen. Despite this criticism, this work also inspired copycat experiments by rival groups in the synthetic biology field. Specifically, investigators of what is now know as the J. Craig Venter Institute, successfully synthesized the ΦX174 DNA bacteriophage from synthetic oligonucleotides [9]. This work was quite similar to the construction of poliovirus; however, it was free of the criticism which attacked the initial work that used PV as the synthetic organism since a bacteriophage isn't a potential bio terrorism agent.

Initially, uniformed politicians and the public immediately dismissed the synthesis of poliovirus as an irresponsible and dangerous stunt, lacking in scientific merit [10], stating that the regeneration of a human pathogen would simply supply terrorists with the

ability to create biological weapons. However, it is the opinion of some scientists that experiments such as the synthesis of poliovirus will actually prove to have great benefits. The synthesis of a replicating organism will allow us to learn the answer to basic biological questions, such as the minimal set of genes required for life, as well as provide the proof of principle experiments that will allow for customizing organisms and genes in the future [7]. These customized organisms could be thoughtfully designed to solve pressing human issues such as energy production or vaccine creation. The possibilities afforded by whole genome synthesis have taken an even larger step forward in light of recent advances. Recently, the J. Craig Venter Institute was able to successfully synthesize and assemble an entire bacterial genome, 582,970 base pairs in size, which dwarfs the original poliovirus synthesis (~10 kilobases) in magnitude and implication [11]. Although this assembled large circular DNA molecule has not shown any sign of biological activity (e.g. the ability to replicate once inside an appropriate environment), this is still truly an astounding feat given the shear size of the genome synthesized. However, especially in light of these results, it is important that the synthesis of novel genetic material receive proper monitoring to ensure that this synthesis and construction are for upright purposes [5].

As evident, the field of synthetic biology is truly burgeoning, with limitless possibilities. The work presented in the following chapters builds upon the ground-breaking study by Cello et al. and investigates the use of genome synthesis as a novel means to produce vaccines, as well as investigates the genetic phenomenon of codon-pair bias (CPB).

Poliovirus Synthesis

The pioneering study using whole genome synthesis was conducted by Cello et al. and it was the first study that demonstrated the applicability of synthetic biology to vaccine creation [8]. Specifically, that one could synthetically create an infectious virus *de novo* and, if so designed, with an attenuated phenotype. This project successfully generated the poliovirus genome by assembling synthetic DNA fragments in a test tube. Specifically, the readily available GenBank sequence for type-1 poliovirus Mahoney strain (PV(M)-wt), originally generated in the Wimmer Laboratory [12], was divided into small oligonucleotides that were stitched together by PCR, yielding a full-length cDNA encoding poliovirus. This cDNA was then inserted into a carrier plasmid, which also had T7 RNA polymerase promoter upstream of the inserted sequence. Then the researchers were able to use an established reverse genetics system to create (+)RNA from this cDNA [13], next incubating this (+)RNA with HeLa S3 cell extract to generate the virus itself [8]. The virus generated (sPV1M) was antigenically identical to the progenitor strain (PV(M)-wt), was able to infect cells, and was able to form visible plaques [8].

While constructing poliovirus during this study, certain synonymous codon changes, among other nucleotide changes in the 5' non-translated region, were inserted into the synthetic genome as genetic 'water marks' so the investigators could distinguish the wild type strain from the synthetically constructed strain. Interestingly, the synthetic virus constructed by Cello et al. showed an attenuated phenotype when administered to CD155 transgenic mice (CD155tg) [8,14]. After further investigation, this phenotype was actually proven to be attributed to the changing of wild type bases in the 5' non-translated region of the genome [15]. However, the idea of manipulating synonymous

codons to attenuate PV, achieved by Cello et al., provided the initial catalyst for the work presented in the following chapters as well as previous studies. Specifically, projects have been conducted to attenuate poliovirus by using synonymous codon changes, implemented via synthetic construction [3,4,16]. These customizations exploited the genetic code's redundancy such that the amino acid sequence of the protein encoded is exactly preserved however it is now alternatively encoded by different, synonymous codons. This is an attractive means to attenuate a virus because it targets translation, a necessary process in the viral life cycle.

The strains constructed in these studies were initially designed for use as a live vaccine for poliovirus; however it is believed that these methods could be applicable of attenuating other, unrelated viruses. Despite the elimination of poliovirus as a scourge effecting modernized countries, save a few isolated incidences of poliomyelitis in the US [17], it is still present in endemic countries in the third world and remains a pathogenic virus capable of infecting and spreading through immunologically naïve human populations.

Poliovirus Replication Cycle and Pathogenesis

Poliovirus (PV) is a member of the large *Picornaviridae* family of non-enveloped, (+)RNA viruses, whose other members that afflict humans include rhinovirus (common cold) and hepatovirus (hepatitis A virus) among others [18]. Picornaviruses are the most common viruses infecting humans. More specifically, PV is a member of the *Enterovirus* genus of this family, which is primarily composed of Coxsackieviruses, Echoviruses, and Human Enteroviruses [19]. Enteroviruses gain entry to the host primarily via fecal-oral contact and the annual rate of enteroviral infections has been estimated to be one billion or more worldwide [20]. PV consists of different strains of three serotypes, for example PV1 (Mahoney or Brunhilde strain), PV2 (Lansing strain), and PV3 (Leon strain). These serotypes are defined by the varied antigenic responses they induce in their host, yet all serotypes proceed with a similar replication cycle [21]. PV establishes an infection by replicating in the gastrointestinal tract of the afflicted individual, usually producing symptoms that are rather mild or even unnoticed [22,23]. The human disease associated with a PV infection, poliomyelitis, is the result of the invasion of the central nervous system (CNS) and subsequent destruction of motor neurons by the virus, resulting in flaccid paralysis [24]. The invasion and destruction of cells in the CNS is actually a rare secondary site of infection, occurring at a rate of 1 per 100 infections for serotype 1. However, due to the vast number of infections observed, poliomyelitis is a common occurrence during PV outbreaks [25].

On the molecular level, a single, icosahedral poliovirus virion is formed from 60 copies of the four structural viral proteins VP1, VP2, VP3, and VP4 [26]. Contained within the capsid is the genomic material in the form of single-stranded, positive-sense

(+)RNA, 7,441 bases in length that is covalently linked to a small peptide VPg at the 5' end and polyadenylated at the 3' end (Figure 1.1) [27,28]. Once the capsid enters and uncoats inside the host cell, the RNA genomic material is released into the cytoplasm. This viral RNA functions both as a template for replication as well as a messenger RNA. Upon entry, the viral RNA is transcribed by cellular ribosomes into a single polyprotein via an Internal Ribosome Entry Site (IRES) present in the 5'-non-translated region (NTR) of all picornavirus genomes including the poliovirus genome (Figure 1.1) [29,30]. The genome is organized as a single open reading frame containing the P1 structural region (encoding the viral capsid) and the P2 and P3 non-structural regions (encoding proteins involved in RNA replication) [12].

For the purposes of the work presented in the following chapters, this process of translation is the aspect of the PV "life" cycle that is most relevant. The IRES is formed by secondary structures of viral RNA and allows for cap-independent translation of the genome [31]. This is important because poliovirus inhibits host cell protein synthesis by cleaving Eukaryotic Translation Initiation Factor-4F (eIF4F), a protein utilized in cap-dependant translation [32,33]. Therefore, once poliovirus has infected a cell, only IRES-containing RNAs will be translated. The PV genome is translated as a single polypeptide, called a 'polyprotein' [34]. This single PV polyprotein contains all viral proteins that are subsequently produced by *cis* cleavage events catalyzed by the viral proteases $2A^{pro}$ and $3C^{pro}$ (or its precursor $3CD^{pro}$). Cleavage events catalyzed by the $2A^{pro}$ occur at the amino acid pair Y-G and $3C^{pro}$ cleavages occur at the amino acid pair Q-G [35]. The initial cleavage event in the replication cycle is catalyzed by the $2A^{pro}$ and occurs at the N-terminus of the P2 region [36]. All other cleavage events yielding viral

proteins required for replication and capsid formation are achieved by $3C^{pro}$ (or its precursor $3CD^{pro}$) and occur at various rates.

Following the initial translation and production of all viral proteins, the process of RNA replication occurs. The replication of RNA inside the host cell transpires on virus-altered membranous structures. These membranous structures are organized from portions of the endoplasmic reticulum via the $2C^{ATPase}$ and the 2BC proteins [37,38]. The process of RNA replication is catalyzed by the viral protein $3D^{pol}$, a single-stranded RNA-dependent RNA polymerase. The $3D^{pol}$ protein synthesizes a complimentary negative strand of RNA using the infecting genomic RNA as the template [39]. This negative strand is then used as a template by $3D^{pol}$ for the synthesis of many copies of (+)RNA, which are transcribed for protein production allowing for logarithmic expansion of this cycle and/or encapsidation of genomes. Interestingly, the $3D^{pol}$ lacks 3'-5' exonuclease proofreading activity, which is actually advantageous for the virus because this property results in a mutation rate of approximately 1 per 10^4 nucleotides incorporated and in turn, results in a virus population that is highly diverse [40]. With the ability to produce genetic diversity, PV can survive and rapidly adapt to hostile environments [41]. Once sufficient quantities of genomic (+)RNA and capsid precursors have been produced, the process of encapsidation and virus release can occur.

Prior to exiting the cell, capsid precursors encapsidate a single (+)RNA comprising its genome. This process of virion assembly is not fully elucidated. The current hypothesis is that first, a noncovalent structure, called the 'protomer', is formed by the capsid proteins VP3, VP1, and VP0. Next, these protomers amass into pentamers and six of these pentamer structures can assemble the procapsid. Association of the

genomic RNA with the virion precursors results in the formation of the icosahedral virion [42]. It is not known whether the genomic RNA is inserted once the icosahedral virion is formed or whether the pentamers form around the RNA [43]. The final event prior to virus release is the maturation cleavage of VP0 into VP2 and VP4, which forms the mature 160S virion [44]. This cleavage is achieved via unknown mechanisms (Figure 1.1) [28,45]. Mature viral particles then exit the cell by lysis and continue the cycle of replication in new, susceptible host cells.

Once released from a host cell, the virus gains entry to a new host cell via attachment to its receptor, CD155. CD155 is a Type I transmembrane glycoprotein that is a member of the immunoglobulin super family of proteins. Normally, membrane–bound CD155 serves the function of a cell-cell adhesion molecule interacting with nectin-3 to form adheren junctions [46,47]. Membrane–bound CD155 also mediates the entry of PV into the host cell [25]. Specifically, studies have shown that CD155 inserts itself into a surface depression on the viral capsid known as the "canyon" [48,49]. This CD155-poliovirus binding causes the myristoylated capsid protein VP4 and the amphipathic helix of the VP1 protein to be exposed and in turn initiate the process of virus entrance into the cell [50]. It is not known if the virus gains entrance via receptor-mediated endocytosis or if the virus injects the RNA into the cell itself [51]. However, once inside, a typical round of replication in a single cell by poliovirus can be completed within eight hours post-infection as seen by *in vitro* growth curves that demonstrate the kinetics of poliovirus growth [25]. A simple diagram of the entire PV replication cycle can be seen in Figure 1.2 [52].

At the tissue level, there is strong evidence to support the hypothesis that poliovirus establishes an infection in the gut-associated secondary lymphatic tissue, called Peyer's patches. Peyer's patches have been shown to express the PV receptor CD155 and contain high concentrations of virions in infected individuals. Specifically, the epithelial cells covering these patches, the follicle-associated epithelium (FAE) cells and microfold (M) cells, are thought to be the primary sites of infection. Also, endothelial cells associated with Peyer's patches express CD155 and have been shown to support viral replication *in vitro* [23,53,54]. Interestingly, not all human tissues expressing CD155, such as the kidney and liver, support PV replication [55,56]. This could be due to splice variants of the receptor resulting in CD155 secretion, cell environments unfavorable to PV replication such as a strong innate immune response (interferon), or sequestration of CD155 in cell-cell junctions [25,57]. From the primary replication site in the gut, poliovirus then gains access to the CNS to cause poliomyelitis and this process occurs via an unknown mechanism. All studies suggest that PV replication, possibly in Peyer's patch-associated endothelial cells of blood vessels, will cause viremia. Once in the bloodstream, PV virions can then access the CNS either by retrograde axonal transport at the neuromuscular junction or by crossing the blood-brain barrier [54,58,59]. Others have hypothesized that poliovirus may, in some instances, bypass the viremic stage of infection and spread from Peyer's patches to the CNS via the vagus nerve [60].

Despite lacking clear evidence of how poliovirus gains access to the CNS, it is certain that poliovirus has the ability to replicate in the CNS once it arrives. The infection of the CNS by poliovirus irreversibly destroys motor neurons, causing the well-

known disease poliomyelitis. Poliomyelitis is characterized by flaccid paralysis of the limbs, as well as possible death in 10% of these paralytic cases, when the virus destroys motor neurons controlling breathing [23,24]. Due to the irreversible effects that poliomyelitis has, i.e. irrecoverable flaccid paralysis, PV is a brutal and crippling virus and fortunately for some a vaccine has been produced.

Global Vaccination Campaign

The development of vaccines has been the most effective means to combat infectious diseases afflicting humanity. Dr. Edwin Jenner developed the first vaccine against the disease smallpox in 1796. Specifically, Dr. Jenner found that scrapings from the blisters on the hands of milkmaids infected with cowpox could be used to inoculate another individual to provide adequate protection against smallpox, a similar virus (symptoms include a severe rash followed by extensive blistering of the skin, similar to cowpox-induced symptoms) [61]. Interestingly, this smallpox vaccine was in actuality a live-attenuated vaccine, which is one of several multiple strategies and types of vaccines that currently are in use to protect humans from infectious disease.

As its name implies, live-attenuated vaccination is the administration of an actual replicating organism into the recipient. This replicating organism has the ability to produce a protective and broad immune response after a single dose [62]. This organism has been altered such that it can replicate to a certain extent within the host, yet not cause (in the vast majority of recipients) the disease the vaccine is intended to prevent. An example of a current live-attenuated vaccine in use today is FluMist® (MedImmune), a vaccine consisting of a strain of influenza that has been adapted, through genetic manipulation and selective pressure, to grow at a temperature significantly below human body temperature [63]. Therefore, once this strain is inoculated into the host, it has a muted replication cycle that is able to elicit a strong anti-influenza immune response, yet the weakened virus is effectively cleared from the recipient. There also exists a live-attenuated vaccine for poliovirus, called the oral poliovirus vaccine (OPV), which was and is still used globally to combat poliovirus [64]. Live-attenuated vaccines have the

greatest efficacy compared to all other forms of vaccination, for example, compared to inactivated vaccines [65]. Namely, the considerable efficacy is due to the replication of the attenuated organism in the host mimicking a natural infection, and therefore the body has a natural, full-spectrum protective response. This response includes the induction of both the innate and adaptive arms of the immune system, as well as the appropriate memory response [62].

Another type of vaccine includes inactivated vaccines, which are produced by either formaldehyde-fixation of the entire, intact pathogenic microorganism, or heat-inactivation of said pathogen. There is also a subset of inactivated vaccines called toxoid vaccines, which are similar in that they are formaldehyde-fixed bacterial toxins that protect the host from disease [66]. The inactivation of the microorganism still allows for maintaining antigenic epitopes, both three-dimensional as well as the primary structure of the microorganism's proteins, such that these epitopes are recognized by the immune system and an immune response is elicited [66]. This immune response is primarily a humoral response resulting in the production of antibodies that are sufficient to prevent the onset of the disease. Since inactivated vaccines are inanimate, the recipient requires multiple inoculations or boosters spaced at intervals of time so that the immune system may raise the proper response and also produce specific memory against this organism [67]. Since multiple doses are required, there is actually a reduced compliance with inactivated vaccines because some individuals, especially children, do not complete the regimen due to the multiple vaccines appearing to be superfluous [68].

A third type of vaccine that is also currently used to fight diseases is referred to as the subunit vaccine. The subunit vaccine consists of a multimeric form of a single

protein from the progenitor pathogen, which is administered to the vaccine recipient. This protein is separated away from the pathogen and administered as a purified product [69]. Like the inactivated vaccine described above, subunit vaccines require multiple doses to garner the proper immune response. An example of a subunit vaccine currently in use is the hepatitis B virus (HBV) vaccine, which is comprised of multimers of the hepatitis B surface antigen (HBsAg). This protein induces the protective antibody response required to clear Hepatitis B. It was initially thought that subunit vaccines would not provide lasting immunity [70]; however it is now believed that, at least for HBV, the subunit vaccine can provide long-term immunity [71]

The progress vaccine development has made in the fight against infectious diseases in the past two centuries is impressive and will continue to evolve with the ever-evolving microorganisms. For example, there are experimental vaccination methods in development that are not discussed here, such as DNA vaccines. Chapters within this thesis dissertation will suggest additional novel approaches for vaccine development. In the case of poliovirus, two of the vaccine types (live-attenuated and inactivated) that are described below have been used in a successful global vaccination campaign that has nearly eradicated poliovirus from the globe.

For residents of industrialized countries, poliovirus has been almost entirely eradicated as a result of effective vaccination campaigns and continued vaccination of infants and children. In the United States, the inactivated polio vaccine (IPV) is administered to children in three doses, the first of which is administered two months after birth [72]. The IPV is produced by formaldehyde fixation of the three poliovirus serotypes, thereby inactivating the virus and providing the intact virion for the immune

system to recognize [73]. The IPV elicits a weak mucosal immunity in the gut (Figure 1.3, [74,75]); however the IPV provides strong humoral immunity, so as to prevent paralytic polio, which has been deemed a sufficient level of protection [76]. A properly IPV immunized population is unlikely to allow PV circulation; however since there are currently still reservoirs of the virus that may persist, vaccination against poliovirus must never cease, even in industrialized countries with an apparent absence of poliomyelitis [77]. Critics of global IPV use state that the IPV is simply too costly to distribute globally [78]. Specifically, that the process of inactivation is nearly five times more expensive than the production costs associated with the oral poliovirus vaccine (OPV), and also requires the use of a syringe to vaccinate [78]. Both of these factors, cost and administration, pose a large hurdle in attempting to distribute IPV in non-industrialized countries such as Bangladesh, Nigeria and Pakistan where wild type poliovirus is still found. There are studies under way to overcome this hurdle by including IPV in the formulation of other vaccines (such as the DPT vaccine) which utilize a syringe [79].

For use outside the industrialized world, an alternative vaccine exists called the oral poliovirus vaccine (OPV) produced by Albert Sabin, which has been utilized by the World Health Organization (WHO) to eradicate poliovirus in non-industrialized countries. Prior to 2000, OPV was used globally (except Sweden and the Netherlands) and in the United States to vaccinate against poliomyelitis; however IPV use was discontinued due to vaccine-associated paralytic poliomyelitis (VAPP), which is when a vaccine recipient develops acute flaccid paralysis caused by the vaccine [80]. OPV consists of all three mutant/attenuated poliovirus serotypes (designated Sabin-1, -2, and -3) and is administered orally, in liquid form, as its name implies. The OPV elicits strong

mucosal and humoral immunity in vaccine recipients due to the nature of a live, attenuated vaccine that grows in the gut (Figure 1.3) [76]. OPV was produced via serial passage of the three serotypes of poliovirus in monkey kidney cells [81]. Serial passaging resulted in the (limited) accumulation of monkey-specific mutations, so that once returned to the human host, the strains showed a vastly neuroattenuated phenotype and thus could be combined and used as a vaccine [82]. Utilizing OPV, the WHO implemented a campaign to globally eradicate poliovirus. This campaign has proven to be a great success towards the goal of eradication; however problems intrinsic to the OPV have proved to be challenging and may result in the inability to achieve full eradication of poliovirus. These problems include the rare possibility of vaccine-induced poliomyelitis in healthy recipients of the OPV, a condition called VAPP, whereby the vaccine itself causes poliomyelitis. Secondly, and the most prevalent drawback, is the generation and circulation of neurovirulent polioviruses derived from OPV, which have been named circulating vaccine derived poliovirus (cVDPV) [83,84]. These novel strains arise via reversions, mutations, and recombination with other enteroviruses while replicating in the gut of the vaccine recipient [85]. These new strains of cVDPV can cause outbreaks of poliomyelitis in immunologically naïve populations [83,86]. Lastly, there is a very rare possibility that an OPV recipient with an immune deficiency develops a 'carrier state,' whereby the virus is never entirely cleared and is continually excreted [87]. These long-term persistent infections again could result in reversion and thus spreading of a virus capable of causing poliomyelitis in naïve communities. In light of this possibility, vaccination must never be stopped [77]. All of these obstacles are due to the genetic instability of the OPV vaccine.

There are actually a limited number of point mutations in each OPV-strain that result in the attenuated phenotype and these have the possibility of reverting via the mechanisms described above [88]. The most debilitating mutation capable of reducing neurovirulence is a single nucleotide change in the IRES of both Sabin-1 and Sabin-3 (Figure 1.4). Therefore, this IRES mutation is the most readily reverted and found changed in excretions from those who experienced VAPP as well as in cVDPV strains [89]. Other point mutations, which encode amino acid changes in the capsid and polymerase, contribute to the attenuated phenotype and are also found reverted [90]. A diagram depicting the attenuating mutations in each Sabin strain, which when reverted, return the strain to neurovirulence, can be seen in Figure 1.4. This problem of reversion is compounded by the ability of *Enteroviruses* to infect their host for lengthy periods of time or even persistently. This extended infection provides OPV strains ample replication cycles to revert the attenuating mutations (Figure 1.4) [91]. Also, the fact that wild type poliovirus or OPV-derived polioviruses are able to pervade through endemic areas, rapidly infecting the entire population compound the drawbacks of OPV [92]. The very nature of these hurdles may prevent the achievement of complete eradication of poliovirus solely using the OPV itself, thus the World Health Organization issued a call for a new vaccine [93]. Now, with the advent of synthetic biology, there may be novel ways to construct a new poliovirus vaccine to answer this call.

A previous study applying synthetic biology as a means to develop a poliovirus vaccine

A previous study which we conducted began the investigation into whether one could attenuate PV by customizing the codons that encode its proteins [4]. In that study,

Mueller and colleagues replaced all codons in the P1 structural region of poliovirus with the synonymous codon that corresponded to the most infrequently utilized codon in human neuronal tissue [94]. These synonymous changes were designed to prevent poliovirus from replicating in the CNS. The P1 region was selected to be replaced by a synthetic fragment, primarily because it has been convincingly demonstrated that the elimination or replacement of this region with exogenous sequence does not effect replication. This strongly suggests an absence of replication regulatory elements within the P1 region [4,95,96]. An example of one amino acid, Proline, of the twenty that were converted to the most rare synonymous codon alternative, can be seen in Figure 1.5. Initially, a full-length P1 encoded by only rare neuronal codons, designated as PV-AB, resulted in a viral construct with a null phenotype [4].

However, after further investigation, we found that the incorporation of only a small fragment of synthetic AB sequence (Figure 1.5) in the P1 region, from nucleotides 2470-2954, was able to significantly attenuate the virus by inhibiting translation [4]. More specifically, a single PV-AB^{2470} virion had a greatly reduced probability, compared to wild type PV, of infecting a single cell. This reduced probability was a result of a 3 orders of magnitude reduction in specific infectivity [4]. The specific infectivity for poliovirus type-1 Mahoney strain (PV(M)-wt) is 1/100, i.e. the probability of a virion productively infecting a cell is only 1 of 100 virions and the PV-AB derivates now had a specific infectivity of 1 of 105,000 [3,4,97,98]. The increase in the number of virions correlating to a single infected cell for the codon usage virus PV-AB^{2470}, is thought to occur because of an increased rate of abortive and/or non-productive infections. The infection of a cell by PV requires efficient, rapid translation and processing of its proteins

upon entry, so that the virus can hijack the cell for its own replication. This hypothesis is supported by the observation that there is a certain threshold of translation required [99]. Mueller et al. showed that by placing a translation hindrance using rare codons, the initial translation of the viral polyprotein is slowed and this leads to an increase in the frequency of abortive infections. Therefore the codon usage virus (PV-AB^{2470}) was significantly attenuated because it is less infectious.

However, in the rare event of the PV-AB^{2470} virus successfully taking over a cell, the inhibitory block of codon usage was no longer as restrictive. It was hypothesized that once a productive infection had been established, all of the cellular resources, specifically tRNAs, were now at the disposal of the virus. Therefore, in this rare event (0.001%) the presence of rare codons was no longer debilitating to the virus, and the virus was able to destroy the cell and produce progeny virions. In the case of poliovirus, this was a detrimental caveat because the virus is neurotropic and specifically targets and destroys motor neurons if it gains access to the CNS [54,58]. The neurotropic nature of PV is especially unfortunate because, unlike epithelial cells, motor neurons cannot be regenerated if destroyed. Thus when we measured the Lethal Dose 50 (LD_{50}) value in susceptible mice for the PV-AB^{2470} virus, there was an attenuation of only ~10^2 particles. This was not a sufficient level of attenuation for a poliovirus vaccine strain due to its retained neurotropic nature [4].

While the utilization of rare codons showed promise and may very well be applicable to many other viruses (especially for non-neurotropic viruses, like dengue virus), the project presented in the following chapters remains focused on poliovirus and a means to produce a new live-attenuated virus strain. The work presented here now

looks to manipulate synonymous encodings of the poliovirus genome independent of codon usage. This was to alter the Codon-Pair Bias (CPB) of poliovirus, a genetic phenomenon thought to influence translation and that has been observed in many organism, yet has had very limited experimental testing. Now, with the advent of synthetic biology and the ability to synthesize genetic material rapidly, the work presented here was able to study both CPB and its ability to attenuate poliovirus.

Codon-Pair Bias

Codon-pair bias (CPB), which is entirely independent of the codon frequency, is the observed fact that within a gene certain codons, corresponding to two amino acids, are found adjacent to one another with frequencies either less or more than expected if these codons were randomly placed next to one another. CPB can be quantified based on statistics and the overall bias of a given gene determined. This genetic phenomena was first observed and analyzed in select open reading frames of *E. coli* [100] and now the phenomenon of CPB has been observed in the genes of many organisms including humans [3,101]. Using methods described in Chapter 2, the CPB has been calculated for all of the 14,795 annotated human genes [3]. Despite the observed evidence of an actual CPB, as well as a means to accurately quantify CPB (refer to Chapter 2), it is not evident to the research community why this bias exists and what selective pressures are governing the 'preference' of codons to be adjacent or apart from one another. The initial work which observed CPB, conducted 20 years ago, was by Gutman and colleagues, and in this work they used an equation very similar to Figure 2.1 to define the CPB of given open reading frames in *E. coli* [100]. In the genes of *E. coli,* they found that indeed there was a bias of over- or under-represented codon-pairs, and that this CPB was statistically significant. Interestingly when they analyzed the extent of this bias and its context in-depth, they also made other observations that helped to define CPB and lead to a hypothesis that CPB influenced the rate of translation [100].

Specifically, they analyzed the codon-pair relationships in these ORFs at different levels (distance and direction) of association between the codons, to ascertain whether this bias occurred only between adjacent codons or whether there were distal biases

between codons. Firstly, they found that this bias no longer existed (i.e. is a random association) when two codons were examined at a distance of two or three codons of separation [100]; thus implying that this bias is only at the codon-codon junction. Secondly, they found that there was no relationship between the directionality of a codon-pair and its bias, i.e. the hypothetical codon-pair A-B did not have the same statistical representation value as pair B-A [100,102]. This directionality implied that bias existed for hypothetical codon-pair (A-B) in the 5' → 3' direction on the mRNA and this bias did not exist in the reverse direction [100]. In sum, CPB was determined to be unidirectional and also to be occurring between nearest neighbor codons, implying the significance of the pair itself.

Since the junction of codons is where the bias occurs it was originally hypothesized that CPB influences translation (since the respective codons will be occupying the A and P site in the ribosome) and the bias has arisen due to selective pressures that govern the rate of translation. The work conducted by Irwin and colleagues was the initial and only wet laboratory follow up [102], save the very recent data from Coleman et al. [3], that experimentally tested the bias observation by Gutman et al. [100]. These experiments by Irwin and colleagues were conducted to investigate the influence CPB has on translation.

This research group conducted what now appear to be cursory experiments; however at the time (1995) there was no ability to synthesize DNA cheaply. Specifically Irwin et al. used PCR and oligonucleotides to change the third and fourth codons (i.e. only one pair at codon positions 3/4) in both the *lacZ* gene and *trp* operon [102]. They replaced the wild type codon-pair with corresponding synonyms pairs that were either

over- or under-represented *E. coli*. The effect this change had on translation was then monitored by modulations in β-galactosidase enzyme activity from the *lacZ* gene and transacetylase enzyme activity from the *trp* operon. This group found that when one placed a synonymous over-represented codon-pair at position 3/4, down-stream enzyme activity encoded by the altered gene was inhibited 2-10 fold. This decrease was modulated based on the degree of over-representation, such that one could suggest a relationship between the degree of CPB and its impact on translation [102]. The inverse of these results were obtained for under-represented codon-pairs, where under–represented pairs increased the rate of translation.

Therefore, building upon the discovery of CPB and using the results obtained in these experiments, the Hatfield group (of which Irwin and Gutman were a part) made a hypothesis that indeed CPB effects translation. Specifically, their hypothesis was that codon-pairs, which are over-represented (at least in the prokaryotic system), are responsible for a reduction in the rate of translation to allow for proper folding of the growing polypeptide [102]. They made the supposition that this decrease caused by over-represented codon-pairs inserted pauses during translation. These pauses were required for proper folding of the nascent polypeptide as it grew and emerged from the traveling ribosome. They also stated directly that these experiments were a limited sample set and that more data was required to make a global claim on the effect of CPB [102]. **Note:** These studies were done in *E. coli,* a prokaryote, and the work presented in the following chapters was done in a eukaryotic system (poliovirus and HeLa R19 cells).

There exists another hypothesis based on computational results regarding CPB that like the Hatfield group considers the codon-codon junction the source of the bias and

its relation to translation. This model was developed and tested, previous to the work presented in the later chapters, solely on computational methods and modeling. The hypothesis is the inverse of the initial findings, such that tRNAs of corresponding pairs are the source of this bias. Specifically, that steric hindrance of certain tRNA pairings will cause an inhibition of translation, thus disrupting the ribosome and having an influence on the rate of translation [103]. Therefore these pairs will be avoided, i.e. under-represented, because their tRNAs interact poorly and thus the efficiency of translation is reduced. Data and observation supporting this hypothesis will be presented in subsequent chapters.

An another important observation that provides further description of CPB suggests that the N3 – N1 nucleotides (i.e. the last nucleotide of codon A and first nucleotide of codon B) between the codon-pair in fact has the greatest influence on CPB and thus strongly influence pairing [104,105,106]. Specifically, that a codon-pair with a C at N3 and G at N1 (designated CpG or CG_{3-1}) are avoided, or strongly under-represented [3]. Actually this dinucleotide is suppressed within codons as well (CG_{12}, CG_{23}) and it is uncertain why this dinucleotide is avoided. The CpG dinucleotide, within an individual codon or between codons, could have an impact on translation, be repressed due to genomic forces, or the presence of CpG is used to distinguish self from non-self in eukaryotes. It is important to note that the rarest individual codons also contain an internal CG (i.e. $CG_{1\text{-}2 \text{ and } 2\text{-}3}$) suggesting this dinucleotide is actively avoided for some purpose.

The first supposition is that CpGs impact translation, firstly because rare codons contain it (i.e. $CG_{1\text{-}2 \text{ and } 2\text{-}3}$) and also as a result of incompatible tRNAs corresponding to

the $CG_{3\text{-}1}$ dinucleotide [105]. Specifically, that this nucleotide pair's high-stacking energy serves to impede the traveling ribosome. Another group supposes that CpG dinucleotides are suppressed within genes as an indirect result of DNA methylation in mammalian genomes [101]; however CpGs still appear in the mRNA and thus is a less likely target of methylation. Lastly, it is known that CpG-containing DNA and CpG-containing single stranded RNA are immune stimulators [107,108,109], thus the under-representation of CpG within a eukaryotic organism's own genes could be a means to prevent self reactivity of a cell's own genes stimulating an innate response. This idea of CpG used as a means for self versus non-self recognition has a few supporting observations. Firstly, that codons themselves containing CpGs (i.e. $CG_{1\text{-}2 \text{ and } 2\text{-}3}$) are rare codons (Figure 1.5, [94]) and codon-pairs with a CpGs ($CG_{3\text{-}1}$) at their junction are under-represented [3]. Also viruses, specifically small RNA viruses infecting eukaryotes, suppress CpGs in their genome [110], possibly because they have evolved to lower their CpG content so as to avoid innate immune recognition and the triggering of a response. Lastly, it has been seen that single-stranded RNA containing CpGs can stimulate monocytes [109]. All of these observations of an effect of nucleic acid composition at the codon-codon junction only serve to further define CPB; however these observations fail to clarify the biological effects of CPB.

The data from the Hatfield group at the time was significant because it elucidated that there was an additional phenomenon in the realm of the genetic code's redundancy. Specifically, that in addition to the well-known codon bias, there was now codon-pair bias. Also, their experiments were the only tangible evidence that demonstrated that the nucleic acid phenomenon, CPB, had an impact on the cellular function of translation.

Looking back, however, one could conclude that these experiments were limited both in the techniques available to study CPB as well as the data produced. Specifically, Irwin et al. may not have provided truly conclusive results that over-represented pairs slow translation [102]. Namely Irwin et al. changed only single pairs in an entire gene. Note: These experiments were conducted in a prokaryotic system, in which the cause and/or effect of CPB could be different from eukaryotes.

In sum, one could now combine the recent progress in whole gene synthesis with computational methods to study and manipulate codon-pair bias on the macro scale to make conclusions on CPB's effects, an impossibility to the discoverers of CPB 20 years ago. The ability to customize a gene and shuffle its codons with a computer was the catalyst of the work presented in subsequent chapters. Specifically, synthetic biology and DNA synthesis allowed for both a large-scale study of CPB, as well as manipulation of poliovirus's CPB with the goal of creating an attenuated virus suitable as a model for vaccine production.

Chapter 1 Figures

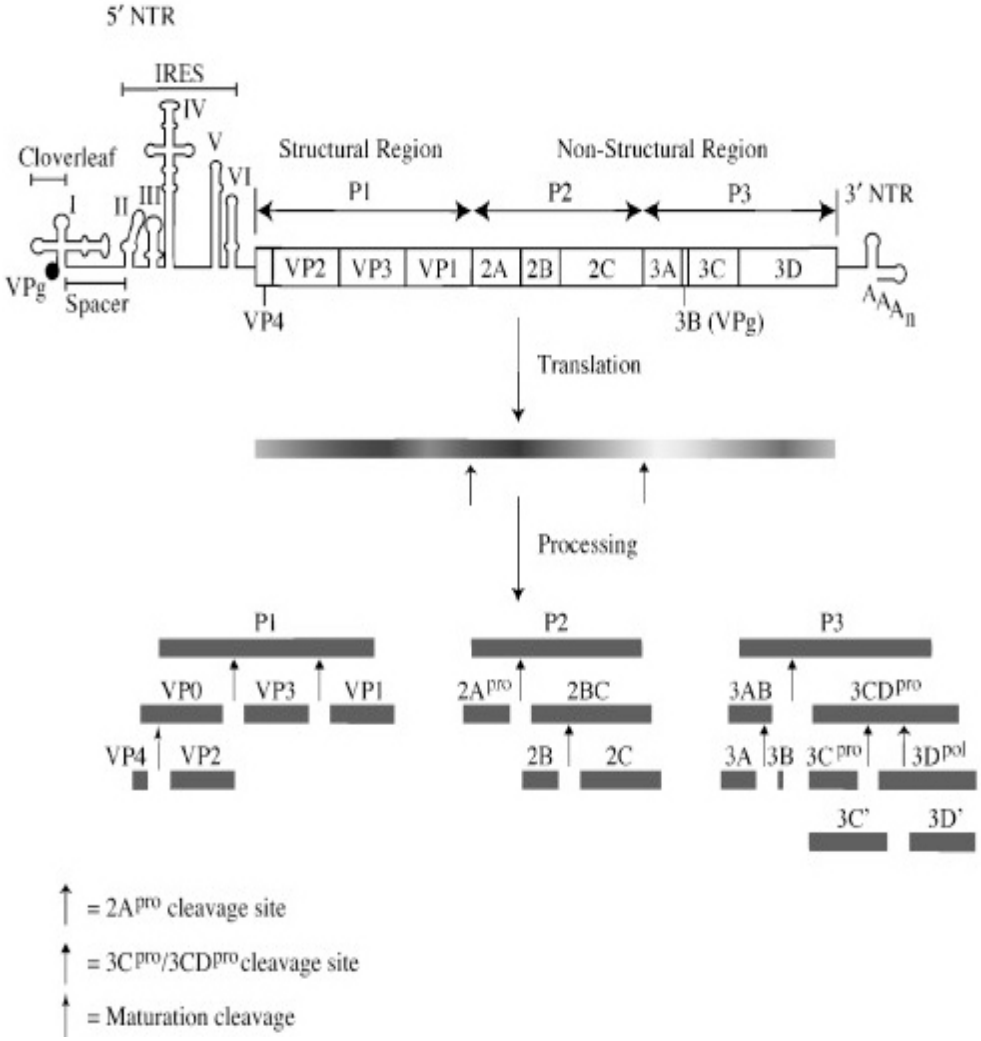


Figure 1.1: Poliovirus genome and polyprotein processing. The poliovirus genome is a positive-stranded RNA molecule with a single open reading frame. The viral genomic RNA is translated via a cap-independent mechanism by a Internal Ribosome Entry Site (IRES) in its 5'NTR. This translation event produces a single polyprotein that subsequently auto-processes via *cis* acting cleavage events with the initial event being the $2A^{pro}$ cleavage of P1. All subsequent cleavages and processing events occur at varying rates catalyzed by either protease $2A^{pro}$ or $3C^{pro}$, to yield all subsequent structural (P1) and non-structural (P2, P3) viral proteins. Figure adapted from DeJesus NH., 2007 [28].

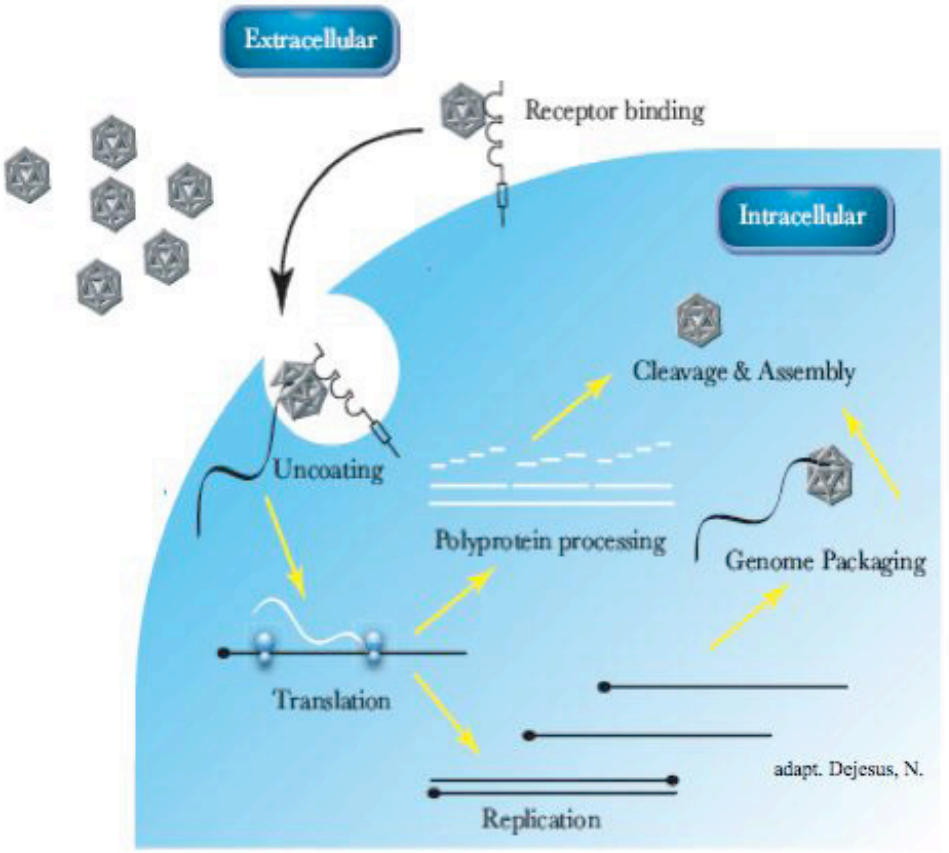


Figure 1.2: Poliovirus cellular replication cycle. Poliovirus binds the transmembrane, cell-surface protein CD155 and this serves as the virus receptor. Once bound, the virus is internalized and the virion uncoats releasing the positive-sense viral (+)RNA into the cytoplasm. Viral RNA is translated, producing a single polyprotein that subsequently auto-processes via *cis* acting cleavage events. The infecting RNA is then used as a template to make (-)RNA via the 3D polymerase. These (-)RNAs are used as a template to synthesize many (+)RNA genomes, which are used to produced additional proteins, thus amplifying the process logarithmically. This RNA replication occurs on membranous vesicles within an infected cell. PV (+)RNA is then encapsidated by structural proteins and the cell is eventually lysed, releasing progeny virus to repeat this cycle of replication.

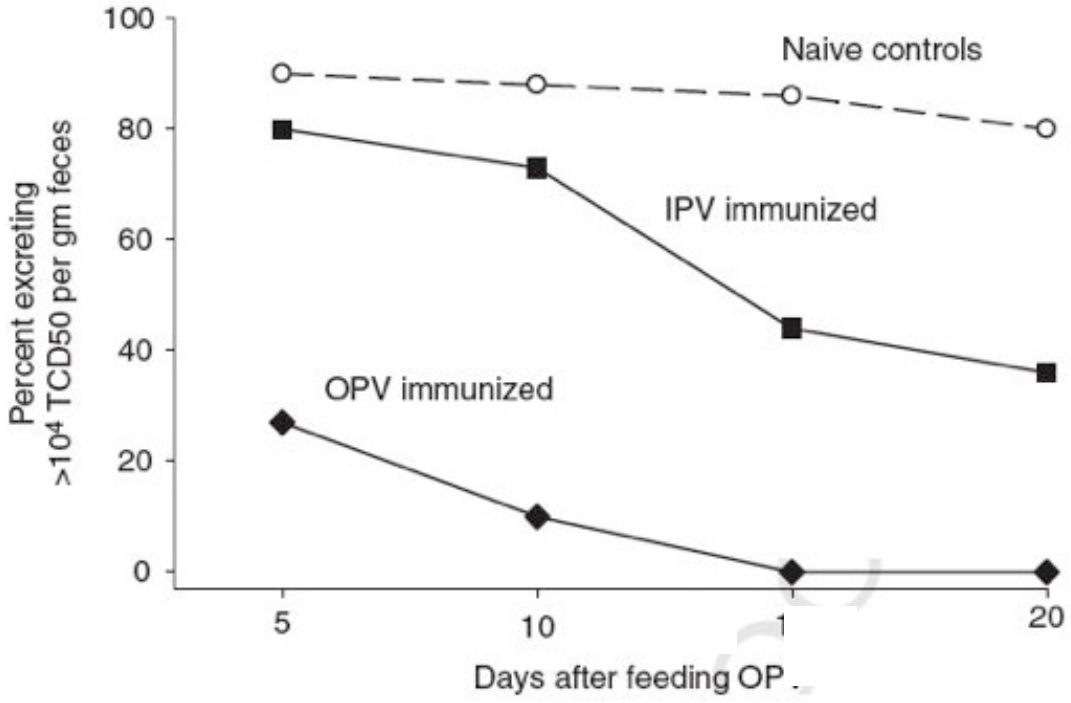


Figure 1.3: Gut immunity provided by alternate methods of poliovirus vaccination. Three groups of immunologically naïve children received either inactivated poliovirus vaccine (IPV), oral-poliovirus vaccine (OPV), or no vaccination. These children were then "challenged" with OPV and viral shedding was monitored in their feces. IPV provided substantially less protection against fecal shedding when compared to OPV, suggesting a decreased production of secretory antibodies [75].

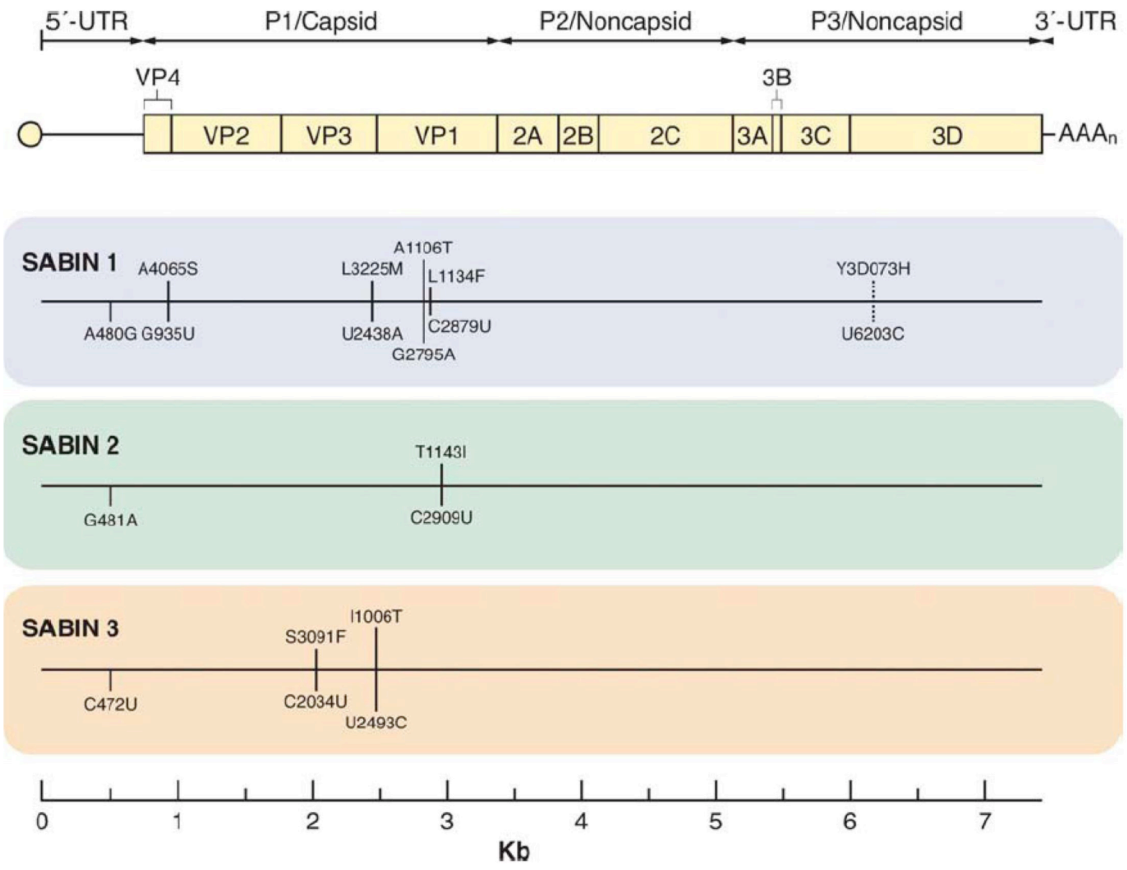


Figure 1.4: The attenuating mutations of the Sabin strains. The current oral-poliovirus vaccine (OPV) is constituted by the three Sabin serotypes displayed above. Each of these strains have nucleotide changes, however only the nucleotide changes contributing to their attenuated phenotypes are depicted. The lines extending upward indicate amino acid changes and the lines extending downward illustrate nucleotide changes. Each Sabin strain possesses a mutation in its IRES and this mutation is thought to be the most debilitating. Adapted from Kew et al. 2005 [88].

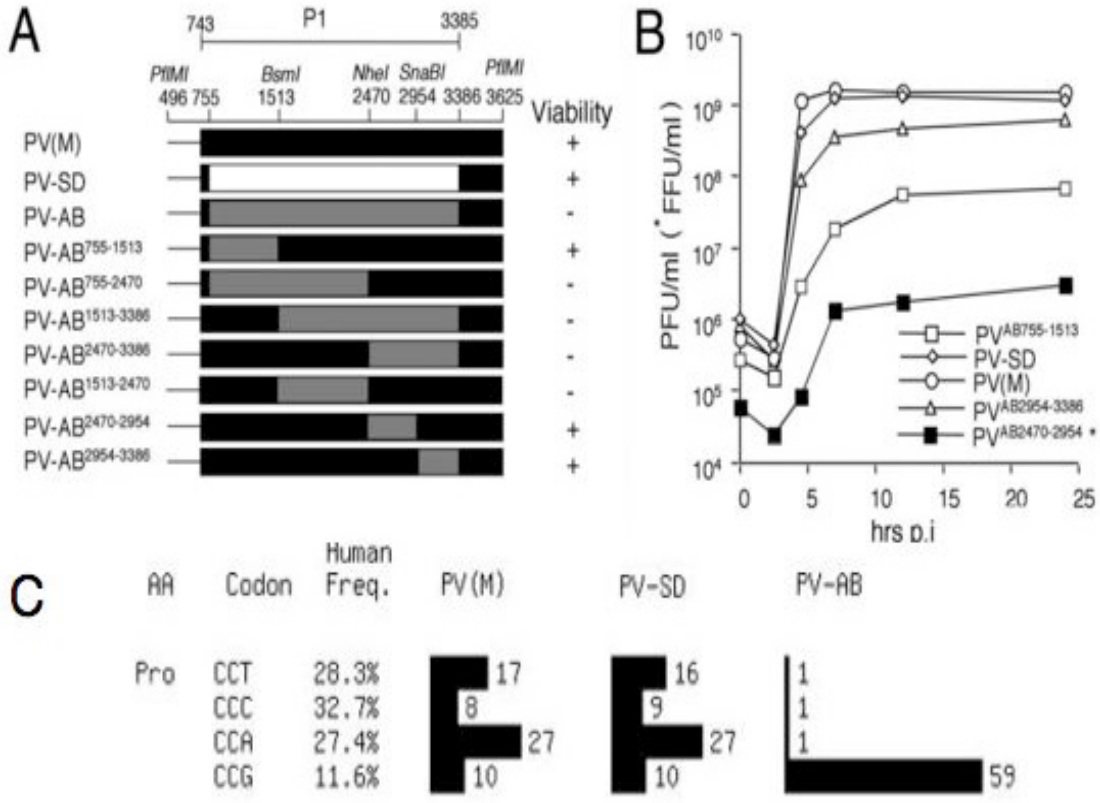


Figure 1.5: Previously synthesized poliovirus customized using synonymous rare codons which had an attenuated phenotype. In a previous study, we constructed a poliovirus whose P1 region had its wild type codons exchanged for synonymous codons that were 'rare' in the human genome [3]. (A) The PV-SD virus, had its codons shuffled but with the usage unchanged and CPB, grew with wild type characteristics. PV-AB corresponds to the rare codon virus and a full-length AB P1 failed to yield a viable virus. It was subsequently subcloned and the viability of these subclones can be seen in panel B. (B) The two most attenuated viruses, PV-AB^{2954} and PV-AB^{2470}, have a markedly attenuated phenotype as seen by the one-step growth kinetics. In addition, these viruses do not yield plaques but rather foci that are developed via immunostaining (hence *FFU/ml). This attenuation was due to a decreased specific infectivity. (C) One example of an amino acid that was changed in PV-AB to the most infrequent human codon.

Chapter 2: Customization of Codon-Pair Bias and Experimental Procedures

Definition and Calculation of Codon-pair bias (CPB)

A computational method was developed, by Dr. Dimitris Papamichail and Dr. Steve Skiena of the Computer Science Department at Stony Brook University, to quantify codon-pair bias. Using this method, first a "Codon-pair Score", or "CPS", was calculated for each of the 3,721 possible codon-pairs (excluding Stop codon-pairs) and each score was assembled into a matrix [3]. The CPS is defined as the natural log of the quotient of the observed occurrences of a given codon-pair divided by the expected number of occurrences of the codon-pair found in all human coding regions (Figure 2.1A). The observed occurrences were calculated by summing all appearances of a given codon-pair in all annotated human genes. The expected number of occurrences of a codon-pair required computation, and is the value that signifies what the pairing frequency should be if the codon-pairing was random (Figure 2.1A). This expected number was calculated such that it was independent of both amino acid frequency as well as of the frequency of each codon. The equation developed was very similar to that of Gutman and Hatfield [111]. Therefore, the expected frequency provides the frequency at which two codons should appear as a pair if their association were random. An example calculation using the codon pair AAA UGC can be seen in Figure 2.2. A positive CPS value signifies that the given codon-pair is statistically over-represented, and a negative CPS value indicates

the pair is statistically under-represented in the human genome. The calculated CPS scores for all 3,721 possible codon-pairs can be found online [3].

Using this assembled matrix of CPSs, the Codon-pair bias (CPB) for an entire open reading frame was then calculated. The CPB of a gene is defined as the arithmetic mean of the individual codon-pair scores within the gene itself (Figure 2.1B, 2.3). The CPB was then calculated for the set of 14,795 annotated human genes (CCDS data set issued by NBCI, release date March 2005) using this computational method and the two equations shown in Figure 2.1. Correlating to the values of the CPSs, a negative CPB indicates the use, on average, of under-represented codon-pairs within a gene, while a positive CPB represents the use, on average, of over-represented codon-pairs within a gene. These values of CPB also indicate the degree to which this average representation extends, such that the more negative the CPB of a gene, the greater the quantity and under-representation of codon-pairs within the gene.

Implementation of a computer-based algorithm to customize the codon-pair bias of a target sequence

Building upon the formulas above, a computer-based algorithm, designated (DPapa), was developed that could customize the CPB of a given input cDNA sequence. This program was constructed by Dr. Dimitris Papamichail and Dr. Steve Skiena of the Computer Science Department at Stony Brook University. DPapa has the ability to manipulate the CPB of any coding region while maintaining the original amino acid sequence. Also, DPapa maintains the codon usage of a gene (i.e. use only the existing codons) but "shuffles" the existing codons. This shuffling creates novel codon-pairs such

that the CPB of the gene itself can be customized, i.e. be increased or decreased (Figure 2.3). DPapa takes an input nucleic acid sequence of interest and then gives an output of a novel nucleic acid sequence that has an altered CPB. The algorithm uses simulated annealing, a mathematical process suitable for full-length optimization, to shuffle all codons [112].

Other parameters of the CPB-altered sequence were also customized by DPapa, such as the corresponding RNA folding free energy of the sequence constructed. This customization was done via an interface between DPapa and the program mFold (a RNA folding program) [113]. The free energy of the RNA corresponding to the novel sequence constructed is designed to be similar to the free energy of the RNA corresponding to the input sequence. This customization thus prevents dramatic changes in secondary structure, which could have resulted from the codon rearrangement. Also, by interfacing DPapa with mFold, the customization process is able to exclude the creation of any regions with large secondary structures, such as stable hairpins or stem loops. DPapa is computer software that takes from the user an input cDNA sequence of a given gene and then the CPB of this gene can be customized to the degree desired. Two poliovirus novel sequences were designed using DPapa both of which were based on the poliovirus type-1 Mahoney strain (PV(M)-wt) P1 region. These two sequences were designated PV-Max and PV-Min. The starting CPB of PV(M)-wt P1 region was calculated to be -0.02, while PV-Min had a CPB of -0.48, and PV-Max had a CPB of +0.25. The designed P1 PV-Max has a positive CPB and thus it utilizes, on average, over-represented codon-pairs. The designed P1 PV-Min has a negative CPB and thus it utilizes, on average, under-represented codon-pairs.

Maintenance of RNA folding energies and exclusion of large, stable secondary structures

To ensure that strong secondary structures that would effect translation efficiency were not created by the shuffling of codons, the designed P1 capsid regions were scanned using the interface between DPapa and mFold [113]. Specifically, 100-base pair long segments, having 80 base pair overlap with each other, were analyzed. Any segment encountered that had a lower binding energy than a threshold of -30 Kcal/mol would incur computer-inserted, random synonymous substitutions at C - G binding locations. The inserted synonymous changes were placed in by the program to alleviate any stable structures encountered during the scan. If a synonymous change was selected, it was done such that the CPB objective was satisfied as best as possible. In actuality, only a few changes were inserted and these changes had no impact on the CPB scores of the customized P1 regions PV-Max and PV-Min. This scan was conducted on the P1 regions of the virus PV(M)-wt, PV-SD [4], PV-Min, and PV-Max and the energy value of each 100-base pair segment was output onto a graph. PV-SD was analyzed because it had its codons shuffled, but its CPB was unchanged.

De novo DNA synthesis and construction of cDNA plasmids

Two poliovirus DNA fragments with altered CPB, PV-Min and PV-Max, corresponding to nucleotides 495 to 3636 of the PV(M)-wt genome, were synthesized *de novo* by Blue Heron Corp. (Bothell, WA) and were provided in a carrier plasmid. The first four codons (12 nucleotides) were not changed in PV-Min and PV-Max and left the same as the wild type poliovirus codons. Subsequent poliovirus cDNA clones/subclones

constructed in this work were based on the cDNA plasmid pT7PVM, which contains a full-length, infectious clone of poliovirus type-1 Mahoney (PV(M)-wt) downstream of a T7 RNA polymerase promoter [13]. The plasmids pPV-Min and pPV-Max were generated by excising the synthetic fragments from Blue Heron's carrier vector via PflM I digestion and using these fragments to replace the respective PflM I fragment taken from the pT7PVM vector. The plasmids pPV-MinXY and pPV-MinZ were constructed by simultaneous digestion of plasmids pPV-Min and pT7PVM with enzymes Nhe I and Bgl II. Then by swapping the excised fragments, the plasmids pPV-MinXY and pPV-MinZ were constructed. The plasmids pPV-MinX and pPV-MinYZ were constructed via Bsm I digestion of plasmid pPV-Min and the exchanging of fragments with a similarly digested pT7PVM. pPV-MinY was constructed by digesting pPV-MinXY with Bsm I and swapping this fragment with the Bsm I fragment from a similarly digested pT7PVM. Plasmid transformations and amplifications were all performed in *E. coli* DH5α.

In vitro transcription and transfection of infectious (+)RNA

Plasmid DNA (1.5 μg) of all viral constructs (PV-Max, PV-Min, PV-MinX, Y, Z, YZ and XY) was linearized via EcoR I enzyme digestion. The linearized DNA was then used as a template and transcribed by purified T7 RNA polymerase (Stratagene), which recognizes and is stimulated by a T7 promoter upstream of the viral cDNA. The transcription reaction occurred for 2 hours at 37°C and yielded single-stranded RNAs [13]. This transcript RNA (~5 μg) was then combined and mixed with 900 μl of a DEAE-Dextran solution (10 mg/ml Diethylaminoethyldextran in Hank's Buffered Salt Solution (HBSS) (Invitrogen). This mixture was then incubated on ice for 30 minutes.

Next, 1 x 10^6 HeLa R19 cells on 35-mm-diameter plates were washed with 1x PBS (Invitrogen) and subsequently the entire RNA/DEAE-Dextran mixture was used to inoculate these cells [13]. These cells were then incubated at room temperature (RT) for 30 minutes after which the transfection supernatant was replaced with Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen) containing 2% bovine calf serum (BCS). The cells were then incubated at 37°C and 5% CO_2 and observed (up to 4 days) for the onset of cytopathic effect (CPE).

S^{35} in vitro translation

The integrity of the P1 regions of PV-Min and PV-Max, and their ability to yield all viral proteins, was determined by performing a S^{35} *in vitro* translation reaction. First, transcript RNA was produced from the plasmids pPV-Max, pPV-Min, and pT7PVM. Specifically, these plasmids were linearized via EcoR I enzyme digestion. The linearized DNA was then used as a template and transcribed by purified T7 RNA polymerase. The transcription reaction occurred for 2 hours at 37°C and yielded single-stranded RNAs [13]. Next, these transcript RNAs were purified via phenol-chloroform extraction, ethanol precipitation, and re-suspension in 25 μl of dH_2O. RNA concentrations were measured using the NanoDrop spectrophotometer (NanoDrop Technologies). Then various concentrations of transcript RNA (100 ng, 200 ng, 500 ng) were mixed with micrococcal nuclease treated HeLa S10 cytoplasmic extracts (prepared by the Molla et al. method [114]) and incubated at 30°C for 4 hours. The HeLa S10 cytoplasmic extracts contained radio-labeled [S^{35}]methionine, which would be incorporated into proteins synthesized during the *in vitro* translation reaction. Lastly, translation products were

subjected to sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) with 12% acrylamide and then analyzed by autoradiography.

Cells and virus amplification

Throughout all experiments uninfected HeLa R19 cells were constantly maintained in DMEM containing 10% BCS, incubated at 37°C and 5% CO_2, and plates were split once confluent. Viruses were amplified on 1 x 10^8 HeLa R19 cell monolayers infected at a multiplicity of infection (MOI) of 1 or 0.1. Cells were infected by first washing the cells once with 1x PBS. Next, cells were inoculated with the appropriate virus solution and placed on a rocker at RT for 30 minutes. Then DMEM with 2% BCS was added to infected cells and these cells were then incubated at 37°C and 5% CO_2 for 3 days or until CPE was observed. Upon onset of CPE, infected plates were frozen at -80°C. To harvest amplified virus, the plates were subjected to 3 freeze/thaw cycles, cell debris was removed form the lysates via low speed centrifugation, and the supernatant containing virus was titered via plaque assay and used for further experiments.

Plaque assays and one-step growth kinetics

Plaque assays were performed on monolayers of HeLa R19 cells. These cells were infected with ten-fold serial dilutions of a given virus sample. Serial dilutions were created by placing 50 μl of the starting virus-containing solution into 450 μl of DMEM . This process was repeated to create ten-fold serial dilutions. These dilutions were then used to infect monolayers of HeLa R19 cells. Specifically, monolayers of cells were first washed once with 1x PBS, then inoculated with 333 μl of each serial dilution and placed

on a rocker at RT for 30 minutes. Next the infected cells were overlaid with a Modified Eagle Medium (MEM) (Invitrogen) solution containing 0.6% tragacanth gum (Sigma-Aldrich) and 2% BCS. Infected, and overlain plates, were then incubated at 37°C and 5% CO_2 for either 2 days for PV(M)-wt and PV-Max, or 3 days for PV-Min (X, Y, XY, or Z) viruses. These were then developed via crystal violet staining and the Plaque Forming Units per milliliter (PFU/ml) titer was calculated by counting visible plaques and multiplying by the dilution factor.

One-step growth curves were obtained by infecting a monolayer of 1 x 10^6 HeLa R19 cells (on individual 35-mm plates) with a MOI of 2 of a given virus. Specifically, cells were washed once with 1x PBS and then inoculated with 2 MOI of virus, placed on a rocker, and incubated at RT for 30 minutes. Then, inoculums were removed and cells were washed twice with 1x PBS. Lastly, DMEM containing 2% BCS was added to infected cells and these cells were then incubated at 37°C and 5% CO_2 for 0, 2, 4, 7, 10, 24, and 48 hours. At the appropriate time points, the corresponding 35-mm plate was frozen at -80°C. Virus produced at these time points was then quantified by plaque assays on HeLa cell monolayers. The low MOI growth curves were produced in a similar fashion, except the initially infecting MOI was reduced to 0.001 MOI.

Capsid heat stability and serial passaging

The thermal stability of the synthetic viruses, PV-MinXY and PV-MinZ, was tested and compared to the wild type virus PV(M)-wt. This was done by heating 1 x 10^8 particles of the respective viruses suspended in 1x PBS to 50°C for 5, 15, 30, and 60

minutes. Then to measure residual infectivity, plaque assays were performed to calculate the decrease of infectious virus at the different time points.

In order to test the genetic stability of the synthetic portions of the P1 region of PV-MinXY and PV-MinZ, these viruses were serially passaged and their genomes were sequenced at given intervals. Briefly, monolayers of 1 x 10^6 HeLa R19 cells were infected with a MOI of ~0.5 (or 100 μl from a passage) of the virus PV-MinXY or PV-MinZ. Cells were infected by first washing the cells once with 1x PBS. Next, cells were inoculated with the appropriate virus solution and placed on a rocker at RT for 30 minutes. Then DMEM containing 2% BCS was added to infected cells and these cells were then incubated at 37°C and 5% CO_2 until CPE was observed. Upon onset of CPE, infected plates were frozen at -80°C. To harvest virus, the plates were subjected to 3 freeze/thaw cycles, cell debris was sedimented from the lysates via low speed centrifugation, and the supernatant containing virus was used to infect a fresh monolayer of 1 x 10^6 HeLa R19 cells. This cycle was repeated 17 times for virus PV-MinXY and 19 times for virus PV-MinZ. Throughout passages, the time interval until CPE developed was noted and remained constant. The nucleotide sequences of the synthetic regions in the viruses that emerged after passages 5, 9, 15, 17(PV-MinXY), and 19(PV-MinZ) were determined via reverse transcriptase PCR (RT-PCR) and sequencing. Also, the titers of the viruses PV-MinXY and PV-MinZ after passages 5, 9, 15, 17, and 19 were determined via plaque assay.

Reverse Transcriptase-PCR and sequencing

To sequence the viruses studied, such as from the passaging experiments, the protocol below was used. Specifically, 250 μl of virus-containing supernatant was gently mixed with 750 μl of the Trizol® Reagent (Invitrogen) and incubated for 5 minutes at RT. Then, 200 μl of chloroform was added and the solution was mixed and incubated at RT for 5 minutes. Next, the solution was centrifuged at 12,000 x *g* for 15 minutes at 4°C and the top supernatant was placed in a fresh tube. Subsequently, this solution containing viral RNA was subjected to ethanol precipitation and the RNA pellet was resuspended in 25 μl dH_2O. This RNA was then converted to DNA via the Titan One-Tube Reverse Transcriptase-PCR (RT-PCR) system (Roche). The DNAs produced from this RT-PCR reaction were then subjected to 1.0% agarose gel electrophoresis to purify it for sequencing. The slices of the gel containing the corresponding DNA were excised, melted, the DNA was ethanol-precipitated, and resuspended in 25 μl dH_2O. Then 100 ng of the resuspended DNA was used in sequencing reactions conducted by the Stony Brook University DNA Sequencing Facility to examine viral sequences.

Virus purification and determination of viral particles via A_{260} absorbance

First, viruses PV(M)-wt, Sabin-1, PV-Max, PV-MinX, PV-MinY, PV-MinXY or PV-MinZ were amplified on monolayers of 1 x 10^8 HeLa R19 cells infected at a multiplicity of infection MOI of 0.5-1. Cells were first washed with 1x PBS. Next, cells were inoculated with the appropriate virus solution and placed on a rocker at RT for 30 minutes. Then DMEM containing 2% BCS was added to infected cells and these cells were then incubated at 37°C and 5% CO_2 until CPE was observed. Upon onset of CPE,

infected plates were frozen at -80°C. To harvest amplified virus, the plates were subjected to 3 freeze/thaw cycles**.** After 3 freeze/thaw cycles, the cell lysates were subjected to 2 centrifugations to remove cell debris. The first centrifugation was at 3,000 x *g* for 15 minutes and the supernatant was then transferred to a new tube. Second, the virus-containing supernatant was subject to a centrifugation at 10,000 x *g* for 15 minutes. After these two low speed centrifugations, RNase A (15 μg, total) was added to the supernatants to destroy free RNA and the samples were incubated at RT for 1 hour. Next, the supernatants were gently mixed with a final concentration of 0.5% SDS and a final concentration of 2 mM ethylenediaminetetraacetic acid (EDTA) and incubated at RT for 30 minutes [4]. Only properly formed, intact virus particles survive this treatment [25]. These supernatants were then overlaid onto a 6-ml sucrose cushion [30% sucrose dissolved in Hank's Buffered Salt Solution (HBSS)] and virus particles were sedimented by ultracentrifugation for 3.5 hours at 28,000 rpm using a SW28 swing-bucket rotor. After centrifugation, the supernatant above the sucrose cushion was removed and the tube was washed twice with HBSS without disturbing the sucrose cushion. The sucrose was then gently removed and the virus pellet at the bottom was resuspended in 1x PBS containing a final concentration of 0.01% SDS [4]. Viral Plaque Forming Units per milliliter (PFU/ml) titers of the purified suspensions were determined via plaque assay as previously described. Virus particle concentrations were determined via the average of three absorbance measurements at 260 nm (A_{260}). These measurements were then converted to particle concentrations by employing the formula 9.4×10^{12} poliovirus particles/ml = 1 A_{260} unit [4,115]. The A_{260} was measured using the NanoDrop spectrophotometer (NanoDrop Technologies).

Virion RNA - quantity and length

The quantity and quality of the virion RNA contained within the synthetic virus PV-MinZ was compared to PV(M)-wt to observe similarity. PV(M)-wt and PV-MinZ were amplified on a monolayer of HeLa R19 cells (1 x10^8 cells) on a 15-cm dish, purified over a sucrose gradient as previously described, and then resuspended in 1x PBS with a final concentration of 0.01% SDS. The final particle concentrations of the resuspended virus solutions were calculated via A_{260} using the formula 9.4 x 10^{12} poliovirus particles/ml = 1 A_{260} unit [4,115]. The A_{260} was measured using the NanoDrop spectrophotometer. Equal quantities of viral particles (3 x 10^{11}) for PV(M)-wt and PV-MinZ were then placed into 300 μl of 1x PBS. The quantity of 3 x 10^{11} particles was used because the virion RNA from 3 x 10^{11} particles = 1 μg of RNA and is thus visible on an agarose gel [116]. The genomic RNA from the particles of either viruses were phenol-chloroform extracted, ethanol-precipitated and resuspended in 20 μl of dH_20. All 20 μl was subjected to a 1.0% TAE agarose gel stained with ethidium bromide. The band intensities were calculated using the NIH:ImageJ program (National Institute of Health).

Quantitative RT-PCR

To analyze the *in vivo* stability of the various viral RNAs (PV(M)-wt, PV-MinXY, and PV-MinZ), monolayers of 1 x 10^6 HeLa R19 cells on 35-mm-diameter plates were infected at a MOI of 5. Cells were first washed with 1x PBS. Next, cells were inoculated with the appropriate virus solution and placed on a rocker at RT for 30 minutes. Then, DMEM containing 2% BCS was added to infected cells and these cells

were then incubated at 37°C and 5% CO_2 for 3 hours. After 3 hours, a final concentration of 4 mM guanidine hydrochloride (GuHCl) was added to the supernatant and then infected cells were returned to the incubator. GuHCl is a potent and specific inhibitor of poliovirus RNA replication [117]; thus only translation was occurring. At times post the addition of GuHCl (0, 15, 30, 60 minutes), the appropriate plates were washed once with HBSS. Then, cell monolayers were lysed using 950 μl Trizol® Reagent (Invitrogen) to isolate total RNA. This solution was incubated at RT for 5 minutes. Next, 200 μl of Chloroform was added and the total RNA solution was mixed and incubated at RT for 5 minutes. The solution was spun at 12,000 x *g* for 15 minutes at 4°C and the top layer was placed in a fresh tube. Subsequently, this solution containing total RNA was subjected to ethanol-precipitation and the RNA pellet was resuspended in 30 μl dH_2O. The concentration of the RNA was measured using the NanoDrop spectrophotometer. In addition, a small aliquot of each was subjected to a 1.0% TAE agarose gel stained with ethidium bromide to confirm quality. Then, using equal quantities of total RNA (200 ng), the number of poliovirus genomes within each sample was measured using SYBR Green I Q RT-PCR (Roche) and primers listed below. The quantity of poliovirus genomes contained within the 200 ng of total RNA was measured by comparing the illumination to that of a control sample. The control was transcript RNA that was run simultaneously in serial ten-fold dilutions (100 fg, 10 pg, 100 pg, 1 μg, 10 μg). Each dilution of this control transcript RNA was supplemented (to equal 200 ng) with non-infected cellular RNA so as to mimic experimental sample conditions. The same primers were used to quantify control RNA.

Oligonucleotides used for Quantitative RT-PCR

PCRE1 (+): 5'GGTAAATCTGTAGCAACCAACC3'

PCRE6 (-): 5'GCACACAGTGACGCGTTAGCC3'

Dicistronic reporter construction, and in vivo translation

The dicistronic reporter replicons were all constructed based upon pdiLuc-PV [4]. PV-Max and PV-Min capsid regions were amplified from pPV-Max and pPV-Min via PCR using the oligonucleotides: P1max-2A-RI (+)/P1max-2A-RI (-) or P1min-2A-RI (+)/P1min-2A-RI (-), respectively (refer to page 44). The DNA produced from these PCR reactions were then subjected to 1.0% agarose gel electrophoresis. To isolate and purify the DNA, slices of the gel containing the corresponding DNA were excised, melted, and the DNA was ethanol-precipitated, and lastly resuspended in 25 μl dH_2O. The PCR fragments were then each inserted into an intermediate vector pCR-®-XL-TOPO® (Invitrogen). Two plasmid pTOPO-PVMin and pTOPO-PVMax were thus constructed. These intermediate vectors were then amplified in One Shot® TOP10 chemically competent cells (Invitrogen). After purification using miniprep columns (Qiagen), the intermediate vectors containing pTOPO-PVMin and pTOPO-PVMax were digested with EcoR I and these fragments were ligated into the equally digested EcoR I pdiLuc-PV[4]. These plasmids were transformed into One Shot® TOP10 chemically competent cells. This created the plasmids pdiLuc-PV-Max and pdiLuc-PV-Min. To construct pdiLuc-PV-MinXY and pdiLuc-PV-MinZ, pdiLuc-PV and pdiLuc-PV-Min each were digested with NheI and BglII and the resulting restriction fragments were

exchanged between the respective vectors. These were then transformed into One Shot® TOP10 chemically competent cells and amplified.

To analyze the *in vivo* translation efficiency of the synthetic regions encoding the poliovirus capsid, transcript RNA from the plasmids pdiLuc-PV, pdiLuc-PV-Max, pdiLuc-PV-Min, pdiLuc-PV-MinXY, and pdiLuc-PV-MinZ was transfected into 2 x 10^5 HeLa R19 cells on 12-well dishes. First, plasmid DNA (1.5 μg) of all reporter constructs was linearized via PvuI enzyme digestion. The linearized DNA was then used as a template and transcribed by purified T7 RNA polymerase. The transcription reaction occurred for 2 hours at 37°C and yielded single-stranded dicistronic RNAs [13]. The RNAs of the dicistronic reporter constructs each were transfected into 2 x 10^5 HeLa R19 cells on 12-well dishes using Lipofectamine 2000 (Invitrogen). Specifically, these cells were washed once with 1x PBS. These cells were then inoculated with transcript RNA in 400 μl of DMEM with 8 μl of Lipofectamine 2000 and then incubated at RT for 45 minutes on a rocker. After this incubation, the transfection supernatant was replaced with DMEM containing 2% BCS and a final concentration of 2 mM GuHCl. GuHCl was added in order to quantify the translation of input RNA only. These cells were then incubated at 37°C and 5% CO_2 for 6 hours. Next, cells were lysed in 200 μl of passive lysis buffer (Promega) and these lysates were analyzed by a dual firefly (F-Luc) Renilla (R-Luc) luciferase assay (Promega) using an illuminometer.

Oligonucleotides used for dicistronic plasmid construction

The following oligonucleotides were utilized to perform PCR in the construction of the dicistronic reporter construct:

P1max-2A-RI(+):

5’CAAGAATTCCTGACCACATACGGAGCTCAAGTATCTTCACAAAAAGTTGG3’

P1max-2A-RI(-):

5’TTCGAATTCTCCGTACGTGGTGAGGTCTTTGGTGGACAAAGG3’

P1min-2A-RI(+):

5’CAAGAATTCCTGACCACATACGGAGCTCAGGTGTCATCCCAAAAAGTAGG

3’

P1min-2A-RI(-):

5’TTCGAATTCTCCGTACGTCGTAAGGTCTTTCGTTGACAGTGG3’.

Neuropathogenicity of synthetic viruses in CD155tg mice and sequencing of extracted virus

To determine neuropathogenicity, groups of four to six CD155tg mice (Tg21 strain) between six and eight weeks of age were injected intracerebrally with purified virus dilutions in 30 μl PBS ranging from 10^2 particles to 10^9 particles [14]. The fifty percent lethal dose (LD_{50}) was calculated according to the method by Reed and Muench [118] to assess pathogenicity. Virus from mice that succumbed to disease was titered and sequenced. Specifically, mice that succumbed to poliomyelitis, were dissected and there spinal cords were removed. Next, the spinal cords were individually weighed, homogenized, and resuspended in 200 μl of DMEM. These tubes were then freeze-thawed three times to ensure viral release from tissue. Virus titers in spinal cord tissues were determined by plaque assay. To sequence viruses from the spinal cords they were

first amplified on 1 x 10^6 HeLa R19 cells and then sequenced using the previously described method.

Testing of PV-MinXY and PV-MinZ as vaccines in CD155tg mice and the antibody response they elicit

To test PV-MinZ and PV-MinXY as vaccines, three doses (10^8 particles in 100 μl of 1x PBS) of these viruses were administered to CD155tg mice between six and eight weeks of age via intraperitoneal injection once a week for three weeks. In parallel, a set of control mice received three mock vaccinations with 100 μl 1x PBS. One week after the final vaccination, 30 μl of blood was extracted from the tail vein and subjected to low-speed centrifugation after which the serum was harvested [119]. Neutralizing antibodies against PV(M)-wt in these sera were analyzed via microneutralization assays performed according to the recommendations of World Health Organization [120]. Specifically, monolayers of 2 x 10^4 HeLa R19 cells were seeded on 96-well plates. These cells were then washed twice with 1x PBS. Next, 200 μl of warm DMEM was added to the cells along with serial dilution of immunized mice serum, which possibly contained anti-PV antibodies. After 5 minutes at RT, 100 plaque forming units (PFU) of PV(M)-wt was added to each well and then these 96-well plates were incubated at 37°C and 5% CO_2 for 6 hours. After six hours, each well was supplemented with 2% BCS. The wells of this 96-well plate were monitored for four days for the onset of CPE. Two weeks after the final intraperitoneal inoculation, the vaccinated and control mice were challenged with 10^6 PFU (a lethal dose) of PV(M)-wt by intramuscular injection and then monitored daily for the onset of paralysis [119]. All experiments utilizing CD155tg mice

were undertaken in compliance with Stony Brook University's IACUC regulations as well as federal guidelines.

Viral titer in infected CD155tg mice tissue

The viral titers in the tissues of infected mice were determined by first infecting mice via intraperitoneal injection with a lethal dose, 10^9 particles, of the respective viruses (PV(M)-wt, PV-MinXY and PV-MinZ). Then after the onset of paralysis, infected mice were sacrificed via cervical dislocation. Next, these mice were dissected, removing various organs including the brain, spinal cord, liver, kidney, lung, and pancreas. After extraction, all organs and tissue were weighed in order to calculate the viral yield per gram of organ tissue. These organs were next homogenized and then resuspended in 400 μl of DMEM. These tubes were then freeze-thawed three times to ensure viral release from tissue. Plaque assays were performed using homogenates and then viral particle concentrations were determined by multiplying PFU by specific infectivity.

Supporting Information

The GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) accession numbers for the codon-pair bias sequences presented are: P1-Max (EU095952) and P1-Min (EU095953).

Chapter 2 Figures

A

$$CPS = \ln\left(\frac{F(AB)_O}{\frac{F(A) \times F(B)}{F(X) \times F(Y)} \times F(XY)}\right)$$

B

$$CPB = \sum_{i=1}^{k} \frac{CPS_i}{k-1}$$

Figure 2.1: Equations used to determine codon-pair scores (CPS) and the codon-pair bias (CPB) of an entire open reading frame. (A) The equation used to calculate the CPS of a given codon-pair independent of codon usage and amino acid bias, thus its relative representation. The codon pair AB encodes for the amino acid pair XY and F denotes frequency (number of occurrences in all confirmed human genes). The CPS score for a given pair determines if the pair is over-represented (+) or under-represented (-) in the human genome. The equation normalizes for amino acid frequency as well as codon frequency. (B) The equation used to calculate the CPB for an entire gene. CPB is the arithmetic mean of the individual codon-pair scores of all pairs making up an ORF.

A

UUU Phe	UCU Ser	UAU Tyr	UGU Cys
UUC Phe	UCC Ser	UAC Tyr	UGC Cys
UUA Leu	UCA Ser	UAA *Stop*	UGA *Stop*
UUG Leu	UCG Ser	UAG *Stop*	UGG Trp
CUU Leu	CCU Pro	CAU His	CGU Arg
CUC Leu	CCC Pro	CAC His	CGC Arg
CUA Leu	CCA Pro	CAA Gln	CGA Arg
CUG Leu	CCG Pro	CAG Gln	CGG Arg
AUU Ile	ACU Thr	AAU Asn	AGU Ser
AUC Ile	ACC Thr	AAC Asn	AGC Ser
AUA Ile	ACA Thr	AAA Lys	AGA Arg
AUG Met	ACG Thr	AAG Lys	AGG Arg
GUU Val	GCU Ala	GAU Asp	GGU Gly
GUC Val	GCC Ala	GAC Asp	GGC Gly
GUA Val	GCA Ala	GAA Glu	GGA Gly
GUG Val	GCG Ala	GAG Glu	GGG Gly

B

The **CPS** for every given codon-pair = **ln (observed/expected)**

Observed Frequency of AAA (1) UGC (2) in all confirmed human genes = **1678**

Expected Frequency is calculated to be:

$$\frac{\text{(Codon 1 Frequency X Codon 2 Frequency)}}{\text{(Amino Acid 1 freq. x AA 2 freq.)}} \times \text{Amino Acid Pair Freq.} = 2156$$

THIS EQUATION NORMALIZES FOR INDIVIDUAL CODON FREQUENCY

Actual Bias:

ln(observed/expected)

ln(1678/ 2156) = **-0.118** = **Codon-Pair Score for** AAA (1) UGC (2)

Figure 2.2: A specific codon-pair example to demonstrate how a CPS is calculated. (A) The amino acid pair K-C can be encoded by four possible codon-pairs. The individual CPS for AAA UGC [K-C] was calculated using Figure 2.1A. This same calculation was conducted for all possible codon-pairs in the human genome and assembled into a matrix. (B) The Observed Frequency is the summed appearance of this pair in all annotated human genes. The Expected Frequency is the calculated value if this pairing was simply random. The Expected Frequency normalizes for codon and amino acid frequencies. The **CPS** in this example was determined to have a negative value, indicating this codon-pair is statistically under-represented.

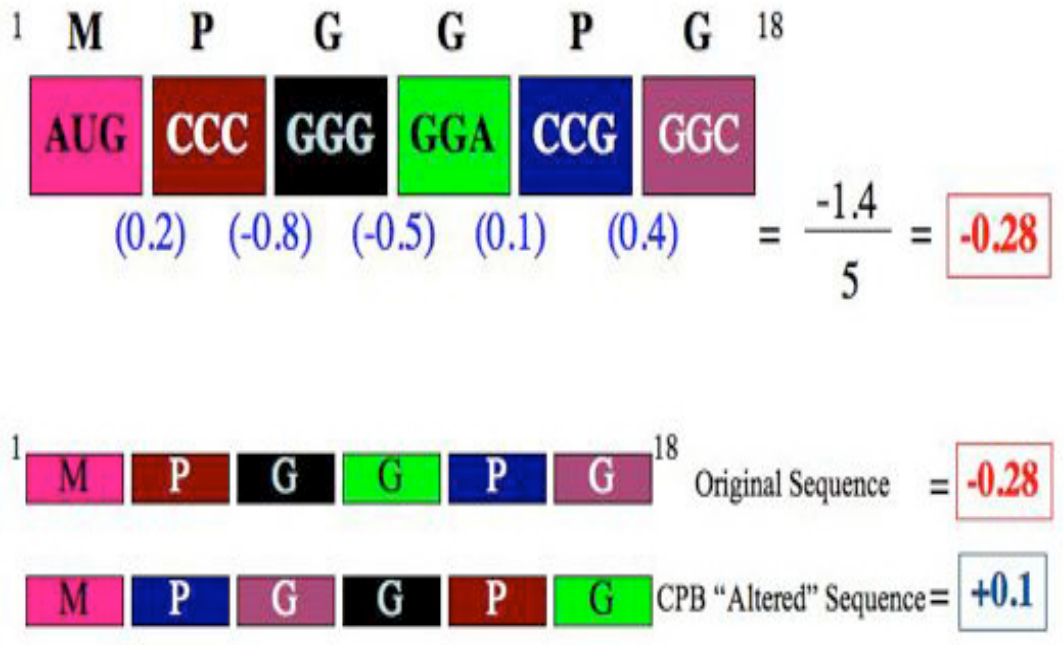


Figure 2.3: Customization of Codon-Pair Bias (CPB). (Top) A hypothetical polypeptide illustrating how CPB is determined for a gene. The values in **Blue** are the individual CPS for each pair as determined by the equation illustrated in Figure 2.1A. The **Red** value is the CPB for this gene as determined by the equation illustrated in Figure 2.1B. (Bottom) The theoretical output of the CPB customizing program, DPapa. This algorithm maintains the individual codon frequency and amino acid frequency, yet produces a sequence with novel pairs. In this hypothetical example, these novel pairings were able to change the CPB from (-0.28) to (+0.1). This indicates the original gene used, on average, under-represented codon-pairs (i.e. the negative score) and now the DPapa output gene, uses over-represented codon-pairs (i.e. the positive score).

Chapter 3: Synthetic Virus Construction *In silico, In vitro,* and *In vivo*

Summary

A computer-based program able to customize a gene's codon-pair bias (CPB) was developed by the Computer Science Department at Stony Brook University. The program, designated DPapa, has the ability to manipulate the CPB of any coding region while maintaining the original amino acid sequence. Additionally, DPapa maintains the codon usage of a gene (i.e. uses only the 'wild type' codons) but "shuffles" these original codons. This shuffling creates novel codon-pairs such that the CPB of the gene itself is altered. DPapa was implemented yielding two cDNA sequences corresponding to the P1 structural region of poliovirus. The two novel encodings of P1 now contained either over-represented (**PV-Max**) or under-represented codon-pairs (**PV-Min**). These novel sequence were produced via large-scale DNA synthesis (by Blue Herron Corp.), and these re-encoded P1s were incorporated into full-length poliovirus genomes via molecular cloning [13]. Then using reverse genetics, the virion form of these novel chimeric viruses was produced. PV-Max grew to the same plaque forming unit titer as PV(M)-wt virus, whereas the viral construct PV-Min had a null-phenotype, even after four blind passages.

Introduction

Calculation of Codon-Pair Bias (CPB)

Codon-pair bias (CPB), which is entirely independent of the codon frequency, is the observed fact that certain codons, corresponding to two amino acids, are found adjacent to one another with frequencies either less or more than expected if these codons were randomly placed next to one another. This genetic phenomenon was first observed and analyzed in select open-reading frames of *E. coli* [100] and now the phenomenon of CPB has been observed in the genes of many organisms including humans [3].

Prior to calculating a gene's CPB, one must first quantify the representation of all possible codon-pairs. For example, the codon-pair Lysine (K, AAA) and Serine (S, AGC) can be used to demonstrate how the representation of each pair within the human genome is calculated. The representation of this specific pair (AAA AGC) is calculated and given a codon-pair score (CPS), which defines the over- or under-representation of this specific pair within the annotated genes of the human genome. The CPS is determined by the natural log of the quotient of the Observed Frequency of this codon-pair divided by the Expected Frequency (Figure 2.1A). The Observed Frequency is simply the total number of appearances of this specific pair (AAA AGC) in all annotated human open reading frames. In turn, the Expected Frequency is the predicted quantity of this pair in all human genes one would expect if all codon pairings were completely random and independent of amino acid and codon frequency, i.e. if no bias existed (Figure 2.1A). The calculation of this Expected Frequency normalizes for both the amino acid frequency for each codon as well as the individual codon frequency to ensure accurate determination of the codon's statistical representation (Figure 2.1A). The

natural log was used in Figure 2.1A because the mathematical sum of logs is equivalent to multiplying each pair, yet the equation yields values that are readily combinable. The output value of the equation illustrated in Figure 2.1A yields the CPS for the individual codon. A negative CPS indicates that the pair is statistically under-represented, where as a positive CPS indicates the pair is statistically over-represented. This example codon-pair (AAA AGC) is illustrated in detail in Figure 2.2. The calculation of this CPS was completed for all possible human pairs and compiled into a matrix.

In order to then calculate the CPB for an individual gene, the matrix of CPSs was used. The CPB of a given gene was calculated by taking the sum (Σ) of all the codon-pair scores that comprise said gene and then dividing by the number of pairs present, i.e. taking the average 'score' for the gene (Figure 2.1B). Therefore, if the gene was then determined to have a negative CPB, it was comprised of, on average, under-represented codon-pairs. The inverse was such that if the gene has a positive CPB, it utilizes, on average, over-represented codon-pairs. In sum, using the calculated scores for every possible pair there was now a means to quantify a gene's CPB. (Chapter 2, Figure 2.1).

Since there was a means to calculate CPB, it could in turn be altered such that a researcher could manipulate a gene's CPB as one sees fit using appropriate computer-based algorithms. The program DPapa was capable of such customization and allows for the shuffling of the codons within a gene. This shuffling was done such that new codon pairings were formed that have novel CPSs. The formation of these novel pairings would therefore alter the genes CPB (Figure 2.3). Note: By shuffling a gene's existing codons to create the novel pairs, any downstream effects of this alteration will only be the result of these novel pairings and not due to effects of codon frequency.

DPapa utilized simulated annealing to achieve this *in silico* manipulation of a gene's CPB. The construction of this program was interesting in its own right; however DPapa was necessarily combined with other biological disciplines in order to study the biological impact of CPB. The model system chosen was poliovirus (PV), which has an established reverse genetics protocol, allowing a researcher to go from RNA to DNA to RNA to the virus, where the first step is superfluous because the cDNA is now chemically synthesized [13]. Additionally, poliovirus was the first organism shown to be completely synthesized *de novo*, thus the protocol for progression from synthesized DNA to viable virus has been established [8,114]. Using this appropriate model system allowed for an accurate study of CPB since poliovirus's genetics and molecular biology have been extensively studied. Also, since PV is in fact a pathogen still afflicting human populations, attenuating it through manipulation of its CPB could yield a synthetic poliovirus, which could serve as a model for the production of vaccines.

Combining disciplines to study the genetic phenomenon CPB and its applicability to vaccine construction

The implementation of this computer-based manipulation of CPB into a functioning biological entity, such as a virus, was only possible because of two other experimental arenas available as tools in the poliovirus model system, namely synthetic biology and reverse genetics. As described in Chapter 1, recent advances in the ability to synthesize large DNA molecules have enabled the macro scale manipulation of a gene's CPB via codon shuffling. The use of previous techniques for genetic manipulation, such as site-directed mutagenesis, would be impractical to utilize to incur said changes, i.e. for

the insertion of 600 synonymous mutations. However, now using the available arena of DNA synthesis, CPB-altered sequences designed on a computer can simply be submitted electronically to a biotechnology company and the corresponding DNA synthesized for experimentation.

In order to use the CPB-altered genetic material created via DNA synthesis, this study also relied on the established reverse genetics of poliovirus. A modification of the reverse genetics system allows for the progression from synthetic cDNA encoding the poliovirus genome to the virus itself. The initial experiments demonstrating this point first constructed, via reverse transcription *in vitro,* a DNA expression plasmid containing the full-length poliovirus cDNA [121]. Subsequently, this cDNA expression plasmid was transfected into poliovirus-susceptible mammalian cells and after a two-day incubation, the infectious virus was produced [122]. This initial advance using the transfection of DNA was a success, yet it proved inefficient. Improvements were made to increase the 'infectability' of this genetic material such that it was easier to initiate the replication cycle in mammalian cells and thus increase the yield of infectious virus following transfection. Namely, a ϕ10 promoter for the T7 RNA polymerase was cloned into the region proceeding the poliovirus genome in the plasmid. This promoter allowed for the production of (+) poliovirus-RNA *in vitro* from the cDNA template plasmid when incubated with purified T7 RNA polymerase. The RNA produced *in vitro* was also infectious and could more efficiently establish a cycle of replication [13].

Therefore, the two disciplines of synthetic biology (DNA synthesis) and an established and efficient reverse genetics system were combined with the computer program DPapa to manipulate portions of the poliovirus coding sequence. This

interdisciplinary collaboration allowed for this study to investigate the phenomena of CPB and especially the applicability of CPB-alteration to vaccine construction.

Results

Large-scale manipulation of CPB in silico

DPapa has the critical ability to maintain the **exact original amino acid sequence** as well as the codon usage of a gene **(i.e. use only the codons that are already present in the gene**) but "shuffle" the codons so that the CPB for the gene can be increased or decreased. Shuffling is done via the process of simulated annealing, which is a mathematical process able to most closely achieve full-length optimization [112]. DPapa was given the input of the DNA coding sequence for the P1 structural region of poliovirus type-1 Mahoney strain (PV(M)-wt), and the CPB of this region was customized (Figure 3.1). After customization using simulated annealing, DPapa produced two novel segments with shuffled codons and thus novel codon-pairs. The first design was PV-Max, which had a positive CPB of +0.246. This positive CPB indicated that PV-Max utilized, on average, over-represented codon-pairs. The second design was PV-Min, which had a negative CPB of -0.474 (Table 3.1). This negative CPB indicated that PV-Min utilized, on average, under-represented codon-pairs (Figure 3.1, 3.2). The first four codons (12 nucleotides) were not changed in PV-Min and PV-Max and left the same as the wild type poliovirus codons. This was done so that translation might initiate equally in all constructs. Also, the P1 region was selected as a suitable region for replacement by a synthetic fragment because it has been demonstrated that the

elimination or replacement of this region with foreign, non-polio derived sequence does not affect replication. This tolerance of foreign sequence in the P1 region strongly suggests the absence of replication regulatory elements within this region [4,95,96].

The computer alteration of the poliovirus P1 region resulted in two novel sequences with extreme CPBs. PV-Max possessed 566 silent mutations when compared to PV(M)-wt and PV-Min possessed 631 silent mutations (Figure 3.1, 3.2). The large number of synonymous mutations resulting from CPB customization is one advantage of this method when attempting to apply CPB-alteration to vaccine production. This is an advantage because it is highly improbable a virus could revert all mutations in order to alleviate their influence (refer to Chapter 3 discussion). This assumes that all mutations have a reasonably similar effect on the virus, a conclusion yet to be confirmed. The calculated CPB scores, using the equation in Figure 2.1B, of the novel P1 regions as well as of all annotated human ORFs can be observed in Figure 3.2. The designed PV-Max and PV-Min had extreme CPBs, +0.246 and -0.474 respectively, which exceeded the outer limit of naturally occurring human sequences (Figure 3.2). This extreme allowed for an enhanced study of the effects of CPB. The CPB of all constructs examine din this and previous studies can be seen in Table 3.1

In addition to the shuffling of codons, other parameters of the CPB-altered sequence can be customized by DPapa, such as the corresponding RNA folding free energy to the sequence constructed. This customization was done via an interface between DPapa and the program mFold (an RNA folding program) [113]. The free energy of the RNAs corresponding to the novel cDNA sequences constructed (i.e. PV-Max) were designed to be similar to the free energy of the RNA corresponding to the

input cDNA sequence (i.e. PV(M)-wt). This customization thus prevented dramatic changes in secondary structure, which could have resulted from the codon rearrangement. Also, by interfacing DPapa with mFold the customization process was able to exclude the creation of any regions with large secondary structures, such as stable hairpins or stem-loops. Specifically, DPapa scanned the RNA in overlapping 100-base pair segments for stable secondary structures with a free energy greater than -30ΔG. If a stable structure was encountered, the DPapa replaced C – G pairings to alleviate folding. The synonymous changes were selected to maintain CPB customization. This scanning was important because it has previously been shown that stable stem-loop structures within an ORF are able to prevent proper translation and subsequently lead to the degradation of the corresponding mRNA [123].

When this free energy constraint was applied to the designed PV-Max and PV-Min sequences, only a few changes were made and these changes did not have an overall impact on their CPBs. As illustrated in Figure 3.3, the folding free energies of the RNA sequences PV-Max and PV-Min were similar to that of PV(M)-wt. The graph in figure 3.3 is the corresponding free energy for every 100-base pair segment scanned. Specifically, PV(M)-wt had a mean folding free energy of -22.478, while PV-Max's mean was -22.032, and PV-Min's mean was -21.443. The construct PV-SD, included in Figure 3.3, was from our previous study and this virus grew with PV(M)-wt-like characteristics [4]. PV-SD, which had a mean folding free energy of -21.540, was similarly analyzed to show that random shuffling of codons to form novel pairs, without altering the CPB, did not effect the RNA folding energy either.

De novo DNA synthesis and construction of synthetic viral cDNA plasmids

The two novel sequences, PV-Max and PV-Min, were submitted electronically to Blue Heron Biotechnology Corporation and synthesized *de novo* (Figure 3.1) [124]. These sequences were then inserted by Blue Herron into a carrier plasmid from which they were excised for further studies. Additionally, Blue Herron provided complete sequence data (gel electrophoresis runs and sequencer output files) to ensure that the sequence ordered was the one provided. Then, using molecular cloning, these novel P1 regions were excised from their carrier plasmids via PflMI restriction enzyme digestion. Subsequently both the PV-Min and PV-Max fragments were exchanged into the plasmid that contains the full-length PV(M)-wt cDNA, pT7PVM [13], replacing the PV(M)-wt P1 region (Figure 3.4). This molecular cloning was successful and yielded plasmids pPV-Max and pPV-Min. These novel viral cDNA constructs with CPB-altered P1 regions were additionally sequenced to ensure proper cloning and to confirm the customization was accurate.

Utilization of reverse genetics to produce infectious virions of synthetic viruses

The plasmids containing the full-length poliovirus cDNA with synthetic capsid replacements, pPV-Min and pPV-Max, have a T7 RNA polymerase promoter upstream of the poliovirus genomic sequence (Figure 3.4). This promoter can stimulate purified T7 RNA polymerase to transcribe the plasmid cDNA into positive-sense RNA. The positive-sense RNA of poliovirus is infectious upon transfection into permissive mammalian cells [125,126]. Specifically, the plasmid was first linearized via enzyme digestion to prevent run-around transcription, thereby ensuring the production of

individual RNA transcripts of the appropriate length. Purified T7 polymerase was next incubated with the linearized genomic cDNA and an *in vitro* transcription reaction yielded (+)RNA for both the PV-Min and PV-Max constructs (Figure 3.4) [13].

Prior to transfection of these RNA transcripts into HeLa R19 cells, the integrity of the RNA was confirmed by performing an S^{35} *in vitro* translation reaction, which produces radio-labeled proteins *in vitro* [114]. This experiment was done to exclude the possibility that these novel RNAs have gross defects (such as mutations or stable structures) and are able to produce all appropriate viral proteins. As stated previously, should these RNAs be wrought with stable, high-order secondary structure, proper translation should be inhibited [123]. This should not be the case considering the mFold data; however biological confirmation was additionally sought. Also, the S^{35} *in vitro* translation was done to ensure the cloning procedures and/or DNA synthesis had not inserted missense and/or nonsense mutations into the PV-Max and PV-Min RNA. According to previous studies [4], this method cannot be correlated to translation efficiency *in vivo* because the HeLa cell extracts are nuclease-treated and supplemented with exogenous tRNAs. Therefore, this is an artificial system that was utilized to confirm RNA integrity.

The PV-Max and PV-Min RNA transcripts were incubated in a HeLa S10 cell extract supplemented with S^{35} radio-labeled methionine and the *in vitro* translation reaction proceeded. This experiment was conducted using a nuclease treated extract that removes all competing cellular mRNAs. Also the extract is supplemented with tRNAs and amino acids. This extract therefore does not mimic the dynamic cellular environment and is only used to study RNA integrity and the ability to produce viral proteins. The

reaction samples were then subjected to SDS-PAGE and analyzed by gel autoradiography. As shown in Figure 3.5, the RNA for each viral construct, all of which encode the exact same amino acid sequence, was capable of producing the correct poliovirus proteins *in vitro*. Importantly, these proteins were able to undergo the appropriate auto-catalytic cleavage events to yield all viral proteins. With these novel synthetic genomes expressing viral proteins properly *in vitro*, reverse genetics was then utilized to yield the infectious viruses (PV-Min and PV-Max).

To produce the virion form of the synthetic viruses PV-Min and PV-Max, the *in vitro* RNA transcripts were transfected into HeLa R19 cells via a modified Hanks DEAE-Dextran method (Figure 3.4) [13]. As previously stated, this RNA is infectious and should produce infectious virions if able to translate *in vivo* above a given threshold of efficiency [4,99]. Since the DEAE-Dextran is slightly toxic to the cells, a control plate was mock transfected in parallel. The PV-Max transcript RNA transfection produced 90% Cytopathic Effect (CPE) in 24 hours, which is comparable to the transfection with PV(M)-wt RNA [13]. The subsequent use of the supernatant from cells subjected to PV-Max RNA transfection also produced 95% CPE in 12 hours, thus indicating that the transfected genomic material successfully produced PV-Max poliovirus virions. The PV-Max virus from this first passage was then titered via plaque assay on HeLa R19 cell monolayers and found to have a comparable titer to PV(M)-wt (Figure 3.6).

Conversely, the transfection of PV-Min transcript RNA yielded no visible CPE after 72 hours and four blind passages of the supernatant, possibly containing extremely low levels of virus, did not produce CPE either. Blind passages were done in order to allow for the possible reversion of attenuating mutations whose reversion could than

allow the virus to regain a viable phenotype. Since the full-length PV-Min P1 had a null-phenotype, the P1 region required subcloning in order to regain a viable phenotype.

Discussion

The results from this chapter illustrate the utility of synthetic biology and its applicability to solve future biological problems. Specifically, the readily available poliovirus DNA sequence (GenBank) was customized via a computer and then synthesized *de novo* in order to address two specific biological aims – the investigation of CPB, as well as developing a model for vaccine development. The use of synthetic biology allowed for the construction of two novel DNA molecules, PV-Max and PV-Min. These molecules use the wild type codons and express the identical wild type amino acid sequence; yet their codons were shuffled to form novel pairings. These novel pairings resulted in 566 silent mutations (PV-Max) and 631 silent mutations (PV-Min). This vast number of mutations could have only been achieved using synthetic biology and DNA synthesis because conventional mutagenesis would have been impractical.

These large numbers of customized mutations was also an aim of this study towards the production of a vaccine model. Specifically, the idea is that many of the 631 mutations caused small, additive defects in PV-Min. These defects caused the PV-Min viral construct to have a null phenotype. Thus, using CPB-alteration to attenuate viruses should create a strain that is genetically stable, since a virus containing this many mutations would require hundreds of reversions to increase its virulence. Genetic stability is an attractive characteristic of new vaccines taking into consideration problems with the Sabin strains in the current oral poliovirus vaccine (OPV). Interestingly, for the Sabin-1 strain in OPV the virus has 51 mutations, but only 5 mutations have been shown to contribute to its attenuated phenotype in susceptible animals [88]. This small number of mutations enables OPV to revert to neurovirulence in vaccine recipients causing

vaccine-associated paralytic poliomyelitis (VAPP). Also, naïve immunologically individuals who encountered mutated OVP strains secreted from vaccine recipients have been shown to develop poliomyelitis. Such secreted viruses have caused small epidemics of poliomyelitis [83,88]. While it remains to be seen if indeed the 631 mutations caused small, additive defects in PV-Min, this model for vaccine production could prove to be attractive.

In sum, the results relating to PV-Max and PV-Min are both interesting. Firstly, the PV-Max results are contrary to the Hatfield groups previous findings, which suggested that over-represented codon-pairs are poor for translation [102]. PV-Max was highly concentrated in over-represented pairs, yet it grew with wild type characteristics. This result leads to the conclusion that previous studies simply required a larger sampling, or that the effects of CPB in the eukaryotic and prokaryotic systems are inversed. Also, the results for the purpose of the production of a vaccine strain, are significant when considering the null phenotype of the PV-Min virus with the P1 utilizing under-represented codon-pairs. Specifically, these results contradicted the previous notion that under-represented codon-pairs enhance translation [102]. These results also suggested that if the null phenotype of PV-Min was partially relieved via subcloning of its P1 region, an attenuated virus could be developed for the possible use as a vaccine.

Chapter 3 Figures

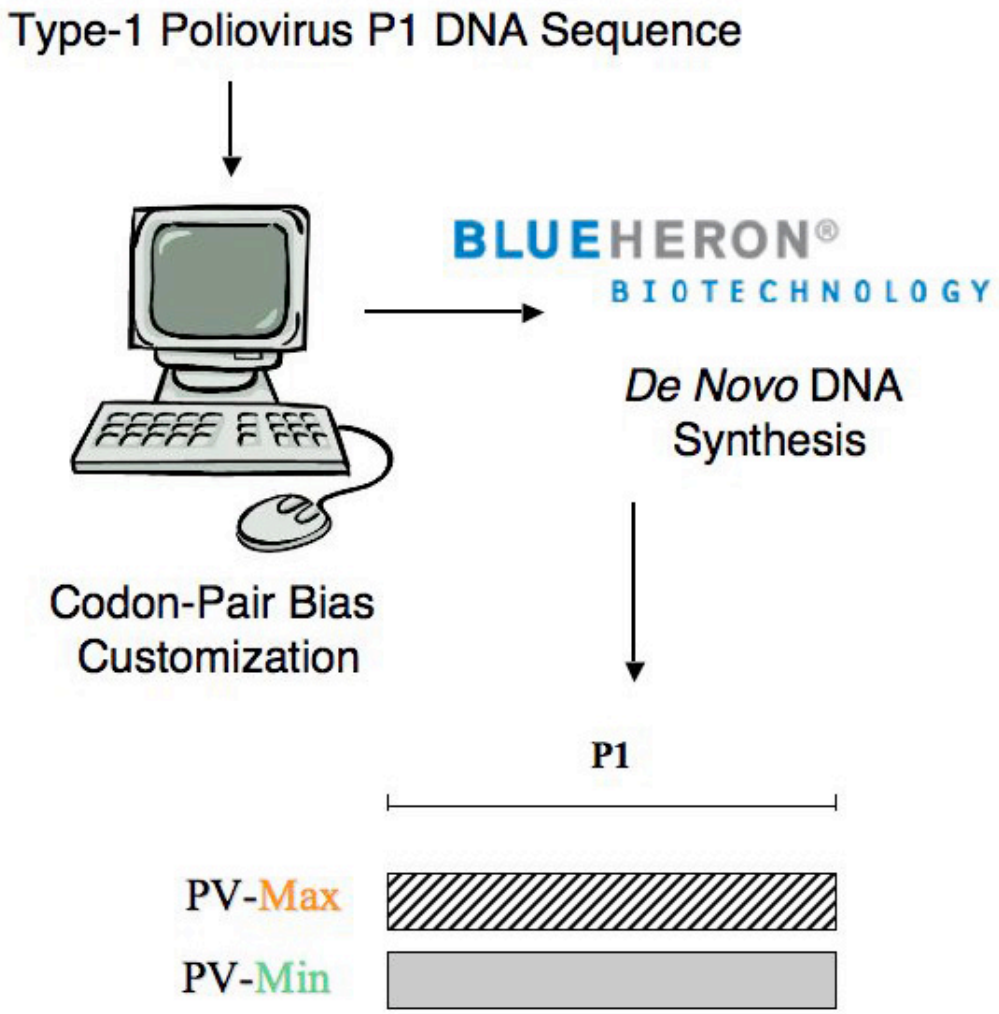


Figure 3.1: Codon-pair bias customization and *de novo* DNA Synthesis. The cDNA sequence encoding the type-1 poliovirus (Mahoney strain) P1 structural region was analyzed and re-encoded by our computer-based program (DPapa). DPapa maintained the codon usage and the amino acid sequence, but shuffled the existing codons to form novel pairs. Others features are controlled by DPapa, such as the folding free energy corresponding to the RNA of the output sequence (refer to Figure 3.3). DPapa was used to design two new P1 regions (encoding the poliovirus capsid, 2643 nucleotides), PV-Min and PV-Max. PV-Min was re-coded to use under-represented codon-pairs compared to the human genome, and contained 631 synonymous mutations. Virus PV-Max was re-coded to use over-represented codon-pairs, and contained 566 synonymous mutations.

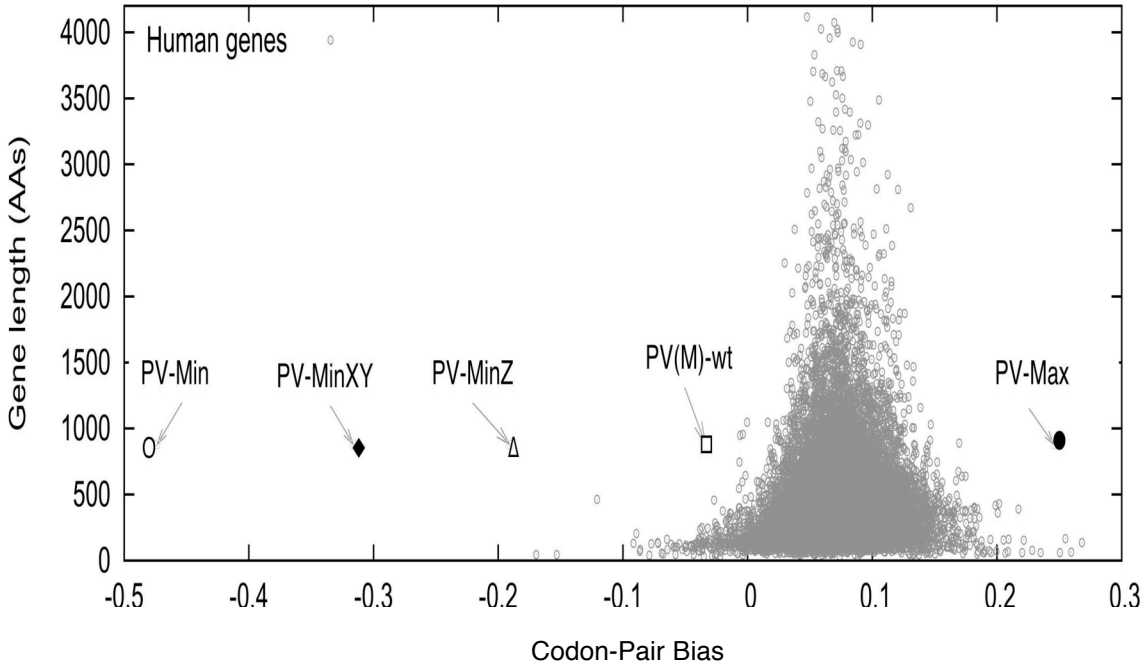
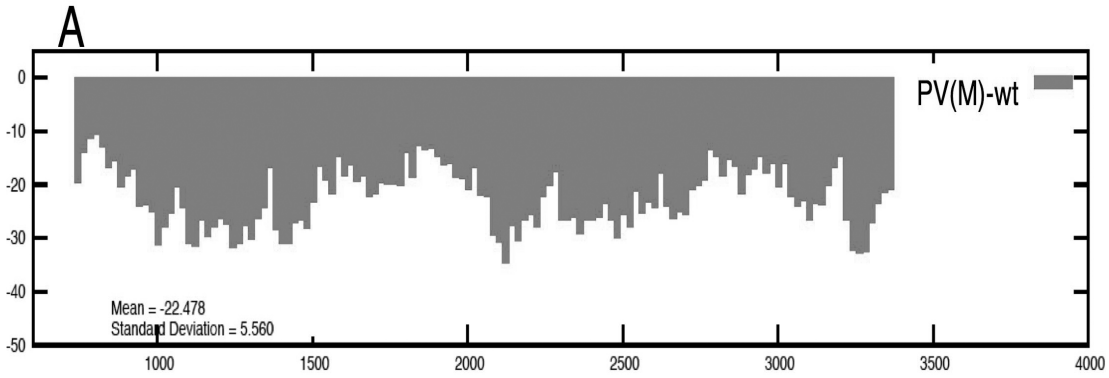
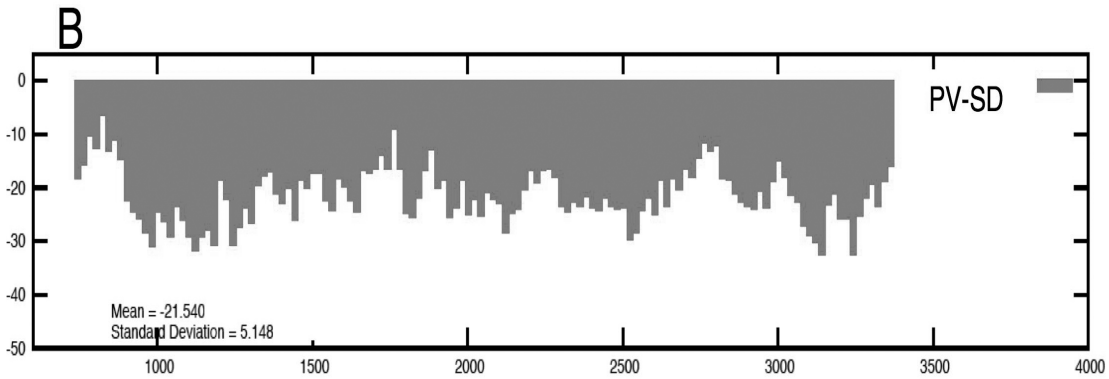
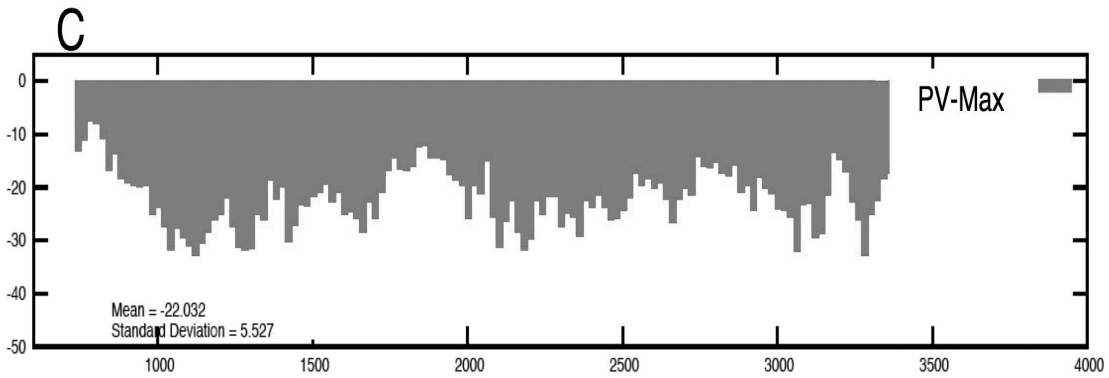


Figure 3.2: Codon-pair bias of all Human ORFs and synthetic P1 regions.
Calculated codon pair bias (CPB) score for all 14,795 annotated human genes. Each gray dot represents the calculated CPB score of a gene plotted against its amino acid length. The CPB of the PV(M)-wt P1 region as well as the CPB of the designed synthetic poliovirus capsids PV-Min, PV-MinXY, PV-MinZ, and PV-Max, are indicated by arrows and labels. The CPB of PV-MinX and PV-MinY can be seen in Table 3.1 along with all other constructs.


A
PV(M)-wt
0
-10
-20
-30
-40
-50
Mean = -22.478
Standard Deviation = 5.560
1000
1500
2000
2500
3000
3500
4000


B
PV-SD
0
-10
-20
-30
-40
-50
Mean = -21.540
Standard Deviation = 5.148
1000
1500
2000
2500
3000
3500
4000


C
PV-Max
0
-10
-20
-30
-40
-50
Mean = -22.032
Standard Deviation = 5.527
1000
1500
2000
2500
3000
3500
4000

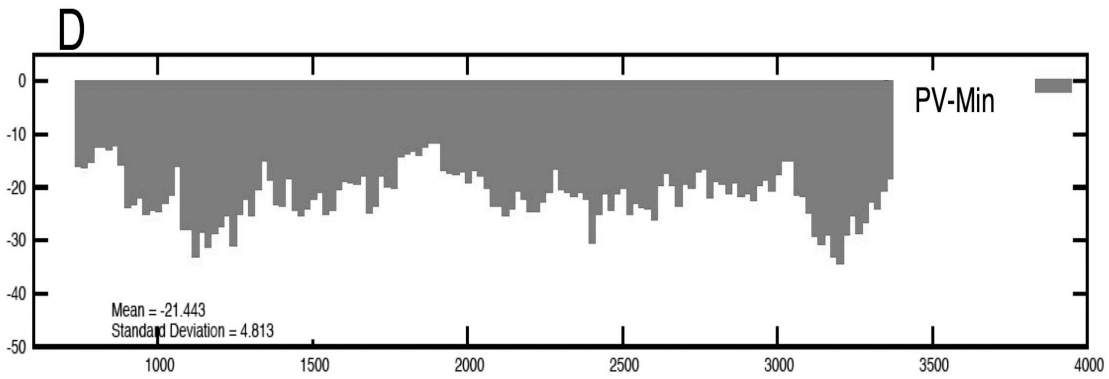

D
PV-Min
0
-10
-20
-30
-40
-50
Mean = -21.443
Standard Deviation = 4.813
1000
1500
2000
2500
3000
3500
4000

Figure 3.3: RNAs of the synthetic and wild type P1 regions have similar folding energies. To exclude stable secondary structures that could have affected translation efficiency, the capsid region of the CPB-altered designs were scanned using the program mFold [113]. Specifically, 100-base long segments, having 80 bases overlap with each other, were analyzed. Any segments with lower binding energy than a threshold of -30Kcal/mol incurred random synonymous substitutions at C - G binding locations to alleviate folding. The synonymous changes were selected to maintain CPB customization. Only a few synonymous changes were made, with no effect on global CPB. The calculated energy value of each scanned 100-base pair segment was output to produce the graph. The mean folding energy for all constructs were: (A) PV(M)-wt, -22.478, (B) PV-SD, -21.540, (C) PV-Max, -22.032, (D) PV-Min, -21.443.

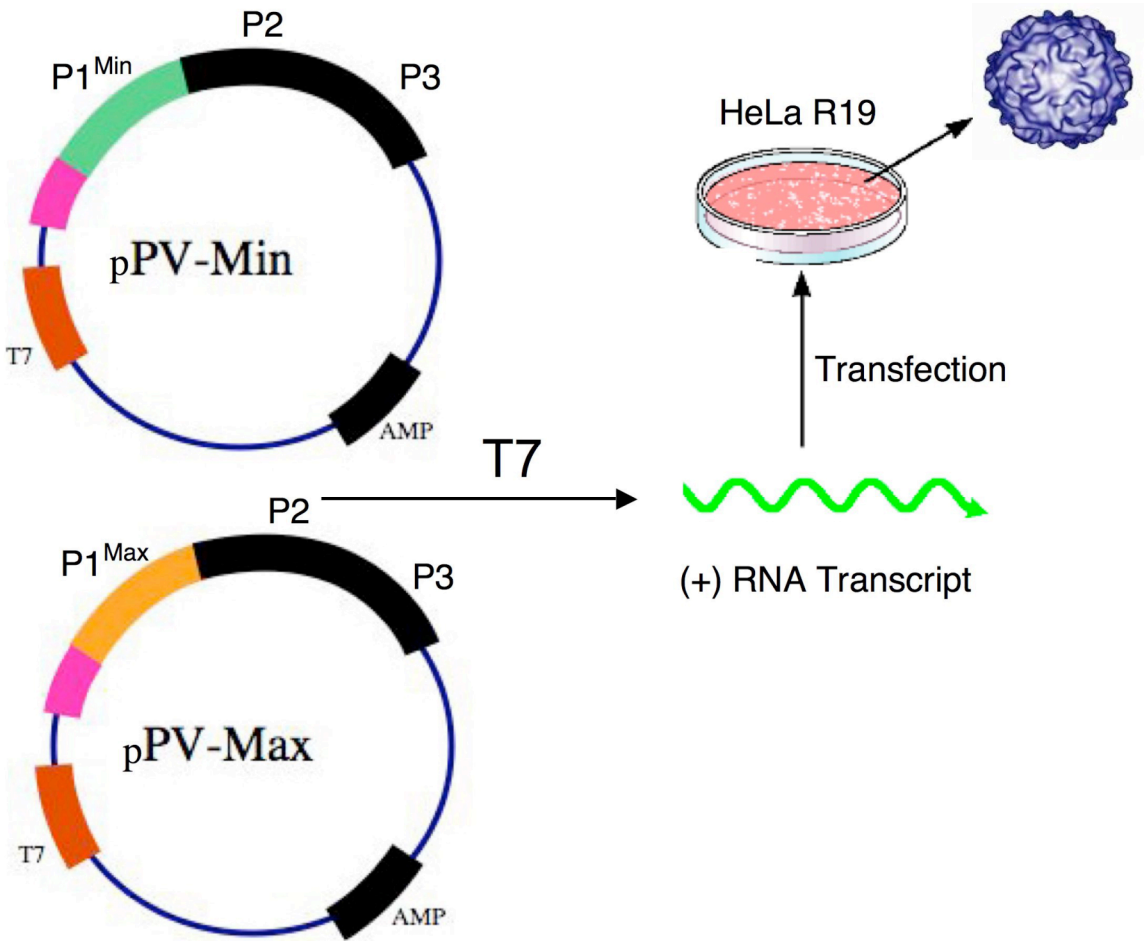


Figure 3.4: *In vitro* construction of synthetic virus cDNAs. The synthetic P1 regions were incorporated, via molecular cloning, into the cDNA plasmid containing the wild type P2 and P3 regions [13]. Then, using the T7 RNA polymerase promoter, (+)RNA was produced *in vitro* and transfected into HeLa R19 cells.

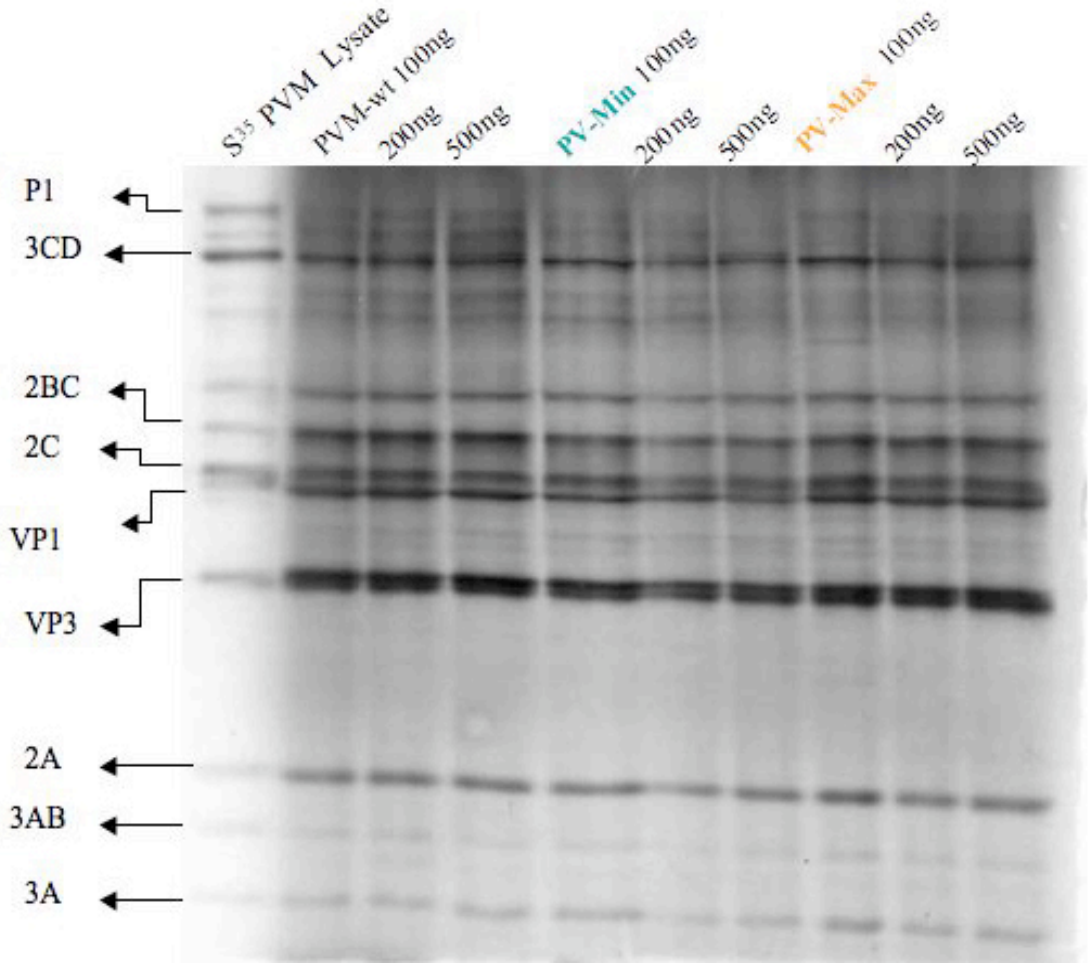


Figure 3.5: RNA integrity displayed by S^{35} *in vitro* translation. In order to assess the ability of the novel RNAs PV-Min and PV-Max to translate properly, *in vitro* transcribed RNA was used in an S^{35}-methionine-labeled *in vitro* translation reaction. PV-Max and PV-Min RNAs were able to produce all viral proteins and these proteins were able to auto-process themselves. This confirms that CPB customization did not create major hairpins, drastic RNA instability, or gross defects that would have inhibited protein production. This experiment was repeated twice with similar results. This experiment was conducted using a nuclease treated extract that removes all competing cellular mRNAs. Also the extract is supplemented with tRNAs and amino acids. This extract therefore does not mimic the dynamic cellular environment.

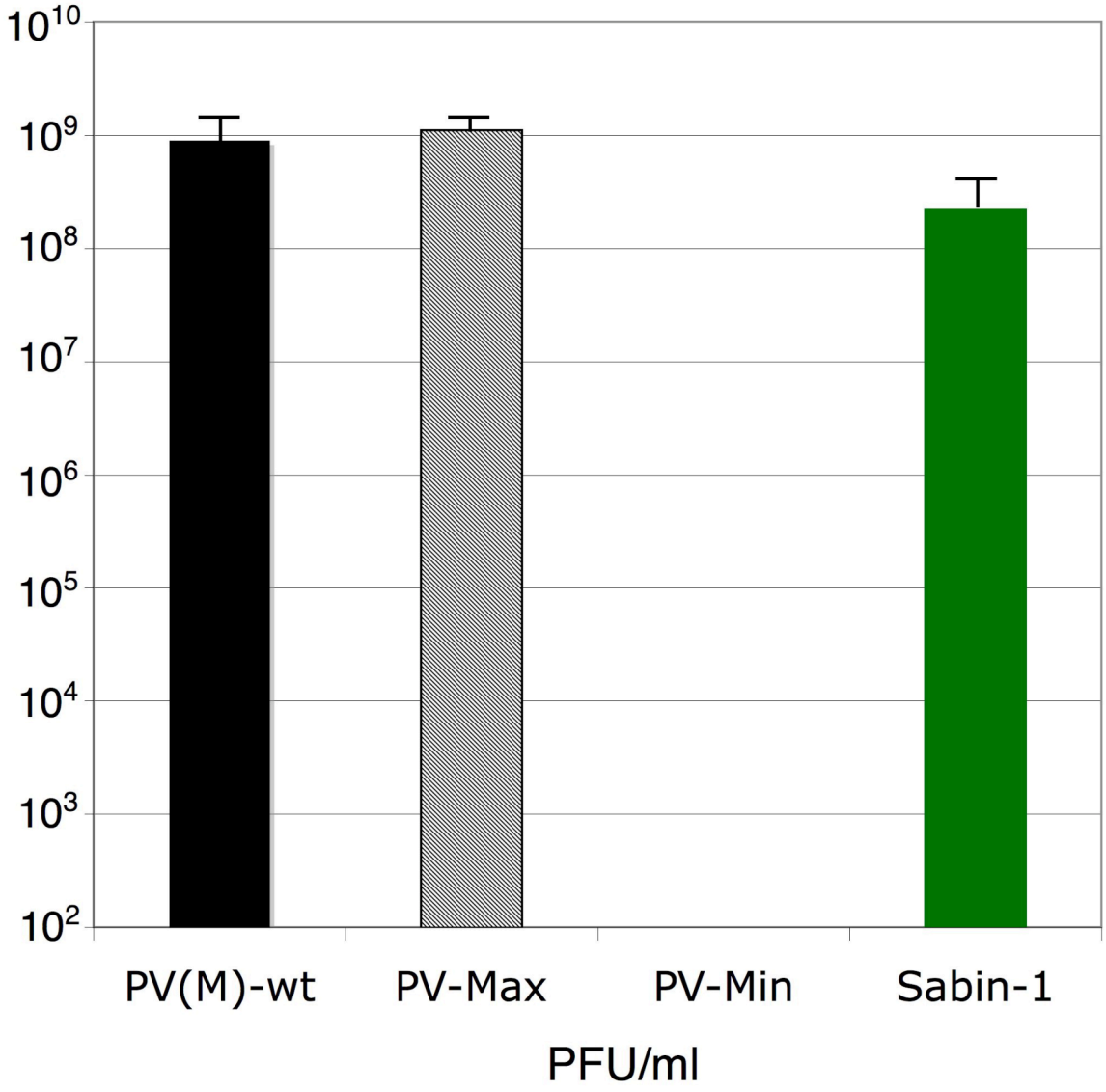


Figure 3.6: Viability of polioviruses with full-length synthetic P1s. After the initial transfection, supernatants were used to infect monolayers of HeLa R19 cells, and the subsequent PFU/ml titers were measured via plaque assay. The novel synthetic poliovirus with the PV-Max-encoded P1 grew to a similar titer as PV(M)-wt, whereas the PV-Min construct failed to yield a viable virus, even after four blind passages. The titer of Sabin-1 was included to demonstrate that an attenuated virus can still grow to robust PFU/ml titers. This graph represents the average of three independent experiments.

Chapter 3 Table

P1-Region	Codon Pair Bias
PV-Max	+0.246
PV(M)-wt	-0.034
PV-SD [4]	-0.095
PV-AB [4]	-0.096
PV-MinX	-0.154
PV-MinY	-0.197
PV-MinZ	-0.205
PV-MinXY	-0.391
PV-Min	-0.474

Table 3.1: Calculated codon-pair bias (CPB) for synthetic viruses. PV-Min had a negative CPB, but the same codon usage as PV(M)-wt, therefore any decrease in its fitness was therefore attributed to the negative CPB. The inclusion of PV-SD was to demonstrate that a random shuffling of codons resulted in the maintenance of a CPB and a phenotype similar to that of PV(M)-wt. In contrast, PV-AB, which used rare codons, had a similar CPB to PV(M)-wt yet was attenuated.

Chapter 4: *In vitro* Characterization of Codon-pair Customized polioviruses

Summary

To rescue viability of PV-Min, wild type poliovirus sequence was reincorporated into its cDNA to regain viability. The incorporation of a full-length synthetic P1 region caused the viral construct PV-Min to a have a null phenotype, therefore The P1 region was subcloned via molecular DNA cloning into five separate P1 variants. Each variant contained only a portion of its P1 encoded by under-represented codon-pairs and the remaining was wild type poliovirus P1 sequence. Then, the *in vitro* growth characteristics of all viruses, the five viable PV-Min subclones and PV-Max, were determined by various methods. One-step growth kinetics showed that growth of PV-Min subclone viruses proceeded with the same kinetics as PV(M)-wt, yet were up to ~2.5 logs attenuated when measured as PFU/ml. The most attenuated viruses were PV-MinXY and PV-MinZ. These similar growth kinetics observed were a result of the virus producing nearly the same number of particles per cell; however, the particles produced were less infectious. This reduction in infectivity was not due to morphological differences in the virion structure. Rather, this reduction in infectivity was shown to be a result of reduced translation efficiency of the corresponding PV-Min P1 regions. Specifically, the efficiency at which the CPB-altered P1s constructed translate *in vivo* was quantified via a luciferase reporter assay. This assay demonstrated that under-

represented codon-pairs decreased the rate of translation and over-represented pairs possibly enhanced the efficiency of translation. It has previously been shown that RNAs that translate poorly are less stable within a cell and are more rapidly degraded. Thus, the *in vivo* RNA stability of the synthetic viruses PV-MinXY and PV-MinZ was analyzed and displayed a reduced stability when compared to PV(M)-wt. The reduction in translation efficiency most likely inhibited the virus from establishing an infection and hence was responsible for the PV-Min subclones' reduced infectivity.

Introduction

In our previous study, we had results that mirrored the null phenotype of the full-length PV-Min viral construct. Previously, we exchanged all codons in the P1 structural region of poliovirus with the most infrequently utilized synonymous codons (Figure 1.4). This virus was designated PV-AB [4]. Interestingly, by swapping back wild type P1 cDNA sequence we restored a viable phenotype. Mueller et al. showed that by causing a translation hindrance using rare codons, the initial translation of the viral polyprotein was too slow and this led to an increase in the frequency of abortive infections [4]. Therefore, the codon-usage virus (PV-AB) was significantly attenuated because it was less infectious. This reduction of infectivity was attributed to a reduced specific infectivity.

Specific infectivity is defined as the number of virus particles that yield a single, productive infectious event. This productive infectious event is usually visible as a plaque on a monolayer of susceptible cells. Therefore, a ratio of Plaque Forming Units (PFU)/ml to particles/ml can be specifically defined for a given virus on a given cell type,

and this ratio yields a virus's specific infectivity. For viruses belonging to the Picornavirus family, of which PV(M)-wt is a member, it is known that this ratio is approximately 1 PFU to 100-1,000 particles [127]. PV(M)-wt is known to have a specific infectivity of approximately 1:100 on HeLa cells [4,97,98]. This ratio of PFU: particles is not a result of cooperativity in a picornavirus infection, i.e. the virus does not need a hundred particles working together to infect a cell. Rather, taking the example of PV(M)-wt, each particle has a 1/100 probability of starting an infection cycle. The fact that cooperativity plays no role in poliovirus infection was previously established due to single hit kinetics, showing that a single particle is capable of initiating an infection [128].

The exact cause of this ratio was previously assumed to be a result of receptor-virus interactions [127]; however, it is now hypothesized to be mainly the result of cell-virus interactions after its entrance into the cell and that receptor-virus interactions play a minor role [4]. Specifically, that the efficiency the entering virus translates its genome once inside the host has a pronounced effect on its ability to establish a productive infection. The hypothesis is that the entering poliovirus must quickly translate a threshold quantity of the $2A^{pro}$ in order to hijack the host cell. This translation inefficiency affecting specific infectivity was seen in our previous viruses [4] and also in Sabin-1 mutant viruses containing IRES mutations, which had an inability to initiate translation [129,130]. The specific infectivity of a virus was previously held as a constant characteristic for a given strain, yet new results suggest this property could vary greatly depending on virus genomic composition as well as the cell type that it is infecting.

A similar phenomenon could have been occurring in the full-length PV-Min viral cDNA construct, therefore a similar approach to elucidating its mechanisms of attenuation was conducted. This possibility of a null phenotype was considered possible during the design of both PV-Min and PV-Max, and by customizing the designs via the inclusion of unique restriction sites viability could be regained with the restoration of wild type sequence. Also, quantification of the exact effects CPB customization has on the rate of translation were elucidated in order to make an initial observation with regards to the effects over- and under-represented codon-pairs have on translation.

Results

PV-Min subcloning to regain viability

Restriction sites were designed during CPB customization within both the PV-Min and PV-Max sequences to rescue viability, in the event either full-length synthetic P1 produced a null virus. This event occurred in the PV-Min viral construct. Therefore, the P1 region was subcloned, restoring portions of the PV(M)-wt P1 sequence, by utilizing these restriction sites. The fragments of PV(M)-wt P1 sequence used for subcloning originated from pT7PVM [13], the original backbone from which the pPV-Min and pPV-Max constructs were created (Figure 3.4). Restriction enzyme digestion, gel purification, ligation, transformation, and sequencing were performed to yield five PV-Min cDNA constructs (Figure 4.1). From each of these subclones, transcript RNA was produced via *in vitro* transcription and then transfected into HeLa R19 cells. The construct PV-MinYZ failed to yield virus following transfection and four blind passages;

all other PV-Min subclone derivatives were able to yield viable viruses following transfection of RNA (Figure 4.1, viability). Importantly, viable viruses were obtained for constructs PV-MinX, PV-MinY and PV-MinZ, which displayed that there are no gross defects in any of these regions (X, Y, or Z) that would have been responsible for the null phenotype of the full-length PV-Min.

After transfecting the five PV-Min derivate RNAs into HeLa R19 cells, each product virus was passaged twice to ensure the virus was properly amplified because RNA transfection is not as efficient as a natural infection by the virus itself. Next, plaque assays were performed to elucidate the apparent PFU/ml titer for each virus (Figure 4.2). These subclones yielded viruses that had varying degrees of attenuation. The viruses containing P1 fragments X and Y were each attenuated by 0.8-1 $\log_{10}$; however, when added together they yielded virus PV-MinXY, which was significantly attenuated by 2.5 orders of magnitude. The virus PV-MinZ was also attenuated on the order of 2.5 $\log_{10}$ like PV-MinXY. Thus, when returning the Y fragment to PV-MinZ, the construct PV-YZ failed to yield viable virus (Figures 4.1 and 4.2). This varying degrees of attenuation therefore was due to an apparent additive effect. Thus, one could conclude the null phenotype of the full-length PV-Min construct was due to the sum of defects in the various three sub-regions.

Plaque phenotypes of the synthetic viruses

The plaque phenotype of the synthetic viruses helped characterize the infectability of these viruses. The synthetic viruses that were most attenuated, PV-MinXY and PV-MinZ, had a small plaque phenotype compared to PV(M)-wt, and the plaques of these

viruses required additional time to develop (Figure 4.3). Briefly, plaque assays were conducted on a monolayer of HeLa R19 cells that were infected with decreasing serial dilutions of the amplified virus stock and then overlaid with tragacanth gum to prevent diffusion of particles throughout the dish. A single infectious event occurs in one cell and then radiates out to neighboring cells, forming a visible plaque that is then enumerated. In my analysis of the synthetic PVs, the number of plaques were then multiplied by the dilution factor to indicate the number of PFU per milliliter of culture medium. The plaque phenotypes of PV-MinXY and PV-MinZ compared to that of PV(M)-wt were significantly smaller. In contrast, the PV-Max virus, which initially grew with wild type-like characteristics, had a large plaque phenotype similar to PV(M)-wt (Figure 4.3). Additionally, under normal conditions PV(M)-wt and PV-Max were able to produce visible plaques after 48 hours at 37°C; however, PV-MinXY and PV-MinZ required 72 hours of incubation at 37°C to yield visible plaques. This indicated a reduced ability for PV-MinXY and PV-MinZ to spread because these viruses required additional time to clear cells and form visible plaques. Also, the small plaque phenotype could have indicated that PV-MinXY and PV-MinZ were more susceptible to interferon. All plaques in Figure 4.3 were developed for 72 hours for equal comparison of plaque phenotypes. As visible, the increased incubation time and the small plaque phenotype was the initial indication, in combination with previous observations [4], that PV-MinXY and PV-MinZ may have had a reduced specific infectivity. To follow this possibility, the growth kinetics of these viruses were elucidated as well as their specific infectivity on HeLa R19 cells.

Growth kinetics of PV-Max and PV-Min derivative viruses

The growth kinetics for all five viable PV-Min virus constructs, as well as PV-Max, were determined via one-step growth curves. Briefly, 1 x 10^6 HeLa R19 cells were infected with 2 MOI of the appropriate virus. The virus titer at each time point over a 48-hour time period was measured via serial dilution of each time point and quantification by plaque assay (Figure 4.4). The endpoint titers determined for each virus in the one-step growth curve mirrored the plaque assay results in Figure 4.2. Also, the concept of the combination of segments having an additive negative effect on growth appeared in the growth curves. Specifically, PV-MinZ grew to a titer 10-fold lower compared to PV(M)-wt and PV-MinY to a titer $10^{1.6}$-fold lower. When these two segments were combined, yielding virus PV-MinXY, the titer was $10^{2.5}$-fold less than PV(M)-wt.

While the magnitude of each endpoint titer was varied, each virus had similar growth kinetics, indicating that the infection by the PV-Min viruses proceeded in a similar manner as PV(M)-wt. Like the wild type virus, the PV-Min chimeric viruses had an eclipse phase followed by an exponential growth phase (albeit muted in the PV-Min derivative viruses). However, as measured by plaque forming units, the final titers of the PV-Min constructs were decreased by up to 1000-fold compared to PV(M)-wt (Figure 4.4). The PV-Min chimeric viruses lowered endpoint titer but logarithmic growth could have resulted from two possibilities. The first possibility was that the PV-Min constructs produced fewer viral particles per cell. The second possibility was that the viral particles produced by these PV-Min viruses had a reduced specific infectivity (i.e. a decreased probability of initiating an infection). Therefore, to determine which of these two

possibilities occurred (or if it was a combination of both), virus purification and specific infectivity determination were examined.

Virus purification to determine specific infectivity

To determine the specific infectivity of the PV-Min chimeric viruses (X, Y, XY, and Z), PV-Max, Sabin-1 and PV(M)-wt, all viruses were purified via ultracentrifugation following amplification on 1 x 10^8 HeLa R19 cells. Prior to the ultracentrifugation, each supernatant was treated with 0.5% SDS and 20mM EDTA for 30 minutes. Importantly, mature poliovirus virions that were formed properly, i.e. properly folded capsids, are resistant to this concentration of SDS [25]. This treatment denatured any cellular proteins remaining in the supernatant that would have precipitated out during the ultracentrifugation. Also, this SDS treatment eliminated immature forms of the poliovirus virion known as 135S A particles. The A particle is a less infectious intermediate structure [131] of the mature 160S virion and is not resistant to SDS [132]. Also, RNase A was added for 1 hour to ensure the degradation of remaining extra-viral or cellular RNAs that might have precipitated out. Considering these two factors, any virus pellet should consist of solely mature virions. Following ultracentrifugation through a sucrose cushion, these purified virus solutions were used in subsequent experiments.

Determination of specific infectivity

Once the virus was purified, the exact amount of particles of a given virus present in solution was calculated by measuring the absorbance (A) at 260nm. This A_{260} value was used to calculate the number of viral particles based on the empirically determined

constant for poliovirus, where 1 A_{260} unit = 9.4 x 10^{12} particles/ml [133]. Using this determined constant, the number of particles present for all purified viruses was determined (Table 4.1). Then, using purified virus stocks, plaque assays were performed to define the PFU titer. Using these two calculated values, the specific infectivity of the given viruses was calculated by dividing the PFU titer by particles present (Table 4.1). The resulting calculated values defined the number of particles that a single PFU represented or the relative specific infectivity. The specific infectivity determined for PV(M)-wt was in accordance with previous studies, where 1 PFU corresponded to 137 virus particles. Previous studies have shown that the ratio for PV(M)-wt is on average 1:100 [4,97,98]. The PV-Max virus, which mirrored all previous growth characteristics of PV(M)-wt, also had a nearly identical specific infectivity of 1/131, indicating a similar phenotype to the wild type virus.

On the other hand, PV-Min viruses (X, Y, XY and Z) had an apparent decrease in their specific infectivity, where now the probability of a single cell being productively infected was reduced from the wild type ratio of 1/100 to 1/15,000 for the most attenuated virus (Table 4.1). PV-MinX and PV-MinY both had a reduction in their specific infectivity (1/305 and 1/1,050, respectively), yet this reduction was not as pronounced as the decrease displayed in PV-MinXY and PV-MinZ (1/9,500 and 1/13,500, respectively). Examining the growth kinetics, cells infected with PV-MinXY or PV-MinZ produced slightly fewer viral particles per cell compared to PV(M)-wt, but the effect was only about 10-fold or slightly less (Figure 4.4). When the number of particles per PFU was measured, it was determined the particles of PV-MinXY and PV-MinZ had a 98.5-fold (~100) reduced probability of generating a plaque (Table 4.1). For

PV(M)-wt, the number of virions represented by a PFU was about 137, whereas for PV-MinZ, a single PFU correlated to 13,500 viral particles. This data showed that the source of attenuation (Figure 4.2, 4.3, and 4.4) was due to a reduced specific infectivity of the virions. The total attenuation from both effects together (specific infectivity $10^{2.3}$ and fewer particles $10^{0.6}$) was about 1000-fold for PV-MinXY and PV-MinZ. Additionally, these results regarding specific infectivity mirrored the previous data, where the attenuation seemed to be a result of an additive effect due to the addition of PV-Min sequence in the P1 region. As calculated in Table 4.1, PV-MinX and PV-MinY both had a lowered specific infectivity, 5- and 10-fold respectively, when compared to PV(M)-wt. Thus, when PV-MinX and PV-MinY were combined to form the virus PV-MinXY, the reduction of specific infectivity was magnified to ~100 fold (from 1/137 to 1/9,500). Note: The Sabin-1 virus, which is currently a component of the oral poliovirus vaccine, was also purified and its specific infectivity determined. This virus is significantly attenuated in animals due to temperature sensitivity [134,135] and other IRES-dependent mechanisms [136]. However, the attenuated phenotype of Sabin-1 does not result in a substantial deviation of specific infectivity from PV(M)-wt. The specific infectivity of Sabin-1 was calculated to be 1/200 (Table 4.1).

The most likely cause of this reduction in specific infectivity was the CPB-altered sequence inserted into the P1 regions of the various PV-Min constructs. Namely, that this CPB-alteration, i.e. use of under-represented codon-pairs, was disrupting the virus's ability to initiate a replicative cycle and this was made evident by their reduced specific infectivity. Since the capsids of these synthetic viruses had the exact same amino acid composition as the PV(M)-wt, as well as had the ability to survive the SDS and EDTA

treatments during purification, the reduction in specific infectivity should not have been a result of gross morphological defects of the capsid. However, to provide additional evidence to rule out this possibility, the capsid resistance to heat denaturation was analyzed.

Capsid heat stability of synthetic viruses

To demonstrate that the reduction in specific infectivity of PV-Min variants was not due to the large scale CPB-modification altering the gross morphology of the virions themselves (as one might expect if capsid proteins were misfolded), the thermal stability of PV-MinXY and PV-MinZ was tested. An equal number of particles were incubated at 50°C and the remaining infectivity quantified after given periods of time via plaque assay. If the capsids of the synthetic viruses were destabilized, they would have been expected to be highly sensitive to inactivation after incubation at 50°C when compared to PV(M)-wt. Since the heat-inactivation profile of the synthetic viruses was unchanged (Figure 4.5), it was concluded that the capsids of PV-MinXY and PV-MinZ were as equally stable as PV(M)-wt. The thermal inactivation kinetics of both synthetic viruses were identical to PV(M)-wt, whereby the solution of the viruses lost their infectivity with the same rate as the wild type poliovirus. In contrast, the Sabin-1 virus, which carries numerous mutations in the genome region encoding the capsid [134,137] causing it to be heat-labile [138], was less stable following heating at 50°C when compared to PV(M)-wt. Sabin-1 was included to provide an example of the thermal inactivation kinetics of a virus with a less stable capsid.

Comparison of virion RNA: length and quantity

Besides discounting the possibility that capsid instability was the reason for the attenuation of PV-MinXY and PV-MinZ, it is possible that the RNA contents of the virions could have been different when comparing the synthetic viruses to PV(M)-wt. Possibly, these synthetic viruses might not have efficiently encapsidated genomic RNA due to the "non-polio" CPB-altered sequence in the P1 region, thus creating empty virions. This was an unlikely possibility because previously "non-polio" sequences in the P1 region have been encapsidated as well as similar A_{260} values calculated in Table 4.1[4,95]. There is also a possibility that this RNA could have been truncated prior to encapsidation due to instability resulting from CPB-alteration. To test this possibility, equal numbers of purified viral particles of PV(M)-wt and PV-MinZ had their virion RNA extracted via phenol-chloroform purification. This RNA was then subjected to a 1% TAE agarose gel to compare both size and quantity of this RNA. Both RNA samples from an equal number of viral particles yielded similar quantities of single-stranded RNA, both of which were the comparable in length (Figure 4.6). Therefore, the RNA encapsidated by the PV-MinZ virus was similar in size and quantity despite CPB-customization of its P1 region. This customization did not grossly effect the RNA's physical properties (Figure 3.3 and Figure 4.6). These synthetic viruses PV-MinXY and PV-MinZ were physically stable in the structure of their capsid and RNA; however, they were attenuated due to the CPB-altered sequence. Aside from physical stability, the genetic stability of these viruses was important to define in order to evaluate their ability to serve as a vaccine. Viruses, RNA viruses in particular, are notorious for mutation and reversion to alleviate an attenuating phenotype and regain fitness [139,140]. To examine

the stability of the synthetic viruses PV-MinXY and PV-MinZ, they were serial passaged, i.e. put through many replicative cycles, to determine the genetic stability of their attenuated phenotype.

Genetic Stability of PV-MinXY and PV-MinZ

Genetic stability was an additional property of customization designed into the CPB viruses due to the large quantity of mutations that were incorporated into PV-Min (631 mutations) and PV-Max (566 mutations). Theoretically, by possessing a large number of mutations, the probability of a single mutation/reversion in PV-MinXY or PV-MinZ having a selective advantage for improved growth is minimal. In order to monitor the genetic stability of PV-MinXY and PV-MinZ, these viruses were extensively serially passaged. Specifically, confluent HeLa R19 cells were infected with a MOI of 0.5 of PV-MinXY and PV-MinZ. A low MOI was used to maximize the number of replication events [141], thus allowing more opportunities for PV-Min to revert. Once the cytopathic effect (CPE) had been observed, the supernatant was harvested and subsequently used to infect a new plate of confluent HeLa R19 cells. This cycle was repeated 17 (PV-MinZ) and 19 (PV-MinXY) times (Figure 4.7). Each virus population in the cell supernatant was checked for reversions in the synthetic region after passages 5, 10, 15 and 17/19 by performing reverse-transcriptase polymerase chain reaction (RT-PCR) and then sequencing the RT-PCR product. No mutations were found in the synthetic regions of PV-MinXY and PV-MinZ. The entire viral genome of each population of the supernatant from final passages, 17 and 19 respectively, was sequenced as well (save a region in the 3' NTR) and compensatory mutations were not found either.

Secondly, the viral titers for each of these milestone passages were measured by plaque assay. These plaques assays did not show an increase in apparent PFU/ml titer indicating no drastic reversion in the specific infectivity of the synthetic viruses had occurred. Also, an increase in the percent of cells displaying CPE did not occur, nor was there a drastic reduction in the incubation time required for CPE to occur. It is important to note that the sequence determined at each step was a consensus, that is the sequences observed were of the virus population. Therefore, taking into consideration the quasispecies phenomenon of PV(M)-wt [142], it was possible that there were underlying strains of virus carrying single mutations at places in the synthetic region yet they did not make through the noise of the consensus. However, if any one of these mutations was significant for a dramatic increase in fitness, it would have come to the forefront and the resulting mutation would have been fixed in the consensus sequence. This did not occur because the consensus sequences of the synthetic regions in PV-MinXY and PV-MinZ remained constant.

In sum, these synthetic viruses were found to yield nearly the same amount of particles per cell as PV(M)-wt, yet these particles were significantly attenuated due to a reduced specific infectivity. The PV-Min chimeric viruses were physically identical to the wild type virus, save the synthetic CPB customized sequence. The capsids of these viruses had identical amino acid composition to PV(M)-wt. These capsids were similar in stability to PV(M)-wt in adverse conditions (heat and SDS), as well as similar in content (virion RNA of equal length and quantity). Therefore, the alteration of CPB towards the use of under-represented codon-pairs must be the source of the PV-Min chimeric viruses' attenuation. Our previous work with the PV-AB virus derivatives displayed markedly similar results to those of PV-MinXY and PV-MinZ. We found the

virus PV-AB^{2470} to be attenuated due to a reduced translation efficiency of its altered P1 region [4]. This reduction in the rate of translation of the PV-AB virus chimeras was measured via a dicistronic reporter construct able to accurately determine the P1 region's rate of translation. Therefore, the P1 regions of the CPB-altered PV-Min constructs were cloned into this dicistronic reporter construct (pdiLuc-PV) to determine their rate of translation.

In vivo characterization of the effect CPB has on the rate of translation

To quantify the *in vivo* effect CPB had on translation, a dicistronic luciferase reporter construct (pdiLuc-PV) was used (Figure 4.8A) [4]. The cDNA of the P1 regions of various viral constructs, PV-Min, PV-MinXY, PV-MinZ and PV-Max, were molecularly cloned into pdiLuc-PV. These DNA constructs were then used to create transcript RNA using the upstream T7 RNA polymerase promoter and the *in vitro* transcription reaction previously described in Chapters 2 and 3. These transcript RNAs were then isolated via phenol-chloroform extraction and transfected into 70% confluent HeLa R19. These cells were then incubated at 37°C for 6 hours in the presence of 2mM guanidine hydrochloride. Guanidine hydrochloride prevents poliovirus replication by inhibiting the 2C^{ATPase} of poliovirus, thus only translation of the initial input RNA was measured [143]. The dicistronic reporter construct encoded both the R-Luc and F-Luc proteins. The R-Luc of this reporter construct served as an internal control based on the Renilla enzyme activity measured by an illuminometer. The Renilla enzyme activity provided a value that normalizes the transfection efficiency. The F-Luc protein served as the read-out for translation efficiency because the F-Luc reporter was translated as a

fusion protein with the proceeding proteins of the P1 region. The rate that the P1 region translated directly controlled the amount of F-Luc protein produced. Thus, the ratio of F-Luc luminescence to R-Luc luminescence provided an accurate quantity that correlated to the rate that the various P1 regions translated.

The translation of variously-encoded P1 regions were tested (Figure 4.8B). PV-MinXY, PV-MinZ, and PV-Min P1s produced 75-50% less F-Luc per unit of R-Luc than did the PV(M)-wt P1 region and the full-length PV-Min P1 region was reduced by 81%. This result strongly suggested that under-represented codon-pairs reduced the rate of translation (Figure 4.8B). The reduced translation is probably sufficient to explain the attenuated phenotype, since smaller reductions in translation caused by other methods have been known to attenuate poliovirus. This data suggests that PV(M)-wt has a fairly high threshold requirement for translation in order to hijack a cell and initiate a replication cycle [4,99]. In contrast, the PV-Max P1 appeared to have an increased rate of translation compared to PV(M)-wt because it produced more F-Luc per R-Luc than the wild type P1 (Figure 4.8B). Since the PV-Max P1 had an increased rate of translation of its P1 region, it may have had an enhanced ability to spread and possibly a faster cycle of replication. This possibility was examined via a low MOI growth curve where there were multiple rounds of replication and thus a virus with improved speed would amplify to a higher titer more rapidly.

Viral growth curves infected with a low MOI to examine the replication speed of PV-Max

The PV-Max and PV(M)-wt viruses were used to infect monolayers of 1 x 10^6 HeLa R19 cells at an MOI of 0.001 to compare the rates of their replication. At each time point, virus was extracted and measured via plaque assay. The determined PFU/ml titer was then converted to particle/ml titer by multiplying the calculated PFU/ml by each virus's respective specific infectivity (Table 4.1). As evident in Figure 4.9, the growth of these two viruses, PV-Max and PV(M)-wt, was quite similar. Therefore, the speed of their replication cycle in HeLa R19 cells was also similar. Despite the P1 region of PV-Max being translated slightly faster than that of PV(M)-wt (Figure 4.8B), this does not correlate to an increased rate of replication (Figure 4.9).

In vivo stability of PV-Min viral RNA

In contrast to PV-Max, the PV-Min viruses XY and Z showed a decreased rate of translation compared to that of PV(M)-wt (Figure 4.8). In a living mammalian cell, mRNAs that have a "block" in their open reading frame, whether it be a stretch of rare codons or a stable hairpin, are translated slowly and therefore are more readily degraded [123]. Specifically, stalling ribosomes alert cellular machinery to degrade the mRNAs that are poorly translating via a process called No Go Decay (NGD) [144]. This degradation event could have been occurring in the PV-MinXY and PV-MinZ viruses because indeed they were translated at a reduced rate compared to wild type poliovirus (Figure 4.8B). To examine this effect, Quantitative Real Time-PCR (Q RT-PCR) was used to quantify the number of genomes per cell infected with PV(M)-wt, PV-MinXY

and PV-MinZ. Also Q RT-PCR was used to determine the rate that these genomes were degraded. The primer pair used was designed to anneal in the P3 region of the genome (Figure 1.1), i.e. closer to the 3', because it has been shown that RNA degradation via NGD occurs at the 3' side of the block first [123,144].

To measure the rate of RNA degradation, HeLa R19 cells were infected at a MOI of 5 with the viruses PV(M)-wt, PV-MinXY, and PV-MinZ. These infected cells were then incubated for 3 hours at 37°C. After 3 hours, cells were treated with 4mM guanidine hydrochloride. By halting replication, via the addition of guanidine hydrochloride, only translation was capable of occurring and therefore the degradation of RNA, due to slow translation, could be measured. Indeed, the viral RNA per cell infected with PV-MinXY and PV-MinZ was degraded more rapidly than PV(M)-wt, 42% and 46% faster, respectively (Figure 4.10). This was a significant difference in degradation on a percentage scale; however, on a logarithmic scale this degradation did not effect the replicative cycle of these viruses once an infection had been established. Therefore, indeed the RNA genomes of PV-MinXY and PV-MinZ may have been degraded due to poor translation, but this effect was inconsequential once the replicative cycle had achieved logarithmic expansion. This point was further illustrated by the copies of viral genomes per cell calculated by Q RT-PCR. There were approximately 5.5 $x10^5$ copies of the PV(M)-wt genome in a single infected cell. Likewise for the CPB-altered viruses PV-MinXY and PV-MinZ, there were approximately 1 x 10^5 copies of their genomes in infected cells (Figure 4.10). These findings correlated with the previous data that the synthetic polioviruses PV-MinXY and PV-MinZ produced nearly the same number of particles per cell as PV(M)-wt during their replicative cycle (Figure 4.4, Table

4.1). However, early in the infection cycle the degradation of the incoming viral genomes during an infection may contribute to the reduction in specific infectivity.

Discussion

The goals of the studies in this chapter were to characterize the effect CPB-alteration had on viral growth and the rate of translation, as well as the possibility of using CPB-alteration as a means to produce a model virus for vaccine production. The most attenuated viruses based on viral titer and growth kinetics were PV-MinXY and PV-MinZ. These viruses therefore were the most compelling candidates for possible use as a vaccine and were carried on in further studies, specifically testing them in an animal model. All PV-Min derivative viruses displayed a reduction in their specific infectivity when compared to PV(M)-wt (Table 4.1).

For the most attenuated viruses, PV-MinXY and PV-MinZ, a single viral particle now had a 100-fold reduced probability of initiating an infectious cycle. The particles of PV-MinXY and PV-MinZ had a lowered specific infectivity (Table 4.1) because their infecting genomic RNA translated with a reduced efficiency (Figure 4.8B) and was degraded more rapidly (Figure 4.10). However, once PV-MinXY and PV-MinZ established an infection, the infection cycle proceeded with the same kinetics as PV(M)-wt and yielded a similar number of viral particles (Figure 4.4, Table 4.1, Figure 4.10).

The reduction in specific infectivity seen in the PV-Min subclones was not due to physical properties of the virus. Since they had the exact same amino acid composition, PV-MinXY and PV-MinZ were as equally resistant to SDS and heat denaturation as

PV(M)-wt (Figure 4.5) and their genomic RNA was similar in length and quantity (Figure 4.6). Rather, the use of under-represented codon-pairs in the P1 regions of PV-MinXY and PV-MinZ caused a reduction in the rate of their genomes' translation and thus a reduced specific infectivity. As stated, specific infectivity of picornaviruses is not cooperative [128] but rather is the probability of a single particle establishing an infection. This is an interesting strategy for replication. For example, PV(M)-wt has evolved to have particles with only a 1% chance of establishing an infection (i.e. 1 PFU: 100 particles). Fortunately for PV(M)-wt, a single cellular infection yields 10^5 particles per infected cell so even with only a 1% success rate for these new particles, that is still 10^3 new infectious cycles initiated from a single infected cell. For the viruses PV-MinXY and PV-MinZ, this probability was reduced to only a 0.01% chance of establishing an infection. Also, these synthetic viruses produced only 6-8 x 10^4 particles per cell as a result of CPB customization. Therefore, these viruses had a reduced ability to spread and initiate new rounds of infections despite producing a similar number of particles per cell as PV(M)-wt.

The PV-Min viruses produced nearly that same number of particles; however since they are less infectious, the plaques were smaller (Figure 4.3). For the creation of a vaccine, using the strategy of CPB customization, this reduction in infectivity could be an attractive characteristic because the virus has the exact same amino acid composition and antigenic structure of the progenitor virus, yet the novel CPB-altered virus simply cannot infect as many cells. This property could allow for proper immune recognition of the virus, yet due to the attenuation, there is less chance of collateral damage caused by the vaccine virus itself. This reduction in infectivity also could be a drawback because a

certain level of infection must be established in order to elicit a complete and proper immune response. If the infecting vaccine virus does not infect sufficiently, an insufficient and our short-term response could be elicited.

The other interesting finding presented in this chapter was the analysis of the effect codon-pairs had on translation. Particularly that, contrary to previous findings [102], under-represented codon-pairs decreased the rate of translation through a given open reading frame and over-represented codon-pairs could increase the rate of translation. Since the codon-usage as well as the amino acid sequence and composition of the experimental reading frames were held constant, these were the first definitive results that codon-pairs indeed affect translation [3].

Therefore, the initial goal of advancing towards a model for vaccine production was achieved, yet an actual vaccine itself was not yet produced. Specifically, that synthetic manipulation of CPB was able to modulate the translational efficiency of the viral genome and this modulation resulted in an attenuated virus whose amino acid identity was maintained but had a reduced specific infectivity. It is hypothesized that this approach of CPB-alteration should be generally applicable to designing new anti-viral vaccines for many types of viruses. This method could have broad applications because CPB-alteration targets an elementary function of all viruses, namely protein translation, as a means to mute viral growth. Since the viruses PV-MinXY and PV-Min were shown *in vitro* to be physically similar to the wild type virus yet attenuated, the next logical step was to test their virulence in an animal model as well as their ability to immunize.

Chapter 4 Figures

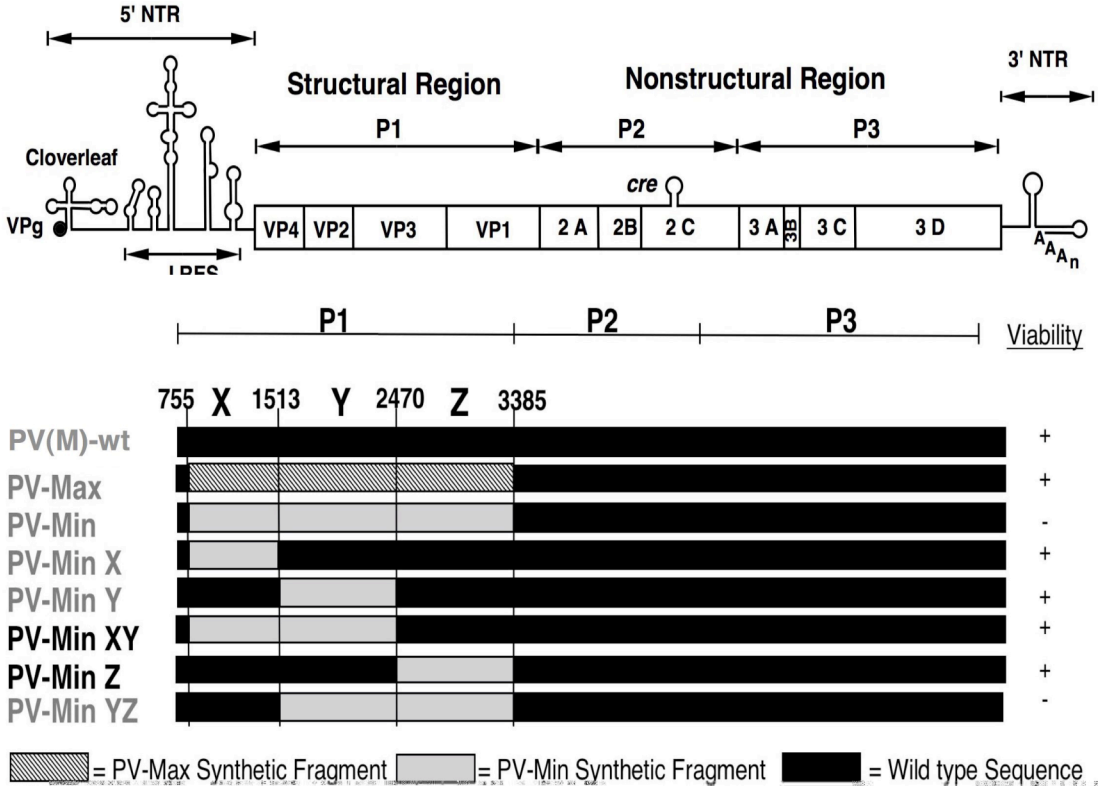


Figure 4.1: Viability of constructed PV-Min subclones. (Top) The poliovirus genome, translated as a single open reading frame. (Bottom) The cDNA of the various chimeric synthetic poliovirus constructs that were constructed via molecular cloning, and their viability in cultured HeLa R19 cells. A virus was obtained containing at least one segment of PV-Min, thus no gross defects were located in any one synthetic region (X,Y, or Z). Nucleotide positions that separate the regions are shown. The first 12 nucleotides (4 codons) in all reading frames are PV(M)-wt sequence to allow proper initiation of translation. Corresponding colors to sequence: Black: PV(M)-wt; Gray: PV-Min; Thatched: PV-Max

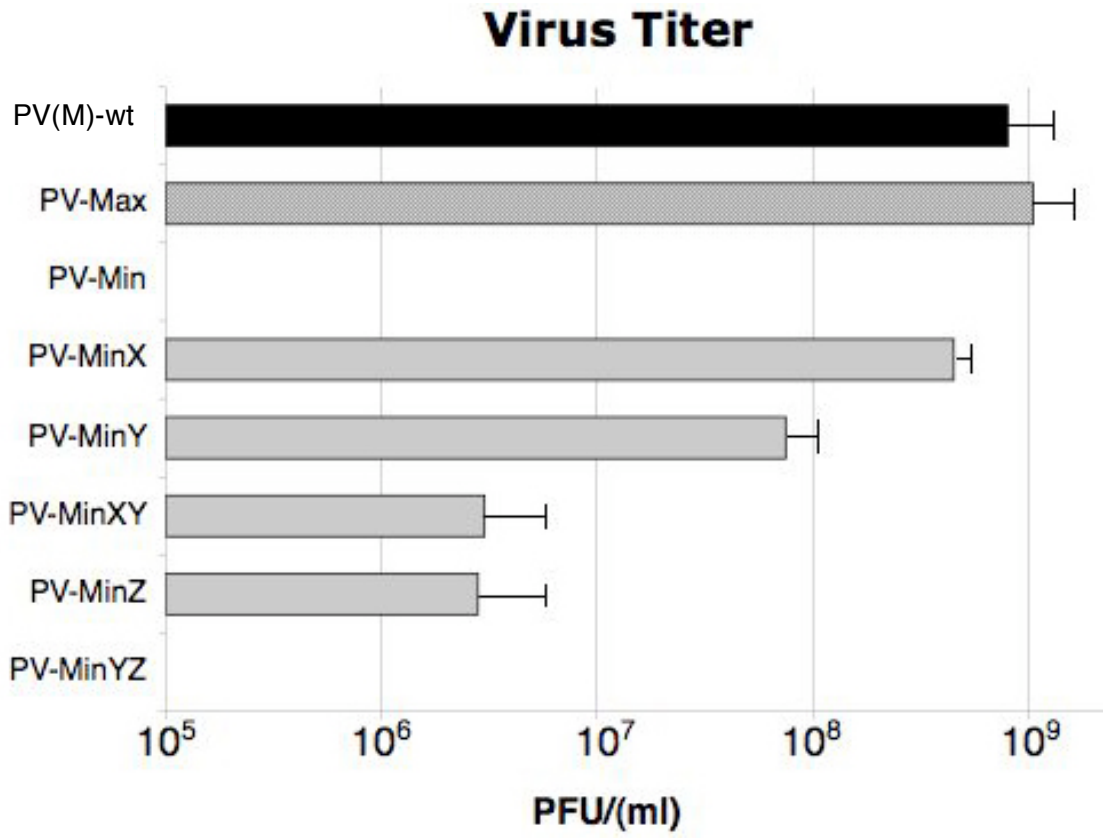


Figure 4.2: Growth titer of synthetic viruses constructed. After the transfection of transcript RNA from all viable viral constructs (Figure 4.1), viruses were amplified via two subsequent passages in HeLa R19 cells. Then the titer that these viruses grew was measured via plaque assay on HeLa R19 cell monolayers. PV-Max and PV(M)-wt viruses grew to a similar titer (~10^9), whereas PV-Min chimera viruses yielded lower PFU/ml titers, with an apparent step-wise decrease in titer. The addition of PV-Min sequence to the P1 region decreased the PFU titer. The graph displays the average titer of three independent experiments with error bars indicating standard deviation values.

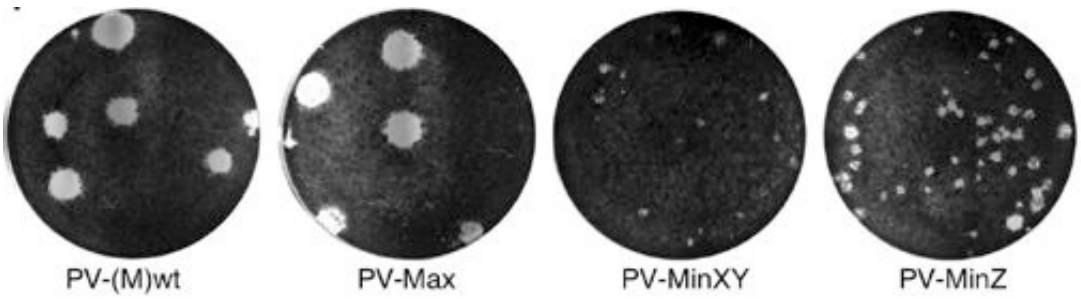


Figure 4.3: Plaque phenotypes of synthetic viruses versus wild type poliovirus. Plaque phenotypes, on HeLa R19 cells, of the viruses PV(M)-wt, PV-Max, PV-MinXY and PV-MinZ developed after 72 hours. PV-Max and PV(M)-wt plaques were visible after 48 hours. The small plaque phenotypes of PV-MinXY and PV-MinZ correlated to the reduced specific infectivity of these viruses.

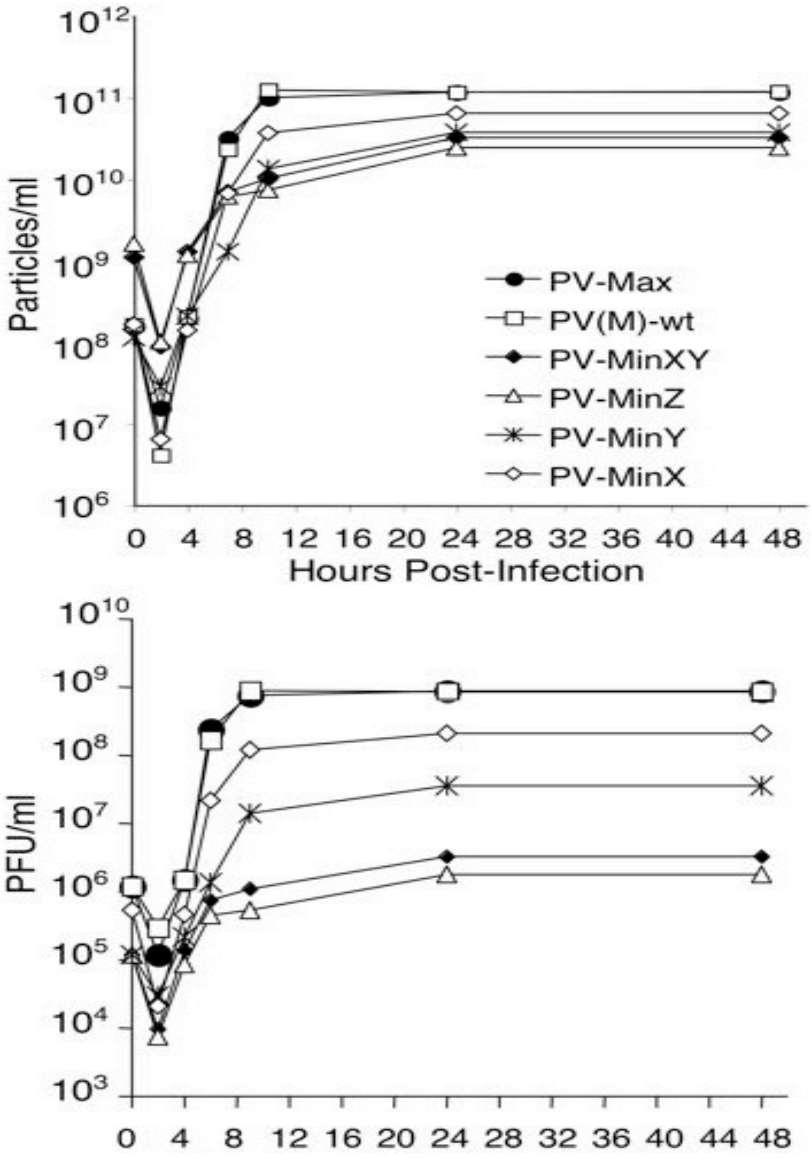


Figure 4.4: Growth kinetics of synthetic polioviruses compared to wild type poliovirus. (Top) The conversion of the calculated PFU/ml growth curve to particles/ml. This conversion was achieved by multiplying the PFU/ml by the respective virus's specific infectivity (Table 4.1). (Bottom) A MOI of 2 was used to infect a monolayer of HeLa R19 cells. The PFU at the given time points (0,2,4,7,10,24,48 hrs) was measured by plaque assay. Each point represents the average of two (PV-MinX, Y, and PV(M)-wt) or three (PV-Max, PV-MinXY, and PV-MinZ) independent experiments. The reduced PFU/ml titer of PV-MinXY and PV-MinZ, yet similar particles/ml titer, correlates to the reduced specific infectivity of these viruses. Symbols: (□) PV(M)-wt, (●) PV-Max, (✧) PV-Min755-1513, (✱) PV-Min1513-2470, (♦) PV-MinXY, (Δ) PV-MinZ.

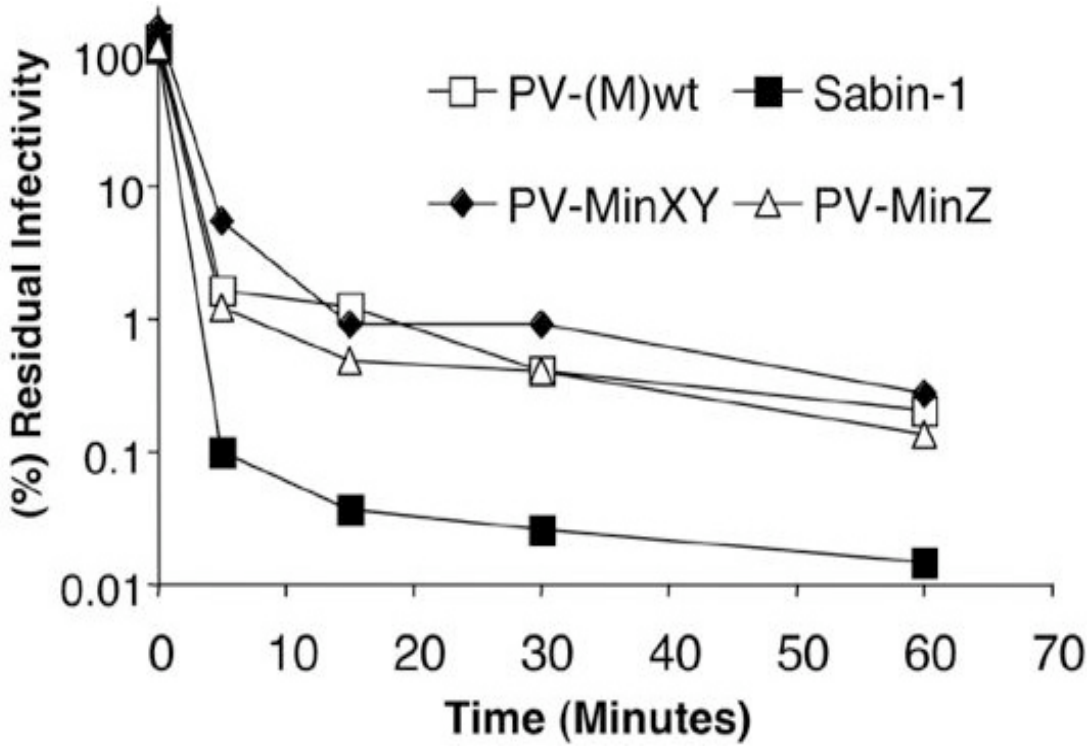


Figure 4.5: Heat stability of synthetic and wild type virions. To rule out that large-scale codon-pair bias modification did not alter the gross morphology of virions, the thermal stability of PV-MinXY and PV-MinZ was examined. An equal number of particles of each virus were incubated at 50°C and the remaining infectivity was quantified via plaque assay after given periods of time (5, 15, 30, and 60 minutes). The capsids of the synthetic viruses had similar stability curves compared to PV(M)-wt, whereas Sabin-1, which has capsid mutations, was less stable. Each point represents the average of two independent experiments.

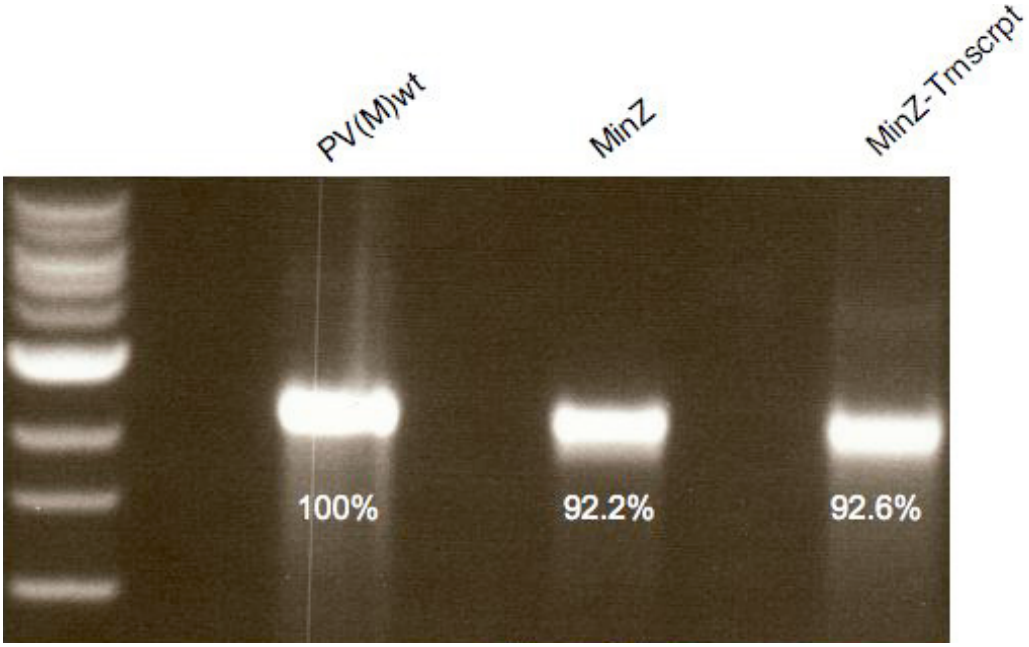


Figure 4.6: Comparison of extracted virion RNA. Purified virion RNA of the respective viruses was isolated from 3 x 10^{11} viral particles and then subjected to a 1% TAE agarose gel. The intensity of each band was measured via the NIH:IMAGE-J program and normalized to the band intensity of the PV(M)-wt RNA. PV-MinZ yielded a similar quantity of RNA as PV(M)-wt. Also, RNA isolated from PV-MinZ was a single molecule of the appropriate length. Transcript RNA (1 μg) was used as a control.

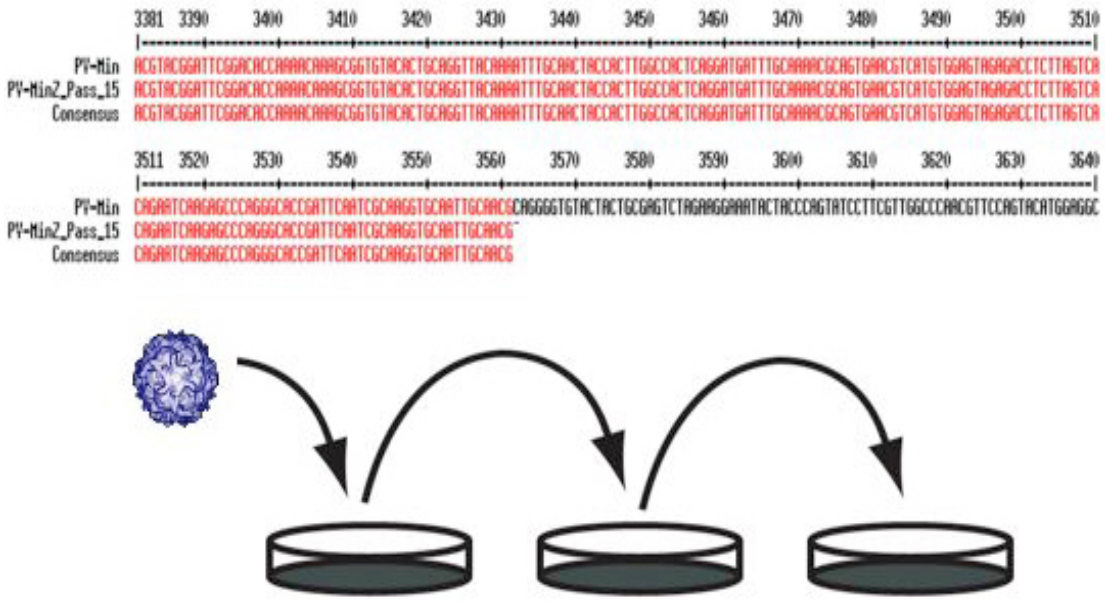


Figure 4.7: Serial passaging of synthetic viruses displayed genomic stability. The genetic stability of the synthetic viruses PV-MinXY and PV-MinZ was assayed by serially passaging the virus at an MOI of 0.5. After passages 5,10, 15, and 17 (XY) or 19 (Z) the virus was sequenced and no fixed mutations were found in the synthetic region. Also, there was no increase in apparent titer (PFU/ml) at each point and no decrease in the time of CPE induction.

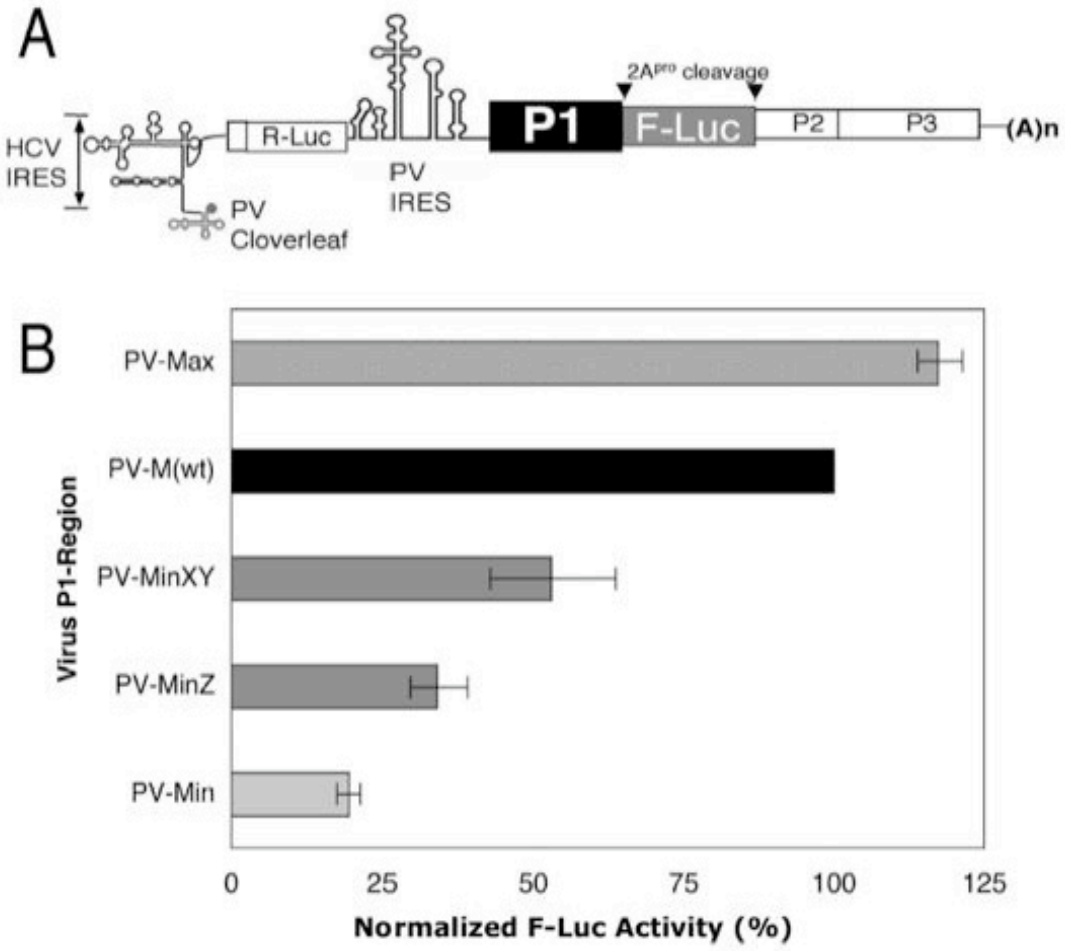


Figure 4.8: The effect of altered codon-pair bias on translation. (A) Structure of a discistronic reporter [3,4]. The first cistron uses the hepatitis C virus (HCV) Internal Ribosome Entry Site (IRES) to initiate translation of *Renilla* luciferase (R-Luc). This first cistron provides an internal control to normalize for the amount of input RNA. The second cistron uses the poliovirus IRES to initiate translation of Firefly luciferase (F-Luc). The region labeled "P1" was replaced by synthetic P1 regions of the CPB-altered viruses. From these constructs (+)RNA was transcribed. (B) Each dicistronic RNA was transfected, in the presence of 2 mM GuHCl, into HeLa R19 cells. After 6 hours, the R-Luc and F-Luc activity was measured. The F-Luc/R-Luc ratio values were normalized to the F-Luc/R-Luc ratio for the PV(M)-wt P1 region, which was set to 100%. The graph displays the average of three independent experiments with error bars indicating standard deviation values. The P1s utilizing under-represented codon-pairs (PV-Min, PV-MinZ, and PV-MinXY) had a decreased rate of translation, whereas the P1 containing over-represented codon-pairs (PV-Max) appeared to have a slight increase in its translation efficiency.

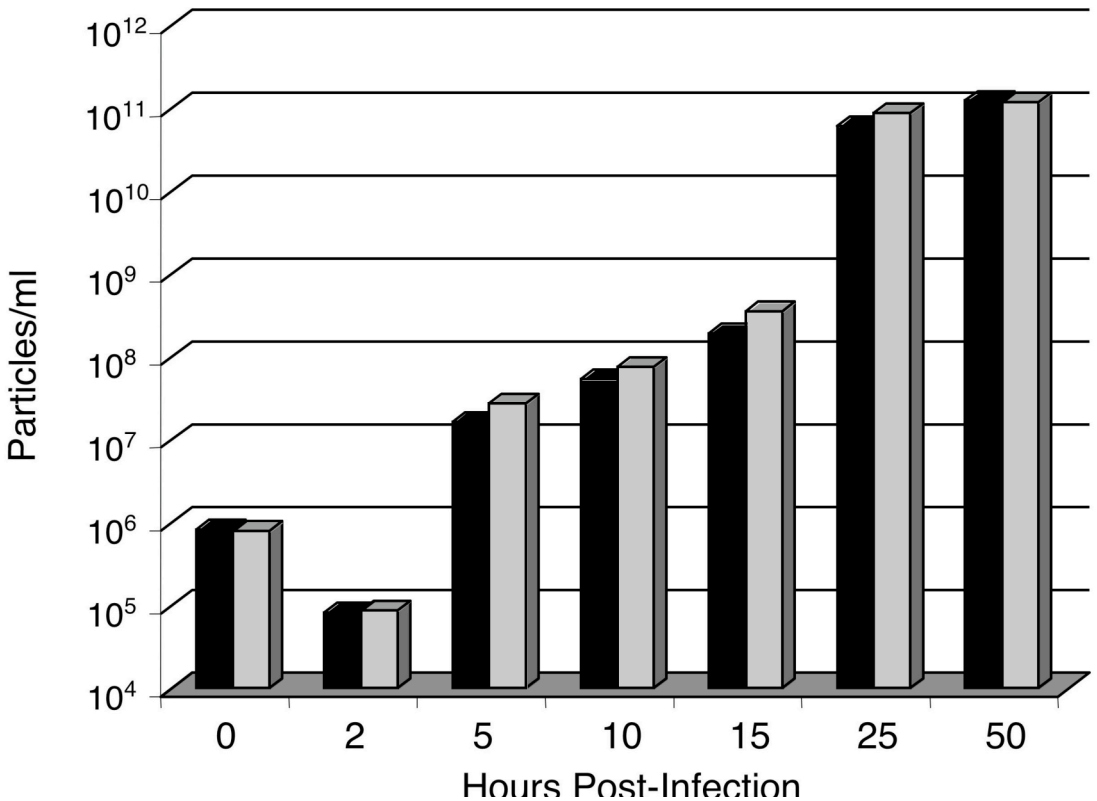


Figure 4.9: Growth curves of PV-Max and wild type poliovirus starting with a low multiplicity of infection. To compare the infection cycle rate and the ability of PV-Max and PV(M)-wt to spread, monolayers of HeLa R19 cells were infected at a low MOI (0.001) and the resulting viral titers were measured at time points post infection via plaque assay. The PFU/ml titer was converted to particles/ml titer via the specific infectivity value of each virus. The two viruses appeared to have identical rates of replication despite the slightly enhanced translation of the PV-Max P1 region. The values presented represent the average of two independent experiments. Black: PV(M)-wt; Gray: PV-Max

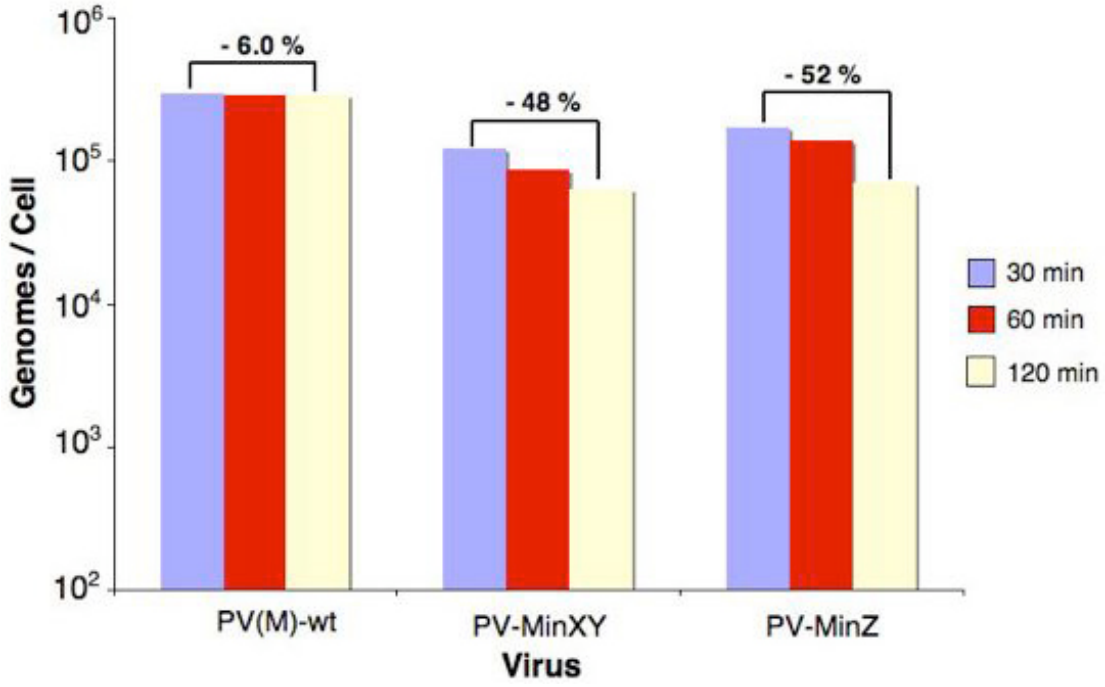


Figure 4.10: *In vivo* RNA stability of synthetic and wild type polioviruses. To assay the *in vivo* stability of viral RNA, a monolayer of HeLa R19 cells was infected with a MOI of 5 and incubated at 37°C for 3 hours. After 3 hours, GuHCl was added to halt replication and then total RNA was extracted at various time points (30, 60, 120 min). The quantity of viral genomes was measured via quantitative RT-PCR. The RNA of the synthetic viruses with altered CPB seemed to degrade with an increased rate compared to PV(M)-wt.

Chapter 4 Table

Virus	A_{260}	Purified Particles/ml[a]	Purified PFU/ml	Relative Specific Infectivity[b]
PV-M(wt)	0.956	8.97×10^{12}	6.0×10^{10}	1 : 137
PV-Max	0.842	7.91×10^{12}	6.0×10^{10}	1 : 131
PV-MinX	0.558	5.24×10^{12}	1.9×10^{10}	1 : 305
PV-MinY	0.509	4.79×10^{12}	9.0×10^{9}	1 : 1,050
PV-MinXY	0.944	8.87×10^{12}	9.6×10^{8}	1 : 9,500
PV-MinZ	0.731	6.87×10^{12}	5.1×10^{8}	1 : 13,500
Sabin-1	0.258	2.40×10^{12}	1.2×10^{10}	1 : 200

Table 4.1: The specific infectivity of synthetic polioviruses compared to wild type poliovirus and Sabin-1.

a) A_{260} - determines particles/ml via 9.4×10^{12} particles/ml = 1 OD_{260} unit [4,73]

b) Calculated by dividing the PFU/ml of purified virus by the Particles/ml

Chapter 5: Examining the Neurovirulence of the Synthetic Viruses and their Ability to Immunize

Summary

To compare the neurovirulence and the extent of *in vivo* attenuation CPB customization achieves, an LD_{50} value was determined for all synthetic viruses by inoculating groupings of CD155tg mice intracerebrally (i.c.) with a range of dilutions of virus ranging from 10^2 to 10^9 particles. These experiments showed that indeed viruses PV-MinXY and PV-MinZ displayed neuroattenuation in mice on the order of 10^3 particles. Since these PV-Min subcloned viruses displayed attenuation in CD155tg mice, they were administered sublethally to immunize mice. Specifically, three doses were injected into CD155tg mice intraperitoneally (i.p.), once a week for three weeks. Fourteen days after the final vaccination, the serum antibody titer was measured and the mice displayed a robust antibody titer. The immunized mice were challenged with a lethal dose of wild type virus and they were protected.

Unfortunately, one mouse from each vaccine cohort died during the vaccinations, which led to an investigation into the sites of viral replication in the mouse. Viral replication was examined in various CD155-expressing organs within the mice. The synthetic viruses, PV-MinXY and PV-MinZ, had limited sites of replication whereas PV(M)-wt could replicate in additional organs. This may explain why these attenuated viruses displayed a smaller LD_{50} range between routes of inoculation compared to

PV(M)-wt. Even though these results (neuroattenuation and immunization) were encouraging, these viruses will need additional modifications to be suitable vaccines.

Introduction

Poliovirus is a human pathogen and the causative agent of the disease poliomyelitis, which is a neurotropic disease causing acute flaccid paralysis. Fortunately, the virus induces these neurological complications in only 1% percent (for Type-1, 0.1% for Type 2 and 3) of those infected, while 90% of individuals infected show no symptoms (Figure 5.1). This virus establishes an infection in the gut of an individual, replicating rapidly to high titers. The virus thus spreads from human to human via fecal-oral route. In the infected individuals, the virus could escape the gut, enter the bloodstream causing viremia, and subsequently spread to the CNS [25]. Since the use of non-human primates, monkeys, as an animal model to study pathogenesis is exceedingly expensive, various mouse models have been developed, each of which have varying degrees of sensitivity [14,145,146]. The one utilized in this study is the Tg21 mouse line developed by Nomoto and colleagues [14]. These mice have been given the human poliovirus receptor CD155 as a transgene. These transgenic mice are sensitive to poliovirus via various inoculation routes; however they do not develop an infection if the virus is administered orally. This inability to infect these mice orally is an intrinsic drawback to using this animal model to study a human pathogen.

The efficient ability that poliovirus can replicate in an infected individual's gut, as well as spread within a population, caused a great drive to develop a vaccine against poliovirus in the first half of the twentieth century. This effort yielded two vaccines, the inactivated poliovirus vaccine (IPV) and the oral poliovirus vaccine (OPV), both of which are able to protect recipients from the onset of poliomyelitis (granted the vaccine itself does not induce disease) [147,148]. Industrialized countries, despite the increased cost of production [78], currently exclusively use IPV because of the risk of vaccine-associated paralytic poliomyelitis (VAPP) associated with OPV administration. In contrast, the WHO has used the OPV as a means to vaccinate the world and attempt to eradicate poliovirus as a scourge causing disease in the third world. As previously stated, multiple drawbacks have arisen with the use of OPV, therefore WHO has initiated a call for a replacement.

Accordingly, the most attenuated CPB-altered viruses PV-MinXY and PV-MinZ were tested as possible poliovirus vaccines in the animal model, CD155tg mice. Should these viruses not yield a suitable live-attenuated strain themselves, the induction of a protective immune response would still be an encouraging step towards utilizing the alteration of CPB to create a live-attenuated virus. Since this means of attenuation targets the genetic code and maintains the exact amino acid sequence as the wild type virus, the customization of CPB could also be combined with other forms of attenuation to yield an attenuated virus as well.

Results

Analyzing the neurovirulence of all CPB-altered viruses

The neurovirulence of the PV-Min chimeras XY and Z and PV-Max was tested in CD155tg mice via intracerebral injection with increasing doses of the viruses. Specifically, groupings of four to six, 6-8 week-old CD155tg mice were inoculated with varying doses and then observed for the onset of poliomyelitis. Control groupings of mice were injected in parallel experiments with PV(M)-wt. Injection doses were based on particles rather than PFU so as to normalize the quantity of virions inserted into the brain. The mice were monitored daily for the onset of flaccid paralysis, the characteristic symptom of poliomyelitis. The standard value used to quantify the virulence of a virus is the Lethal Dose 50 (LD_{50}). This value indicates the dose of inoculating virus that fifty percent of the animals live and fifty percent die. The synthetic viruses PV-MinXY and PV-MinZ had a higher LD_{50} than PV(M)-wt and therefore were, 1,500-fold based on particles or 20-fold based on PFU, less pathogenic (Table 5.1). In contrast, PV-Max virulence was identical to that of PV(M)-wt, despite having a P1 region that was translated slightly more efficiently (Figure 4.8). The neuroattenuation of PV-MinXY and PV-MinZ was not as substantial as that of Sabin-1 that was 300,000 times less pathogenic than PV(M)-wt (Table 5.1). However PV-MinXY and PV-MinZ were more neuroattenuated than PV-AB^{2470} [4].

Growth of synthetic viruses in target tissue and the appearance of reversions

The spinal cords of mice infected with PV(M)-wt, PV-MinXY, and PV-MinZ were removed from mice that had succumbed to the infection. The viral loads in this

target tissue was calculated via plaque assay and then converted to particles per gram of tissue. These titers were measured in order to quantify the ability of the virus to replicate in neuronal tissue (Table 5.2). Similar to viral growth in the tissue culture experiments in Chapter 4, the synthetic viruses PV-MinXY and PV-MinZ produced similar quantities of particles per gram of neuronal tissue as PV(M)-wt. This result indicated that once an infection was established in the CNS by PV-MinXY or PV-MinZ, they were able to grow well, and thus cause poliomyelitis in the infected mouse (Table 5.2).

The viruses extracted from homogenized spinal cords were also sequenced to look for reversion in the viruses, which may have enabled this growth in the CNS. First, the virus extracted from target tissues were amplified on HeLa cells in order to obtain enough template for sequencing. Upon sequencing, it was found that indeed reversions appeared in the synthetic regions of both the PV-MinXY and PV-MinZ viruses extracted from mouse spinal cords. Two separate isolates from mice infected with PV-MinZ, mice MZ1 and MZ2, each had a single synonymous mutation. Mouse MZ1 had a synonymous mutation at nucleotide 3116 and mouse MZ2 had a mutation at nucleotide 2860. These synonymous mutations in fact changed the designed PV-Min codon-pairs from ones with high degrees of under-representation to ones with neutral or over-representation (Figure 5.2). The mutation at nucleotide 2860, in mouse MZ2, resulted in new codon-pair scores that were changed from the original -0.2 and -0.3 to +0.1 and -0.02. The values of the synonymous mutation at nucleotide 3116 in mouse MZ1 can be seen in Figure 5.2.

For mice infected with, and killed by PV-MinXY, the same mutation was found in the synthetic region from two isolates. Interestingly, the exact same mutation was found at nucleotide 987 for both isolates. This mutation resulted in an amino acid change

that actually resulted in the creation of a codon-pair with a greater degree of under-representation. This mutation at position 987 was found in different mice inoculated via different routes of infection, either intracerebrally or intraperitoneally (Figure 5.3). This result was puzzling in that the mutation resulted in an amino acid change but no improvement or alleviation of the CPB customization. The amino acid change that was observed was a Valine to Alanine change, two non-polar amino acids (Figure 5.3). All elucidated mutations may have been a result of the quasispecies effect in poliovirus or provided a selective advantage. The quasispecies hypothesis suggests that a single virus from the inoculum's population makes it through the CNS 'bottleneck' [149] and this single virus results produces the strain that destroys the entire CNS [41]. This quasispecies theory suggests that there is a pool of viruses all with slight (i.e. 1-2 nucleotide) differences. When this pool is injected into a mouse, only one strain randomly makes it through the bottleneck, and this becomes the founder strain that eventually infects the entire CNS. In order to ascertain whether these mutations were a result of the known bottleneck or if they had provided a selective advantage, additional experiments were conducted. It cannot be ruled out that any of these mutations found in mice that succumbed to PV-MinXY or PV-MinZ did not arise prior to inoculation because the infecting virus stocks were not sequenced.

Since the synthetic viruses were shown to be genetically stable in tissue culture (Figure 4.7), it was interesting that reversions were observed after a single inoculation into mice. There is, however, a greater selective pressure in the mouse as compared to the tissue culture dish. To investigate whether these single point mutations could be capable of increasing virulence, extracted viruses were readministered via i.c. injection to

groupings of naïve mice along with a parallel group of naïve mice receiving PV(M)-wt. At the dose of 10^5 particles, a dose 100-fold below the original LD_{50} of PV-MinXY and PV-MinZ, all injected PV(M)-wt mice died, whereas the mice infected with PV-MinXY and PV-MinZ mutants survived. This data indicated these mutant viruses did not have a substantial increase in their virulence as a result of the mutations. This process was repeated twice, i.e. the viruses were passaged through mice, and they were not found to gain an increase in virulence after 3 passages.

Secondly, sequenced viral stocks of PV-MinXY and PV-MinZ, which were the 'wild type' version of these two viruses, were administered again to groups of CD155tg mice to determine if the original mutations were required for the viruses to induce disease in the mice. Should these mutations be required for virulence they would reappear again [15]. After these mice succumbed to poliomyelitis and the virus was extracted, amplified, and sequenced, new single independent mutations were found. The new mutations were not at the same positions as the original mutations and did not increase virulence (data not shown). Therefore, despite having the appearance of reversions, these highlighted reversions seemed not to play a significant role in an increased virulence in the mouse CNS. This result was slightly discouraging in that these viruses began to show slight genetic instability, suggesting that the perceived genetic stability provided by CPB-alteration may require further experimentation (for example single-molecule sequencing of viral populations) to be fully elucidated.

Despite the appearance of reversions, the fact remained that these synthetic viruses, PV-MinXY and PV-MinZ, were significantly attenuated by 10^3 particles, even when administered directly to neuronal tissue. It was possible that PV-MinXY and PV-

MinZ could therefore be suitable for use as live-attenuated viruses capable of immunizing mice. These viruses were administered sub-lethally and were the first model CPB-altered viruses that could serve to function as a vaccine. Since these were pioneering viruses, it is recognized that there are great possibilities for additional modifications before the production of suitable vaccines.

PV-MinXY and PV-MinZ as vaccine candidates

To test the functionality of PV-MinXY and PV-MinZ as vaccine candidates, a dose of 10^8 particles were administered via i.p. injections to groups of eight CD155tg mice, with an interval of one week between each of the three administrations. The regimen used followed a previously determined 'vaccine' regimen for an attenuated poliovirus [119]. All mock-vaccinated mice survived the regimen. However, one mouse from each 'vaccine' cohort succumbed to poliomyelitis after the initial injection, therefore seven mice remained. Seven days following the final vaccination, the anti-poliovirus antibodies raised by these immunized mice were measured to quantify the humoral response elicited.

The presence of neutralizing antibodies was tested via the collection of blood from the tail vein of each mouse. The control serum samples were extracted from two mock-vaccinated CD155tg mice. Titers of poliovirus-neutralizing antibodies in mouse serum samples were determined by a microneutralization assay [119]. This assay used serial dilutions of extracted serum added to the supernatant of HeLa R19 cells. Then, 10^2 PFU of PV(M)-wt was used to infect these HeLa R19 cells. The infected cells were then observed for the onset of CPE for four days. The last dilution that viral growth was

inhibited defined the anti-PV(M)-wt antibodies in the serum. The viruses PV-MinXY and PV-MinZ were able to elicit a robust immune response (Figure 5.4A), whereas the mock-vaccinated mice had anti-PV(M)-wt antibodies below the level of detection. Importantly, capsids of these viruses had the exact same amino acid composition as the wild type virus so the antibodies they produced were able to readily identify the PV(M)-wt virus in the microneutralization assay.

Since these mice produced a robust immune response, these groups of seven mice were then challenged fourteen days following the final vaccination via intramuscular injection with 10^8 particles of PV(M)-wt, a lethal dose [119]. This was performed to examine whether the robust immune response raised by these immunized mice was capable of protecting them from invasion of the CNS by PV(M)-wt. Following the challenge, mice were monitored for the onset of disease symptoms, such as paresia, flaccid paralysis, or death throughout the length of the experiment. All immunized mice in both 'vaccine' cohorts, PV-MinXY and PV-MinZ, were protected after challenge with PV(M)-wt, whereas all mock vaccinated mice developed poliomyelitis and did not survive (Figure 5.4B).

This immunization experiment was a partial success because indeed these synthetic viruses could initiate a robust immune response when administered as a 'vaccine,' and this immune response was able to protect against PV(M)-wt challenge. Unfortunately, one mouse from each cohort developed disease during the i.p. dosing regimen. This result led to the hypothesis that the LD_{50} by different routes of inoculation for the viruses PV-MinXY and PV-MinZ did not vary as much as with PV(M)-wt. PV(M)-wt had a $10^{3.8}$-fold difference between i.c. (10^4) and i.p ($10^{7.8}$) inoculation routes

(Table 5.1 and 5.3). The LD_{50} values when calculated for PV-MinXY were, i.c. ($10^{7.1}$) and i.p ($10^{8.7}$). The LD_{50} values when calculated for PV-MinZ were, i.c. ($10^{7.3}$) and i.p ($10^{8.7}$). Interestingly, the LD_{50} for the viruses PV-MinXY and PV-MinZ had only one log discrepancy between i.c. and i.p. injection, unlike PV(M)-wt (Table 5.1 and 5.3). Since these synthetic viruses maintained a level of neurovirulence via i.p. this indicated that the synthetic viruses were replicating and destroying neuronal tissues. It was possible, however, the infection of other tissues may have been severely impaired.

To examine extra-neuronal replication, groupings of three CD155tg mice were injected i.p. with 10^9 particles of PV-Min, PV-Max, and PV(M)-wt. Mice were sacrificed after the development of poliomyelitis and extra-neuronal tissues were extracted, homogenized, and measured for virus replication by plaque assay. Outside of the CNS, the synthetic viruses did not infect the kidneys, like PV(M)-wt. If the synthetic viruses did infect extra-neuronal tissue, such as the pancreas, their replication cycle was muted or also absent (Figure 5.5). In the case of the pancreas is important to note that due to PV(M)-wt's lowered specific infectivity, PV(M)-wt yielded 50-fold more particles per gram of tissue, yet 500-fold more infectious virus than the synthetic viruses. The extra-neuronal replication results for PV(M)-wt correlate with a previous study [150].

One hypothesis for this result was that there may have been intrinsic properties of the CNS that overcome the translation hindrance induced by the CPB-alteration (refer to Chapter 5 discussion). Also, the synthetic viruses PV-MinXY and PV-MinZ could be affected by the immune status of the various organs that they might infect since they have a restricted tissue tropism. It has previously been shown that PV(M)-wt itself was affected by interferon expression in transgenic mice and that CD155tg : interferon(IFN)

α/β knockout mice permitted PV(M)-wt replication in many organs compared to CD155tg mice [150]. The initial assay examining the sensitivity of PV-MinXY and PV-MinZ to IFN was conducted *in vitro* by comparing growth on HeLa R19 cells to that of Vero cells, which lack the IFNα gene and thus IFNα cannot be secreted.

To test the susceptibility of the synthetic viruses to IFN, known dilutions of the viruses PV-MinXY, PV-MinZ, and PV(M)-wt were plaque assayed simultaneously on HeLa R19 cells and Vero cells. The wild type virus grew 7.9-fold less efficiently on Vero cells (Table 5.4). The synthetic viruses PV-MinXY and PV-MinZ grew approximately 2.6- or 3.2-fold less efficiently on Vero cells, thus they had a 3.1- or 2.4-fold increase in growth compared to PV(M)-wt (Table 5.4). The growth of the synthetic viruses PV-MinXY and PV-MinZ in Vero cells was not as poor when compared to PV(M)-wt, indicating the absence of IFNα may have benefited their growth. If IFN played no role, the synthetic viruses should have grown with the same deficiency on Vero cells as PV(M)-wt (i.e. 1 : 7.9). They still grew with decreased efficiency but this reduction was not as significant as PV(M)-wt. Therefore, the CPB-altered virus may have had an increased sensitivity to the innate immune response, possibly due to increased immune stimulation resulting from CPB-alteration. This avenue is an interesting future direction when exploring the *in vivo* effects of CPB. Specifically, investigating the impact of the CpG dinucleotide on increased immune stimulation because this dinucleotide is present between the most under-represented codon-pairs and may play a role in the innate response. It has been shown that single-stranded RNA containing CpG motifs were able to stimulate monocytes [109]. Therefore, these CPB-

altered viruses containing under-represented codon-pairs (i.e. a large number of CpGs) could upregulate the immune response.

Discussion

Indeed the results presented in this Chapter are encouraging for using CPB-alteration as a means to produce a live-attenuated virus capable of serving as a vaccine, even if the two candidate viruses, PV-MinXY and PV-MinZ, failed to achieve this goal. The viruses PV-MinXY and PV-MinZ that had a negative CPB, i.e. used under-represented codon-pairs, were attenuated when used to infect CD155tg mice. When administered sublethally in a 'vaccine' type regimen, these viruses were able to successfully immunize susceptible mice and protected these mice upon challenge with PV(M)-wt (Figure 5.4). These synthetic viruses, however, are not suitable in their current state to be used as vaccines. They were not as attenuated as the current vaccine strain Sabin-1 (Tables 5.1 and 5.3). Secondly, these viruses also had an interesting characteristic where they seemingly targeted the CNS for replication. This CNS targeting was inferred because the synthetic virus PV-MinXY and PV-MinZ had a reduced LD_{50} range between i.p. and i.c. injections and also seemed to grow selectively in neuronal tissue (Figure 5.5).

These synthetic viruses may therefore only grow in motor neurons due to intrinsic properties of these cells and cannot replicate efficiently in any other organs that express the CD155 receptor. This may occur due to various characteristics of the CNS and motor neurons in particular. Firstly, motor neurons are under-surveyed by the immune system,

which is evident by the marked decreased expression of MHC-I molecules on the cell surface. This decreased expression is because these cells cannot be regenerated if destroyed, such as by $CD8^+$ cells [151]. Secondly, motor neurons are the largest cells in the human body as well as the largest cells in the CNS. This massive size correlates to an increased metabolic rate and, in all likelihood, the highest metabolic rate of any other cell type in the body [152]. Specifically, as the motor neuron increases in size, the rate of protein synthesis increases [153]. This is logical because the motor neuron must have a high rate of cellular processes in order to maintain homeostasis in such a large cell, as well as be continually receiving and responding to impulses. Thus, the rate of translation in motor neurons may be markedly higher than that of other cell types. This high rate of translation may be able to overcome the method of attenuation (i.e. down-modulating translation) that the CPB deoptimization implements.

Therefore, the neurotropic nature of PV(M)-wt, possessing the ability to replicate in motor neurons, would allow these apparently attenuated viruses to still induce paralysis. Since PV-MinXY and PV-MinZ show decreased replication in other organs, the only place they may have been able to replicate was the CNS. This was a disastrous possibility for the construction of a live-attenuated virus with an intended use as a poliovirus vaccine. However, the results presented are encouraging for the use of applying CPB-alteration as a method for constructing live-attenuated vaccines for non-neurotopic viruses.

Chapter 5 Figures

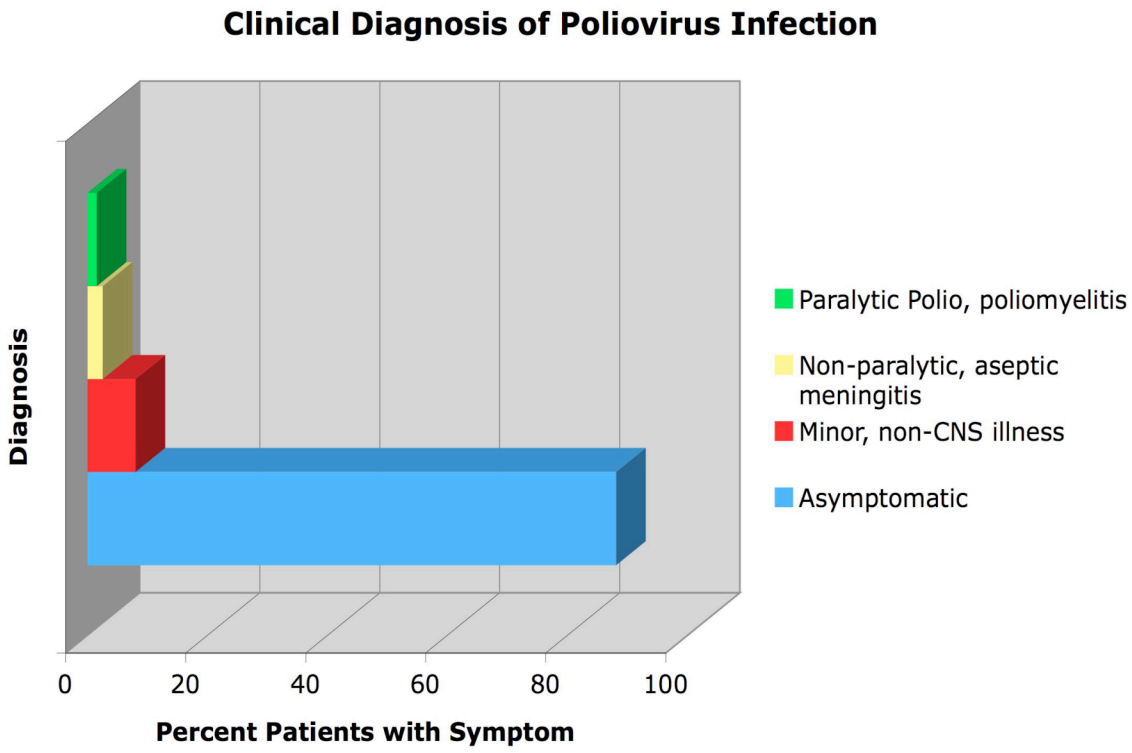


Figure 5.1: Diagnosis associated with Type-1 poliovirus infection. The majority of infected individuals are asymptomatic, whereas about 1-8% of infected individuals develop neurological complications. Graph constructed form publicly available data from the Center for Disease Control.

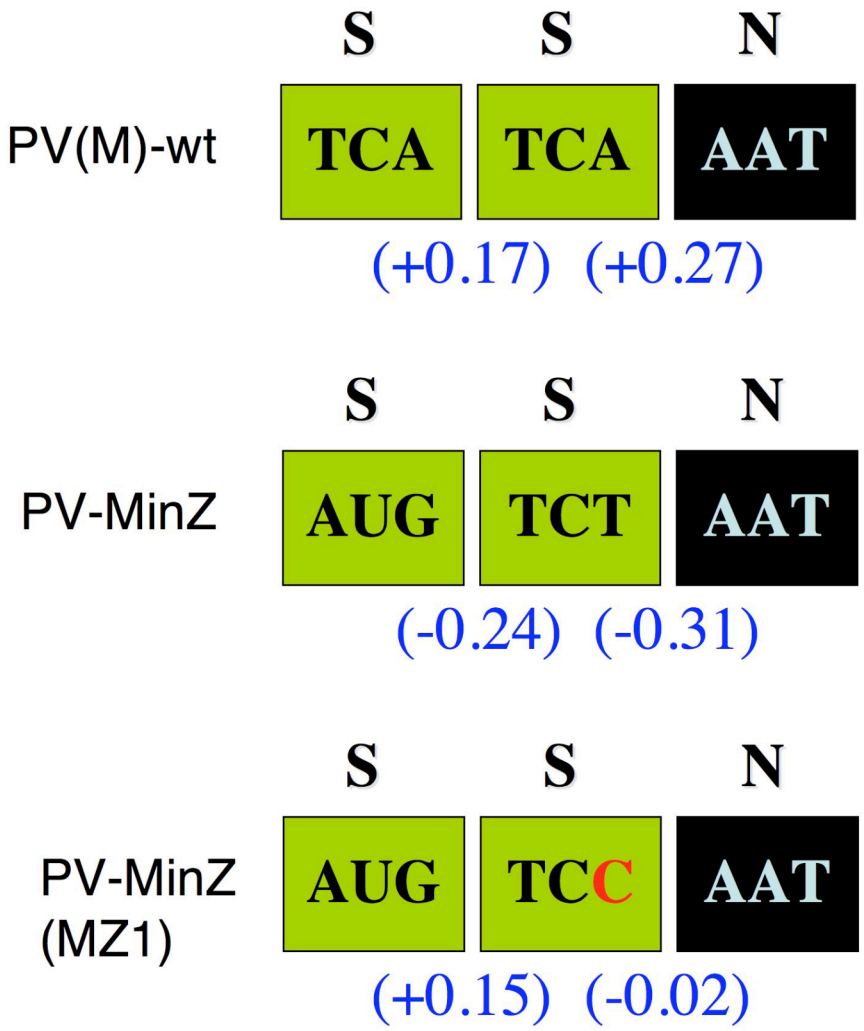


Figure 5.2: Example of a mutation found in the virus extracted from the spinal cord of a CD155tg mouse infected with PV-MinZ. The spinal cord from a mouse that succumbed to poliomyelitis was extracted and the infecting virus was isolated. The mouse was infected via i.c. injection with PV-MinZ at a dose of 10^9 particles. The mutation is indicated in **Red** and occurred at nucleotide 3116. This mutation did alleviate two negative CPSs, however this mutation did not increase the virulence of the virus upon re-administration to naïve mice. Negative values correspond to under-represented codon-pairs, whereas positive values correspond to over-represented codon-pairs.

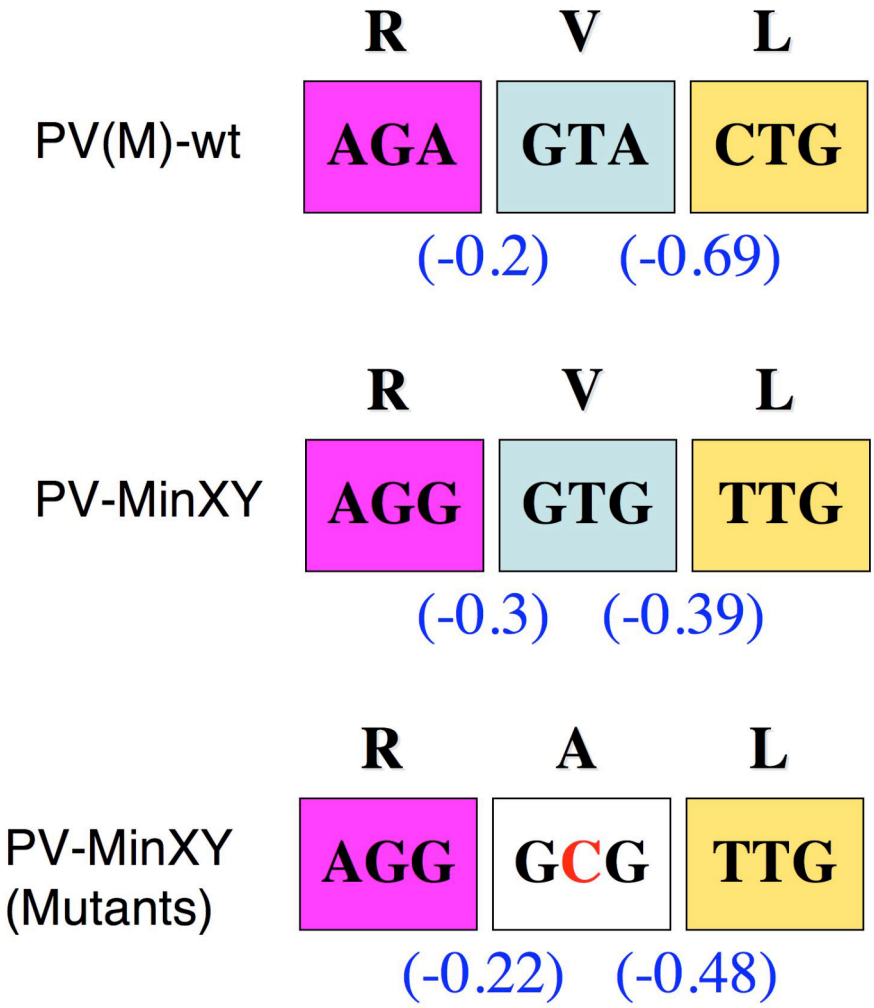
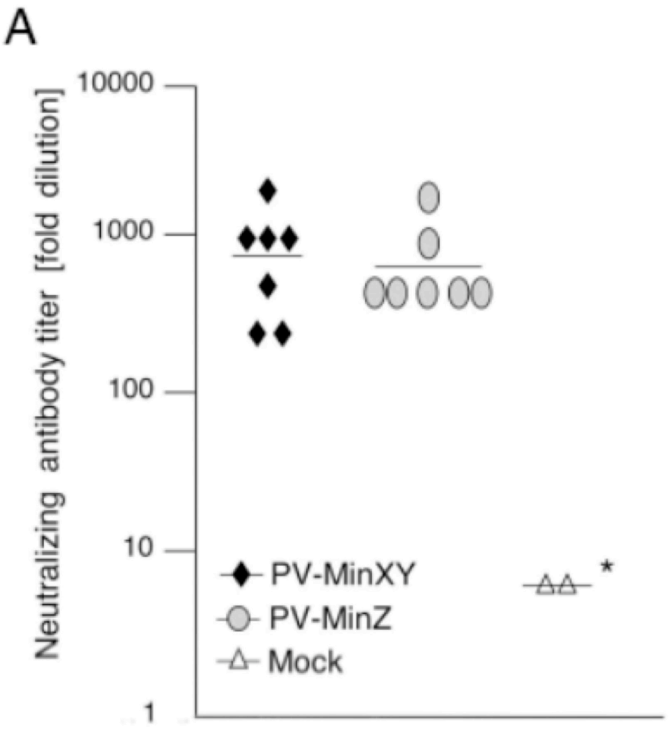


Figure 5.3: The mutation of PV-MinXY found in the spinal cords of infected CD155tg mice. The spinal cord from a mouse that succumbed to poliomyelitis was extracted and the infecting virus was isolated. The mouse was infected via i.c. injection with PV-MinXY at a dose of 10^9 particles. The mutation is indicated in **Red** and occurred at nucleotide 987. This mutation did not alleviate the negative codon-pairing and resulted in an amino acid change, from Valine to Alanine. This mutation did not increase the virulence of the virus upon re-administration to naïve mice. Negative values correspond to under-represented codon-pairs, whereas positive values correspond to over-represented codon-pairs.



B

Virus	Mice Protected (out of 7)
PV-MinZ	7
PV-MinXY	7
Mock vaccinated	0

Figure 5.4: Induction of neutralizing antibodies by PV-MinZ and PV-MinXY.
(A) Groups of eight CD155tg mice, seven of which completed the regimen, were each inoculated by intraperitoneal injection three times, at weekly intervals, with 10^8 particles of PV-MinZ and PV-MinXY. The serum conversion was measured 10 days after the final vaccination. The average neutralizing antibody titer for each virus construct was measured via microneutralization assays. The neutralizing antibody titers were 878 for PV-MinXY and 805 for PV-MinZ. (*) No virus neutralization for mock-vaccinated animals was detected at the lowest tested dilution of 1:8. (B) Immunized mice survived a challenge with a lethal dose of PV(M)-wt, whereas mock-vaccinated mice succumbed to infection. The challenge was administered via intramuscular injection of 10^8 PV(M)-wt particles.

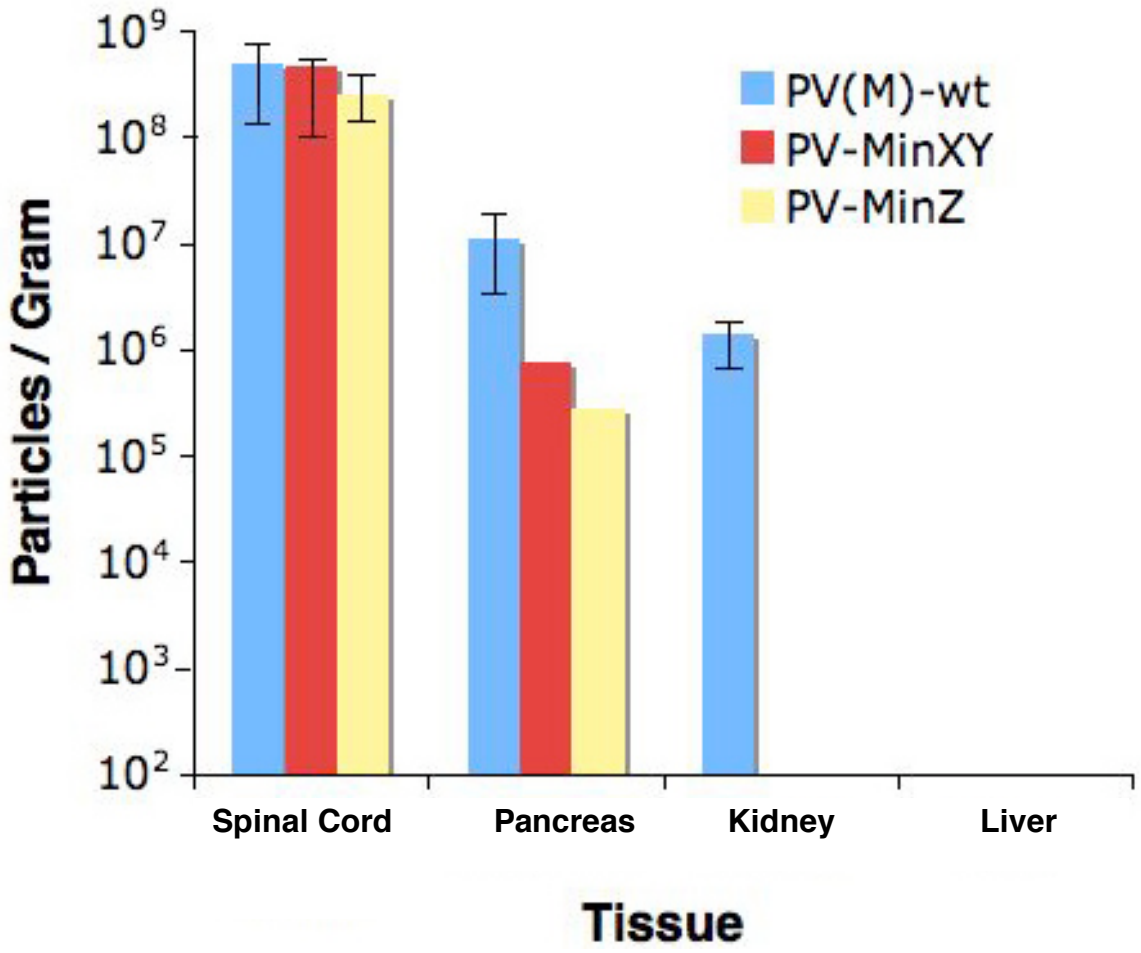


Figure 5.5: Replication of PV(M)-wt, PV-MinXY, and PV-MinZ in neuronal and extra-neuronal tissues. To compare the pathogenic tissue tropism of the synthetic viruses with PV(M)-wt, CD155tg mice were infected via i.p. and their organs removed and homogenized. Viral titers were measured via plaque assay. The values are expressed as particles/gram, which were calculated by multiplying the PFU titer by each virus's specific infectivity. Note: The wild type virus had a lower specific infectivity (Table 4.1) so, in the pancreas for example, PV(M)-wt yielded 50-fold more particles per gram yet 500-fold more infectious virus. The extra-neuronal replication results for PV(M)-wt correlate with previous studies [150].

Chapter 5 Tables

Virus	LD_{50} (i.c.) (Particles)	LD_{50} (i.c.) (PFU)
PV(M)-wt	$10^{4.0}$	$10^{1.9}$
PV-Max	$10^{4.1}$	$10^{1.9}$
PV-MinY	$10^{5.0}$	$10^{2.2}$
PV-MinXY	$10^{7.1}$	$10^{3.2}$
PV-MinZ	$10^{7.3}$	$10^{3.2}$
Sabin-1	$10^{9.5}$	$10^{7.4}$

Table 5.1: Calculated LD_{50} for wild type poliovirus, Sabin-1, and synthetic viruses via intracerebral (i.c.) injection.

Virus	Viral Particles per gram of s.c.
PV(M)-wt	5.1×10^{11}
PV-MinXY	9.7×10^{10}
PV-MinZ	4.8×10^{11}

Table 5.2: Calculated viral particles per gram of spinal cord (s.c.) for poliovirus and synthetic viruses via i.c. injection. Virus recovered from the spinal cord of infected mice at the time of death or paralysis, expressed in particles per gram of tissue. The values were derived by multiplying the PFU per gram of tissue titer by the virus's specific infectivity value in Table 4.1. The determined values are the average of three independent mice that succumbed to poliomyelitis. All mice were injected i.c. with 10^8 particles.

Virus	LD_{50} (i.p.) (Particles)
PV(M)-wt	$10^{7.8}$
PV-MinXY	$10^{8.7}$
PV-MinZ	$10^{8.7}$
Sabin-1	$>10^{10}$

Table 5.3: Calculated LD_{50} for poliovirus and synthetic viruses via intraperitoneal (i.p.) injection

Virus	Relative Growth Vero : HeLa
PV(M)-wt	1 PFU : 7.9 PFU
PV-MinXY	1 PFU : 3.2 PFU (2.4-Fold Increase)
PV-MinZ	1 PFU : 2.6 PFU (3.1-Fold Increase)

Table 5.4: Relative growth of wild type poliovirus and synthetic polioviruses on Vero cells versus HeLa R19 cells. A known solution of virus was plaqued in parallel on Vero cells and HeLa cells and the resulting PFU titer on each cell type was compared. The results indicated are the average of five independent experiments.

Chapter 6: Summary and Future Directions (Influenza A virus and the CpG dinucleotide)

Summary

Construction and characterization of codon-pair customized polioviruses

Synthetic Biology, specifically combining the synthesis of large DNAs with new methods for customizing living systems [4,16,154], has allowed for the re-engineering of organisms for particular purposes. These purposes included studying biological phenomena as well as the development of vaccines [3]. In this study, the codons within the coding sequence of the poliovirus's P1 region were shuffled via the computer-based algorithm DPapa to produce novel encodings. These novel encodings maintained the exact same amino acid sequence and codon usage yet new codon-pairs were formed. The viral construct with a P1 region consisting, on average, of under-represented codon-pairs was called PV-Min and the virus with a P1 region utilizing over-represented codon-pairs was PV-Max. Although it has been recognized that codon-pairing is statistically biased [100], this phenomenon has previously been studied primarily by informatics [101,104].

The study presented here found that a poliovirus recoded to use under-represented codon-pairs was attenuated. In tissue culture, this attenuation took the form of a reduced specific infectivity, whereby particles of the PV-Min chimeric viruses had a reduced probability of initiating an infectious cycle; however, once a successful infectious cycle was achieved, the PV-Min chimeric viruses could yield only 4- to 8-fold less viral

particles per cell when compared to PV(M)-wt. Secondly, these viruses displayed attenuation in a susceptible animal model, CD155 transgenic mice, and were able to provide protective immunization of these mice when administered as a 'vaccine.' Despite this immunizing ability, the synthetic viruses in their current construction are not suitable vaccines. This is because they still caused disease at too low a dose to be safe. It is still encouraging that CPB-alteration was able to substantially attenuate the PV-Min chimeric viruses.

The cause of this reduced specific infectivity, and thus attenuation *in vivo* and *in vitro*, was not due to gross defects of physical properties of the virus capsid or its genome. Rather, the attenuation was due to the CPB customization of the PV-Min chimera viruses. Specifically, this alteration of CPB, such that the PV-Min chimera viruses now had portions of their P1s utilizing under-represented codon-pairs, caused a reduction in the rate of translation. One theory for the existence of codon-pair bias and its influence on translation rates is that certain tRNAs interact poorly on the ribosome [105], and thus the codon-pairs causing the juxtaposition of such tRNAs are under-represented. The translation data was consistent with this theory. Importantly, the data presented in this work was the first *in vivo* exploration examining the effect CPB has on translation in the eukaryotic system.

An important corollary to this work is that the resulting attenuation of the PV-Min chimera viruses was not simply caused by random changes in synonymous codons if those changes do not systematically reduce or codon-pair bias. Previously, we constructed the virus PV-SD, with 937 mutations in synonymous codons in the P1 region [4]. In PV-SD, neither codon bias nor codon-pair bias was changed, and that virus was

not attenuated (Table 3.1). An additional proof of this principle, that the attenuation of the PV-Min viruses was due to under-represented codon-pairs and not random changes, is the PV-Max virus that was constructed with 566 synonymous mutations and was also not attenuated (Chapters 3, 4, and 5). Despite containing over-represented codons and translating its P1 with an improved rate, PV-Max was not more virulent than PV(M)-wt. This result may be due to the fact that evolution has already selected an optimized encoding of PV(M)-wt. Also, it is possible there is an upper limit to the speed that a virus can grow. Namely, a virus needs to be in equilibrium with the host cell, such that too rapid a growth rate could be detrimental to the viral replication cycle. For example, if a virus kills the cell too fast, it may not make as many viral particles as it could during its replication cycle. Therefore, the idea of creating a more virulent virus than PV(M)-wt is discredited because the wild type already has a replication cycle that kills a cell in eight hours, so a 'super-virus' that would replicate faster would actually be unlikely.

For the PV-Min virus, the data presented suggests that many of the 631 mutations in PV-Min caused small, additive defects that in turn debilitated the virus by reducing its translation efficiency. The correlation between the degree of codon-pair deoptimization and the degree of viral attenuation, i.e. the more negative the alteration of CPB the more attenuation, in combination with the lack of viral reversion upon passaging *in vitro*, both lend support to this hypothesis. Reversion however did appear when extracting and sequencing the virus from animals that were killed by the synthetic viruses, suggesting this hypothesis may be incomplete. These mutations, however, were not shown to increase the virulence of the virus. Also, the exact same mutations did not reappear when readministering the virus, suggesting these mutations were a result of the quasispecies

phenomenon [149]. Such that there are many underlying 'strains,' each carrying a point mutation or two, and one of these strains slides through the bottleneck, amplifies in the CNS, and kills the animal. This idea of a quasispecies cloud of viruses can be assumed because the poliovirus polymerase has an error rate of approximately 1 per 10^4 nucleotides copied [155]. When sequencing the *in vitro* passaging experiments, these single changes washed away after each round of replication and the consensus sequences remained while these transient underlying strains were not seen. This suggests that in fact the synthetic region reasonably was stable *in vitro* and no mutation could have alleviated the attenuating effect of CPB-alteration. When this population of viruses were administered to an animal, a weeding out happened in the form of a bottleneck, which has been observed for poliovirus's entrance to the CNS [149]. Therefore, only one of these 'strains' made it through and established the infection of the CNS, which eventually killed the mouse. If these mutations were required for virulence, they would have reappeared every time the virus was administered [15]; however, since that did not happen in the case of PV-MinXY and PV-MinZ, the results suggested the mutations seen were a result of the phenomenon of quasispecies. A future experiment that would aid in confirming this hypothesis would be to plaque purify the virus and sequence the individual virus clones in each plaque. If indeed there are single changes scattered throughout the P1 regions of the clones then this population, like PV(M)-wt, is heterogeneous under the consensus sequence.

The most significant results relating to the data presented is the possibility of CPB-alteration playing a role in creating new vaccines for various types of viruses. By de-optimizing CPB, one could produce a virus with a custom level of attenuation. There

are multiple benefits to using CPB-alteration as a means for producing a vaccine, some of which were explored and defined in the work presented. Firstly, a vaccine must induce an immune response to a pathogen, which is dependent on the three-dimensional structure of the pathogen as well as its primary amino acid sequence. When implementing CPB customization, via synonymous changes, one constructs an attenuated virus that encodes the exact amino acid sequence as the progenitor wild type virus. Since this virus is identical in structure, it is possible this virus could elicit the appropriate immune response. Next, it can be combined with other attenuating changes, such as the adaptation of the virus to growth at low temperatures or other synthetic biology approaches [6,119]. This combination would therefore take advantage of additional modes of attenuation, while providing the unique advantages afforded by CPB customization. The strategy of CPB customization is in contrast to existing methods of attenuation, which typically depend on a small number of mutations that can revert. These attenuating changes are inserted randomly by passaging, whereas the CPB approach specifically targets the coding sequence allowing for directed manipulation.

In sum, since CPB-alteration targets an elementary function common to all viruses, namely protein translation, this approach could be generally applicable to designing new anti-viral vaccines for many types of viruses. With this future direction ever present, CPB-alteration has already been applied to the genome segments of influenza A virus with the goal of attenuating the virus and thus creating a live vaccine.

Future Directions

Influenza A virus

The specific hurdle for producing a live-attenuated vaccine for poliovirus (PV) using CPB-alteration is that the virus selectively targets motor neurons for destruction once the PV infection has spread to the CNS [156]. These cells cannot be regenerated if destroyed, such as during a PV infection that lyses the host cell. However, if one were to apply CPB-alteration to a virus that targets cells capable of regeneration, it could prove extraordinary successfully. Since CPB-alteration targets protein translation, this approach could be applied to viruses that currently have reverse genetics systems. The initial work investigating the application of CPB customization to another virus has begun using the virus Influenza A PR/8/34.

Influenza is a well-known human pathogen that, between 1990 and 1999, was responsible for an average of 35,000 deaths annually in the US [157]. The impact of Influenza A virus on the US economy has been estimated to exceed $23 billion due to approximately 200,000 hospitalizations as well as decreased work efficiency and productivity [158]. Globally, 300,000 to 500,000 people die each year due to influenza [159]. Although the virus causes disease amongst all age groups, the rates of serious complications are highest in children and persons >65 years of age. Another significant aspect of influenza is its potential to mutate or re-assort gene segments such that the virus is now highly pathogenic and the population is immunological naïve to this novel strain [160]. The most pertinent example of this was the 1918 Spanish Flu pandemic, which is estimated to be responsible for 30 million deaths [161]. With the prospect of another influenza pandemic [162] and a possible high mortality rate of a resulting pandemic due

to the pathogenicity of the Influenza A virus [163], there is an increased need for a novel flu vaccine that is effective and could be rapidly produced.

The experiments investigating the applicability of CPB-alteration to Influenza A virus have begun, under the direction of Dr. Steffen Mueller of the Department of Molecular Genetics and Microbiology. The computer-based algorithm DPapa was used to shuffle the codons of the coding region of the nucleoprotein (NP) segment of Influenza A virus, thus creating a segment that now has a negative CPB (i.e. uses under-represented codon-pairs). Then using molecular cloning and the established reverse genetics system for Influenza A PR/8/34, a viable and slightly attenuated virus was obtained named PR/8/34/NP^{MIN}, which had a small plaque phenotype (data not shown). This virus's overall titer was not reduced dramatically compared to the wild type strain; thus the possibility to alter the CPB of more segments exists and this is the present avenue of future investigations.

Investigating the impact of the CpG dinucleotide

The use of synthetic biology now permits the investigation of biological phenomena that were formerly impractical using previous experimental techniques. The ability to synthesize DNA has allowed researchers to alter gene structure to desired extents. Take for example the work presented here where the coding sequence for the poliovirus P1 region was shuffled so that the phenomena of CPB could be investigated. An interesting out growth of this study was the fact that codon-pairs that are statistically under-represented in the human genome contained the dinucleotide C – G, where the C is the third nucleotide of the first codon and G is the first nucleotide of the second codon.

The reason why this dinucleotide is under-represented is not specifically known. The CpG dinucleotide could have an impact on translation [3], be repressed due to genomic forces, or the presence of CpG could be used to distinguish self from non-self [164]. The two viral constructs used in this study, PV-Max and PV-Min, were the first poliovirus genomes that had their CpG content altered dramatically as compared to PV(M)-wt (Figure 6.1). Interestingly, the greater the CpG content over the wild type value, the more substantial the level of attenuation (Figure 6.1, Figure 4.2). However, in the case of the PV-Min chimeras, they have both an increased CpG content as well as a negative CPB. To understand which phenomena is, in fact, responsible for the attenuation (or it could be a combination of both), genome synthesis has been employed again to investigate the biological phenomenon of CpG.

Based upon the result presented in the previous chapters, two additional viral constructs have been synthesized and cloned. Initially, the virus PV-Max had a positive CPB and small CpG content, and PV-Min had a negative CPB and large CpG (Figure 6.1). The two new synthetic P1 regions synthesized altered these two characteristics. The first new synthetic P1 region is PV-wtMaxCpG, which has a PV(M)-wt CPB and a large appearance of CpGs. The second region is designated PVMin-wtCpG, which has a negative CPB but a wild type appearance of CpGs. These synthetic constructs were synthesized and molecularly cloned into the pT7PVM vector, exchanging the wild type P1 region. This work was conducted by Molly Arabov of Dr. Eckard Wimmer's laboratory. In the future, it will be seen if these constructs produce viable viruses and if these viruses have any level of attenuation.

Both examples here present promising future directions of the initial project of my dissertation work, which investigated the construction and characterization of codon-pair customized polioviruses. This work is an illustration of the utility and promise of synthetic biology and how, in the future, it can aid in the beneficial solution to pressing biological concerns.

Chapter 6 Figure

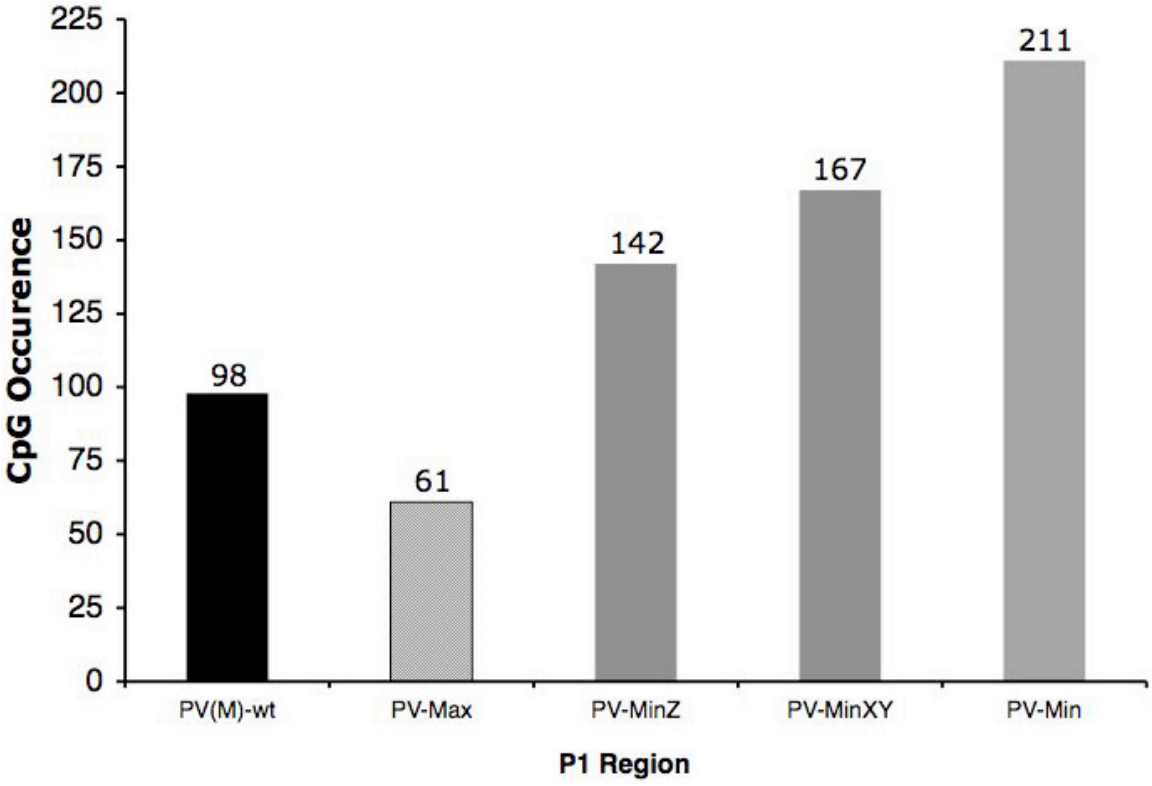


Figure 6.1: Occurrence of the CpG dinucleotide in PV(M)-wt and CPB-altered synthetic viruses. The appearance of the CpG dinucleotide in the P1 region of the wild type virus compared to the synthetic viruses. The greater the number of CpGs, the more substantial the level of attenuation.

References

1. Kaznessis Y (2007) Models for synthetic biology. BMC Syst Biol 1: 47.
2. Endy D (2005) Foundations for engineering biology. Nature 438: 449-453.
3. Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, et al. (2008) Virus attenuation by genome-scale changes in codon pair bias. Science 320: 1784-1787.
4. Mueller S (2006) Reduction of the Rate of Poliovirus Protein Synthesis through Large-Scale Codon Deoptimization Causes Attenuation of Viral Virulence by Lowering Specific Infectivity. Journal of Virology 80: 9687-9696.
5. Bügl H, Danner JP, Molinari RJ, Mulligan JT, Park HO, et al. (2007) DNA synthesis and biological security. Nat Biotechnol 25: 627-629.
6. Flanagan EB, Zamparo JM, Ball LA, Rodriguez LL, Wertz GW (2001) Rearrangement of the genes of vesicular stomatitis virus eliminates clinical disease in the natural host: new strategy for vaccine development. Journal of Virology 75: 6107-6114.
7. Forster A, Church G (2006) Towards synthesis of a minimal cell. Mol Syst Biol 2: 45.
8. Cello J (2002) Chemical Synthesis of Poliovirus cDNA: Generation of Infectious Virus in the Absence of Natural Template. Science 297: 1016-1018.
9. Smith HO, Hutchison CA, Pfannkoch C, Venter JC (2003) Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. Proc Natl Acad Sci USA 100: 15440-15445.
10. Couzin J (2002) Bioterrorism. A call for restraint on biological data. Science 297: 749-751.
11. Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, et al. (2008) Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. Science 319: 1215-1220.
12. Kitamura N, Semler BL, Rothberg PG, Larsen GR, Adler CJ, et al. (1981) Primary structure, gene organization and polypeptide expression of poliovirus RNA. Nature 291: 547-553.
13. van der Werf S, Bradley J, Wimmer E, Studier FW, Dunn JJ (1986) Synthesis of infectious poliovirus RNA by purified T7 RNA polymerase. Proc Natl Acad Sci USA 83: 2330-2334.
14. Koike S (1994) Cellular receptors for animal viruses; Wimmer E, editor. Cold Spring Harbor: Cold Spring Harbor Laboratory Press. 463-480 p.
15. De Jesus N, Franco D, Paul A, Wimmer E, Cello J (2005) Mutation of a single conserved nucleotide between the cloverleaf and internal ribosome entry site attenuates poliovirus neurovirulence. J Virol 79: 14235-14243.
16. Burns CC, Shaw J, Campagnoli R, Jorba J, Vincent A, et al. (2006) Modulation of poliovirus replicative fitness in HeLa cells by deoptimization of synonymous codon usage in the capsid region. Journal of Virology 80: 3259-3272.
17. (2005) Poliovirus infections in four unvaccinated children--Minnesota, August-October 2005. MMWR Morb Mortal Wkly Rep 54: 1053-1055.
18. Pfister T, C. Mirzayan, E. Wimmer (1999) In: Webster RG, A. Granoff, editor. The Encyclopedia of Virology. Second ed. London: Academic Press Ltd. pp. 1330-1348.

19. Melnick JL (1995) Enteroviruses: Polioviruses, Coxsackieviruses, Ecoviruses, and Newer Enteroviruses; Fields BN, Knipe DM, Chanock RM, Hirsch MS, Melnick JL et al., editors. New York: Raven Press. 549-605 p.
20. Oberste MS, Maher K, Flemister MR, Marchetti G, Kilpatrick DR, et al. (2000) Comparison of classic and molecular approaches for the identification of untypeable enteroviruses. J Clin Microbiol 38: 1170-1174.
21. Paul J (1971) A History of Poliomyelitis: Yale studies in the history of science and medicine. New Haven, Conn.: Yale University Press. 496 p.
22. Bodian D, Hortsmann D (1965) Polioviruses. In: Horsfall F, Tamm I, editors. Viral and rickettsial infections of Man. 4th ed. Philadelphia: J. B. Lippincott Co. pp. 430–473.
23. Melnick JL (1996) Current status of poliovirus infections. Clin Microbiol Rev 9: 293-300.
24. Landsteiner K, Popper E (1909) Ubertragung der Poliomyelitis acuta auf Affen. Z ImmunnitatsForsch Orig 2: 377-390.
25. Mueller S, Wimmer E, Cello J (2005) Poliovirus and poliomyelitis: A tale of guts, brains, and an accidental event. Virus Research 111: 175-193.
26. Rossmann MG, Johnson JE (1989) Icosahedral RNA virus structure. Annu Rev Biochem 58: 533-573.
27. Lee YF, A. N, Detjen BM, Wimmer E (1977) A protein covalently linked to poliovirus genome RNA. Proc Natl Acad Sci U S A 74: 59-63.
28. De Jesus N (2007) Epidemics to eradication: the modern history of poliomyelitis. Virol J 4: 70.
29. Jang SK, Krausslich HG, Nicklin MJ, Duke GM, Palmenberg AC, et al. (1988) A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J Virol 62: 2636-2643.
30. Pelletier J, Kaplan G, Racaniello VR, Sonenberg N (1988) Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5' noncoding region. Mol Cell Biol 8: 1103-1112.
31. Pelletier J, Kaplan G, Racaniello VR, Sonenberg N (1988) Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5'-noncoding region. Mol Cell Biol 8: 1103-1112.
32. Kräusslich HG, Nicklin MJH, Toyoda H, Etchison D, Wimmer E (1987) Poliovirus proteinase 2A induces cleavage of eukaryotic initiation factor 4F polypeptide p220. J Virol 61: 2711-2718.
33. Sonenberg N (1987) Regulation of translation by poliovirus. Adv Vir Res 33: 175-204.
34. Kitamura N, Adler C, Wimmer E (1980) Structure and expression of the picornavirus genome. Ann N Y Acad Sci 354: 183-201.
35. Krausslich HG, Wimmer E (1988) Viral proteinases. Annu Rev Biochem 57: 701-754.
36. Palmenberg AC (1987) Picornaviral processing: some new ideas. J Cell Biochem 33: 191-198.
37. Bienz K, Egger D, Troxler M, Pasamontes L (1990) Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. J Virol 64: 1156-1163.

38. Cho MW, Teterina N, Egger D, Bienz K, Ehrenfeld E (1994) Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells. Virology 202: 129-145.
39. Kok CC, McMinn PC (2008) Picornavirus RNA-dependent RNA polymerase. Int J Biochem Cell Biol 7: 7.
40. Drake JW (1999) The distribution of rates of spontaneous mutation over viruses, prokaryotes, and eukaryotes. Ann N Y Acad Sci 870: 100-107.
41. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439: 344-348.
42. Rossman M (2002) Picornavirus strucutre overview. In: BL. S, Wimmer E, editors. Molecular biology of picornaviruses. Washington, D.C.: ASM Press.
43. Ansardi DC, Porter DC, Anderson MJ, Morrow CD (1996) Poliovirus assembly and encapsidation of genomic RNA. Adv Virus Res 46: 1-68.
44. Lee WM, Monroe SS, Rueckert RR (1993) Role of maturation cleavage in infectivity of picornaviruses: Activation of an infectosome. J Virol 67: 2110-2122.
45. Kräusslich HG, Nicklin MJH, Lee CK, Wimmer E (1988) Polyprotein processing in picornavirus replication. Biochimie 70: 119-130.
46. Mueller S, Wimmer E (2003) Recruitment of nectin-3 to cell-cell junctions through trans-heterophilic interaction with CD155, a vitronectin and poliovirus receptor that localizes to alpha(v)beta3 integrin-containing membrane microdomains. J Biol Chem 278: 31251-31260.
47. Sato T, Irie K, Ooshio T, Ikeda W, Takai Y (2004) Involvement of heterophilic trans-interaction of Necl-5/Tage4/PVR/CD155 with nectin-3 in formation of nectin- and cadherin-based adherens junctions. Genes Cells 9: 791-799.
48. Bernhardt G, Harber J, Zibert A, deCrombrugghe M, Wimmer E (1994) The poliovirus receptor: identification of domains and amino acid residues critical for virus binding. Virology 203: 344-356.
49. Harber J, Bernhardt G, Lu HH, Sgro JY, Wimmer E (1995) Canyon rim residues, including antigenic determinants, modulate serotype-specific binding of polioviruses to mutants of the poliovirus receptor. Virology 214: 559-570.
50. Li Q, Yafal AG, Lee YM, Hogle J, Chow M (1994) Poliovirus neutralization by antibodies to internal epitopes of VP4 and VP1 results from reversible exposure of these sequences at physiological temperature. J Virol 68: 3965-3970.
51. Fricks CE, Hogle JM (1990) Cell-induced conformational change in poliovirus: externalization of the amino terminus of VP1 is responsible for liposome binding. J Virol 64: 1934-1945.
52. DeJesus N (2007).
53. Sabin AB (1956) Pathogenesis of Poliomyelitis. Science 123: 1151-1157.
54. Iwasaki A, Welker R, Mueller S, Linehan M, Nomoto A, et al. (2002) Immunofluorescence analysis of poliovirus receptor expression in Peyer's patches of humans, primates, and CD155 transgenic mice: implications for poliovirus infection. J Infect Dis 186: 585-592.
55. Freistadt MS, Kaplan G, Racaniello VR (1990) Heterogeneous expression of poliovirus receptor-related proteins in human cells and tissues. Mol Cell Biol 10: 5700-5706.

56. Mendelsohn CL, Wimmer E, Racaniello VR (1989) Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. Cell 56: 855-865.
57. Koike S, Horie H, Ise I, Okitsu A, Yoshida M, et al. (1990) The poliovirus receptor protein is produced both as membrane-bound and secreted forms. EMBO J 9: 3217-3224.
58. Yang WX, Terasaki T, Shiroki K, Ohka S, Aoki J, et al. (1997) Efficient delivery of circulating poliovirus to the central nervous system independently of poliovirus receptor. Virology 229: 421-428.
59. Ohka S, Yang WX, Terada E, Iwasaki K, Nomoto A (1998) Retrograde transport of intact poliovirus through the axon via the fast transport system. Virology 250: 67-75.
60. Morrison LA, Sidman RL, Fields BN (1991) Direct spread of reovirus from the intestinal lumen to the central nervous system through vagal autonomic nerve fibers. Proc Natl Acad Sci U S A 88: 3852-3856.
61. Stewart AJ, Devlin PM (2006) The history of the smallpox vaccine. J Infect 52: 329-334.
62. Goldsby R, Kindt T, Osborne B (2000) Kuby Immunology, 4th Edition; Kuby J, editor: W.H. Freeman & Co.
63. Billich A (2000) Technology evaluation: FluMist, University of Michigan. Curr Opin Mol Ther 2: 340-344.
64. Dowdle WR, De Gourville E, Kew OM, Pallansch MA, Wood DJ (2003) Polio eradication: the OPV paradox. Rev Med Virol 13: 277-291.
65. Belshe RB, Edwards KM, Vesikari T, Black SV, Walker RE, et al. (2007) Live attenuated versus inactivated influenza vaccine in infants and young children. N Engl J Med 356: 685-696.
66. Stauffer F, El-Bacha T, Da Poian AT (2006) Advances in the development of inactivated virus vaccines. Recent Patents Anti-Infect Drug Disc 1: 291-296.
67. CDC (2008) National Immunization Program. http://www.cdc.gov/vaccines/default.htm.
68. Jackson LA, Neuzil KM, Baggs J, Davis RL, Black S, et al. (2006) Compliance with the recommendations for 2 doses of trivalent inactivated influenza vaccine in children less than 9 years of age receiving influenza vaccine for the first time: a Vaccine Safety Datalink study. PEDIATRICS 118: 2032-2037.
69. McAuliffe VJ, Purcell RH, Gerin JL (1980) Type B hepatitis: a review of current prospects for a safe and effective vaccine. Rev Infect Dis 2: 470-492.
70. Chan CY, Lee SD, Tsai YT, Lo KJ (1992) Long-term follow-up of hepatitis B vaccination in susceptible hospital personnel. J Gastroenterol Hepatol 7: 266-269.
71. (2000) Are booster immunisations needed for lifelong hepatitis B immunity? European Consensus Group on Hepatitis B Immunity. Lancet 355: 561-565.
72. Beale AJ (1991) Efficacy and safety of oral poliovirus vaccine and inactivated poliovirus vaccine. Pediatr Infect Dis J 10: 970-972.
73. Salk JE, Krech U, Younger JS, Bennett BL, Lewis LJ, et al. (1954) Formaldehyde treatment and safety testing of experimental poliomyelitis vaccines. Am J Pub Health 44: 563.

74. Nathanson N (2008) The Pathogenesis of Poliomyelitis: What We Don't Know. Advances in Virus Research 71.
75. Henry JL, Jaikaran ES, Davies JR, Tomlinson AJ, Mason PJ, et al. (1966) A study of poliovaccination in infancy: excretion following challenge with live virus by children given killed or living poliovaccine. J Hyg (Lond) 64: 105-120.
76. Fescharek R, Budde RK, Arras C (1997) OPV vs IPV--could placental immunity reduce the number of vaccine-associated paralytic poliomyelitis? Vaccine 15: 1707-1708.
77. Chumakov K, Ehrenfeld E, Wimmer E, Agol VI (2007) Vaccination against polio should not be stopped. Nat Rev Microbiol 5: 952-958.
78. Lui X, Levin A, Makinen M, Day J (2003) OPV vs IPV: Past and Future Choice of Vaccine in the Global Polio Eradication Program
. Bethesda, Maryland: Partners for Health Reformplus. 1-46 p.
79. Heininger U, Sanger R, Jacquet JM, Schuerman L (2007) Booster immunization with a hexavalent diphtheria, tetanus, acellular pertussis, hepatitis B, inactivated poliovirus vaccine and Haemophilus influenzae type b conjugate combination vaccine in the second year of life: safety, immunogenicity and persistence of antibody responses. Vaccine 25: 1055-1063.
80. Chezzi C, Dommann CJ, Blackburn NK, Maselesele E, McAnerney J, et al. (1998) Genetic stability of oral polio vaccine prepared on primary monkey kidney cells or Vero cells--effects of passage in cell culture and the human gastrointestinal tract. Vaccine 16: 2031-2038.
81. Sabin A, Ward R (1941) Nature of non-paralytic and transitory paralytic poliomyelitis in rhesus monkeys inocculated with human virus. J Exp Med 73: 757-770
82. Sabin A (1957) Spec Publ NY Acad Sci 5: 113–127.
83. Kew O, Morris-Glasgow V, Landaverde M, Burns C, Shaw J, et al. (2002) Outbreak of poliomyelitis in Hispaniola associated with circulating type 1 vaccine-derived poliovirus. Science 296: 356-359.
84. Kew OM, Wright PF, Agol VI, Delpeyroux F, Shimizu H, et al. (2004) Circulating vaccine-derived polioviruses: current state of knowledge. Bull World Health Organ 82: 16-23.
85. Jiang P, Faase JA, Toyoda H, Paul A, Wimmer E, et al. (2007) Evidence for emergence of diverse polioviruses from C-cluster coxsackie A viruses and implications for global poliovirus eradication. Proc Natl Acad Sci USA 104: 9457-9462.
86. Shimizu H, Thorley B, Paladin FJ, Brussen KA, Stambos V, et al. (2004) Circulation of type 1 vaccine-derived poliovirus in the Philippines in 2001. J Virol 78: 13512-13521.
87. Halsey NA, Pinto J, Espinosa-Rosales F, Faure-Fontenla MA, da Silva E, et al. (2004) Search for poliovirus carriers among people with primary immune deficiency diseases in the United States, Mexico, Brazil, and the United Kingdom. Bull World Health Organ 82: 3-8.
88. Kew O, Sutter R, De Gourville E, Dowdle W, Pallansch M (2005) Vaccine-derived polioviruses and the endgame strategy for global polio eradication. Annu Rev Microbiol 59: 587-635.

89. Stanway G, Hughes PJ, Mountford RC, Reeve P, Minor PD, et al. (1984) Comparison of the complete nucleotide sequences of the genomes of the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin vaccine derivative P3/Leon 12a1b. Proc Natl Acad Sci U S A 81: 1539-1543.
90. Guillot S, Otelea D, Delpeyroux F, Crainic R (1994) Point mutations involved in the attenuation/neurovirulence alternation in type 1 and 2 oral polio vaccine strains detected by site-specific polymerase chain reaction. Vaccine 12: 503-507.
91. Fine P (1999) Transmissibility and persistence of oral polio vaccine viruses: implications for the global poliomyelitis eradication initiative. Am J Epidemiol 150: 1001-1021.
92. Sutter R, Kew O, Cochi S (2004) Poliovirus vaccine - live. In: Plotkin S, Orenstein W, editors. Vaccines. Philadelphia: Sanders. pp. pp. 651-705.
93. Heymann DL, Sutter RW, Aylward RB (2005) A global call for new polio vaccines. Nature 434: 699-700.
94. Plotkin JB, Robins H, Levine AJ (2004) Tissue-specific codon usage and the expression of human genes. Proc Natl Acad Sci U S A 101: 12588-12591.
95. Kaplan G, Racaniello VR (1988) Construction and characterization of poliovirus subgenomic replicons. J Virol 62: 1687-1696.
96. Johansen LK, Morrow CD (2000) The RNA encompassing the internal ribosome entry site in the poliovirus 5' nontranslated region enhances the encapsidation of genomic RNA. Virology 273: 391-399.
97. Schwerdt CE, Fogh J (1957) The ratio of physical particles per infectious unit observed for poliomyelitis viruses. Virology 4: 41-52.
98. Joklik WK, Darnell JE, Jr. (1961) The adsorption and early fate of purified poliovirus in HeLa cells. Virology 13: 439-447.
99. Zhao WD, Wimmer E (2001) Genetic analysis of a poliovirus/hepatitis C virus chimera: new structure for domain II of the internal ribosomal entry site of hepatitis C virus. Journal of Virology 75: 3719-3730.
100. Gutman GA, Hatfield GW (1989) Nonrandom utilization of codon pairs in Escherichia coli. Proc Natl Acad Sci USA 86: 3699-3703.
101. Moura G, Pinheiro M, Arrais J, Gomes A, Carreto L, et al. (2007) Large Scale Comparative Codon-Pair Context Analysis Unveils General Rules that Fine-Tune Evolution of mRNA Primary Structure. PLoS ONE 2: e847.
102. Irwin B, Heck JD, Hatfield GW (1995) Codon pair utilization biases influence translational elongation step times. J Biol Chem 270: 22801-22806.
103. Buchan JR, Aucott LS, Stansfield I (2006) tRNA properties help shape codon pair preferences in open reading frames. Nucleic Acids Res 34: 1015-1027.
104. Fedorov A, Saxonov S, Gilbert W (2002) Regularities of context-dependent codon bias in eukaryotic genes. Nucleic Acids Research 30: 1192-1197.
105. Buchan J (2006) tRNA properties help shape codon pair preferences in open reading frames. Nucleic Acids Research 34: 1015-1027.
106. Buchan J, Stansfield I (2005) Codon pair bias in prokaryotic and eukaryotic genomes. BioSysBio: Bioinformatics and Systems Biology Conference Edinburgh, UK, 14–15 July 2005: 1.
107. Hacker H, Vabulas RM, Takeuchi O, Hoshino K, Akira S, et al. (2000) Immune cell activation by bacterial CpG-DNA through myeloid differentiation marker 88 and

tumor necrosis factor receptor-associated factor (TRAF)6. J Exp Med 192: 595-600.
108. Bauer S, Kirschning CJ, Hacker H, Redecke V, Hausmann S, et al. (2001) Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition. Proc Natl Acad Sci U S A 98: 9237-9242.
109. Sugiyama T, Gursel M, Takeshita F, Coban C, Conover J, et al. (2005) CpG RNA: identification of novel single-stranded RNA that stimulates human CD14+CD11c+ monocytes. J Immunol 174: 2273-2279.
110. Karlin S, Doerfler W, Cardon LR (1994) Why is CpG suppressed in the genomes of virtually all small eukaryotic viruses but not in those of large eukaryotic viruses? Journal of Virology 68: 2889-2897.
111. Gutman GA, Hatfield GW (1989) Nonrandom utilization of codon pairs in Escherichia coli. Proc Natl Acad Sci U S A 86: 3699-3703.
112. Park S, Yang X, Saven JG (2004) Advances in computational protein design. Curr Opin Struct Biol 14: 487-494.
113. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406-3415.
114. Molla A, Paul AV, Wimmer E (1991) Cell-free, de novo synthesis of poliovirus. Science 254: 1647-1651.
115. Ruckert R (1990) Picornaviridae and their replication; Fields BN KD, Chanock RM, Hirsch MS, Melnick JL et al., editor. New York: Raven Press. 507-548 p.
116. Ohlbaum A, Figueroa F, Grado C, Contreras G (1970) Target molecular weight of foot-and-mouth disease virus and poliovirus. J Gen Virol 6: 429-432.
117. Pfister T, Wimmer E (1999) Characterization of the nucleoside triphosphatase activity of poliovirus protein 2C reveals a mechanism by which guanidine inhibits poliovirus replication. J Biol Chem 274: 6992-7001.
118. Reed LJ, Muench M (1938) A simple method for estimating fifty percent endpoints. Am J Hyg 27: 493-497.
119. Toyoda H, Yin J, Mueller S, Wimmer E, Cello J (2007) Oncolytic treatment and cure of neuroblastoma by a novel attenuated poliovirus in a novel poliovirus-susceptible animal model. Cancer Research 67: 2857-2864.
120. Wahby AF (2000) Combined cell culture enzyme-linked immunosorbent assay for quantification of poliovirus neutralization- relevant antibodies. Clin Diagn Lab Immunol 7: 915-919.
121. Racaniello VR, Baltimore D (1981) Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. Proc Natl Acad Sci U S A 78: 4887-4891.
122. Racaniello VR, Baltimore D (1981) Cloned poliovirus complementary DNA is infectious in mammalian cells. Science 214: 916-919.
123. Doma M, Parker R (2006) Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature 440: 561-564.
124. www.blueheronbio.com (2006).
125. Alexander H, Koch G, Mountain I, Van DO (1958) Infectivity of ribunucleic acid from poliovirus in human cell monolayers. J Exp Med 108: 493-506.
126. Colter JS, Bird HH, Moyer AW, Brown RA (1957) Infectivity of ribonucleic acid isolated from virus infected tissues. J Virol 4: 522.

127. Rueckert RR (1985) Fields Virology: Picornaviruses and their replication; B. N. Fields DMK, R. M. Chanock, J. L. Melnick, B. Roizman, and R. E. Shope, editor. New York, N.Y.: Raven Press. p. 705–738 p.
128. Molla A, Jang SK, Paul AV, Reuer Q, Wimmer E (1992) Cardioviral internal ribosomal entry site is functional in a genetically engineered dicistronic poliovirus. Nature 356: 255-257.
129. Svitkin YV, Maslova SV, Agol VI (1985) The genomes of attenuated and virulent poliovirus strains differ in their in vitro translation efficiencies. Virology 147: 243-252.
130. Svitkin YV, Alpatova GA, Lipskaya GA, Maslova SV, Agol VI, et al. (1993) Towards development of an in vitro translation test for poliovirus neurovirulence. Dev Biol Stand 78: 27-32.
131. Hogle J (2002) POLIOVIRUS CELL ENTRY: Common Structural Themes in Viral Cell Entry Pathways. Annu Rev Microbiol 56: 677-702.
132. Huang Y, Hogle JM, Chow M (2000) Is the 135S poliovirus particle an intermediate during cell entry? J Virol 74: 8757-8761.
133. Ruckert RR (1985) Picornaviruses and their replication; Fields BN, editor. New York: Raven Press. 705-738 p.
134. Georgescu MM, Tardy-Panit M, Guillot S, Crainic R, Delpeyroux F (1995) Mapping of mutations contributing to the temperature sensitivity of the Sabin 1 vaccine strain of poliovirus. J Virol 69: 5278-5286.
135. Paul AV, Mugavero J, Yin J, Hobson S, Schultz S, et al. (2000) Studies on the attenuation phenotype of polio vaccines: poliovirus RNA polymerase derived from Sabin type 1 sequence is temperature sensitive in the uridylylation of VPg. Virology 272: 72-84.
136. McGoldrick A, Macadam AJ, Dunn G, Rowe A, Burlison J, et al. (1995) Role of mutations G-480 and C-6203 in the attenuation phenotype of Sabin type 1 poliovirus. J Virol 69: 7601-7605.
137. Christodoulou C, Colbere-Garapin F, Macadam A, Taffs LF, Marsden S, et al. (1990) Mapping of mutations associated with neurovirulence in monkeys infected with Sabin 1 poliovirus revertants selected at high temperature. J Virol 64: 4922-4929.
138. Shiomi H, Urasawa T, Urasawa S, Kobayashi N, Abe S, et al. (2004) Isolation and characterisation of poliovirus mutants resistant to heating at 50 degrees Celsius for 30 min. J Med Virol 74: 484-491.
139. Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, et al. (1982) Rapid evolution of RNA genomes. Science 215: 1577-1585.
140. Drake JW, Holland JJ (1999) Mutation rates among RNA viruses. Proc Natl Acad Sci U S A 96: 13910-13913.
141. Vignuzzi M, Wendt E, Andino R (2008) Engineering attenuated virus vaccines by controlling replication fidelity. Nat Med 14: 154-161.
142. Domingo E, Ruiz-Jarabo CM, Sierra S, Arias A, Pariente N, et al. (2002) Emergence and selection of RNA virus variants: memory and extinction. Virus Res 82: 39-44.
143. Murray KE, Nibert ML (2007) Guanidine Hydrochloride Inhibits Mammalian Orthoreovirus Growth by Reversibly Blocking the Synthesis of Double-Stranded RNA. J Virol.

144. Buchan J, Stansfield I (2007) Halting a cellular production line: responses to ribosomal pausing during translation. Biol Cell 99: 475.
145. Crotty S, Hix L, Sigal LJ, Andino R (2002) Poliovirus pathogenesis in a new poliovirus receptor transgenic mouse model: age-dependent paralysis and a mucosal route of infection. J Gen Virol 83: 1707-1720.
146. Khan SW, E. (2008) Unpublished Data, Tag4 mice susceptible to poliovirus.
147. (1997) Paralytic poliomyelitis--United States, 1980-1994. MMWR Morb Mortal Wkly Rep 46: 79-83.
148. Nathanson N, Langmuir AD (1995) The Cutter incident. Poliomyelitis following formaldehyde-inactivated poliovirus vaccination in the United States during the Spring of 1955. II. Relationship of poliomyelitis to Cutter vaccine. 1963. Am J Epidemiol 142: 109-140; discussion 107-108.
149. Pfeiffer JK, Kirkegaard K (2006) Bottleneck-mediated quasispecies restriction during spread of an RNA virus from inoculation site to brain. Proc Natl Acad Sci U S A 103: 5520-5525.
150. Ida-Hosonuma M, Iwasaki T, Yoshikawa T, Nagata N, Sato Y, et al. (2005) The alpha/beta interferon response controls tissue tropism and pathogenicity of poliovirus. Journal of Virology 79: 4460-4469.
151. Joly E, Mucke L, Oldstone MB (1991) Viral persistence in neurons explained by lack of major histocompatibility class I expression. Science 253: 1283-1285.
152. Savage VM, Allen AP, Brown JH, Gillooly JF, Herman AB, et al. (2007) Scaling of number, size, and metabolic rate of cells with body size in mammals. Proc Natl Acad Sci U S A 104: 4718-4723.
153. McIlwain DL, Hoke VB (1994) Radiolabeling motoneuron proteins in the isolated frog spinal cord preparation. J Neurosci Methods 52: 197-202.
154. Chan L, Kosuri S, Endy D (2005) Refactoring bacteriophage T7. Mol Syst Biol 1: 2005.0018.
155. Ward CD, Flanegan JB (1992) Determination of the poliovirus RNA polymerase error frequency at eight sites in the viral genome. Journal of Virology 66: 3784-3793.
156. Hodes R, Peacock SM, Jr., Bodian D (1949) Selective destruction of large motoneurons by poliomyelitis virus; size of motoneurons in the spinal cord of rhesus monkeys. J Neuropathol Exp Neurol 8: 400-410.
157. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, et al. (2003) Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA 289: 179-186.
158. Cram P, Blitz SG, Monto A, Fendrick AM (2001) Influenza. Cost of illness and considerations in the economic evaluation of new and emerging therapies. Pharmacoeconomics 19: 223-230.
159. Influenza Report 2006; BS Kamps CH, W Preiser, editor. Cologne, Germany: Flying Publisher. 225 p.
160. Russell CJ, Webster RG (2005) The genesis of a pandemic influenza virus. Cell 123: 368-371.
161. Kilbourne ED (2006) Influenza pandemics of the 20th century. Emerg Infect Dis 12: 9-14.

162. Stephenson I, Democratis J (2005) Influenza: current threat from avian influenza. British Medical Bulletin 75-76: 63-80.
163. Coleman JR (2007) The PB1-F2 protein of Influenza A virus: increasing pathogenicity by disrupting alveolar macrophages. Virol J 4: 9.
164. Barton GM, Kagan JC, Medzhitov R (2006) Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. Nat Immunol 7: 49-56.