Stony Brook University

Recoding of Vesicular Stomatitis Virus Genes by Computer-aided design provides a live attenuated vaccine candidate

A dissertation presented

by

Bingyin Wang

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Molecular and Cell Biology

Stony Brook University

May 2015

Stony Brook University

The Graduate School

Bingyin Wang

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

Patrick Hearing, Ph.D-Chairperson of Defense, Professor

Department of Molecular Genetics and Microbiology

Laurie Krug, Ph.D-Assistant Professor

Department of Molecular Genetics and Microbiology

Carol A. Carter, Ph.D-Professor

Department of Molecular Genetics and Microbiology

This dissertation is accepted by the Graduate School

Charles Taber

Dean of the Graduate School

Abstract of the Dissertation

Recoding Vesicular Stomatitis Virus Genes by Computer-aided design to develop a live attenuated vaccine candidate

by

Bingyin Wang

Doctor of Philosophy

in

Molecular and Cell Biology

Stony Brook University

2015

Codon pair bias (CPB), which has been observed in all organisms, is a neglected genomic phenomenon that substantially affects gene expression. CPB results from synonymous codons that are paired more or less frequently in ORFeomes regardless of codon bias. The effect of an individual codon pair change is usually small, but when amplified by large-scale genome recoding (simultaneously generating hundreds of synonymous mutations without altering codon usage or amino acid sequence), strikingly altered biological phenotypes are observed. Although the molecular mechanism underlying codon pair bias still remains enigmatic, the utility of codon pair bias in the development of live attenuated vaccines was recently demonstrated by recodings of poliovirus (a positive-strand RNA virus) and influenza virus (a negative-strand segmented RNA virus).

In an attempt to expand the attenuation spectrum of codon pair bias, I recoded the L and G genes of vesicular stomatitis virus (VSV), a non-segmented negative strand RNA virus. For L gene recoding, I introduced a total of 858 and 623 silent mutations into a 5'-terminal segment of the viral L gene (designated as L1), to create sequences containing either overrepresented (sdmax design) and underrepresented (min design) codon pairs, respectively. These two sequences were designated as $L1^{sdmax}$ and $L1^{min}$, respectively. Recombinant VSV containing the $L1^{min}$ sequence could not be recovered via a well-established reverse genetics system whereas the introduction of the $L1^{sdmax}$ sequence into the viral genome conferred a modest level of attenuation in cell culture. The phenotype of the $L1^{sdmax}$ virus was unexpected, which might result from a discoordination

between translation elongation rate and protein folding kinetics. More strikingly, in mice the $L1^{sdmax}$ virus was almost as immunogenic as the parental strain but highly attenuated. Despite this, the recombinant virus showed a temporarily productive replication in the infected tissues while neither histopathological change nor severe morbidity has ever been found in $L1^{sdmax}$-infected animals.

As for the G gene recoding, I introduced 323 silent mutations into the G gene open-reading frame. The newly generated recombinant virus was designated as G^{min}, which contained hundreds of underrepresented codon pairs. Although during the early stages of virus infection significantly less G gene products were produced in infected cells, the G^{min} virus replicated just as well as its parental strain in tissue culture and showed neurovirulence in a mouse model.

Taken together, my research suggests that $L1^{sdmax}$, but not $L1^{min}$ or G^{min}, is a promising live attenuated vaccine candidate for VSV infection. In addition, these results demonstrate a novel approach to attain a balance between VSV virulence and immunogenicity, which could serve as a paradigm for the attenuation of other negative-strand, non-segmented RNA viruses (NNS viruses).

Table of contents

List of Figures

LIST OF TABLES

Chapter 1: Introduction

1.1 Vesicular Stomatitis Virus

Vesicular stomatitis virus (abbreviated as VSV) is a prototype of the non-segmented, negative strand RNA viruses (NNS viruses) in the family of *Rhabdoviridae* (1, 2). It is an arthropod-borne pathogen that primarily causes acute vesicular disease in livestock (3, 4). Cases of human infection occasionally occur, especially in individuals who are directly exposed to infected animals or laboratory environments. Individuals infected with VSV usually develop mild flu-like symptoms but a single case of encephalitis in a 3-year-old child has been related to VSV infection (5, 6). Nevertheless, in general VSV is considered to be innocuous in human beings, and serves as an attractive model to study the other 75 or more mammalian pathogens in the *Rhabdoviridae* family (1).

VSV naturally infects a wide range of farm animals, such as cattle, horse and swine. It causes epidemics every ten to fifteen years in parts of the Southwestern United States and Central America (2). The infection of livestock is characterized by vesicular lesions and ulceration on mucosal surfaces, usually around the mouth, hoof and teat, all of which resembles another notorious agricultural pathogen foot-and-mouth disease virus (FMDV). Natural VSV infection in the United States is associated with two serotypes, Indiana (VSV_I) and New Jersey (VSV_{NJ}) (4).

However, in experimental animals, VSV (either the naturally isolated strain or the parental strain used in the laboratory setting) appears to be neurotropic and can cause fatal encephalitis instead of vesicular disease, if inoculated intranasally (7-10). In mice VSV first replicates in the nasal epithelial cells and then spreads to olfactory neurons (11). By disseminating via retrograde transport along the axon, the virus gains entry into the central nervous system (CNS) and causes brain disease. In mouse models the infection via this route usually leads to severe morbidity and mortality in 6-10 days, depending on the initial infection dose (7-10, 12).

In primates VSV also causes histopathological lesions in brains when the virus is given via an intrathalamic route (13). In rats, intravascular injection of high-dose VSV Indiana serotype appears to be neurotoxic as well (14, 15). It is also noteworthy that a significant fraction

of animals tested in that study, died of acute liver failure within several days, most likely due to the cytopathic nature of the virus towards hepatocytes and the following focal inflammation in the liver (15). For newborn and young mice that are no more than 6 weeks old, high doses of virus injected intraperitoneally can be lethal (16). These studies conducted in diverse rodent systems not only indicate that VSV is a neurotropic virus but also suggest that immunocompetent mouse models are applicable for the safety assessment of VSV vaccine candidates.

VSV has been extensively studied for decades, as a laboratory tool to probe aspects of molecular virology and cell biology, but its great potential in clinical application has not been appreciated until recently. VSV could serve as an ideal vaccine delivery platform since it exhibits the following advantages (2). First, this virus is not prevalent in the general population thus there is no existing immunity against a VSV-based vaccination strategy. Second, VSV can accommodate a large foreign gene fragment and its genome is relatively easy to manipulate. Third, VSV replicates incredibly fast and lyses the infected cells within a short time. Its strong cytopathic effect serves as a natural adjuvant for vaccination and a single dose of inoculation can induce a potent adaptive immune response against the foreign antigen it delivers. Fourth, the virus replicates exclusively in the cytoplasm. It never gains entry to the nucleus; therefore there is no risk of genome integration. The last but not the least, VSV naturally infects at the mucosal surfaces. Therefore, the VSV-based vaccine vector can be given via a mucosal route. This property appears to be especially attractive for the *in vivo* delivery of antigens of certain paramyxoviruses, exemplified by human respiratory syncytial virus (hRSV) and human metapneumovirus (hMPV). These viruses, the natural infection of which is via the mucosal route, account for the majority of respiratory diseases in infants and elderly (17, 18). The development of VSV-based vaccine delivery will not only eliminate the need for needles but also afford a better mucosal immunity than the vaccines administered parenterally.

In 1998, it was reported that vaccination with a recombinant VSV expressing influenza HA protein could protect mice against a lethal flu virus challenge, although the strain they used appeared to be neuropathogenic (5, 19). Two years later, the same group showed that primary vaccination with VSV vectors expressing simian immunodeficiency virus (SIV) *env* and *gag* genes could protect rhesus monkey from a pathogenic SIV infection (20). After their seminal work, greater efforts have been invested to improve the safety profile as well as efficacy of the initial VSV vectors, leading to the development of a large number of VSV-based vaccine

candidates against other mammalian viral pathogens, such as measles, respiratory syncytial virus, papillomavirus, hepatitis B virus, SARS (severe acute respiratory syndrome) virus and some extremely contagious filoviruses, such as Ebola virus (EBOV) and Marburg virus (MARV) (21-33).

VSV also demonstrates a distinguished oncolytic property because of its broad tissue tropism and exquisite sensitivity to innate immunity. In 2004, two recombinant VSV strains, both of which had defects in antagonizing host innate immunity due to single amino acid substitutions in the viral M protein, served well as potent anti-cancer agents when systemically injected into immunocompromised mice (34). In cell culture VSV is extremely sensitive to type I IFN pre-treatment, which causes a dramatic reduction of virus yield (up to 10,000 folds). This intrinsic sensitivity to IFN eventually limits virus replication as well as dissemination in an immunocompetent host (2, 35). Many cancer cells, however, are not responsive to IFN signaling. They either contain inactive genes that are specifically associated with IFN response or have over-active proliferative signaling cascades that inhibit IFN response. It is reasoned that in the presence of a type I IFN "cloud" that is induced by VSV infection, normal cells are protected given that they are responsive and capable of entering an antiviral state. In contrast, cancer cells, which abrogate their IFN signaling in exchange of unlimited growth, become more susceptible to virus infection (36-38). Currently a Phase I clinical trial is being conducted at the Mayo Clinic to evaluate the safety profile of a recombinant IFN-β expressing VSV for the treatment of human liver cancer (39).

Taken together, the previous studies suggest that VSV has great potential for prophylactic and therapeutic applications in clinics. However, the wild type strain of VSV is an infectious agent with neurotoxicity in experimental animals, which poses serious safety concerns for direct human use. Therefore, an attenuated VSV strain that replicates robustly in tissue culture but lacks virulence *in vivo*, needs to be developed.

1.2 VSV genome structure

As a prototypic NNS virus, VSV has a negative sense, non-segmented RNA genome of 11,162 nucleotides in length. It is composed of five genes and two untranslated regions, in the order of a Leader sequence, N, P, M, G, L and a Trailer sequence from the 3' to the 5' end **(**Fig1.1**)**. Its simple genome composition and replication strategy allow for a rapid and productive infectious cycle once the infection is established (2, 40).

The viral genome is flanked by two regulatory sequences, the Leader sequence at the 3'end, which is involved in the initiation of positive strand RNA synthesis, and the 5' Trailer sequence that is involved in starting negative strand RNA synthesis from the other end. Highly conserved sequences are present at the beginning (3' AUACUUUUUUU 5') and the end (3' UUGUCNNUAG 5') of each viral gene, which are recognized by the viral encoded RNA polymerase. These cis-acting elements as well as the highly conserved dinucleotide junctions (3' C/GA 5') between viral genes collectively specify where to terminate transcription of the upstream gene and reinitiate transcription of the downstream one **(**Fig1.2**)** (40).

The viral genes are transcribed into a series of discrete monocistronic mRNAs and the abundances of these transcripts gradually decrease from the 3'end to the 5'end, with about 30% transcriptional attenuation at each intergenic junction. The "stop-and-start" model, which is most widely accepted, was developed to account for the sequentiality and polarity of VSV transcription. The polymerase has to stop at the end of the upstream gene and add the poly-A tail to the newly synthesized mRNA in a "stuttering" manner before it restarts at the beginning of the downstream gene (40). At the end of each viral gene up to 30% of RNA polymerase molecules fail to reinitiate, which leads to a gradient in the abundance of viral mRNAs from the 3'end to the 5'end **(**Fig1.1**)**.

1.3 VSV virion and the viral proteins

VSV forms a bullet-like virion (180nm in length and 75nm in diameter), which consists of two functionally relevant components: the ribonucleoprotein complex (RNP) and the outer membrane derived from the previously infected cells. The virus genome is located within the RNP complex and is nearly always tightly encapsidated by nucleocapsid (N) protein during the intracellular life cycle. This RNA-N complex forms a helical structure that is resistant to RNase digestion and serves as the template for viral RNA synthesis. The viral RNA dependent RNA polymerase (RdRP) that consists of the L protein (Large catalytic subunit) and P (Phosphoprotein, co-factor for L protein) recognizes the RNA-N complex and catalyzes RNA synthesis (genome replication and viral gene transcription). VSV is the first virus identified to package the viral encoded RdRP inside the virion. These four components, nucleocapsid protein (N), phosphoprotein (P), large catalytic subunit (L) and viral RNA collectively form the ribonucleoprotein complex (RNP), which is infectious once inside the cell (40).

VSV L protein

The VSV L protein provides all of the enzymatic activities involved in nucleotide polymerization, mRNA capping and cap modification (40-47). Except for the members of the *Orthomyxoviridae* (e.g., influenza viruses), whose RNA polymerase is a tripartite complex, all of the negative-strand RNA viruses encode a single large polymerase L (48, 49). Amino acid sequence comparison of L proteins from multiple NNS viruses revealed six conserved regions (CR I-VI), which led to the notion that the L protein may be organized into a series of domains that retain independent catalytic activities (50). However, the large size of the L protein and the presence of flexible "linker" regions between functional domains greatly impede the progress of high-resolution structural studies. The overall molecular architecture of the L protein has not been elucidated until very recently. Single particle electronic microscopy (EM) provides evidence that the L protein is organized into a core ring structure and an appendage containing three additional globular domains (51). It is hypothesized that the RNA polymerization activity of the L protein resides within the ring domain because a similar structure has been adopted by the RdRP molecules of other viruses, like rotavirus. One the other hand, the flexible appendage presumably contains the domains that are essential for the formation of the mRNA cap, and this structure undergoes a significant conformational rearrangement upon binding to the virally

encoded polymerase co-factor P. Although the P protein is much smaller (~29kDa in comparison to ~250kDa of L), its interaction with L is not through bivalent binding to either termini of the L protein. Instead, the interaction depends on the tertiary structure of an intact L protein, which supports the hypothesis that the proper stepwise assembly of the L protein is indispensable for its polymerase activity (51). Consistent with this idea, accumulating evidence demonstrates a causal link between viral mRNA capping and polyadenylation, suggesting that the structural arrangement (possibly also rearrangement) of multiple functional domains of the L protein is playing a regulatory role during viral RNA synthesis (52).

VSV N protein

The nucleocapsid proteins (N) are produced in the greatest abundance since the N gene is closest to the 3'end of the viral genome. As mentioned above, the N proteins tightly encapsidate the viral RNA genomes (negative strand) as well as the replication intermediates (positive strand). The viral RNA is only temporarily un-encapsidated and becomes available to the proceeding RNA polymerase during transcription or replication. Since the polymerase moves along the RNA template, the previously unveiled region of viral RNA will become re-encapsidated (53, 54). A soluble fraction of N-RNA oligomers can be isolated *in vitro*. The molecular weight of the purified complex suggests that it is composed of 10 N molecules and a 90-nt RNA. When co-expressed in *E.coli*, the N and VSV P proteins, another component of the RdRP, form multiple types of protein complexes with different molar ratios. It is proposed that during natural infection, the N: P ratio is approximately 2:1, considering that a single P protein can induce the correct curvature on a N/N dimer and facilitate the binding to viral RNA. The N protein may undergo aggregation *in vitro* if not provided with the appropriate amount of P protein (55). The P-binding region of the N protein is at the C-terminus, since the recent structural studies showed that the C-lobes of two adjacent N proteins formed a unique binding site for the P protein (55).

VSV P protein

Phosphoprotein (P) comprises three domains: domain I containing some highly acidic amino acids and 3 phosphorylation sites, domain III that is critical for N protein binding and contains a cluster of basic residues and domain II, a linker to bridge the other two domains. P

protein functions as a dimer, which delivers the L protein of the RdRP to the N-RNA template (56). The L protein alone does not efficiently bind to the N-RNA complex and therefore the presence of the P protein is critical for efficient template recognition. Recently, it has been suggested that the P protein plays an even more significant role in regulating polymerase processivity than a mere physical bridge between the N-RNA template and the polymerase (51). The dynamic interaction between P and L appears to induce a conformational rearrangement of the viral RNA polymerase, which should play a regulatory role in viral RNA synthesis.

VSV M protein

Matrix protein (M) is a major component of the VSV virion. They "glue" the encapsidated viral genome to the outer membrane. As the major virulence factor in VSV, the viral matrix protein potently shuts off the host protein translation, enabling the virus to evade the cellular antiviral response (2). During the early stage of infection, M protein interacts with Nup98, a major component of nuclear pore complex (NPC), and inhibits the nuclear export of host mRNAs (57). A similar inhibitory effect was also observed when the M protein was expressed alone (58). Recently, M protein was reported to interact with the chromatin fraction of host nuclei thus it may regulate host gene expression at the transcription level (59, 60).

Besides its role in antagonizing the host antiviral program, the M protein also plays a role in viral genome packaging and virion assembly. During the late stages of infection, the accumulation of M proteins close to the plasma membrane aids the inactivation of the RNP complexes and prepares them for packaging. Two motifs, PPPY and PSAP in the VSV M protein, have been identified to interact with host factors and facilitate virus budding (61, 62). Interestingly, these motifs appear to be conserved in other enveloped viruses, for instance, HIV and Ebola virus (63, 64).

VSV G protein

Glycoprotein (G) forms non-covalently associated trimers that are embedded on the virion surface and induces virus entry. When the VSV G protein is synthesized in polarized cells, such as epithelial cells, the newly synthesized G proteins are targeted towards the basolateral plasma membrane (65). The G protein confers a remarkably broad cell tropism to the virus, although the identity of the host receptor or receptors has been a decades-old mystery. The latest

studies suggest that low-density lipoprotein receptor (LDLR) may serve as a major cellular receptor for VSV, while other LDLR family members could also contribute to the virus uptake (66). Because of the wide range of cells that VSV can infect, VSV G has been extensively studied and widely used to "pseudotype" other viral vectors, such as lentivirus, to introduce exogenous genetic material into a variety of cells or animals (67).

1.4 Intracellular life cycle of VSV

Virus replication occurs exclusively in the cytoplasm. First, VSV enters the host cells via clathrin-mediated endocytosis. The virus engages the host cell through an interaction between its glycoprotein and an unidentified host membrane component. After attachment, the virus is entrapped by a clathrin-coated pit. However, as mentioned above, VSV forms a bullet-shape virion that measures 180x75nm, which exceeds the normal size of a clathrin-coated vesicle. Therefore, the formation of an endocytic carrier can not be mediated solely by clathrin (68). The elongated shape of the WT virion prevents the full closure of clathrin-coated pit and therefore the actin machinery nearby has to be triggered to finalize the internalization process. Interestingly, a defective interfering VSV particle (DI) that is shorter than the WT virus can be easily internalized by clathrin-coated vesicles, without the assembly of actin filaments (69).

The low pH environment within the endosomes results in the fusion between viral membranes and host vesicles (70). The infectious RNP complex, which is competent for transcription, is released and readily initiates the infection. This is accompanied by the dissociation of M proteins, which are subsequently delivered to the proximity of the nuclear pore complex and inhibit the export of host mRNAs from the nucleus.

The input viral genome (negative strand) is first transcribed into a 47-nt leader sequence and a series of viral mRNAs by the RNA-dependent RNA polymerases that are packaged into the virion. RNA synthesis switches from transcription to replication, only after enough nucleocapsid (N) proteins have been synthesized. Genome replication depends on ongoing protein synthesis, because unlike the naked viral transcripts, the full-length viral RNA species (including both negative-strand RNA genome and positive-strand replication intermediate) have to be encapsidated immediately (40).

Similar to positive strand RNA viruses and double strand RNA viruses, VSV also compartmentalizes its replication machinery in order to favor RNA synthesis and evade host recognition. During the early stage of virus infection, primary transcription occurs throughout the host cytoplasm. Later on, the viral proteins accumulate and form the inclusion bodies that are the predominant sites for the following RNA synthesis (71). Interestingly, the newly synthesized viral transcripts are transported away from the inclusion bodies, which are dependent upon the microtubule network. In all probability, the redirected viral mRNAs facilitate protein expression (71).

Just as most other enveloped viruses, VSV completes its replication cycle by forming membranous vesicles and budding on the surface of the plasma membrane (72). Prior to virus budding, the RNP complex condenses upon binding to the M protein. The initiation of budding occurs when the condensed RNP complex approaches the inner leaflet of the plasma membrane, probably depending on the interaction between the VSV M protein and the G protein. As the only spike protein encoded by the virus, the G protein has a C-terminal cytoplasmic tail that is essential for the interaction with other viral components during virus assembly (73, 74). The truncation of this cytoplasmic tail or even residue substitutions significantly attenuate the virus (38).

A number of host factors are hijacked to facilitate virus budding. For instance, the PPPY motif within M protein interacts with Nedd4, a host-encoded E3 ubiquitin ligase involved in the endocytic pathway and vacuole sorting (75, 76). Therefore, it is not surprising that some host proteins have been identified inside the VSV virion. These factors may play significant roles in either virus replication or virion assembly, and are accidentally enclosed by the newly-formed virus particles (77, 78).

1.5 Codon pair bias and codon pair bias score

The degeneracy of codons exhibits as the multiplicities of codon choice to specify one amino acid. For example, in humans Arginine can be encoded by six codons with different frequencies: CGT (42%), CGC (37%), CGA (5%), CGG (8%), AGA (4%) and AGG (3%). Similarly, any given protein can be encoded by a large number of nucleotide sequences. For example, a pentapeptide Gly-Ala-Met-Phe-Leu, can be encoded in almost 200 ways (4x4x1x2x6=192). However, these nucleotide sequences are not equally represented in nature. In other words, only a limited number of sequences are selected during evolution (regarded as wild type sequences). In addition, for different organisms, divergent preferences evolved in terms of how to best "encode" a protein.

One of the most well-studied encoding "preferences" is codon bias, which mostly reflects the codon adaptation to the tRNA pool inside the cells. A strong codon bias exists within almost any given organism. It is hypothesized that the codon usage directly affects the rate of ribosome traffic on the mRNA since it presumably takes longer to incorporate the tRNAs pairing with rare codons into the ribosome A-site. In prokaryotic cells substitution of the synonymous but frequently used codons by the infrequently used ones causes a reduction of protein expression, and the replenishment of those tRNAs corresponding to the infrequently used codons rescues the gene expression deficiency (79, 80). First supported by findings in *E.coli*, the premise that codon bias affects both the efficiency and accuracy of translation by controlling translation kinetics, was later confirmed in eukaryotic organisms and has been applied to virus attenuation. These pioneer studies served as good examples for the development of live attenuated vaccine strains (81, 82).

However, accumulating evidence indicates that codon bias is not the sole driving force for the evolution of gene composition. Codon pair bias (CPB) is another important but highly underappreciated genetic phenomenon that affects gene expression. In this bias, certain synonymous codons are found to be adjacent to each other more or less frequently than the prediction based on the genome-wide codon usage of the two individual codons that make up the pair. In other words, it results from the observation that some codons "dislike" or "favor" each other. For example, in human genomes the codon pair GCC-GAA that encodes the amino acid pair Ala-Glu has an expected frequency of 16.8% (40.3% x 41.9%), however, the observed frequency of this codon pair is only 3%, much less than the expected frequency (83, 84).

In a given genome all the codon pairs can be mainly categorized into two groups: 1) overrepresented codon pairs with higher observed frequencies than the expected frequencies and 2) underrepresented codon pairs whose observed frequencies are less than the expected frequencies. It is worthwhile to note that similar to codon bias, codon pair bias also differs in diverse organisms and is expected to be an accumulating effect. The differences between single codon pairs should be small but when amplified by large-scale recoding, robust phenotypes are scored. During the recoding, neither codon usage nor amino acid sequence is altered (85, 86).

For the convenience of the study, Codon Pair Bias Score (CPS) has been developed to evaluate the codon pair bias of a given codon pair or a gene fragment based on the genome-wide analysis of the reference genome, similar to the Codon Adaptation Index (CAI) widely used in codon bias analysis. Briefly, the codon pair bias score of a given codon pair is the natural logarithm of the ratio of the observed codon pair frequency and the expected codon pair frequency in a given genome. The frequency of the corresponding amino acid pair in the particular genome has also been taken into account. The codon pair bias score for a given gene is the arithmetic mean of the individual codon pair bias scores of all pairs that constitute the gene. Therefore, a high codon pair bias score of a given gene usually indicates an enrichment of overrepresented codon pairs while a low codon pair bias score means a higher frequency of underrepresented codon pairs (Fig 1.3) (87).

1.6 Large-scale genome recoding strategies

As mentioned above, the nature of codon pair bias suggests that in order to produce significant phenotypes, the use of computer-aided design and chemical synthesis are necessary since codon pair bias is expected to be an accumulating effect and obviously it is not feasible to mutate hundreds of codon pairs using conventional PCR-mediated mutagenesis. In 2002 our group generated the first infectious poliovirus by using chemically synthesized viral cDNA, in the absence of natural template (88, 89). Since then, the rapid development of chemical DNA synthesis technology have conferred us the capability of modifying a large portion of the viral genome in a short time and at an affordable cost (90, 91).

Using a unique computer algorithm, we have applied multiple codon pair bias-based recoding strategies to a variety of viral pathogens, including poliovirus, influenza A virus, dengue virus and respiratory syncytial virus (87, 92-95). In our first endeavor to recode the poliovirus genome, we altered the expression of the viral P1 gene by rationally introducing hundreds of silent mutations to the corresponding region of the viral genome. Underrepresented codon pairs (Min design) exerted a substantially negative effect upon gene expression, as expected. Interestingly, the extent of expression deficiency as well as the virus attenuation was proportional to the length of the recoded gene segments, which supported the notion that codon pair bias functions as an accumulating effect. Additionally, a slight translation augmentation was achieved by codon pair optimization of the same P1 region (Max design) while a "neutral" recoding (termed as Scramble or SD design) that shuffled the synonymous codons without changing the overall codon pair bias of the same region did not significantly interfere with poliovirus growth (87). Taken together, this pivotal study suggested that viral gene expression could be deliberately controlled by a rational design of coding sequences. More importantly, the change of gene expression level probably results from numerous synonymous substitutions, which makes phenotypic reversion nearly impossible. This unique property of virus attenuation resulting from codon pair bias is of great value for the development of live attenuated vaccine strains against RNA viruses, since the spontaneous mutation rates of RNA viruses are much higher than those of DNA viruses and other microorganisms (96).

Later on, a similar recoding approach was applied to the influenza A virus PR8 strain, which served as a paradigm of rapid-response vaccine development against the threatening of emerging flu strains. A number of influenza virus genes, including NP, PB1, HA and NA, have

been recoded separately or jointly, and as expected, all the "min" constructs have led to dramatic reductions of viral gene expression accompanied with different levels of virus attenuation (93, 95). For one of the recombinant strains, $(NA+HA)^{min}$, levels of NA and HA proteins were considerably reduced in cell culture. This recombinant virus exhibited a remarkable attenuation phenotype and generated long-lasting protective immunity in an animal model. The extremely low PD_{50} (medium protective dose) achieved by $(NA+HA)^{min}$ virus, is as low as 2.4, which is perhaps the lowest PD_{50} ever reported in a mouse model (93).

In addition to modulating virulence in an array of existing and/or emerging viral pathogens, we also utilized large-scale genome recoding as a tool to scan an RNA virus genome and search for cis-acting elements that were essential for virus replication. In these designs, which are termed as "scrambled", synonymous codons within the viral genome are randomly shuffled to introduce the maximal number of nucleotide changes, without altering the amino acid sequence, codon bias or codon pair bias. Since synonymous mutations are introduced to almost every codon, theoretically any functional RNA sequence signatures that are longer than 3-4 nucleotides would have been destroyed, regardless of whether they are linear nucleotide sequences or have formed a higher order structure. By taking advantage of this unique recoding strategy "scramble", we have successfully identified two functionally redundant RNA elements in the 3'-terminal coding sequence of poliovirus RNA polymerase $3D^{pol}$ (97). Specifically in that study, the "scramble" strategy functioned well as an easy-to-use but powerful platform approach to unveil viral RNA elements that could not be located using the conventional mutational analysis.

Recently, a revised recoding strategy termed as "Scramble-max" (Sdmax) has been developed, based on the working model of codon pair bias. This strategy can be applied to genome scanning in a similar manner, in search of RNA sequence signatures that are critical for virus replication. Relative to the WT or "SD" sequence, the "sdmax" sequence has an overall higher codon pair score, which in turns avoids the inadvertent generation of local "deoptimized" stretches. Table 1.1 summarizes the characteristics of these three distinct recoding strategies that are described above.

1.7 Figures and Tables

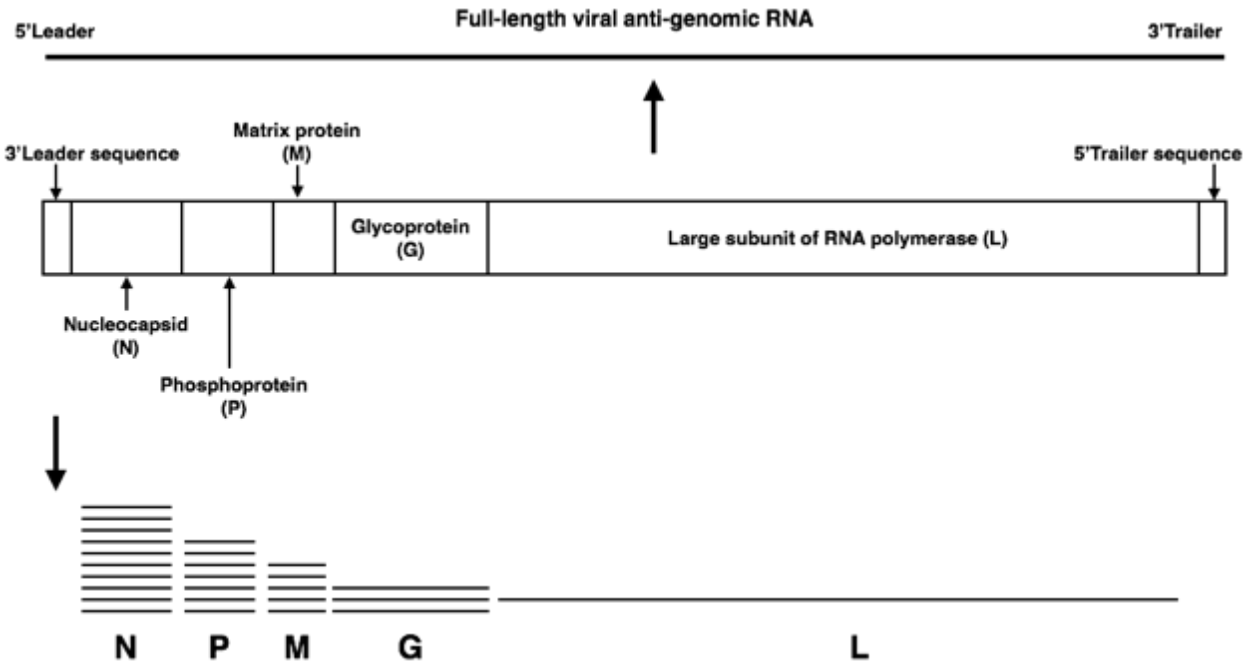


Figure 1.1 VSV genome structure and the products of viral RNA synthesis.
From the 3'end to the 5'end is the Leader sequence, the genes encoding N, P, M, G and L, and the Trailer sequence. The full-length viral anti-genomic RNA, which is the replication intermediate, is shown above the template. The five viral gene transcripts are shown below the template.

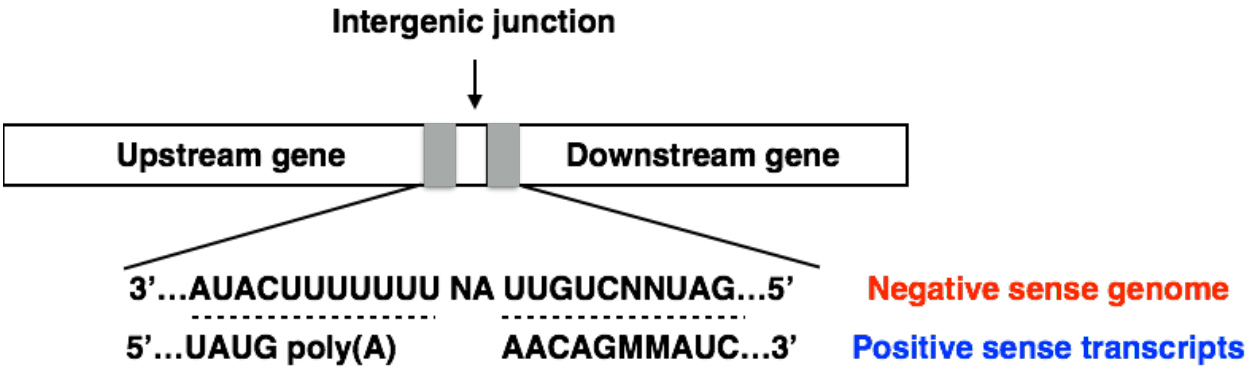


Figure 1.2 Conserved sequences near the intergenic junctions of VSV.
Shown from the 3'end to the 5'end there are the sequences presented at the end of the upstream gene (3' AUACUUUUUUU 5'), the intergenic junction sequence (3' NA 5') and the conserved sequence at the beginning of the downstream gene (3' UUGUCNNUAG 5'). The sequences are involved in signaling transcription termination, polyadenylation and re-initiation. The sequence shown from the 5'end to the 3'end is the corresponding sequence presented on the transcription products.

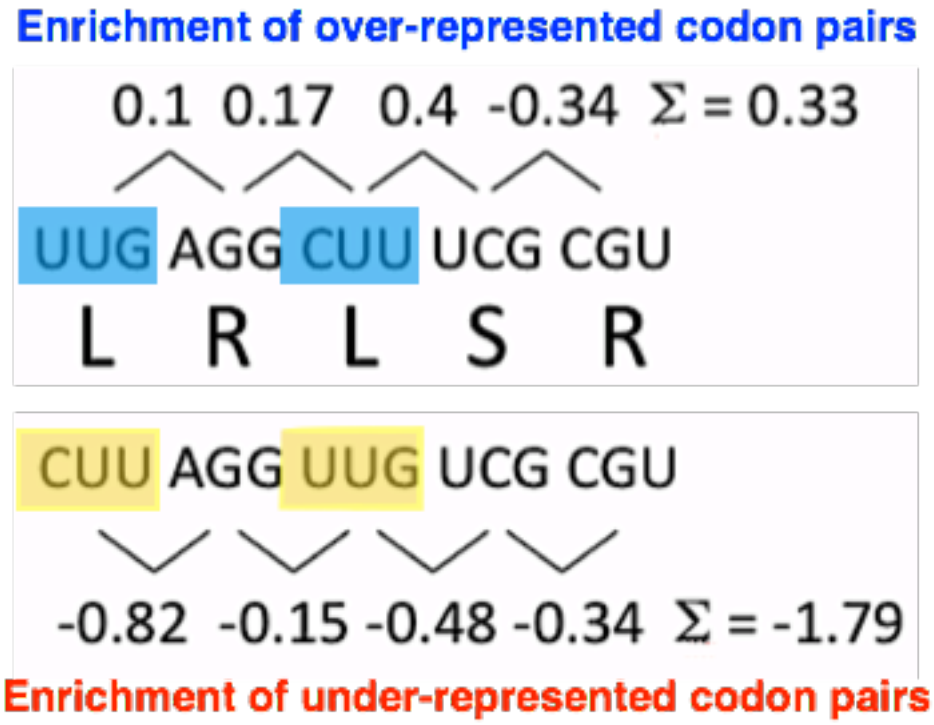


Figure 1.3 Codon pair bias score of a given sequence.
A pentapeptide, Leu-Arg-Leu-Ser-Arg, can be encoded by 6x6x6x6x6=7776 different ways. The upper sequence, UUG AGG CUU UCG CGU, has a codon pair bias score of 0.33, which suggests an enrichment of overrepresented codon pairs in the sequence. However, if the positions of the two synonymous codons, UUG and CUU, are exchanged, several underrepresented codon pairs are generated, which leads to a significantly lower codon pair bias score (-1.79, the bottom sequence). During the shuffling, neither codon usage nor amino acid sequence is altered.

	Deoptimize/Min	Scramble/SD	Scramble-Max/ Sdmax
Codon pair bias score	Lower than WT	Same as WT	Higher than WT
Codon usage	Same	Same	Same
Amino acid sequence	Same	Same	Same
Codon pair	Under-represented	Neutral	Over-represented
Phenotype expected	Highly attenuated	WT-like	WT-like
Application	Virus attenuation	Genome scanning	Genome scanning

Table 1.1 A summary of the three distinct recoding strategies.
Min/deoptimized design, which contains a higher amount of underrepresented codon pairs, usually leads to a reduction of viral gene expression accompanied with virus attenuation. Scramble/SD design that retains the average codon pair bias score of WT sequence is not expected to attenuate the virus. Scramble-max/Sdmax design, a revised version of SD design, has a relatively higher codon pair bias score in comparison to WT or SD sequence.

Chapter 2 Reverse genetics of VSV

2.1 Introduction

VSV can be developed into an antigen delivery platform or an oncolytic virus. However, WT VSV is not safe enough for direct human use because of its neurotoxicity when given intranasally. As mentioned before, the "deoptimized/min" strategy built on codon pair bias theory has shown great potential in developing live attenuated vaccine strains for poliovirus (a positive strand RNA virus) and influenza virus (a segmented, negative strand RNA virus). It is of interest to test a similar strategy in VSV, which is a prototype of negative strand, non-segmented RNA viruses (the superfamily of *Mononegavirales*). To my knowledge, this is the first virus in the family of *Rhabdoviridae* that was attenuated by codon pair bias. In addition, by utilizing a novel recoding strategy, "sdmax", which resembles the "SD" design that we have tested in the poliovirus study, it is also the first time to search for RNA sequence signatures in the ORFs of VSV genes by large-scale genome recoding. The identification of unknown nucleotide sequence signatures, which may be involved in virus replication or pathogenesis, could also contribute to the development of live vaccines against VSV infection. Considering that a number of human-threatening pathogens, such as human respiratory syncytial virus, are also grouped into *Mononegavirales* and share a similar genome structure with VSV, my study may be useful for the development of live, replication competent, attenuated vacccine strains for these viruses.

To study a recombinant RNA virus with an altered growth phenotype, the complete viral genome has to be reverse transcribed into a cDNA clone that can be manipulated in the test tube and then converted back into an infectious RNA virus. Unlike a positive strand RNA virus, the reverse genetics of negative strand non-segmented RNA viruses is usually problematic. The full-length genomic or anti-genomic viral RNAs are not infectious, owning to a lack of the signals that can be recognized by the host transcription apparatus. Therefore, transfection of the nascent viral genome never leads to a successful virus recovery. To circumvent this problem, the virus-encoded proteins that are necessary for the assembly of the RNA-dependent RNA polymerase (for VSV these components are N, P and L proteins, which are packaged into the virion during natural infection) have to be co-expressed in the cells that are transfected with the full-length viral RNA(98, 99). Specifically for the reverse genetics of VSV Indiana serotype, transcription

of the N, P and L genes as well as viral anti-genomic RNA have to be initiated by bacteriophage T7 polymerase that is expressed from a supporter vaccinia virus. The successful rescue event independent of supporter virus appears to be very rare (personal communication with Dr. Harty, University of Pennsylvania), and the vaccinia-free recovery strategy was also highly inefficient in my own experiments (data not shown)(100).

2.2 Results

2.2.1 Recovery of WT VSV from a cDNA clone

In close collaboration with Dr. SP Whelan at Harvard Medical School, I chose the vaccinia virus-mediated strategy to recover WT VSV, which was invented by Dr. SP Whelan in 1995(99). Briefly, I infected BHK21 cells with a recombinant vaccinia virus strain vTF7-3, to express T7 polymerase in the cytoplasm, followed by co-transfection of three supporter plasmids pN, pP and pL as well as the anti-genomic VSV cDNA clone pVSV1+.

The supernatants were collected 48-72 hours after transfection. The vaccinia virus strain used herein is replication competent, and therefore the vaccinia contamination has to be physically removed. Vaccinia virus is a very large, complex enveloped virus belonging to the *poxvirus* family, the dimension of which is roughly 360x270x250 nm. Most of the viruses can be removed by filters with a pore size of 0.2 μm(101). After clarification and filtration, I passaged the supernatants into a fresh monolayer of BHK21 cells. For WT VSV, a typical cytopathic effect caused by VSV was expected after 1-2 days incubation at 37°C. The supernatants were then harvested and the viruses were plaque purified in Vero cells (Fig 2.1). The virus recovery was confirmed by three different protein assays: immunocytochemistry, immunoblotting and ^{35}S-Met trans-labeling of infected cells (data not shown).

2.2.2 Optimization of the VSV recovery system

The ultimate goal of the project is to develop a live attenuated VSV, whose replication may be significantly less competent than that of the WT virus in cell culture. As described above, the reverse genetics of VSV is relatively complicated. The efficacy of rescue depends highly on the quality of the supporter plasmids, their molar ratios as well as the infection efficiency of the helper virus (personal communication with Dr. SP Whelan at HMS). Therefore I utilized a VSV-based reporter plasmid pVSV (-) GFP (kindly provided by our collaborator) to optimize the whole recovery system (Fig 2.2A).

The pVSV (-) GFP plasmid contains a GFP ORF, which is flanked with the two VSV untranslated regulatory regions (3'Leader sequence and 5'Trailer sequence), providing an easy and quantitative readout to evaluate the recovery efficiency. I infected the BHK21 cells with the vaccinia virus vTF7-3, co-transfected three supporter plasmids and the GFP reporter, followed by the measurement of the percentage of GFP positive cells using flow cytometry under different

transfection and infection settings. In this system, the percentage of GFP positive cells (GFP+%) in the transfected cell population reflects the efficiency of authentic virus recovery (Fig 2.2B).

The percentage of GFP positive cells increased with the dose of vTF7-3 at a given transfection setting. The transfection efficiency reached a peak when the cells were infected at a MOI of 10 prior to co-transfection. Under my experimental conditions, the dose of vaccinia virus should not exceed a MOI of 10, given the finding that high MOI infection led to rapid cell death.

I also tested and compared the recovery efficiency by deliberately adjusting the molar ratios of the three supporter plasmids (pN, pP and pL). The highest GFP+% was achieved when the co-transfection was carried out at a ratio of pN: pP: pL=6.5μg: 1.5μg:1.25μg. Taken together, after optimization, the estimated successful rate for virus recovery could exceed 20%, which made it possible to recover highly attenuated VSV strains (Fig 2.3).

2.3 Discussion

The genomes of NNS viruses are less amenable to genetic manipulation for several reasons. First, the precise 5' and 3'ends of the viral genome are required for the recognition of viral RNA polymerase, given the fact that cis-acting elements that are essential for virus replication are located at both ends. This has been circumvented by the combinational use of T7 polymerase and an autolytic activity of a ribozyme sequence(103). Second, the viral RNA polymerase is required for the initiation of viral gene transcription. During natural infection, the polymerase is packaged into the virion and released to the cytoplasm immediately after virus entry. Therefore similarly, during the artificial recovery, the polymerase has to be provided *in trans*. Third, the viral RNA polymerase can only recognize the encapsidated viral genome. The naked RNA genome or anti-genome is not infectious and not able to initiate a successful infection. The viral N protein has to be provided *in trans* as well.

Recombinant vaccinia virus expresses T7 polymerase efficiently in the cytoplasm, which drives the RNA synthesis of the viral polymerase genes as well as the full-length viral anti-genome. The use of vaccinia virus is indispensable because besides the T7 polymerase it offers, it also expresses cap-specific mRNA methyl-transferase (mRNA capping enzyme) and poly-A polymerase(104). In the absence of these mRNA modification enzymes, the viral N, P and L genes can not be expressed robustly since the mRNAs transcribed by T7 polymerase lack 5' cap and 3'poly-A tail, which are critical for mRNA stability and translation enhancement(105).

Provided that all the components of the viral RNA polymerase are expressed at an appropriate molar ratio and assembled properly, the viral full-length anti-genomic RNA, which is transcribed by T7 polymerase as well, is recognized by the RdRP and serves as a template for viral RNA synthesis. The negative sense, full-length viral genomes are firstly synthesized and subsequently transcribed into a series of viral mRNAs. As discussed above, the successful rate of virus recovery is dependent upon the transfection efficiency(106).

The reporter plasmid pVSV (-) GFP contains a GFP ORF that is flanked by the two VSV regulatory sequences, the 3' Leader sequence and the 5' Trailer sequence. The transcription of the GFP gene is initiated by T7 polymerase and then rapidly expanded by the newly synthesized viral transcription apparatus, since all the GFP transcripts contain the cis-acting elements for the recognition of the viral RNA polymerase. Therefore, after transfection, each GFP positive cell represents at least one successful transfection event, offering an estimation of the recovery rate.

It is noteworthy that the 0.2 μm filtration does not completely remove vaccinia virus contamination in the collected supernatants. During the following passages, however, VSV replicates much faster than the helper virus. Thus, VSV becomes predominant in the population, which enables it to rapidly infect and deplete all the cells before vaccinia virus does.

2.4 Materials and Methods

Cells

BHK21 cells (baby hamster kidney, obtained from ATCC #CCL-10) and Vero cells (monkey kidney epithelial, obtained from ATCC #CCL-81) were used for VSV recovery and titration, respectively. The cell lines were maintained in Dulbecco's Modified Eagle's Medium (DMEM) plus 10% fetal bovine serum (FBS) at 37°C.

Virus:

Parental vesicular stomatitis virus strain (abbreviated as WT VSV in this study) was recovered from an infectious cDNA clone of VSV Indiana serotype pVSV1+ by a well-established reverse genetics system (99).

Recombinant vaccinia virus strain vTF7-3 was a kind gift from Dr. SP Whelan at Harvard Medical School. The plasmid pTF7-3 containing the vaccinia virus P7.5 promoter and the bacteriophage T7 RNA polymerase gene was used to transfect CV-1 monkey kidney cells infected with wild type vaccinia virus. The resultant recombinant vaccinia virus containing the T7 polymerase gene under the control of the vaccinia P7.5 promoter was isolated and designated as vTF7(107).

Reverse genetics of VSV

The reverse genetics of WT VSV was performed as described before(99). Briefly, 80% confluent BHK21 cells were infected with vTF7-3 at a MOI of 10 for 1 hour at 37°C. Then the virus inoculum was removed and the infected cells were co-transfected with 6.5 μg pN, 1.5 μg pP, 1.25 μg pL and 6 μg pVSV1 (+). After 5 hours, the culture media was discarded and the transfected cells were supplemented with fresh 2% FBS DMEM. Following 48-72 hours incubation, the culture supernatants were collected and clarified at 3000 rpm (revolutions per minute) for 2 minutes. The supernatants were then filtered though a 0.2 μm filter and passaged onto a fresh monolayer of BHK21 cells.

After 1-2 days, if a typical cytopathic effect caused by VSV infection could be observed, the supernatants were harvested, aliquoted and stocked at -70°C.

Plaque-forming assay

VSV plaque assays were performed on 90-100% confluent Vero cells in 6-well plates, as described before (108). Virus samples were prepared by a series of 10-fold dilutions, by transferring 100 µl of the original virus stock into 900 µl DMEM, and repeating the procedure until the desired dilutions were reached. The Vero cells were then infected with those 10-fold serial dilutions of WT VSV. Specifically, cells in each well were first washed three times with 2 ml pre-warmed phosphate buffered saline (PBS), and then incubated with 300 µl of each diluted sample at 37°C and 5% CO_2 for 60 minutes. The plates were shaken every 10 minutes to prevent the cells from drying out. The inoculum was subsequently removed and the cells were overlaid with a low-melting agar gel containing 0.25% FBS. The cells were cold shocked at 4°C for 5 minutes and then incubated at 37°C overnight or for prolonged periods. The cells were stained with 1% crystal violet for at least 10 minutes. The number of visible plaques on each well were counted and multiplied by the dilution factor to obtain the plaque forming units per milliliter (PFU/ml).

Immunocytochemistry

BHK21 cells were infected with WT VSV at a low MOI of 0.001. At 8 hpi (hours post infection), the culture media was removed and the cells were washed extensively with pre-warmed PBS for 4 times. Cells were fixed with 3.7% Paraformaldehyde (PFA) for 30 minutes and then permeabilized with 0.2% Saponin/ 2% BSA/ PBS for 15 minutes.

For VSV G protein staining, cells were stained with mouse derived anti-VSV G antibody [P5D4] (purchased from Abcam, Inc, Cambridge, MA) at a dilution of 1:500, followed by a 1:500 dilution of secondary anti-mouse antibody conjugated to AlexaFluor-594 (Green color, Invitrogen Molecular Probes). The stained cells were then carefully washed three times and imaged using de-convolution microscope.

Immunoblotting

BHK21 cells were infected with WT VSV at a MOI of 3. After overnight incubation, the cells were washed with PBS once and then resuspended in RIPA lysis buffer on ice for 30 minutes. The formulation of RIPA buffer is as below: 25mM Tris-HCl (pH7.6), 150mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS and 1x protease inhibitor cocktail (purchased

from Sigma). The suspension was transferred to a new micro-centrifuge tube and spun for 10 minutes at 12,000x g to remove cell debris. The VSV proteins were then separated on 7.5% SDS-PAGE gel and transferred to a nitrocellulose membrane (Whatman, Minnetonka, MN). Blots were blocked for 1 hour in 5% dry milk powder dissolved in 1x Tris-buffered saline (TBS, pH7.4). After being blocked, blots were then washed three times with 1x TBS containing 0.1% Tween-20 and probed with both anti-VSV G antibody [P5D4] (purchased from Abcam, Inc, Cambridge, MA) at a dilution of 1:10000, and anti-VSV M antibody (purchased from KeraFAST, [23H12]) at a dilution of 1:2000 overnight at 4°C.

Blots were then washed with 1x TBS containing 0.1% Tween-20 for three times and incubated with blocking buffer containing 1:10000 diluted goat anti-mouse horseradish peroxidase-conjugated antibodies (Pierce Antibody products, Inc, USA) at room temperature for 1 hour. Blots were washed three times with 1x TBS containing 0.1% Tween-20 and analyzed using chemiluminescent Western Blot substrate (Pierce Antibody products, Inc, USA).

^{35}S-Met trans-labeling of VSV infected cells

BHK21 cells were plated in 35mm culture dishes on the day prior to infection in order to reach an optimal density (80% confluent when infection was conducted). Cells were then infected with WT VSV at a MOI of 100, to completely block the host gene expression. At 3.5 hpi the culture media was removed. Cells were then washed with pre-warmed PBS three times and incubated in DMEM lacking methionine and cysteine for 30 minutes, in order to deplete the pools of free Met/Cys inside the cells. At 4 hpi ^{35}S-labeled Met/Cys was added to the culture media when the host protein synthesis was virtually shut off by the viruses. At 6 hpi the whole cell lysates were collected using RIPA lysis buffer and analyzed on a 7.5% SDS-PAGE gel.

Flow cytometry (FACS) analysis to quantitate GFP+% in the transfected population

Monolayers of BHK21 cells were initially infected with recombinant vaccinia virus vTF7-3 at a given MOI (three MOIs tested in this study: MOI=2, MOI=5, MOI=10), followed by the co-transfection of three supporter plasmids (pN, pP and pL) as well as the negative-sense VSV minigenome plasmid pVSV(-) GFP. After 4-5 hours, the media was replaced with DMEM containing 2% BCS.

The transfected cells were incubated at 37°C overnight. The culture media was removed and cells were washed with cold PBS 3 times. The cells were then detached from the culture plates using Trypsin-EDTA and concentrated by centrifugation for 10 min at 800x g. The supernatants were carefully removed by aspiration and the cell pellets were resuspended in 500 µl cold PBS. 500 µl of ice cold, buffered 2% formaldehyde solution was added and mixed again (approximately 10^6 cells per ml) . The fixed cells were subsequently incubated at 4°C for 15 min.

After fixation the cells were spun down for 5min at 300x g at 4°C. The supernatants were removed by aspiration and the cell pellets were carefully resuspended in cold PBS. The samples were filtered through a nylon mesh to remove cell clumps before acquisition on the flow cytometer. The GFP positive cells were counted by flow cytometry and the collected data were further analyzed using FlowJo software.

2.5 Figures

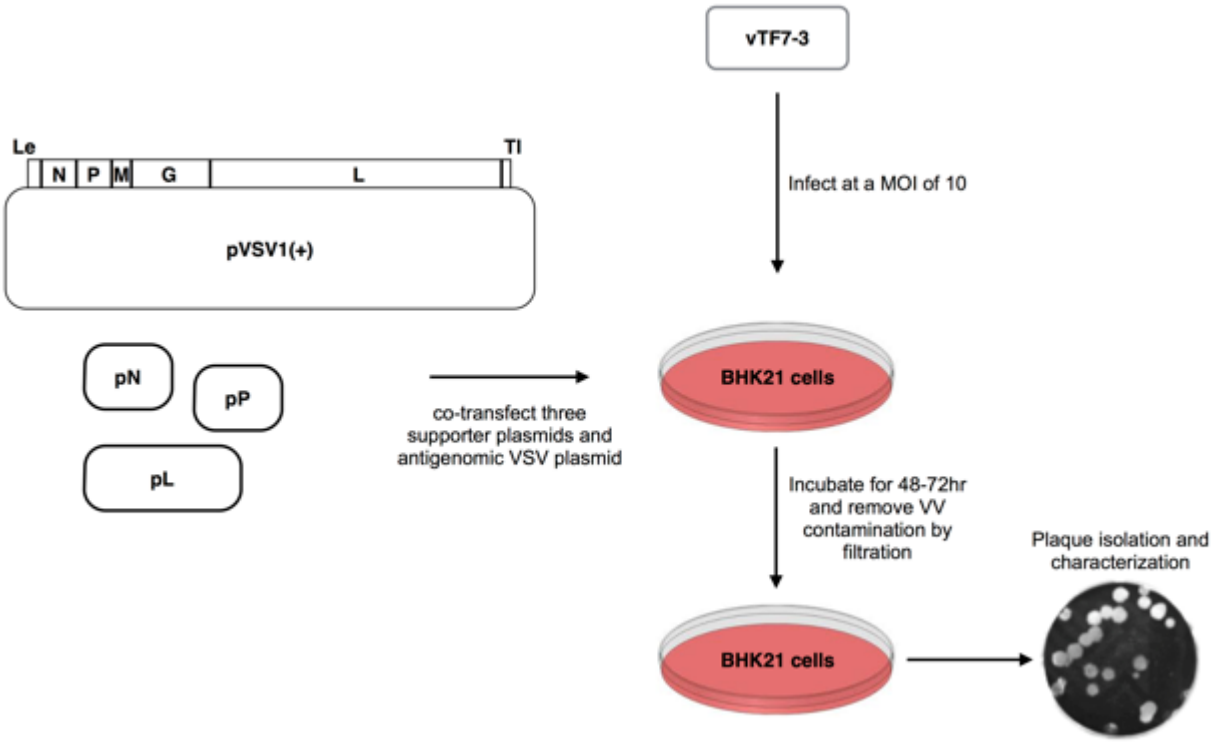


Figure 2.1 Reverse genetics of WT VSV.
BHK21 cells were first infected with T7 polymerase-expressing vaccinia virus vTF7-3 for 1 hour, followed by the co-transfection of the three supporter plasmids as well as the parental antigenomic plasmid pVSV1 (+). At 48-72 hpt, the supernatants were harvested, filtered through a 0.2 μm filter to remove most vaccinia contamination and passaged onto a fresh monolayer of BHK21 cells. When the typical cytopathic effect of VSV was observed, the supernatants were collected and allowed to form plaques in Vero cells (99).

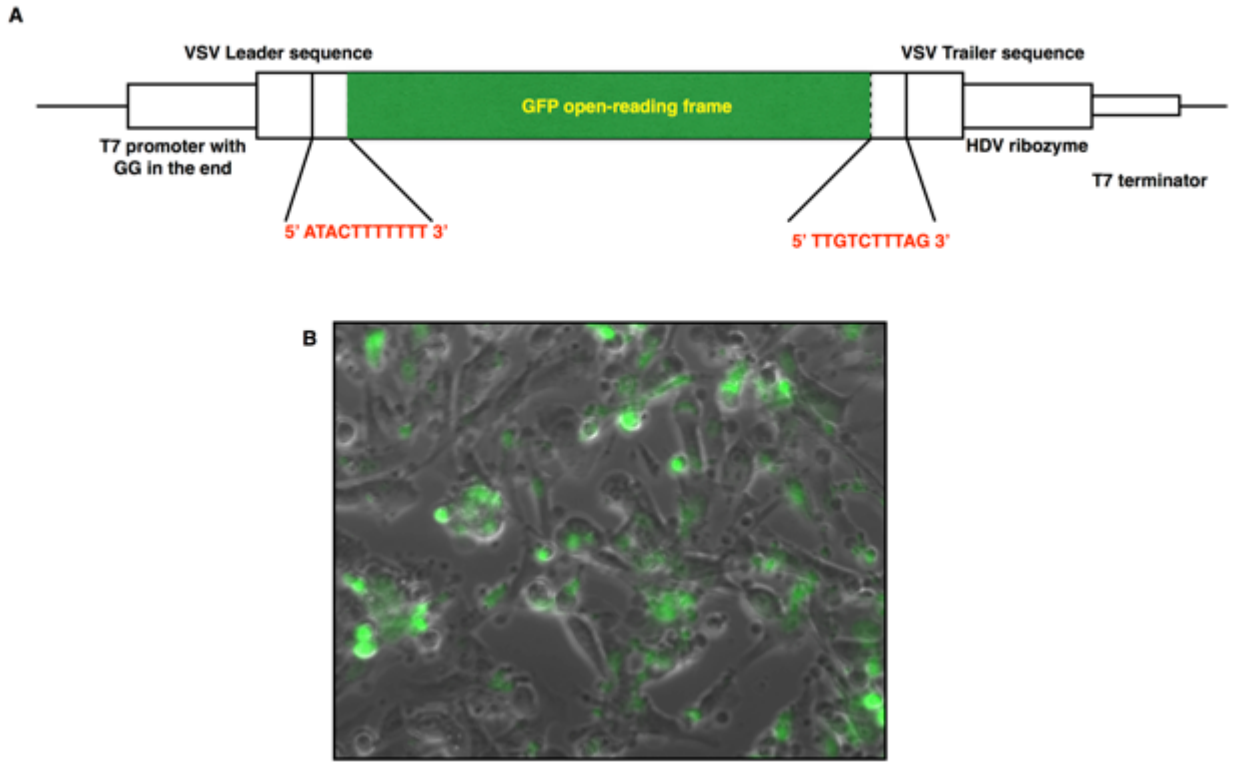


Figure 2.2 VSV minigenome-based reporter pVSV (-) GFP.
(A) A schematic diagram of VSV minigenome-based reporter pVSV (-) GFP. The GFP gene ORF is flanked by the cis-acting elements that can be recognized by VSV RdRP. (B) The image taken 18hr after transfection showed the robust expression of GFP.

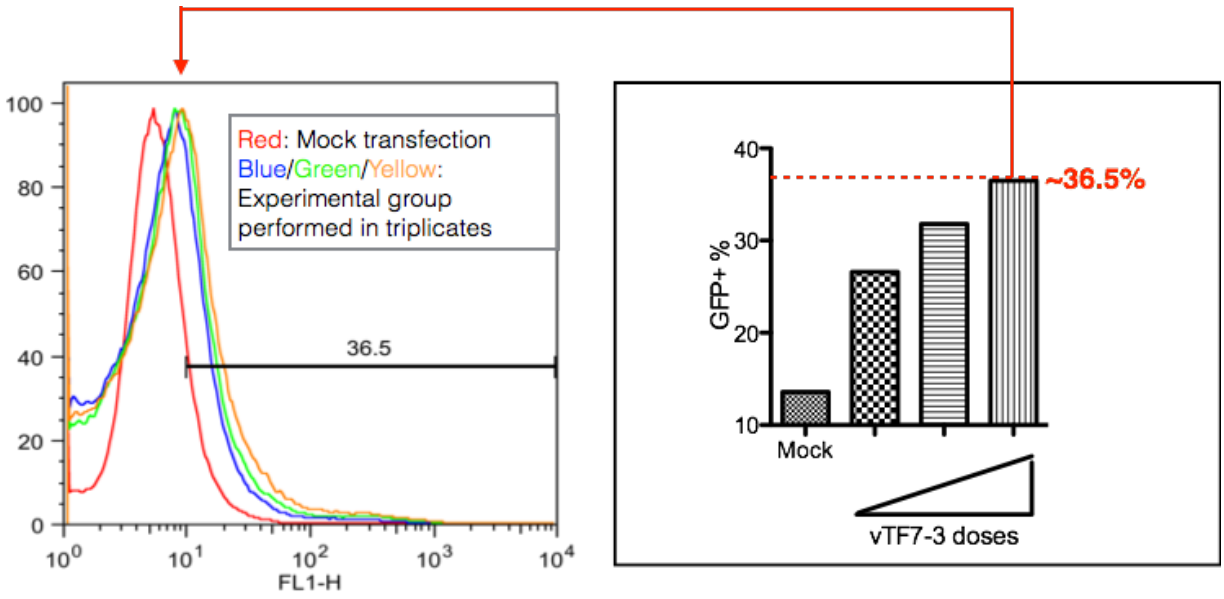


Figure 2.3 Detailed feasibility study of virus recovery.
An incremental enrichment of GFP+% in the transfected population correlated well with a higher dose of vTF7-3. After the molar ratio of co-transfected plasmids and the dose of vaccinia virus were deliberately adjusted, the percentage of GFP positive (GFP+%) cells could be as high as ~25%.

Chapter 3 Design, construction and recovery of recombinant VSVs containing partially recoded L genes

3.1 Introduction

The combinational use of computational biology and synthetic virology has made it easier and more affordable to achieve large-scale, codon pair bias-based genome recoding, which iteratively results in strikingly altered biological phenotypes. The viruses we have previously tested, including poliovirus, influenza virus and dengue virus, all appeared to be highly attenuated when a number of underrepresented codon pairs were introduced into their genomes(87, 92, 93, 95, 109). Since all viruses depend on the protein synthesis machinery of their host cells, the principle of codon pair bias-based recoding should be applicable to most, if not all, types of viruses.

As a prototype of non-segmented, negative sense RNA viruses and maybe the most rapidly growing virus in cell culture, VSV has been extensively studied in the past few decades. Its simple genome organization, the existence of a well-established reverse genetics system and the relatively innocuous phenotype in healthy human beings, collectively make it an attractive model to study the effect of codon pair bias on mammalian pathogens that are closely related to VSV. On the other hand, VSV can be easily engineered and subverted for a variety of prophylactic and/or therapeutic applications. It is therefore of great interest to develop attenuated VSV strains, which presumably function as superior vectors for the delivery of antigens or anticancer regimens.

Nearly all negative strand RNA viruses encode a single large RNA polymerase L (except for the family member of *Orthomyxoviridae*), which contains all the catalytic activities involved in virus genome replication and mRNA synthesis (51, 110). The amino acid sequence alignment and low-resolution structural studies suggest that a conserved molecular architecture of the L polymerase may be adopted by many NNS viruses, which makes it an attractive target for gene recoding (50).

In addition, in our previous endeavor to recode the poliovirus genome, we realized that viral gene expression could be deliberately altered by a rational design of coding sequences, given the fact that the “Max” design of P1 led to a considerable translation augmentation

whereas the “Min” design of the same region resulted in a non-viable phenotype(87). Therefore it is also of interest to test multiple recoding strategies in this study.

3.2 Results

The well-established computer algorithm to recode viral genes on a large scale was adopted in my study(87, 92, 93, 95, 97, 109). Basically, the synonymous codons in the gene ORF were shuffled to create either underrepresented or overrepresented codon pairs, resulting in two distinct nucleotide sequences that encode the same amino acid sequence. Previous studies indicated that the genes in which the numbers of underrepresented codon pairs were increased were generally poorly expressed whereas the genes containing more overrepresented codon pairs were not negatively affected.

In an attempt to examine the effect of codon pair bias upon VSV replication, the 5'-terminal segment of the VSV L gene ORF (nt 4961-7534, numbers referring to the complete VSV Indiana serotype genome sequence GenBank#: J02428.1) was chosen for recoding and designated as L1(Fig 3.1). The first 228 (nt 4733-4961) nucleotides of the VSV L gene were left untouched, provided that the cis-acting elements that were critical for the recognition of the viral transcription machinery exist at both ends of the viral genes(40).

The recoding strategy "sdmax" was initially employed to "scan" the target region, to rule out the possibilities that certain unidentified RNA sequences that were involved in virus replication were accidentally removed during the following "deoptimization". The cDNA corresponding to the recoded region L1 were synthesized *de novo* (GenScript, NJ, USA) and designated as $L1^{sdmax}$. The "sdmax" sequence, which contained hundreds of overrepresented codon pairs but encoded the same amino acid sequence as WT, were subsequently cloned into the parental VSV antigenomic plasmid pVSV1+ to replace the corresponding WT sequence.

The reverse genetics of the recombinant virus $L1^{sdmax}$ was carried out using the vaccinia virus-mediated system as described before (99). The $L1^{sdmax}$ design yielded infectious virus after one blind passage, characterized by a typical VSV CPE observed at the beginning of P2 (passage 2) whereas the CPE of virus recovered from WT sequence usually appeared during P1 (passage 1).

The observation that a viable virus could be successfully recovered from $L1^{sdmax}$ suggests that the L1 region is suitable for recoding since it is devoid of cis-acting RNA elements that are essential for virus replication. Encouraged by the finding, I recoded the same region L1 using the "min" strategy, which resulted in an extensive enrichment of underrepresented codon pairs. Again, neither codon usage nor amino acid sequence was altered. The $L1^{min}$ sequence was

similarly synthesized and incorporated into the parental cDNA plasmid. The characteristics of these recoded sequences were summarized in Table 3.1 and the distribution of overrepresented/underrepresented codon pairs in $L1^{min}$/$L1^{sdmax}$ sequences were graphed in Figure 3.2.

The $L1^{min}$ design failed to generate a viable virus. Up to the end of Passage 5 (P5), I did not observe any CPE typically caused by VSV, although during the passages the vaccinia virus that were not completely removed slowly caused cell death. A variety of assays, including plaque-forming assay, immunocytochemistry and RT-PCR did not lead to any positive results signaling the existance of a viable $L1^{min}$ virus (data not shown). In other words, no trace of a replicative virus could be detected during the continuous passage of $L1^{min}$ design.

3.3 Discussion

By shuffling the positions of synonymous codons to generate more "favored" (higher codon pair scores) or more "disliked" codon pairs (lower codon pair scores) in a segment of the VSV L gene (L1 region), two L encoding sequences with opposite codon pair bias, $L1^{sdmax}$ and $L1^{min}$, were artificially generated. The L1 region described here (nt 4961-7534), comprises the coding sequence of domains I-III and part of domain IV of VSV L protein, which primarily forms the "core ring" structure of the viral RdRP and catalyzes RNA polymerization(51). I chose L1 as the target region for recoding, for several reasons: 1) The recombinant virus containing $L1^{sdmax}$ sequence was successfully recovered, suggesting that L1 region does not contain any RNA sequence signatures (for instance, cis-acting elements or highly structured RNA segment) that are critical for virus replication in tissue culture. As shown in Table 3.1, for the $L1^{sdmax}$ design, approximately 33% of the nucleotides in the L1 region have been mutated, thus theoretically any RNA signal or highly structured element longer than 3-4 nt would have been destroyed. I was aware that no RNA element had ever been identified in any ORF of the VSV genome, but the "sdmax" scanning was intentionally included as a proof-of-principle experiment; 2) the restriction sites at both ends (HpaI and AvaI) provided convenience for molecular cloning; 3) the L1 segment had a proper length for chemical synthesis, which assured a phenotype after recoding. In addition, recent genome-wide studies identified a conserved "ramp" of rare codons located within the first ~150 nt of the mRNA, which is believed to play a significant role in translation initiation(111). Therefore I left the first ~200 nt of the L gene ORF identical to the WT sequence, to avoid potential complications in data interpretation.

The lethal growth phenotype of $L1^{min}$ design was not surprising since the "min" design of poliovirus P1 domain, whose recoded region was almost as long as the L1 described here, also resulted in a non-viable virus (87). This observation also supports the hypothesis that codon pair bias could be an accumulating effect; if the "min" region is too long, it could become too detrimental to gene expression, which imposes a bottleneck during virus recovery.

3.4 Materials and methods

Viruses

The recombinant vaccinia virus strain vTF7-3 was used to deliver T7 RNA polymerase. The recombinant $L1^{sdmax}$ virus was the mutant VSV (Indiana serotype) containing a codon pair-optimized L1 segment. The recoded region was between nucleotide 4961 (HpaI site) and nucleotide 7534 (AvaI site) in the VSV L gene (the numbers referred to the complete genome sequence of VSV Indiana serotype, GenBank#J02428.1).

Synthesis and recovery of the VSV derivatives

The same algorithm described before was utilized to design the nucleotide sequences that encode the same peptide as WT but have distinct codon pair bias. The entire recoded segments were chemically synthesized by GenScript (NJ, USA) and cloned into the parental VSV antigenomic plasmid pVSV1+ to replace the corresponding WT sequence. The same reverse genetics described above was used to recover recombinant VSVs. Briefly, BHK21 cells were infected with vTF7-3 at a MOI of 10 and then co-transfected with the three supporter plasmids pN, pP and pL as well as the recombinant antigenomic plasmid containing the L1 derivatives. At 48-72 hpt (hours post transfection), the supernatants were collected and filtered, in order to remove most vaccinia virus contamination. The blind passages were performed every 2-3 days, until a typical CPE caused by VSV was observed. The recovered $L1^{sdmax}$ virus was subsequently plaque purified in Vero cells.

3.5 Figures and Tables

	$L1^{sdmax}$	$L1^{min}$	$L1^{WT}$
Codon Pair Bias Score	0.184	-0.412	0.027
Mutations/Length(nt)	858/2574	623/2574	-
Mutations%	33.3%	24.2%	-
Recoded regions(nt)	4961-7534	4961-7534	-
#U_3A_1 in ORF	144	241	144
#C_3G_1 in ORF	42	149	52
AAAUU(Au-rich)	21	20	20

Table 3.1 Characteristics of $L1^{sdmax}$ and $L1^{min}$ sequences

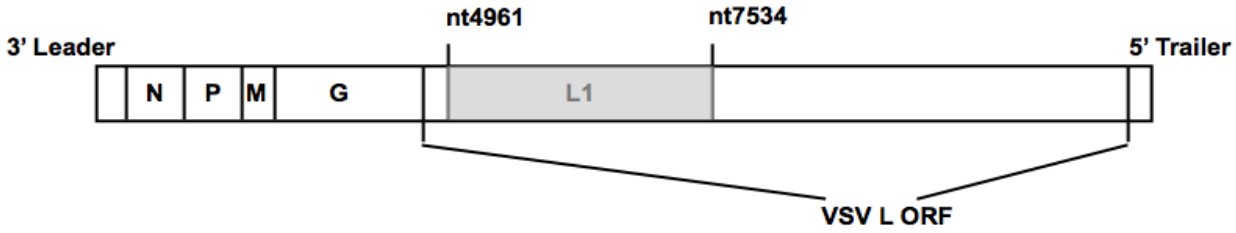


Figure 3.1 Genome structure of VSV and construction of the two L gene derivatives.
$L1^{sdmax}$ and $L1^{min}$ sequences were designed using the unique computer algorithm built in-house and generated by chemical synthesis. The two synthesized gene segments $L1^{sdmax}$ and $L1^{min}$, were used respectively to replace the corresponding region of WT VSV.

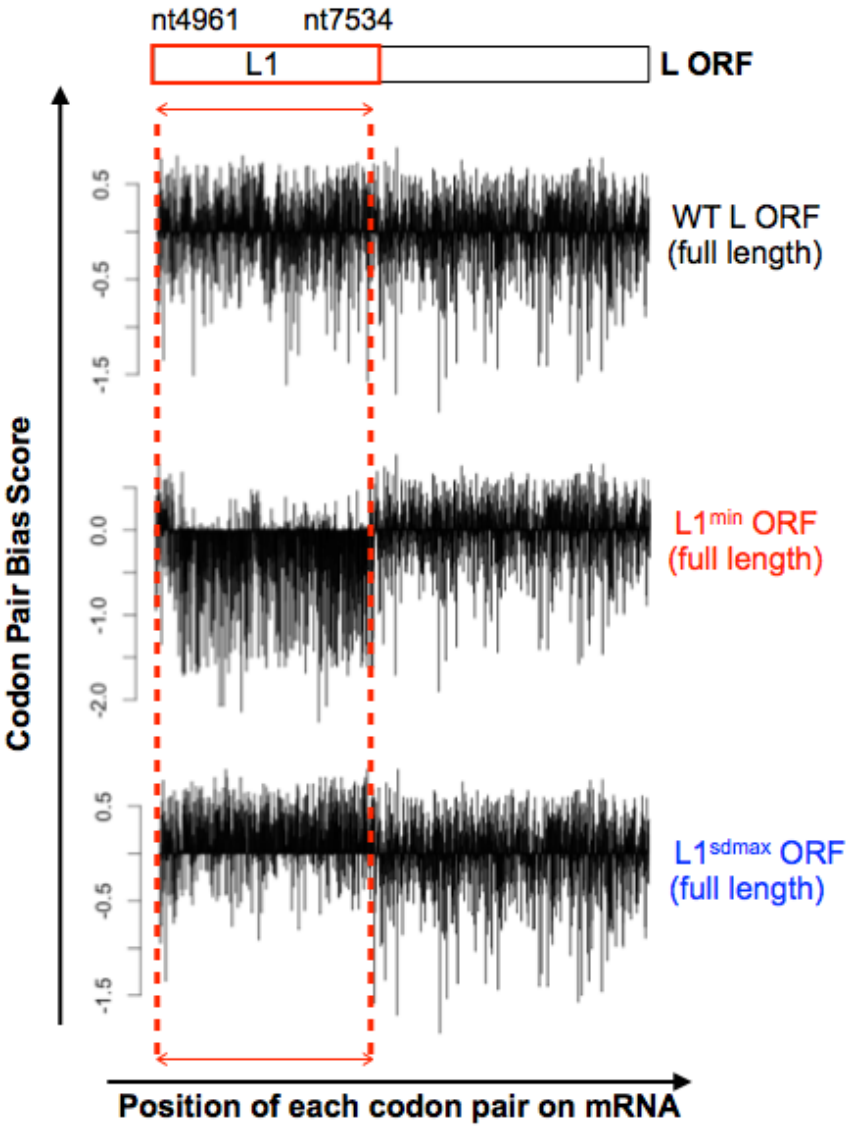


Figure 3.2 Distribution of codon pair bias in WT/L1^{sdmax}/L1^{min} encoding sequences.
The red dash lines highlighted the location of L1 region in the full-length L gene open-reading frame. The position of each codon was marked on the X-axis and the codon pair score (the bar graph) represented the frequency of individual codon pair in the host genome. The non-bias codon pair score was set to 0.0. More or less frequently used codon pairs were scaled up or down accordingly.

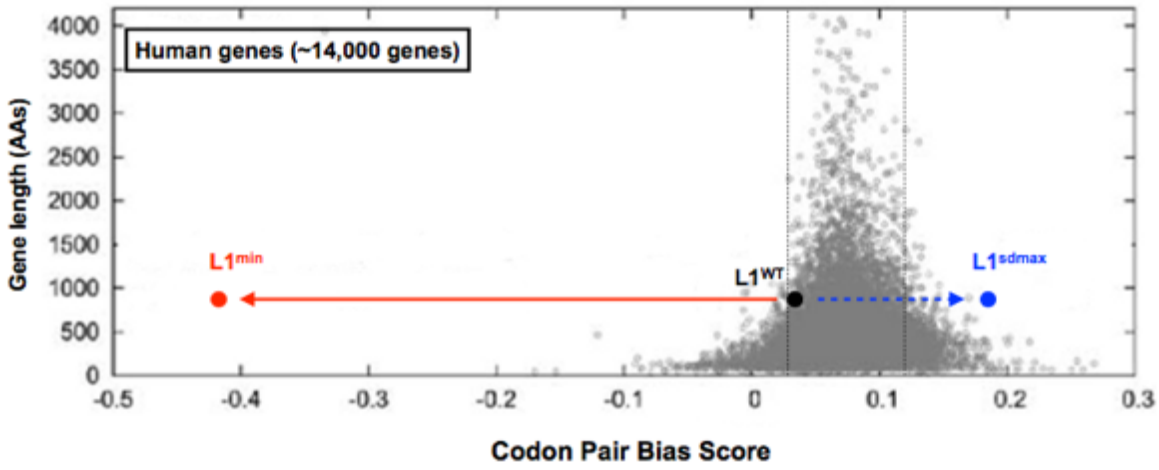


Figure 3.3 Codon pair bias of human genes plotted against their amino acid lengths. The two VSV L1 derivatives and the WT L1 sequence were marked as well.

Chapter 4 Characterization of the L1^{sdmax} virus in tissue culture

4.1 Introduction

The L1^{sdmax} design yielded infectious virus after one blind passage whereas L1^{min} failed to produce a viable virus. Based on the experience in "maximizing" the poliovirus P1 region, the "sdmax" design described herein was not anticipated to negatively affect gene expression (87). Nevertheless, It was noteworthy that the L1^{sdmax} variant did not cause CPE until the second blind passage while WT virus usually led to a significant cell death during the first passage, suggesting that L1^{sdmax} might not grow as competent as its wild-type counterpart.

4.2 Results

4.2.1 Plaque-forming assay

The growth of the L1^{sdmax} virus was analyzed in tissue culture cells. Provided that VSV has a broad cell tropism, plaque formation was examined in a variety of cell types. In all the cell lines I tested, including BHK21 cells, Vero cells, HeLa cells and A549 cells, the plaque phenotypes of L1^{sdmax} revealed slightly smaller sizes in comparison to WT VSV (Fig4.1).

4.2.2 Multi-step growth kinetics

The growth curves of L1^{sdmax} virus was delineated in multiple cell lines. BHK21 cells were infected with WT or L1^{sdmax} virus at a MOI of 0.1 and the virus titers in the culture media were determined by plaque forming assay in Vero cells. Similar to WT VSV, L1^{sdmax} virus also achieved the peak titer between 24 hpi and 48 hpi, which was approximately $5x10^8$ PFU/ml. However, viral replication of this variant, was slightly delayed (Fig4.1A).

Since the L1^{sdmax} virus replicated relatively well in BHK21 cells, it was of interest to characterize this variant further, given that an attenuated virus with ideal growth properties in cell lines would be highly preferred for the purpose of manufacturing. In addition, the ultimate goal of this study was to develop a vaccine candidate applicable to humans and it would be necessary to confirm that the attenuation phenotype of L1^{sdmax} could be reproduced in human cell lines.

Unlike the BHK21 cells that I used for virus recovery and initial characterization, A549 cells (human lung epithelial carcinoma cell line) are also susceptible to VSV infection and possess a robust cellular innate immunity (38, 112). Multiple-step growth kinetics analyzed in A549 cells by inoculation at low MOIs of 0.1 or 0.001 of the viruses showed the maximal titer of $L1^{sdmax}$ virus between 24 and 48 hpi, which was comparable to that of the WT virus. However, similar to the finding in BHK21 cells, before the titer reached plateau, the replication of the mutant virus in A549 cells was also slightly delayed (Fig4.1B).

4.2.3 Immunocytochemistry.

To characterize the growth properties of the $L1^{sdmax}$ virus in more detail, I further performed immunocytochemistry to measure the yield of viral proteins and expand the range of cell types that were examined. HeLa cells were infected at a MOI of 10 and the production of VSV N proteins were visualized by antibody staining. In both WT- and $L1^{sdmax}$-infected cells, the primary infection sites were clearly stained with anti-VSV N antibody about 30 min after infection, an observation suggesting that $L1^{sdmax}$ was not deficient in either host entry or primary transcription/translation initiation. However, at 4 hpi, much less N protein was detected in $L1^{sdmax}$-infected cells, which was consistent with the delayed growth phenotype of $L1^{sdmax}$ virus observed in BHK21 and A549 cells (Fig4.2).

4.2.4 Quantitative RT-PCR to measure viral RNA synthesis

Real-time RT-PCR was performed to quantitate viral RNA products in infected cells. For viral genome quantitation, BHK21 cells were infected with viruses at a MOI of 3 and total RNA was extracted at various time points. Following first strand cDNA synthesis, real-time PCR was conducted to exclusively measure the levels of the full-length viral RNA, most of which represented negative sense genomic RNA, by using a primer pair spanning the intergenic region between the VSV G and L genes (Fig4.3A). As shown in Fig4.3B, the synthesis of the full-length viral genome was significantly delayed in $L1^{sdmax}$-infected cells. With the time increased, it became more revealing that $L1^{sdmax}$ virus did not replicated as well as WT VSV.

In a similar manner, real-time RT-PCR was performed to analyze viral gene transcription in infected cells. BHK21 cells were infected with viruses at a MOI of 3 and total RNA was extracted at the indicated time points. The abundance of viral gene transcripts was determined by

using primers mapping to the individual viral gene ORF and then normalized to the mRNA level of the GAPDH gene. As shown in Fig4.5A, the synthesis of VSV N and L gene transcripts was significantly reduced in L1^{sdmax}-infected cells compared to the WT-infected ones.

4.2.5 Western Blot to measure viral protein synthesis

Viral protein synthesis was measured in BHK21 cells by Western Blot analysis. BHK21 cells were infected with WT or L1^{sdmax} virus at a MOI of 3. At multiple time points whole cell lysates were harvested and probed with anti-VSV antibodies. Rabbit polyclonal antiserum against the C-terminus of the VSV L protein was used to recognize the full-length viral polymerase in infected cells while anti-VSV G staining was performed as an internal control(71). Consistent with the findings that the yields of infectious particles (shown by growth curve) and full-length viral genomes (shown by real-time RT-PCR) decreased, the glycoprotein (G) synthesis was also modestly reduced in L1^{sdmax}-infected cells. The band intensities were analyzed and quantitated using ImageJ. At 6 hpi, the level of G protein in WT-infected cells was 2.5 fold of that in L1^{sdmax} -infected ones. At 9 hpi, the ratio turned to 2.6 (Fig 4.4).

Interestingly, the reduction of L protein synthesis was greater than that of the G protein in L1^{sdmax}-infected cells. At 6 hpi, the level of L protein in WT-infected cells was 8-fold more than that of L1^{sdmax}-infected ones. At 9 hpi, the ratio changed to 15.8: 1.4. Since the L protein is the catalytic subunit of the VSV RdRP, its function presumably determines the kinetics of virus replication. It is therefore possible that the substantial enrichment of the overrepresented codon pairs in the L1^{sdmax} sequence might somehow have a negative effect upon gene expression. Strikingly, this observation was contradictory to the effect of “codon pair maximizing” upon poliovirus replication (87).

4.2.6 Northern Blot to measure the level of viral L mRNA

A reduction of protein synthesis could either result from a less efficient translation (less protein produced per mRNA) or an unstable mRNA, which leads to less available transcripts for translation. To distinguish between the two possibilities, northern blot analysis was performed to directly measure the level of full-length viral L gene transcript. BHK21 cells were infected at a MOI of 3 with either WT or L1^{sdmax} virus and total RNA was extracted at multiple time points.

As shown in Fig 4.5B, at 3, 6 and 9 hpi the viral L mRNA was significantly less abundant in $L1^{sdmax}$ virus-infected cells than the WT-infected ones. At later time points in WT-infected cells (6 and 9 hpi), a band with a higher molecular weight could be observed, which presumably represented the full-length antigenomic viral RNA (positive sense, replication intermediate). The band intensities were quantitated using ImageJ. At 6 hpi, the ratio of WT and $L1^{sdmax}$ was 1: 0.34. At 9 hpi, the ratio turned to be 2.32: 0.82 (approximately 1:0.35). Although the quantitation could be arbitrary, it offered an estimation of the relative amount of full-length L mRNA in the infected cells.

4.3 Discussion

By using a unique recoding strategy of codon pair "sdmax", I designed, constructed and recovered a VSV derivative $L1^{sdmax}$ that contained 858 synonymous mutations in its L gene ORF. With the existing synonymous codons extensively shuffled to create new "favored" codon pairs, the L1 segment of the $L1^{sdmax}$ virus has a significantly higher codon pair bias score than its WT counterpart.

It was unexpected that "sdmax" recoding resulted in an attenuation phenotype, provided that the codon pairing in "sdmax" sequence has presumably been designed to yield mRNA favorable for translation. However, the delayed occurrence of viral CPE during the $L1^{sdmax}$ recovery led us to characterize the growth phenotype of this virus variant in tissue culture in more detail. Indeed, $L1^{sdmax}$ virus formed slightly smaller plaques in a variety of cell lines. Plaque forming assay, which was devised by Dulbecco in 1952, localizes the destruction of infected cells and the size of plaque positively correlates with the rate of virus replication (113). Although VSV itself is well known for high sensitivity to the cellular innate immunity, the plaque phenotype of $L1^{sdmax}$ is independent of the cell type it infects. The two cell lines I used for plaque forming assay, A549 and HeLa cells, both retain an intact signaling transduction network to mount an innate immune response against virus infection and are both responsive to Type I interferon pre-treatment (38, 112, 114). Nevertheless, the plaque phenotype of $L1^{sdmax}$ is not more pronounced in these two cell lines relative to that in BHK21 cells (a cell line well known for its blunted innate immunity), suggesting that the "sdmax" virus variant is not more sensitive to the immune response.

Multi-step growth kinetics was measured to assess virus replication in this study. Low MOIs were used for the assay because VSV is notorious for the rapid generation of defective interfering (DI) particles when a high multiplicity of infection is used for infection(115). The infectious viruses that were released to the supernatants increased exponentially from 3 hpi to 9 hpi, and reached the plateau by 24 hpi. In BHK21 cells, at the early time points, the $L1^{sdmax}$ virus replicated slightly slower than WT VSV. However, at 48 hpi, when the full CPE occurred, the end-point virus titer of $L1^{sdmax}$ was close to that of WT virus, indicating that in tissue culture $L1^{sdmax}$ could produce almost the same number of progenies as WT did by fully exploiting the host cell resources. Similarly, in A549 cells the replication of $L1^{sdmax}$ virus was also delayed.

Taken together, all these data further support the notion that L1^{sdmax} replicates less competent than the WT and this attenuation phenotype is independent of the status of cellular immunity.

The next question to address was whether L1^{sdmax} was deficient in host cell attachment and entry, considering that the reduction of virus titers in L1^{sdmax}-infected cells appeared always during the early stage of virus infection. As shown in Fig 4.2, the positive staining of VSV N protein at 0.5 hpi indicated the location of virus internalization. The infected cells were extensively washed with PBS to remove any unattached viruses, before they were subject to immunostaining. When the cells were infected with either WT or L1^{sdmax} virus at the same MOI, equal number of virus particles that had been internalized by the cells were identified at the early time point. Therefore it is reasonable to assume that L1^{sdmax} is not deficient in host cell entry or the initiation of primary transcription/translation. It also suggests that equal amounts of infectious virus particles were added as input at the beginning of the experiment. At 2 and 4 hpi, in WT-infected cells, a robust expression of viral N protein was found while significantly less N protein was detected in L1^{sdmax}-infected ones. It should be noted that an exposure time of 200 ms (instead of 1500 ms at 0.5 hpi) at the later time points was used to adjust the fluorescence signal intensity.

Real-time RT-PCR measures the levels of viral RNA products in the infected cells. In order to exclusively quantitate the full-length viral RNA species, a primer pair that spans the intergenic junction between the G and L genes was deliberately designed. In VSV the five viral genes are transcribed into a series of discrete monocistronic mRNAs and none of these viral transcripts contain the intergenic region. Thus only the full-length viral RNAs (either positive sense anti-genome or negative sense genome) are expected to be amplified and quantitated in this experimental setting. At the same time, the positive sense anti-genome, which functions as the virus replication intermediate, should be much less abundant than the negative sense viral genome in the infected cells, due to so-called “asymmetric RNA replication”. Therefore, viral genome replication could be easily and precisely quantitated in this manner.

When the cells were infected at the same MOI with either WT or L1^{sdmax} virus, a similar level of full-length viral RNA was found in the cells at both time 0 and 0.5 hpi, consistent with the finding in immunocytochemistry assay that comparable amount of virus particles was identified if the cells were infected at the same MOI. Therefore, it can be assumed that the specific infectivity of the L1^{sdmax} virus, which represents the number of virus particles per

infectious unit, should be close to the one of WT VSV. However, with time increased, it became more apparent that L1^{sdmax} was producing fewer viral genomes than the WT virus. At 3 hpi, 6 hpi and 9 hpi, 4-6 fold less viral genomes were identified in L1^{sdmax}-infected cells.

The same batch of RNA samples was further analyzed for the quantitation of individual viral gene transcript. As shown in Fig 4.5A, WT virus was producing more viral mRNA (both N and L gene transcripts) than L1^{sdmax}. Interestingly, during the infection of L1^{sdmax} virus, the same kinetics of mRNA accumulation was found for both the L and N genes (ranging from 1 hpi to 6 hpi). Since all the viral transcripts are sequentially synthesized by the same amount of viral RNA polymerase, it is reasoned that the transcription of the "sdmax" sequence should not be negatively affected by the substantial nucleotide substitution.

To make the L1^{sdmax} virus, more than 800 synonymous mutations have been incorporated into the L gene ORF, which could significantly alter the local secondary structure of the viral RNA and affect polymerase processivity. The truncated or pre-terminated forms of viral transcripts are not readily identified by PCR since the amplicon for real-time PCR assay is usually less than 150 bp in length. Thereby, Northern Blot analysis is advantageous here given that the full-length viral transcripts can be directly quantitated, enabling us to rule out the possibility that portion of the L transcripts in L1^{sdmax}-infected cells were not transcribed to full length. As expected, WT virus produced more (~3 fold) intact L mRNA than L1^{sdmax}, which correlates very well with the result obtained by real-time RT-PCR (cross-compare Fig 4.3B and Fig 4.5B). More importantly, no RNA species that migrated faster than the full-length L transcript could be identified in either WT or L1^{sdmax}-infected samples, suggesting that "sdmax" recoding does not result in a higher tendency to pre-terminate transcription.

Finally, as shown in Fig 4.4, less viral protein was present in L1^{sdmax}-infected cells as expected. The quantitation of band intensities revealed a modest reduction (~2.5 fold) of G protein synthesis, which probably resulted from the delayed growth of L1^{sdmax} virus in cell culture. Interestingly, a greater reduction of L protein synthesis was observed when the same sample was probed with anti-L antibody, suggesting that the substantial enrichment of overrepresented codon pairs (in L1^{sdmax} sequence) somehow leads to a deficiency in protein synthesis, a phenomenon not expected.

4.4 Materials and methods

Multi-step growth kinetics

BHK21 and A549 cells were cultured on 6-well plates with DMEM/10% FBS until 80% confluent. The cells were infected with either WT or $L1^{sdmax}$ virus at a MOI of 0.1 or 0.001. After one hour absorption at 37°C, the virus inoculum was removed and the cells were supplemented with DMEM containing 2% FBS. At the indicated time points, the culture media was collected and virus titers in the supernatants were then determined by plaque forming assay in Vero cells.

Immunocytochemistry

HeLa cells were infected with WT or $L1^{sdmax}$ virus at a MOI of 10. At the indicated time points (0.5 hpi, 2 hpi and 4 hpi), the media was removed and the cells were washed extensively with pre-warmed PBS for 4 times. Cells were fixed with 3.7% Paraformaldehyde (PFA) for 30 minutes and permeabilized with 0.2% Saponin/ 2% BSA/ PBS for 15 minutes.

For VSV N protein staining, cells were stained with monoclonal antibody 10G4 (purchased from KeraFAST, Inc) at a dilution of 1:300 followed by a 1:500 dilution of a secondary anti-mouse antibody conjugated to AlexaFluor-594 (Invitrogen Molecular Probes). The stained cells were then washed three more times and imaged using a deconvolution microscope.

Real-time RT-PCR to measure viral genome replication and viral gene transcription

BHK21 cells were infected with viruses at a MOI of 3. At the indicated time points, total RNA was extracted and reverse transcribed by random hexamer priming using SuperScript III (purchased from Invitrogen). Quantitative PCR was carried out with a primer pair spanning the intergenic region between the VSV G and L gene. For the quantitation of individual viral mRNA, total RNA extracted from infected cells was reverse transcribed with Oligo $(dT)_{20}$. Primers were designed accordingly to target viral genes N and L (see Table 4.1).

Western Blot analysis to measure viral protein synthesis

BHK21 cells were infected with WT or L1^{sdmax} virus at a MOI of 3. At the indicated time points, the cells were washed with PBS for once and then resuspended in RIPA lysis buffer on ice for 30 minutes. The suspension was transferred to a new micro-centrifuge tube and spun for 10 minutes at 12,000x g to remove cell debris.

The VSV proteins were then separated on 7.5% SDS-PAGE gel and transferred to a nitrocellulose membrane (Whatman, Minnetonka, MN). Blots were blocked for 1 hour in 5% dry milk powder dissolved in 1x Tris-buffered saline (TBS, pH7.4). After being blocked, blots were then washed three times with 1x TBS containing 0.1% Tween-20 and probed with anti-VSV G antibody [P5D4] (purchased from Abcam, Inc, Cambridge, MA) at a dilution of 1:10000 for the detection of G protein.

Similarly, rabbit polyclonal antisera raised against the VSV L protein was used to recognize full-length L polymerase protein in the cell lysates. The blots were probed with anti-VSV L antisera at a dilution of 1:1000 at 4°C overnight, followed by the secondary antibody probing.

Northern Blot analysis to measure viral mRNA synthesis

RNA probe (negative sense) that recognizes the VSV L mRNA was generated using the DIG RNA labeling kit (purchased from Roche Diagnostics, USA). Briefly, the DNA template that contained T7 promoter at the 5'end and the first 400nt of VSV L gene was amplified and PCR purified. 1 μg of the purified PCR product then served as a template for *in vitro* transcription, during which the DIG-labeled UTP was added to the reaction. Before the Northern Blot experiments, the labeling efficiency of the RNA probe was determined as follows. Briefly, a series of dilutions of DIG-labeled RNA was applied to a small strip of positively charged nylon membrane. The nylon membrane was then subjected to immunological detection with the anti-DIG-AP and CDP-star solution.

The total RNA samples collected from WT or L1^{sdmax}-infected cells were separated on formaldehyde agarose gels. Before running, the samples was mixed with 2-3 volume of RNA loading buffer and denatured at 65°C for 10 min. After chilling on ice for 1 min, approximately 1 μg total RNA was loaded per lane. After running, the gel was stained with ethidium bromide to examine the quality as well as the integrity of the RNA samples.

After electrophoresis, the gel was rinsed with 20x SSC for 2x 15 min. The RNA was subsequently transferred from the gel to the membrane by gravity at room temperature overnight, followed by the UV cross-linking on the wet membrane without prior washing.

The nylon membrane was pre-hybridized with DIG Easy Hyb for 30 min at 68°C with gentle agitation. At the same time, the DIG-labeled RNA probe was denatured by boiling for 5 min and rapid cooling in ice/water, and mixed with pre-warmed DIG Easy Hyb. The membrane was then incubated with probe/hybridization mixture at 68°C overnight, with gentle agitation.

After probe binding, the membrane was washed 2x 5 min in 2x SSC supplemented with 0.1% SDS at room temperature under constant agitation, followed by 2x 15 min in 0.1x SSC/0.1%SDS at 68°C. After stringent washes, the membrane was rinsed with washing buffer for 3 min, followed by a 30 min blocking at room temperature, 30 min antibody binding and 2x 15 min washing. The membrane was equilibrated in detection buffer for 5 min, then placed in a development folder and soaked with CDP-Star solution (provided with the RNA labeling kit). The membrane was exposed to X-ray film for at least 15 min and developed.

4.5 Figures and Tables

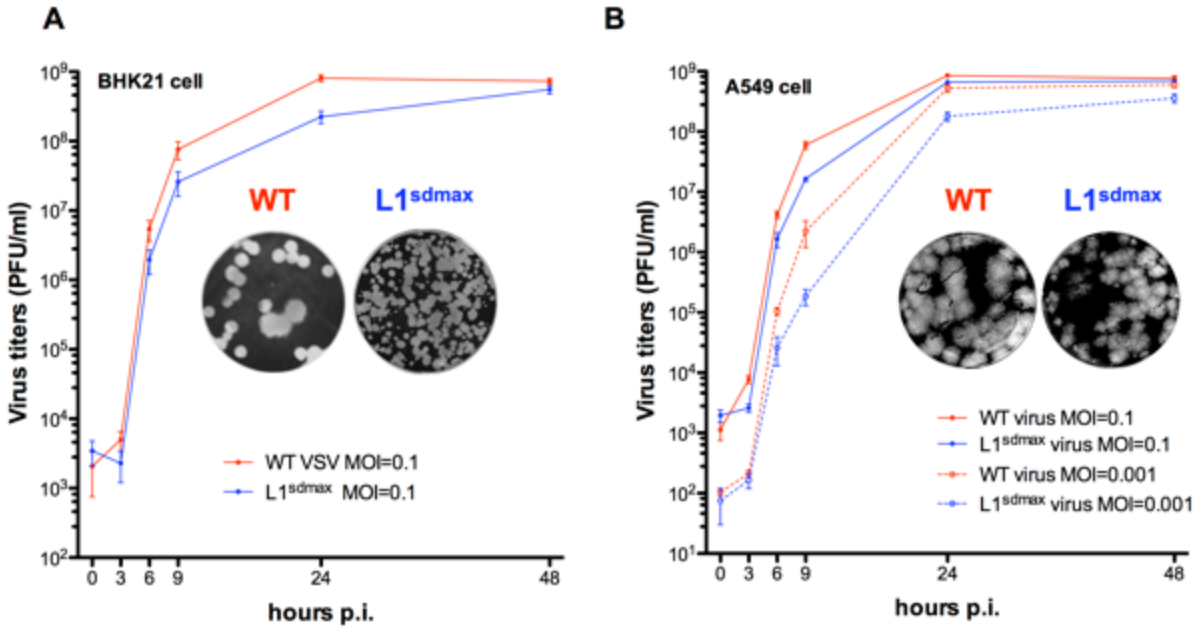


Figure 4.1 Plaque morphologies and multi-step growth kinetics of WT and L1^{sdmax} virus. Plaques formed in BHK21 cells (A) were developed at 18 hpi while those in A549 cells (B) were developed at 48 hpi, as described in Materials and Methods. For the measurement of growth kinetics, at the indicated time points, the culture media was collected and virus titers were determined by plaque forming assays in Vero cells.

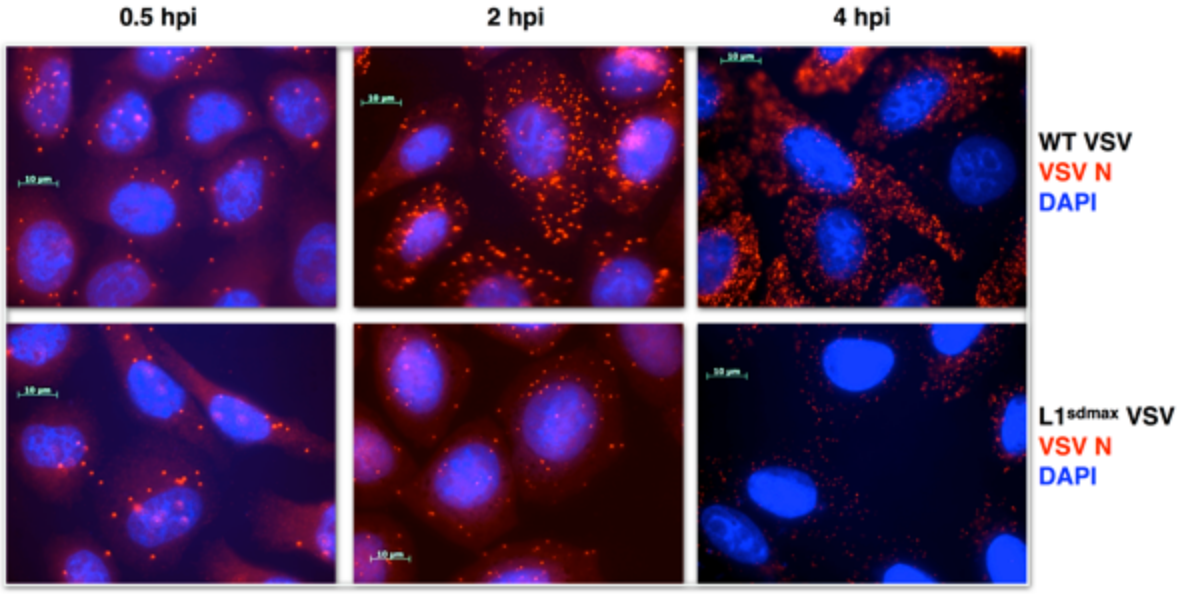


Figure 4.2 Immunocytochemistry of WT/L1^{sdmax}-infected cells.
HeLa cells were infected with WT or L1^{sdmax} virus at a MOI of 10 and the synthesis of VSV N protein was examined by fluorescence microscopy at multiple time points.

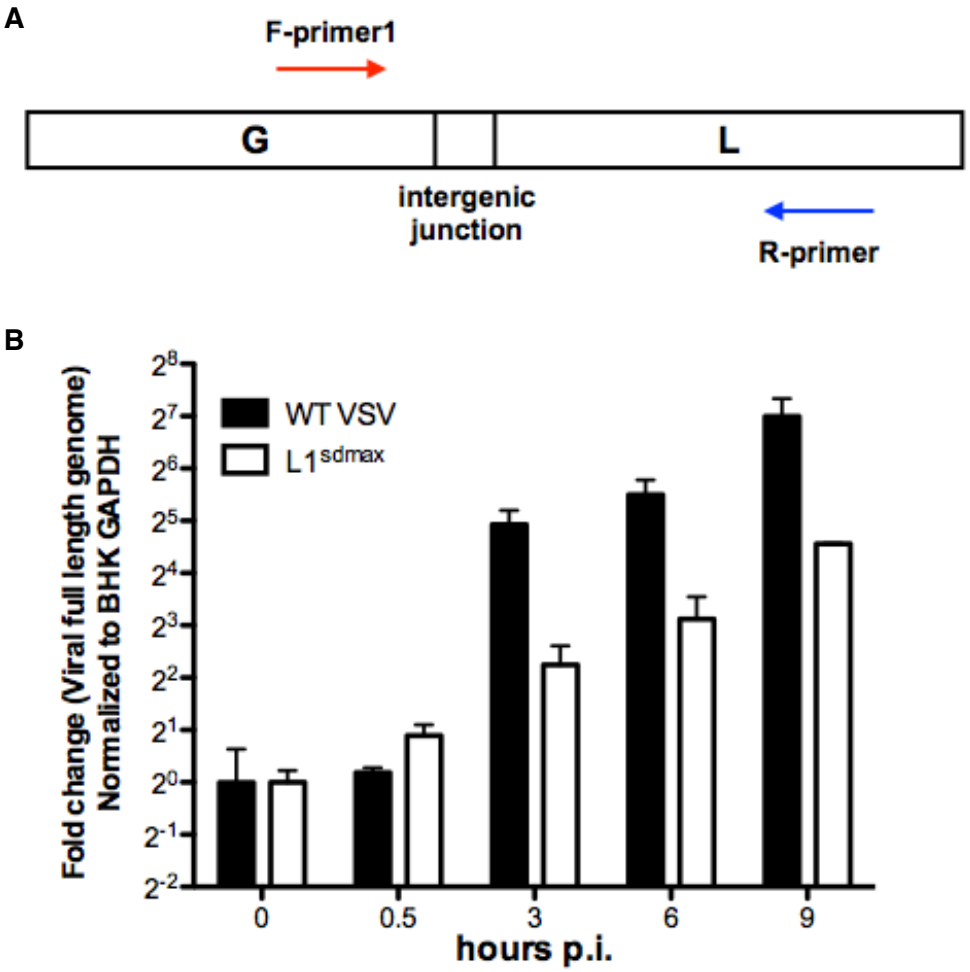


Figure 4.3 Real-time RT-PCR to measure virus replication.
(A) A schematic diagram to show the position of the primer pair that exclusively amplifies full-length viral genome. (B) BHK21 cells were infected with viruses at a MOI of 3 and total RNA was extracted at the indicated time points. Real-time PCR was conducted to exclusively quantitate the full-length viral RNA (mainly negative sense genomic RNA), as described in Materials and Methods.

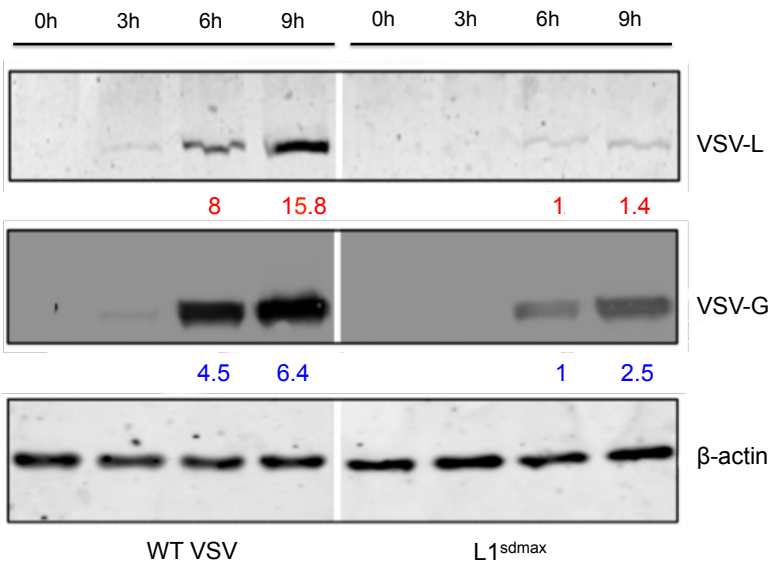


Figure 4.4 Western Blot analysis to measure viral protein synthesis in infected cells.
BHK21 cells were infected at a MOI of 3 and the expression level of viral proteins were determined by Western Blot at the indicated time points. The quantitation of VSV L and G proteins was analyzed with ImageJ and expressed as the fold increases over the band intensities identified in L1^{sdmax}-infected samples at 6 hpi.

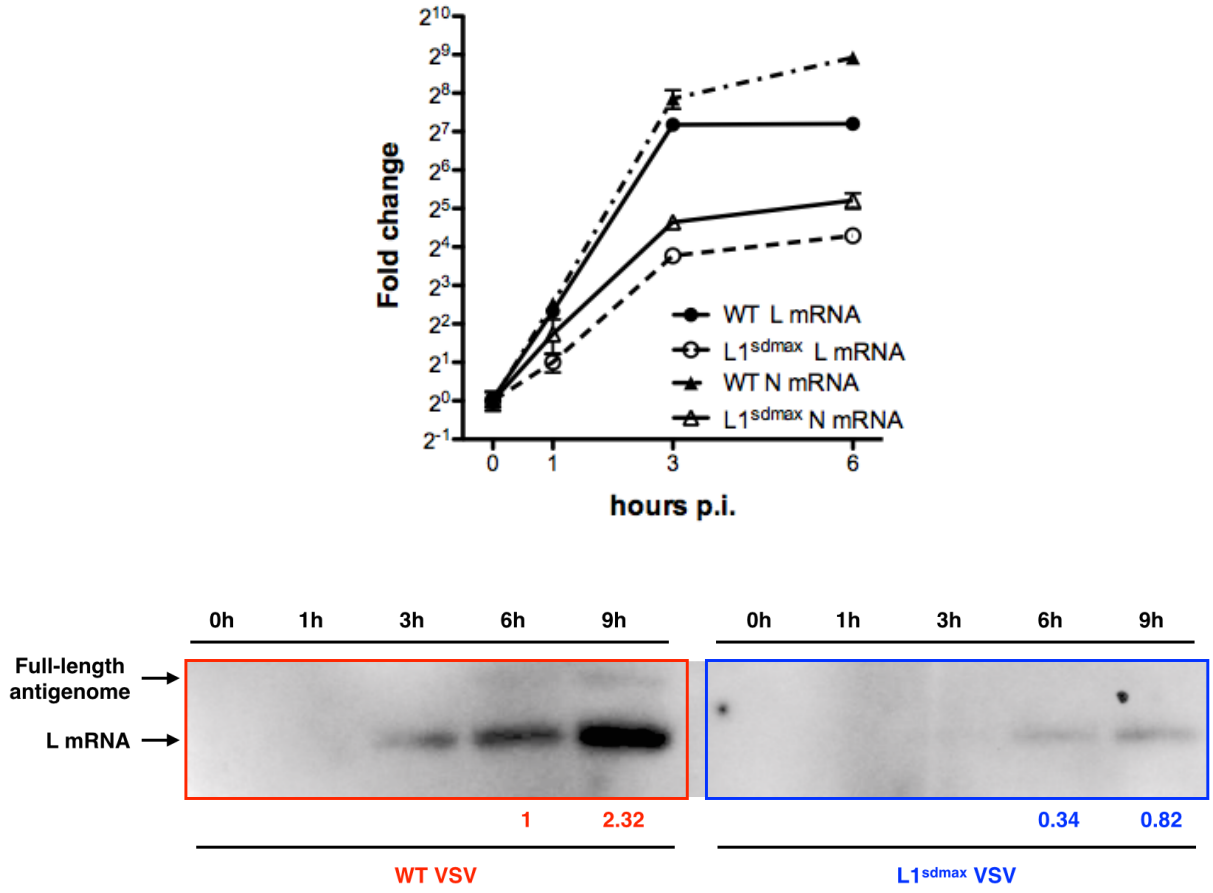


Figure 4.5 QRT-PCR and Northern Blot to examine viral RNA synthesis in infected cells. (A) Quantitative RT-PCR to measure the level of viral N and L gene transcripts in infected cells. BHK21 cells were infected with viruses at a MOI of 10 and total RNA was extracted at the indicated time points. The levels of viral mRNAs were determined by qRT-PCR and normalized to the mRNA level of the GAPDH gene in BHK21 cells. (B) Northern Blot to quantitate the full-length L gene transcript in infected BHK21 cells. The relative amount of the viral L gene mRNA was analyzed with ImageJ and expressed as the fold changes over the band intensity identified in WT-infected samples at 6 hpi.

	Forward primer	Reverse primer
Primer for VSV genome quantification	5' TTGTGATATCATGCTCAAAGAGGCC 3'	5' AAGAACTCCATCCCAGTTCTTACTATCC 3'
VSV L primer	5' TTCAATGAAGATGACTATGCCACAAGAG 3'	5' AAGAACTCCATCCCAGTTCTTACTATCC 3'
VSV N primer	5' ATGACAAATGGTTGCCTTTGTATCTACTT 3'	5' ACGACCTTCTGGCACAAGAGGT 3'
BHK21 GAPDH primer	5' TCTACTGGCGTCTTCACCACCATGGAG 3'	5'ATGCTAATACGACTCACTATAGGGTGGGCCCTCAGAT GCCTGCTTCACCAC 3'

Table 4.1 A summary of the primers used for quantitative RT-PCR.

Chapter 5 The molecular mechanism underlying codon pair bias-based virus attenuation

5.1 Introduction

The unexpected attenuation phenotype of $L1^{sdmax}$ is intriguing. In contrast to the "min" strategy, the "sdmax" design leads to a substantial enrichment of the "favored" codon pairs, the ones that are overrepresented in the host genome and are presumably designed to yield mRNA favorable for protein synthesis. The recombinant poliovirus $P1^{max}$ that contained extra overrepresented codon pairs in part of its genome replicated identical to the WT virus and demonstrated a modest translation augmentation(87). According to the working model of codon pair bias, which was proposed by us and others(87, 92, 93, 95, 109, 116), the codon pairs that are overrepresented in a given genome are likely to facilitate translation elongation, which is frequently beneficial to gene expression.

The "sdmax" design was initially included in this study to scan the viral genome in search of RNA sequence signatures that are critical for virus replication(94, 97). Both the $L1^{sdmax}$ and $L1^{min}$ sequences contain hundreds of synonymous mutations, of which on average one out of every three to four nucleotide is altered. Given the fact that $L1^{sdmax}$ is just slightly attenuated in tissue culture, it is less likely that the L1 region contains certain unidentified nucleotide-based, cis-acting elements that are longer than 3-4 nt. Therefore, it seems that the "sdmax" recoding might negatively affect the expression of the VSV L gene, which came to us as a surprise.

The delayed replication of $L1^{sdmax}$ virus led to a reduced yield of all the viral products, which made it difficult to explain the decreased level of L protein synthesis by simply comparing the yields of viral protein to those of viral mRNA. Therefore, to circumvent this problem, a transient expression assay was developed, to examine viral protein synthesis without the influence of the virus replication cycle.

Recent studies have shown that clustered rare codons are frequently distributed between defined protein domains, especially for large proteins containing multiple functional domains, in order to facilitate a stepwise assembly of the folding units (117-123). The distribution pattern of the "slow" codons collectively reduces the read-through activity of the ribosomes between domains, which turns out to be desirable for proper protein folding. Interestingly, the WT L1

sequence shows a relatively low codon pair score (CPS=0.027, lower than the scores of most human genes that are in the same length, as shown in Fig 3.3), implicating that there might be a compromise between translation accuracy and folding efficiency.

5.2 Results

5.2.1 Construction of the L expression plasmids

Two L gene expression plasmids were constructed by replacing the cognate WT L ORF of the pL vector with the synthetic L derivatives, $L1^{sdmax}$ and $L1^{min}$, respectively. These two plasmids were designated as $pL^{L1sdmax}$ and pL^{L1min}, and used for the subsequent transient expression assay (Fig5.1A).

The transcription of the recoded L variants is under the control of the cognate T7 promoter of the pL plasmid. For transient expression, BHK21 cells were infected with vTF7-3, the recombinant vaccinia virus that had been used for VSV recovery. The infected cells were then co-transfected with the pP and pL derivatives. At the indicated time points, whole cell lysates and total RNA were harvested and the yields of viral products were measured. Particularly, in this experimental setting, the L gene derivatives were transcribed by T7 polymerase provided *in trans*. Unlike the authentic virus infection, the transcription of the L-specific mRNA was no longer influenced by the cycle of virus replication (Fig5.1B).

The full-length L protein could barely be detected in pL^{L1min}-transfected cells whereas the transfection of the WT expression plasmid led to a robust L protein expression at 24 hpi (Fig5.2B). This finding strongly supports our initial assumption that codon pair deoptimization directly influenced gene expression, no matter what mechanism (or mechanisms) is/are involved. It also provides the explanation with regard to the lethal phenotype of the $L1^{min}$ design.

On the other hand, relative to its WT counterpart, the accumulation of the L protein in $pL^{L1sdmax}$-transfected cells was also significantly delayed (Fig5.2A). In addition, the reduction of L gene expression did not result from the changes at the transcription level, since equal amounts of L transcripts were found in pL^{WT}-and $pL^{L1sdmax}$-transfected cells (Fig5.3B).

5.2.2 *In vitro* transcription of the $L1^{sdmax}$ sequence

In the following *in vitro* transcription assay, it was found that $L1^{sdmax}$ was transcribed by T7 polymerase just as well as WT sequence: equal amounts of pL^{WT} and $pL^{L1sdmax}$ were linearized and used as templates for *in vitro* transcription. As shown in Fig5.3A, similar levels of the transcription products were identified. It was not surprising since RNA polymerase is not anticipated to have sequence preference.

5.2.3 Virus replication in cells concomitantly treated with 17AAG

Previous studies have shown that the VSV L protein is not stable *per se.* When expressed alone, L protein exhibits a short half-life in cells (124). A host chaperone molecule Heat Shock protein 90 (Hsp90) is essential for the proper folding of the newly synthesized L protein. Pharmaceutical deprivation of Hsp90 gives rise to a selective degradation of L protein in infected cells in a proteasome-dependent manner (125). For the $L1^{sdmax}$ virus, an exquisite sensitivity to the Hsp90 inhibition was observed, suggesting that an enhanced engagement of host chaperones might be required for the $L1^{sdmax}$ virus replication (Fig5.6).

5.2.4 ^{35}S trans-labeling of newly synthesized viral proteins

In an attempt to test the hypothesis that "sdmax" recoding leads to a less stable protein, I measured the half-life of the newly synthesized VSV proteins during authentic virus infection by pulse-chase analysis(102). HeLa cells were infected with either WT or $L1^{sdmax}$ virus at a MOI of 10. At 4 hpi they were labeled with ^{35}S-Met/Cys for 30 min and the radiolabeled cells were then "chased" for additional periods. Whole cell lysates were harvested at the indicated time points and further analyzed by SDS-PAGE and phosphorimaging. As shown in Fig5.5, at 4 hpi VSV proteins were synthesized virtually exclusively, as evidenced by the appearance of four major bands (L, G, co-migrating of N and P, M). The newly synthesized L proteins, as well as the other viral proteins expressed in WT-infected cells, were stable throughout the experiment (up to 8 hpi). In contrast, in the cells infected with $L1^{sdmax}$ virus, at later time points there was a slight but gradual decrease of the labeled L proteins, suggesting that the L proteins expressed in $L1^{sdmax}$-infected cells might have an altered stability even though they were assembled with the identical amino acid sequence.

5.3 Discussion

5.3.1 Impact of codon pair deoptimization on RNA stability and protein translation

In this section I attempted to unveil how the two opposite large-scale recoding strategies influence gene expression. As predicted, the $L1^{min}$ sequence only yielded trace amounts of full-length L proteins after transient expression in tissue culture cells. This observation validated our previous findings in codon pair "deoptimized" poliovirus and influenza A virus, yet the underlying mechanism(s) still remained inconclusive (87, 92, 93, 95, 109). It might be difficult for the ribosome to read through a string of "rare" codon pairs, which essentially rendered the corresponding mRNA less efficient in translation (less protein produced per mRNA). Another explanation to account for the insufficient expression could be the instability of the "deoptimized" transcripts, although it still remained unclear in which cellular compartment this destabilization primarily occurred. Moreover, the creation of rare codon pairs, by rearranging existing codons, inadvertently leads to an enrichment of UpA and CpG dinucleotides in the mRNA sequences. For instance, the $L1^{min}$, but not the $L1^{sdmax}$ design, contains much more U_3A_1 and C_3G_1 dinucleotides than the WT sequence (herein the number of U_3A_1 or UpA dinucleotides accounts for all the codon pairs XXU_3-A_1XX in the L gene open-reading frame). In general, the frequency of UpA and CpG dinucleotides is suppressed in the ORFs of nearly all the mammalian viruses, first reported for the poliovirus genome and later observed in other viral genomes as well(126, 127). Recent studies suggested that the dinucleotides composition in RNA virus genomes influences virus replication, possibly because of an unclear host-pathogen interaction(128). Although I can not simply exclude the possibility that strengthened host recognition of U_3A_1/C_3G_1 partially contributed to the attenuation phenotype of the codon pair-deoptimized viruses, in this study the reverse genetics of $L1^{min}$ was exclusively conducted in BHK21 cells, a cell line well known for its deficiency in innate immunity response (129). Hence the nonviable phenotype of $L1^{min}$ design cannot be simply attributed to an enhanced engagement of host immune response.

As mentioned, the presence of CpG and UpA dinucleotides is suppressed in the gene ORFs for reasons not well understood. One hypothesis is that the underrepresentation of UpA is related to the instability of mRNA molecules. This is supported by an observation that in

eukaryotic cells a reduction of UpA frequency in coding region results in a prolonged mRNA half-life accompanied with increased protein expression (130). In a study with hepatitis C virus, a higher frequency of UpA/UpU dinucleotides in the HCV genome enhances the susceptibility to RNase L cleavage(131). The existence of another mRNA quality control system adds extra layer of complexity. Message RNAs with translation stalling, possibly caused by codon pair deoptimization, are subject to endonucleolytic cleavage termed as "no-go decay" (132). Taken together, the poor protein expression observed in codon pair deoptimization may result from a slow translation elongation rate, which might further destabilize the corresponding mRNA. Perhaps, frequently stalled translation elongation and rapid mRNA degradation collectively contribute to the phenotype of the deoptimized genes.

5.3.2 Impact of codon pair "scramble-max" on protein translation and co-translational folding

The most striking phenotype in this study is the attenuation of $L1^{sdmax}$ that obviously results from an enrichment of the overrepresented codon pairs in the L coding sequence. The attenuation of $L1^{sdmax}$ is not host range restricted or immunity status-dependent. Considering that the transcriptional regulatory sequences that have been identified in the untranslated regions of the L gene were never altered, it was less likely that the $L1^{sdmax}$ sequence was inadequately transcribed by the viral RNA polymerase. The *in vitro* transcription assay also supported the notion, although in that experiment the full-length L gene was transcribed by T7 RNA polymerase instead of the viral RdRP. Indeed, the "sdmax" gene was less efficiently translated than the WT sequence and this is not likely due to a destabilized L mRNA because 1) a high level of L mRNA was detected when the full-length $L1^{sdmax}$ sequence was placed under the control of T7 promoter and transfected into the cells expressing T7 polymerase; 2) during the $L1^{sdmax}$ virus infection, the same kinetics of the mRNA accumulation was found for both L and N genes. Taken together, the expression deficiency of the $L1^{sdmax}$ sequence does not seem to occur at the transcription level.

Our previous experiments in which we "maximized" the poliovirus P1 region (increase of codon pairs with positive codon pair score) suggested that this genetic alteration enhanced the expression of the corresponding viral genes without increasing virulence (87). The P1 region in

the poliovirus genome is almost as long as the L1 region described here. Thus, the substantial accumulation of the "favored" codon pairs in $L1^{sdmax}$ sequence may be sufficient to accelerate the translation elongation.

In contrast to the arrangement of transcription machineries shared by mammalian cells and many other RNA viruses, VSV assigns all the catalytic activities associated with RNA synthesis to a single large protein L (~250kDa), which inevitably contains multiple independently functioning domains (110, 133). The VSV L protein forms a highly organized tertiary structure, which is essential for the coordination of its different RNA synthesizing activities (51). It would not be surprising that the virus has evolved certain mechanisms to ensure proper folding of the newly synthesized polymerase protein.

As mentioned before, the L protein folding highly depends on the activity of host chaperone Hsp90(125). Provided that enhanced chaperone engagement might be required for $L1^{sdmax}$ virus, I tested the hypothesis that an altered translation elongation rate (presumably caused by the "sdmax" recoding) interferes with the protein folding of VSV L. In HeLa cells, at physiological temperature the $L1^{sdmax}$ virus formed smaller plaques than WT did. However, the plaque phenotype of $L1^{sdmax}$ became indistinguishable from its WT counterpart at 33°C. The deleterious effect of the "sdmax" recoding seems to disappear if the plaques are allowed to form at a lower temperature that leads to a global slow-down of the ribosome translocation rate. Furthermore, the L protein synthesized by the $L1^{sdmax}$ virus showed a reduction of protein half-life, which was approximately 90 min, but the L protein expressed in WT-infected cells remained stable up to 210 min after pulse-labeling. An early study conducted in a bacterial system suggested that codon pair utilization influenced the folding or other functions of nascent polypeptide chains by altering translational efficiency, although only a small number of codon pairs were included and analyzed at that time (116). Collectively, these findings strongly suggest that the folding efficiency of a large multi-domain protein, like the VSV L protein described here, might be severely perturbed by a "faster-translating" region that contains hundreds of "good" codon pairs.

Several recent studies performed in different biological systems support the hypothesis that translational kinetics influences co-translational folding. A pioneer study published in 2007 showed that a synonymous single-nucleotide polymorphism (SNP) identified in the human MDR1 gene, which should not alter peptide sequence, led to a different protein conformation and

altered enzymatic activity (134). It was hypothesized that the presence of a rare codon, marked by the SNP, affected the timing of co-translational folding and thereby altered the protein structure. Similarly, a bacterial multi-domain protein SufI was later identified to contain a natural "slow-translating" region, which resulted in a "transient ribosomal arrest". If translation elongation was artificially accelerated by providing excess amount of the corresponding rare tRNAs or mutating rare codons into the frequently used ones, the folding efficiency of the SufI protein was severely perturbed (122). The biological significance of the non-optimal codon usage was also demonstrated in *Neurospora*, a fungi that had been extensively studied as a model organism. The open-reading frame of the *Neurospora* frequency protein gene (FRQ), which was essential for the circadian clock function, appeared to have a high level of non-optimal codons (123). The reversal of codon deoptimization in this gene abolished its enzymatic activity, stability as well as the phosphorylation profile. Last but not the least, a genome-wide analysis of ten yeast species unveiled a distribution pattern of non-optimal codons in yeast genes, which often formed clusters, associated with the secondary structure of the translated polypeptides. It was therefore proposed that the rare codons played a critical role at the "sensitive" positions in regulating the rhythm of translation elongation (121). Taken together, these studies provide compelling evidence that genome composition, such as codon bias, affects protein folding by controlling translation elongation rate. Although currently direct evidence is lacking that codon pair bias influences translation in a similar manner, my findings are consistent with the proposed model that the artificial "sdmax" sequence may alter the "rhythm" of protein translation. Altogether, enhanced ribosome reading-through might not always be beneficial for gene expression.

5.4 Materials and methods

Construction of the L expression plasmids

The corresponding region of the WT L gene in the expression plasmid pL was replaced by two recoded derivatives $L1^{sdmax}$ and $L1^{min}$, and the two new constructs were designated as $pL^{L1sdmax}$ and pL^{L1min}, respectively.

Transient expression assay

70-80% confluent BHK21 cells were infected with vTF7-3 at a MOI of 10 for 1hr and then co-transfected with the pP and pL derivatives. The VSV P protein was co-introduced to stabilize the newly synthesized L protein(124). To measure L gene transcription and translation, whole cell lysates and total RNA were harvested at 6 hr, 12 hr and 24 hr, before the appearance of morphology abnormality and significant cell death caused by vaccinia virus replication. The yields of L gene products were determined by Western Blot analysis and quantitative RT-PCR.

In vitro transcription assay

1 μg pL^{WT} or $pL^{L1sdmax}$ plasmid was linearized by *Nhe*I and PCR purified prior to *in vitro* transcription. Aliquots of the transcribed products were analyzed on acidic agarose gels.

Virus yields in 17AAG treated cells

17AAG or Tanespimycin, purchased from Sigma-Aldrich, Inc, is a derivative of geldanamycin and a well-defined potent Hsp90 inhibitor(135). HeLa cells were pretreated with increasing concentrations of Hsp90 inhibitor 17AAG for 15 min and then infected with either WT or $L1^{sdmax}$ virus at a MOI of 0.01. Incubation was continued in the presence of the indicated concentrations of 17AAG until 12 hpi, when the supernatants were harvested. Plaque forming assays were then performed in HeLa cells to measure the production of infectious viral progenies in the supernatants with or without Hsp90 inhibition.

Pulse-and-chase assay

HeLa cells were plated in 35 mm dishes on the day prior to infection, in order to reach an optimal density (~80% confluence) when the pulse-chase experiment was carried out. Cells were

infected with WT/L1^{sdmax} viruses at a MOI of 10 and at 3.5 hpi the culture media was removed. Cells were then washed with PBS three times and incubated in DMEM lacking methionine and cysteine for 30min, to deplete the pools of free Met/Cys inside the cells. ^{35}S-labeled Met/Cys was added to the culture media at 4 hpi, when the host protein synthesis was virtually shut off by the virus infection. After a 30min-labeling period, the media was aspirated and the cells were extensively washed with warm PBS three times. The radiolabeled cells were then chased with standard infection media (complete DMEM supplemented with 2% FBS) for 30 min, 60 min, 90 min, 150 min and 210 min. Whole cell lysates were collected at the indicated time points and analyzed on a 7.5% SDS-PAGE gel. The fate of viral proteins synthesized during the labeling period was further analyzed with ImageJ.

5.5 Figures and Tables

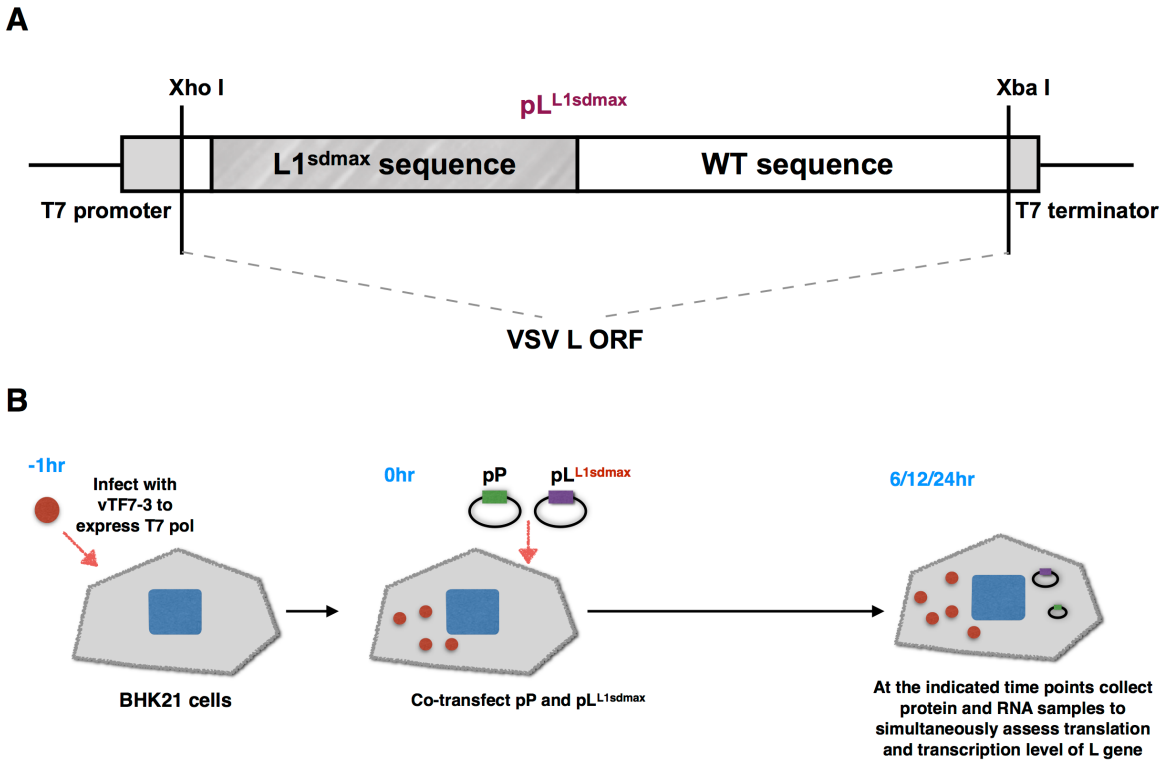


Figure 5.1 Expression vectors containing VSV L derivatives with altered codon pair biases. (A) The full-length L gene containing the synthetic $L1^{sdmax}/L1^{min}$ sequence was placed under the control of the T7 promoter and the newly constructed expression plasmids were designated as $pL^{L1sdmax}$ and pL^{L1min}, respectively. (B) A schematic diagram to show the experimental procedure. The expression of $L1^{sdmax}/L1^{min}$ sequences was examined in BHK21 cells that were infected with T7 pol-expressing vaccinia virus.

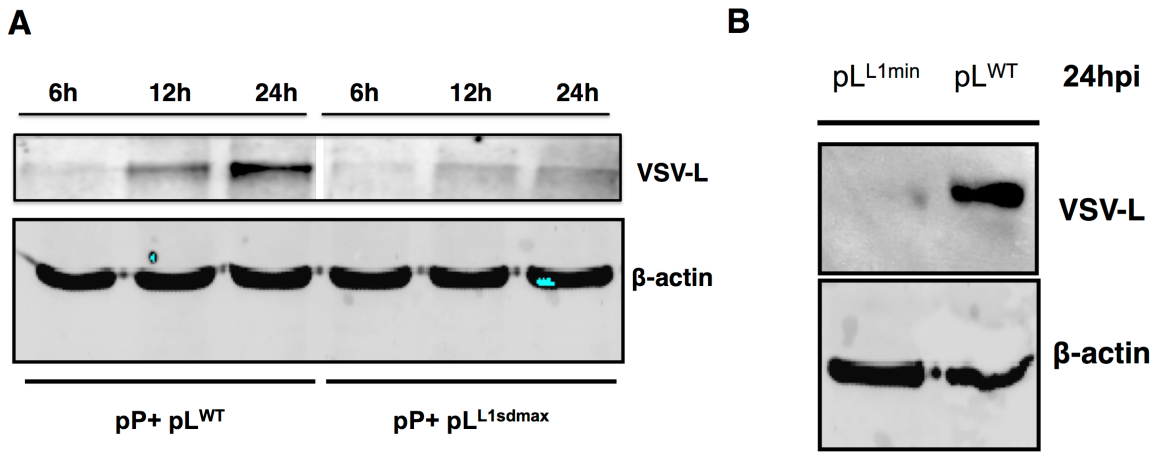


Figure 5.2 The L protein synthesis in pL$^{L1sdmax/L1min-}$transfected BHK21 cells.
Whole cell lysates were collected at the indicated time points and probed with polyclonal antisera against the full-length VSV L protein. (A) the L gene expression in pL^{WT}- and pL^{L1sdmax}-transfected cells over the time course were examined and compared by Western Blot. (B) the L protein synthesis in pL^{WT}- and pL^{L1min}-transfected cells were examined at 24 hpi.

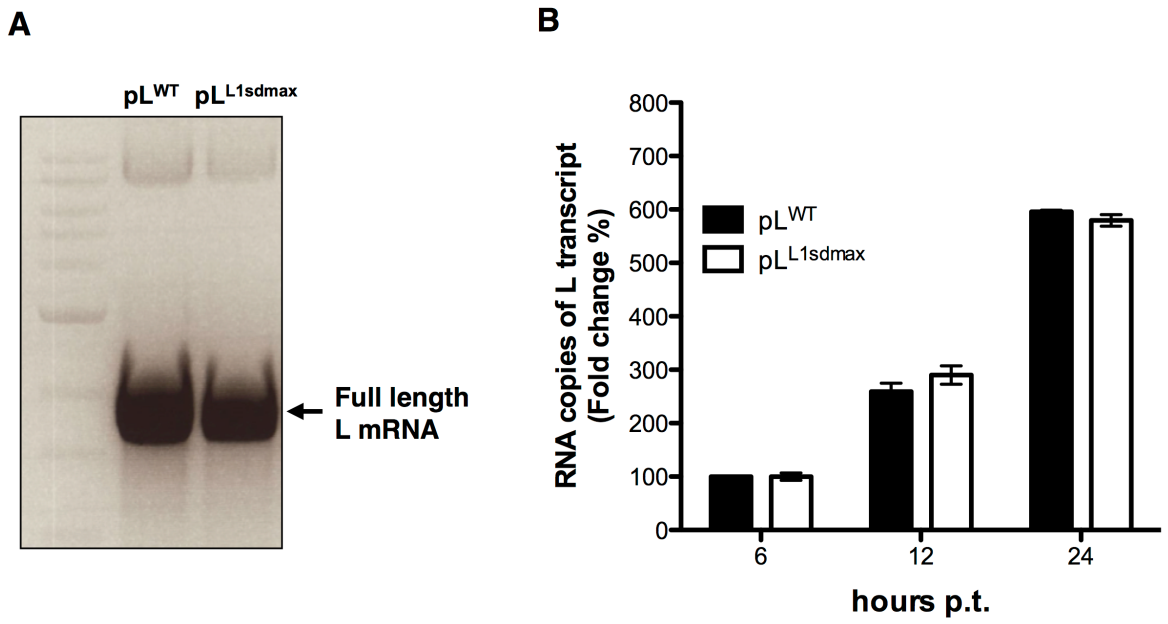


Figure 5.3 The transcription of the L1^{sdmax} sequence.
(A) *In vitro* transcription of the L1^{sdmax} sequence. pL^{L1sdmax} and pL^{WT} plasmids were linearized by *Nhe*I prior to *in vitro* transcription. Aliquotes of the transcription products were analyzed on acidic agarose gels. (B) After transfection, L mRNA synthesis was examined by qRT-PCR and normalized to the transcription level of the GAPDH gene.

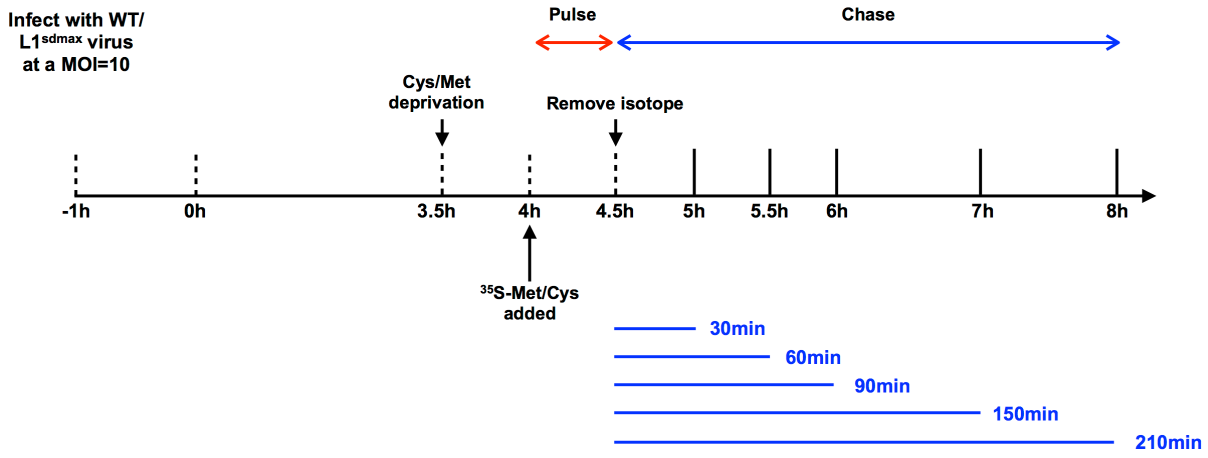
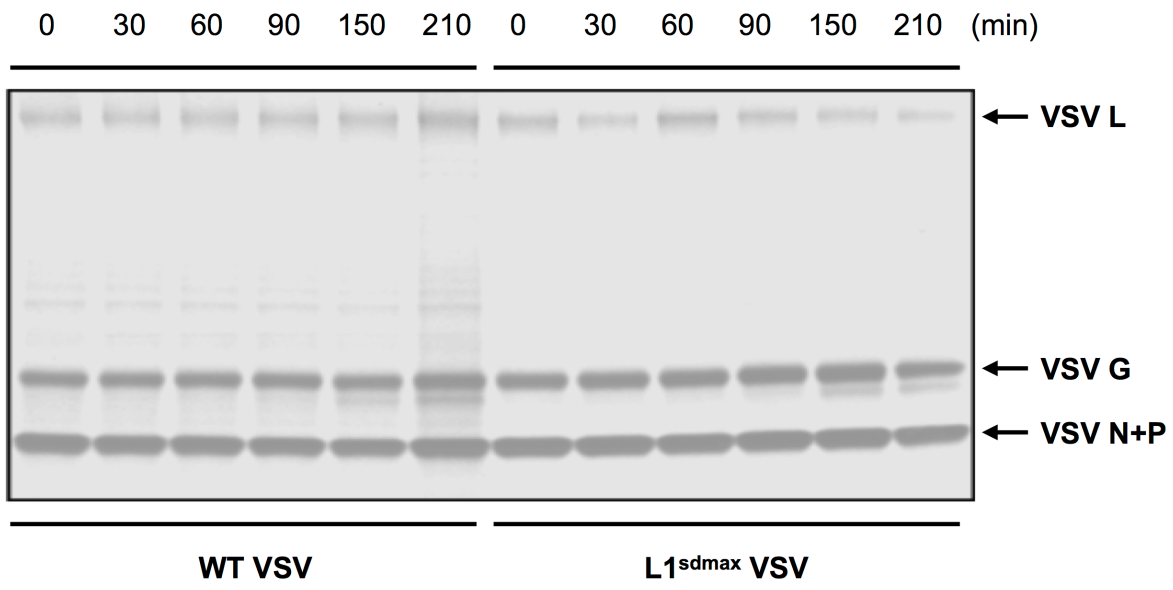


Figure 5.4 The schematic diagram of the pulse-chase assay.
Briefly, HeLa cells were infected with viruses for 4 hr and viral proteins were then radiolabeled (pulsed) for 30 min, followed by additional "chase" periods with a standard infection media. The whole cell lysates were collected and separated on a 7.5% SDS-PAGE gel.



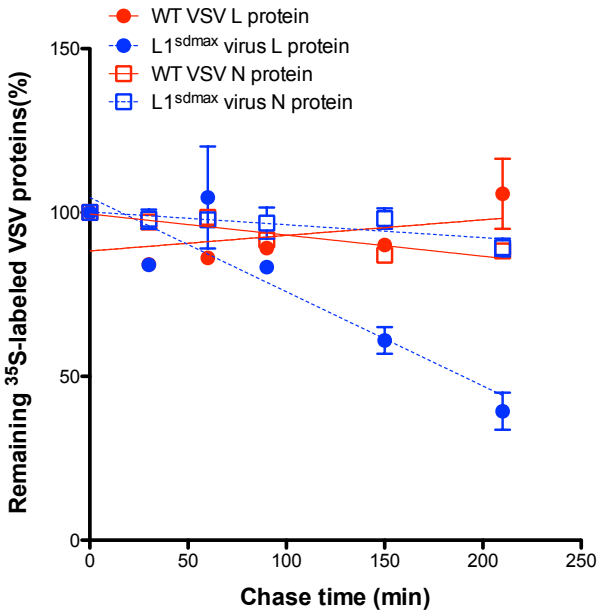


Figure 5.5 Reduced stability of the L protein synthesized in L1^{sdmax}-infected cells.
(A) The ^{35}S-labeled virus-specific proteins were separated by SDS-PAGE electrophoresis and examined by phosphorescence imaging. (B) Quantitation of band intensities from multiple independent pulse-and-chase experiments.

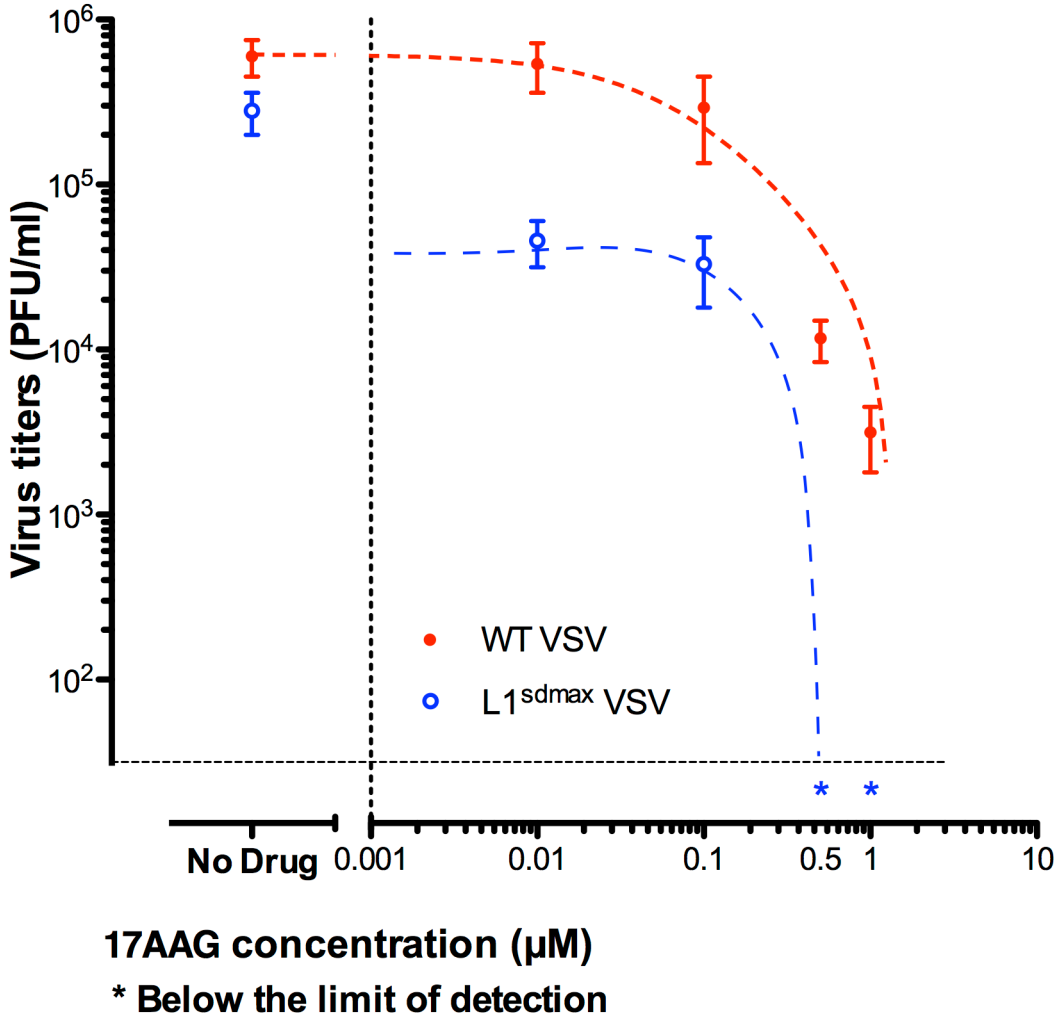


Figure 5.6 The inhibitory effect of Hsp90 inactivation upon VSV replication.
HeLa cells were pretreated with increasing concentrations of Hsp90 inhibitor 17AAG for 15min and then infected with WT or L1^{sdmax} virus at a MOI of 0.01. At 12 hpi, the supernatants were harvested and virus titers were determined in HeLa cells by plaque-forming assay. The * asterisks indicates that the titers of L1^{sdmax} virus were undetectable (<30 PFU/ml) in cultures when the drug concentration was 0.5 μM or higher.

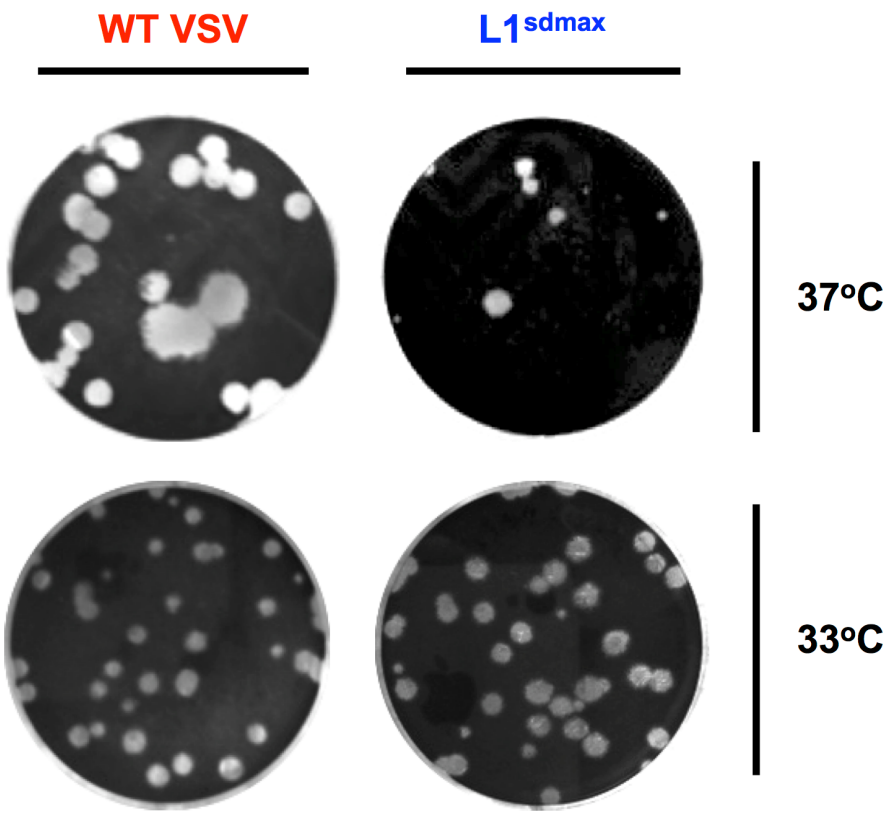


Figure 5.7 Plaque phenotype of the L1^{sdmax} virus at a lower temperature.
HeLa cells were infected with viruses at physiological temperature (37°C) for 30min, overlaid with 0.25% agar gel and incubated at the indicated temperatures (37°C or 33°C) for additional 48hr. The plaques were then developed with 1% crystal violet.

Chapter 6 Characterization of the L1^{sdmax} virus in an animal model

6.1 Introduction

The L1^{sdmax} virus exhibits a modest attenuation in cell culture. Based on our previous experience in working with influenza virus and dengue virus, the phenotypes observed in tissue culture with a codon pair-recoded virus do not necessarily predict phenotypes of the same recoded virus when it is tested in experimental animals. For instance, the recombinant influenza strain (NA+HA)^{min} contained deoptimized hemagglutinin/neuraminidase genes and replicated robustly in MDCK cells. However, this virus variant was extremely attenuated in mice(93). Therefore, it was of interest to evaluate the virulence of L1^{sdmax} virus in an animal model.

Inoculation of VSV intranasally is one of the most convenient and sensitive ways to examine the virulence. In nature VSV infects a wide range of livestock and causes epidemics in the southwestern United States(2). Acute VSV infection in farm animals is frequently characterized by vesicular lesions and ulceration on mucosal surfaces(3). However, in experimental animals, rodents and primates included, WT VSV is neurotropic when given intranasally and causes fatal encephalitis instead. In mice VSV first replicates in the nasal epithelial cells and then spreads to the olfactory neurons. By disseminating via the axons, the virus gains entry to the central nervous system, which usually leads to severe morbidity and mortality in 7-10 days(6).

In this study the virulence, pathogenicity and immunogenicity of the L1^{sdmax} virus were examined in a well-defined mouse model. In BALB/c mice, the L1^{sdmax} virus is highly attenuated yet almost equally immunogenic as WT VSV. Despite a temporary productive replication in the infected organs, neither histological change nor severe mobility was observed in L1^{sdmax}-infected mice.

6.2 Results

6.2.1 Diminished virulence of the L1^{sdmax} virus after intranasal infection

Five-to-six weeks old male BALB/c mice were infected with different doses of WT or L1^{sdmax} viruses intranasally. Their body weights and symptoms of neuronal disorder were closely monitored as indicators of virus pathogenesis in the following 14 days. Consistent with previous studies, WT-infected mice lost weight rapidly and exhibited typical symptoms of VSV infection starting around 3 dpi, including ruffled fur, lethargy, prostration and occasionally, hind limb paralysis(38)(Fig6.1). At a titer of $3x10^8$ PFU per animal, all the mice (N=5) succumbed to virus infection within 5 days post infection. In low titer groups (e.g. 10^3 and 10^4 PFU), 20% of the infected mice (N=10) died before day 10, but they also showed a relatively delayed onset of disease. Under my experimental conditions, the LD_{50} for WT VSV is approximately 10^5 PFU if the virus is given intranasally(Fig6.2).

In contrast, neither mortality nor severe morbidity was observed in any mice infected with the L1^{sdmax} virus. In the group of mice infected with $2x10^8$ PFU, some animals showed ruffled coats and transient weight loss between day 4 and day 7. However, these mice recovered quickly and their weights increased steadily afterwards (up to 21 dpi, Fig6.3). Taken together, the L1^{sdmax} virus is highly attenuated in BALB/c mice (at least 2,000 fold increase of LD_{50}, Fig6.4). This observation is surprising but it strongly implies that L1^{sdmax} virus qualifies as a potential vaccine candidate.

6.2.2 Virus burdens in the infected organs

The efficacy of live attenuated vaccines depends on a limited degree of viral replication *in vivo*, e.g. for a successful stimulation of immunological memory(95). Moreover, virulence and the virus load identified in infected organs are usually positively correlated. Therefore, I continued to examine virus burden in the infected animals by both plaque forming assay and quantitative RT-PCR. At 3 and 6 dpi, three mice intranasally infected with either WT or L1^{sdmax} virus at a dose of $5x10^6$ PFU were euthanized and the virus titers in the tissue homogenates (both brains and lungs) were measured by plaque forming assay (Fig6.5A). As expected, on day 3 the virus load in the L1^{sdmax}-infected organs was virtually two orders of magnitude less than that in their WT counterparts. The average titers of WT-infected organs were 5-7 $x10^5$ PFU/gram while

in the L1^{sdmax} groups they were 1.3-1.8x10^3 PFU/gram. On day 6, the infectious L1^{sdmax} virus was not recoverable from any lung samples that I examined and could only be detected in one brain sample (N=3). This observation strongly suggests that even though the L1^{sdmax} virus replicates *in vivo*, it is readily cleared and the infection is self-limiting. In contrast, in WT-infected animals, a heavy burden of virus was found in brain tissue at 6 dpi (approximately 10^4 PFU/gram), which ultimately resulted in the death of all the remaining animals (around 9 dpi, data not shown).

I also examined the viral RNA load in the brains using quantitative RT-PCR. Briefly, total RNA was extracted from the infected brains (including olfactory bulbs) and a 150 bp fragment corresponding to 5'-terminus of the VSV N gene was amplified. The VSV-specific RNA were identified in all of the infected samples at both 3 dpi and 6 dpi. However, the viral RNA load in L1^{sdmax}-infected organs was considerably lower, especially at the later time point (Fig6.5B). These results suggest that in comparison to WT VSV, the L1^{sdmax} virus is less competent to replicate and spread in mouse tissues. Interestingly, at 3 dpi, active viral RNA synthesis occurred in the L1^{sdmax}-infected brains without a robust production of infectious particles (cross-comparison of Fig6.5A and 6.5B). The reason for this discrepancy is unknown but in any event, the significantly less virus burden and viral RNA products offer an explanation for the diminished neurovirulence of the L1^{sdmax} virus in mice.

6.2.3 Disparate histopathology of WT- and L1^{sdmax}-infected tissues

The limited replication of the L1^{sdmax} virus in the brains of infected animals presents with safety concerns should such a viral variant serve as a promising vaccine candidate. Therefore I conducted further histological studies. BALB/c mice were infected intranasally with 10^7 PFU of WT or L1^{sdmax} virus, and their brains were harvested for hematoxylin and eosin staining on day 5. The brain tissues infected with WT VSV appeared to be dark red, a signature of severe inflammation(136) (Fig6.6 D-F) while the tissues infected with L1^{sdmax} or inoculated with PBS showed a pink appearance(Fig6.6 G-I). Brain sections from WT-infected mice exhibited severe immunopathogenic characteristics, evidenced by multifocal perivascular cuffing, mononuclear cell infiltration in the parenchyma, reactive astrogliosis, prominent meningeal mononuclear cell infiltration and hemorrhage(137) (Fig6.6 E and F). In contrast, tissues collected from the L1^{sdmax} group only exhibited mild, if any, inflammatory changes or other clinical signs of virus infection. Lymphocyte adherence and infiltration were rare occurrences (Fig6.6 I).

It has long been known that as a neurotropic virus, VSV infection compromises the integrity of the blood brain barrier (BBB)(37, 138, 139). The destruction of the BBB allows for the permeabilization of anti-viral immune cells that promotes virus clearance, but an intensive immunological infiltration also frequently causes severe immunopathogenesis that directly leads to the mortality of experimental animals. Therefore I used a well-established Evans blue assay to assess the integrity of the BBB in infected animals, which should provide an additional indicator of virus pathogenesis(140). I injected intraperitoneally 2% Evans blue in PBS 1 hour before the infected mice were euthanized and transcardially perfused. Evans blue has a high affinity to the albumin in blood and only enters the CNS when the BBB is compromised(140). At 6 dpi the brains of animals infected with $L1^{sdmax}$ showed no sign of BBB damage whereas a widely spread, intensive parenchymal staining was apparent in WT-infected brains, a clear sign of diminished BBB integrity(Fig6.7B). Experiments were conducted only when animals exhibited severe weights loss because VSV pathogenesis developed with variable kinetics. When the experiment was carried out, WT-infected mouse with the apparent staining was moribund whereas the $L1^{sdmax}$-infected animal appeared to be healthy.

6.2.4 Immunogenicity of the $L1^{sdmax}$ virus in mice

An ideal vaccine candidate should present with weak virulence and strong immunogenicity. WT VSV is remarkably immunogenic because it replicates robustly *in vivo* and is highly cytopathic(6). Therefore it was of interest to find out whether the restricted replication of this virus variant was sufficient to provoke an effective immunological protection in infected animals.

Five-to-six weeks old BALB/c mice were immunized with 10^5, 10^4, 10^3 PFU of WT virus or 10^6, 10^5, 10^4, 10^3 PFU of $L1^{sdmax}$ virus intranasally. After three weeks, all mice surviving from the primary infection received 10^7 PFU of WT virus, representing an approximately 100-fold LD_{50} of WT VSV, also via an intranasal route. None of the $L1^{sdmax}$-immunized mice exhibited any signs of disease and their body weights actually increased after the lethal challenge whereas the age-matched virus naive mice succumbed to the virulent infection after 6-8 days (Fig6.8).

I next asked whether a B cell-mediated humoral immunity could be elicited by the $L1^{sdmax}$ vaccination. I collected sera from tail vein blood of the vaccinated animals at 21 dpi and measured the titers of virus specific neutralizing antibodies (nAb) in them. $L1^{sdmax}$ was capable

of inducing a robust protective antibody response given that a similar level of nAb was detected in mice that were immunized with the same amount of either WT or $L1^{sdmax}$ virus, whereas the control group inoculated with PBS did not develop any humoral response against VSV infection (Fig6.9). Taken together, these data indicate that a single dose vaccination of the $L1^{sdmax}$ virus can elicit a potent immune response that fully protects mice from the following virulent challenge.

6.2.5 Prolonged induction of pro-inflammatory cytokines and chemokines in WT- but not in $L1^{sdmax}$-infected mice

Inflammatory mediators, such as cytokines TNF-α and interleukin-6, as well as leukocytes recruiters, such as chemokines MCP-1 and RANTES, are actively involved in the clearance of neurotropic viruses. However, an abnormally extended activity of these pro-inflammatory factors usually leads to an irreparable damage to neurons and contributes significantly to virus-associated neuronpathogenesis. In addition, during VSV infection, mice susceptibility correlated positively with the extensive cytokine/chemokine expression in the central nervous system(136). To determine whether a different inflammatory response occurred during $L1^{sdmax}$ infection, I profiled the transcription levels of several pro-inflammatory cytokines and chemokines in virus-infected tissues. Under my experimental conditions, all the infected mice exhibited enhanced cytokine/chemokine expressions (~1000 fold) relative to the PBS-inoculated group (Fig6.10 and 6.11). In comparison to their WT-infected counterparts, MCP-1 and RANTES were induced to a lesser extent in $L1^{sdmax}$-infected mice, which correlated well with the observation of significantly fewer infiltrated lymphocytes in $L1^{sdmax}$-infected brain parenchyma (Fig6.10). In addition, at 3 dpi the $L1^{sdmax}$-infected mice showed similar levels of TNF-α and IL-6 relative to those of WT-infected ones, also consistent with the finding of a similar level of viral RNA at the same time point (Fig6.11 and Fig6.5B). These findings reiterate the immunogenicity of the $L1^{sdmax}$ virus, because a strong cytokine induction during the early phase of infection is not only essential to mount immune response for the virus clearance, but is also critical for the development of immunological memory for long-term protection(141). WT-infected mice retained a high level of cytokine production up to 6 dpi while $L1^{sdmax}$-infected ones showed a significantly declined cytokine synthesis by then (Fig6.11). It strongly suggests that in contrast to the highly attenuated $L1^{sdmax}$ virus, WT virus infection in the CNS cannot be

controlled without sustained inflammation, which eventually leads to immunopathogenesis and irreversible neuronal damage.

6.3 Discussions

In this study I identified and characterized a highly attenuated recombinant virus strain $L1^{sdmax}$, which retained most, if not all of the immunogenicity of WT VSV. The $L1^{sdmax}$ virus underwent a limited replication and dissemination *in vivo* with reduced morbidity and undetected pathological changes, and this self-limited infection could elicit a high level of virus-specific neutralizing antibody and fully protect mice from a subsequent lethal WT challenge. The $L1^{sdmax}$ virus appears to be an attractive live vaccine candidate because of its remarkable safety margin and strong efficacy. In addition, it is likely that the same recoding strategy can be applied to the other viruses in the order of *Mononegavirales* that encode very large polymerases. My study extends previous approaches to attenuate VSV by adding large-scale polymerase gene recoding. Compared to the conventional attenuation strategy that usually involves a few point mutations, one of the advantages of my approach is the genetic/phenotypic stability of the recoded virus. In indeed, I continuously passaged the $L1^{sdmax}$ virus for 5 times in BHK21 cells at low MOIs and did not find any phenotypic reversion (data not shown). Nonetheless, its genetic stability *in vivo* remains to be further investigated.

My study may not only facilitate the development of a VSV-based oncolytic therapy and antigen delivery, but also serves as a useful example for the attenuation of other NNS RNA viruses. A live attenuated version of a viral pathogen is generally a good choice for vaccine, because it expresses the full spectrum of viral antigens and frequently elicits a broad, long-lasting systemic immune response. Moreover, the production of live virus vaccines is usually more cost effective than that of "killed" virus vaccines. Since the L proteins of the viruses in the order of *Mononegavirales* share a similar molecular and structural architecture(133), it will be of interest to test whether my approach to attenuate VSV could be extended to other related species.

Previous studies have shown that during an infection with Venezuelan equine encephalitis virus (VEEV), a well-defined neurotropic arbovirus, the infected cells in the nasal mucosa initiated a robust cytokine response. This activation signal was thought to propagate along the olfactory nerves and set the CNS into an alerted state even before the virus reached the brain(142). Therefore in my experiment, strong cytokine induction in the CNS during the early phase of $L1^{sdmax}$ infection was probably, at least, due in part to its noticeable propagation in nasal epithelial cells and olfactory sensory neurons (OSN) in the periphery. Besides, a "bystander effect" observed in VEEV infection also suggested that the uninfected "bystander" cells in the

CNS invoked cytokine production even without virus replication at that particular site(142). This might offer explanations for a seemingly debatable result presented in my study: at 3 dpi, relative to the WT-infected mice, a significant lower virus titer found in $L1^{sdmax}$-infected brains was accompanied by a similar transcription level of pro-inflammatory cytokines (TNF-α and IL-6). As the temporary induction of pro-inflammatory cytokines helped to mount an effective immune response, the invaded $L1^{sdmax}$ virus was rapidly brought under control. In contrast, WT VSV showed a better ability to replicate and spread *in vivo* even when the mice developed a sustained activation of cytokine signaling. From 3 dpi to 6 dpi, accompanied by augmented chemokine response, viral load decreased but was still at a high level, approximately 10^4 PFU/gram at 6 dpi, the time when cells of adaptive immunity typically entered the brain and compromised the integrity of the blood brain barrier.

6.4 Materials and methods

Virus infection and LD_{50} determination in mice

Five-to-six weeks old male BALB/c mice were anesthetized with ketamine/xylazine and inoculated intranasally with different doses of viruses in a total volume of 50 µl. Animal body weights and the symptoms of neurological disorder were monitored daily for 14 days post infection. Mice that lost up to 25% of their original weights were euthanized for animal welfare. The virus LD_{50} was calculated based on the Reed-Muench method(143).

Virus load in mouse tissues

Groups of male BALB/c mice were intranasally infected with $5x10^6$ PFU of either WT or $L1^{sdmax}$ virus. For virus titer determination, at 3 and 6 dpi, 3 mice of each group were euthanized. Their brains and lungs were harvested and homogenized. The supernatants clarified from the homogenates were then serially diluted and the virus titers were determined by plaque forming assay in Vero cells. For viral RNA load analysis, six mice of each group were euthanized at both 3 and 6 dpi. Total RNA was extracted from brains using Trizol following the manufacturer's instructions and reverse transcribed into the first strand cDNA with random hexamer using SuperScript III reverse transcriptase (Invitrogen). The cDNA samples were subjected to real-time PCR and the level of viral RNA was measured using primer pair specific to VSV N gene. In parallel, real-time PCR directed against mouse β-actin mRNA was performed to control for any differences in the amount of total RNA inputs.

Primary vaccination and virus challenge

Five-to-six weeks old BALB/c mice were immunized with various doses of WT or $L1^{sdmax}$ virus intranasally at day 0. After 21 days, all the mice that survived from the primary infection were challenged with a lethal dose of WT VSV ($1x10^7$ PFU i.n.). Their body weights and disease symptoms were further monitored until 32 dpi. A group of age-matched virus-naive mice were infected at 21 dpi as the negative control.

Neutralizing antibody assay

Tail-vein blood was collected 21 days after vaccination. Blood was allowed to clot after overnight incubation at 4°C and the sera were isolated after centrifugation, followed by a two-fold serial dilution in DMEM starting at a dilution of 1:20. 50 µl of the diluted sera was incubated with 100 PFU WT VSV in 50 µl DMEM at 37°C for 1hr. 100 µl Vero cell suspension was combined with the mixture and then seeded on 96-well plates. Triplicates were performed for each serum sample. Cell morphology was checked daily and after 48-hour incubation at 37°C, the remaining cells were stained with 1% crystal violet. The reciprocal of the dilution that gives 100% protection from VSV CPE was recorded as virus-specific neutralizing antibody titer.

Histopathology

Mice were infected with $5x10^6$ PFU virus by intranasal inoculation and euthanized at 5 dpi. Mice were transcardially perfused with 10% PFA/PBS and their brains were fixed in 10% PFA/PBS at room temperature overnight, followed by 50% ethanol processing and paraffin embedding. The brain tissues were then sectioned at 5 µm and stained with Hematoxylin & eosin. The histopathological results were examined independently by two neuropathologists in the school of medicine, Stony Brook University.

Evans blue assay

Mice were infected with a dose of $5x10^6$ PFU intranasally. At 6 dpi, the mice were injected with 300 µl 2% Evans blue/PBS intraperitoneally and the dye was allowed to circulate for 1 hr. After that, the mice were euthanized and transcardially perfused with 10 ml normal saline. Their brains were subsequently dissected and photographed.

Quantitative RT-PCR to analyze cytokine/chemokine induction

Total RNA was extracted from mice brains that were inoculated with PBS/WT/$L1^{sdmax}$ at both 3 and 6 dpi, then treated with RNase-free DNase I for 30 min before a reverse transcription was conducted with Oligo $(dT)_{20}$. Quantitative PCR was subsequently performed to assess the induction of cytokines (TNF-α and interleukin 6) and chemokines (MCP-1 and RANTES) using gene specific primers. All reactions were subsequently normalized to the RNA levels of mouse β-actin and reported as increases in induction over PBS-inoculated animals.

6.5 Figures and Tables

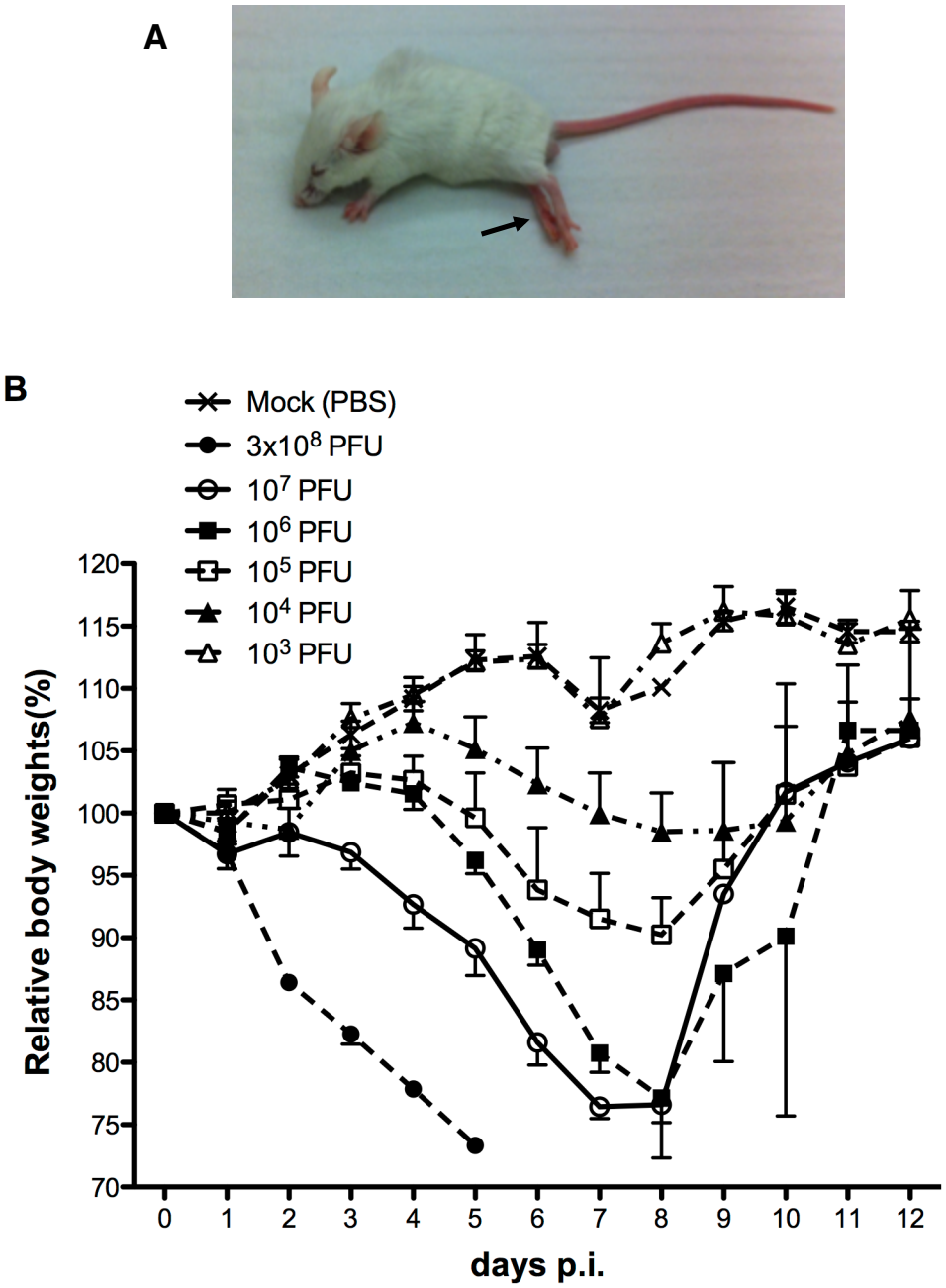


Figure 6.1 Pathogenesis of WT VSV in BALB/c mice.
(A) Five-to-six weeks old male BALB/c mice (N=5) were intranasally infected with WT VSV at different doses. WT-infected mice developed hind limb paralysis (indicated by arrow) and other symptoms of neuronal disorder. (B) Dynamics of mouse body weights after WT infection.

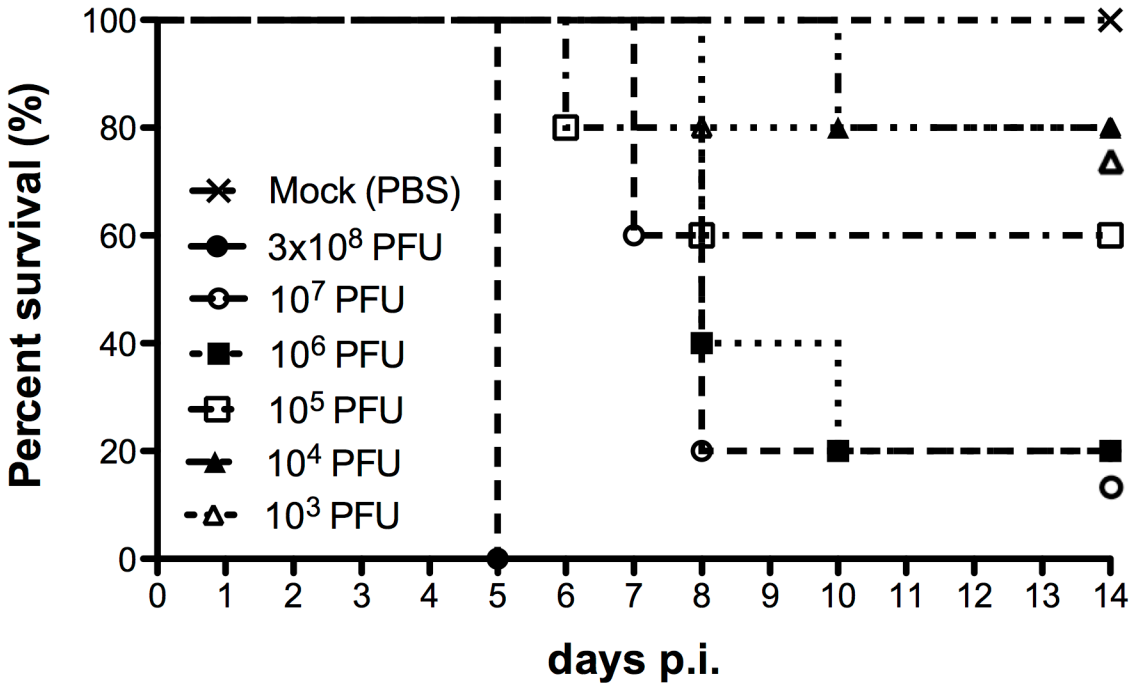


Figure 6.2 LD_{50} of WT VSV.
Survival rate of WT-infected animals (i.n) were continuously monitored for 14 days following infection. Under my experimental conditions, the LD_{50} for WT VSV is approximately $1.5x10^5$ PFU/animal.

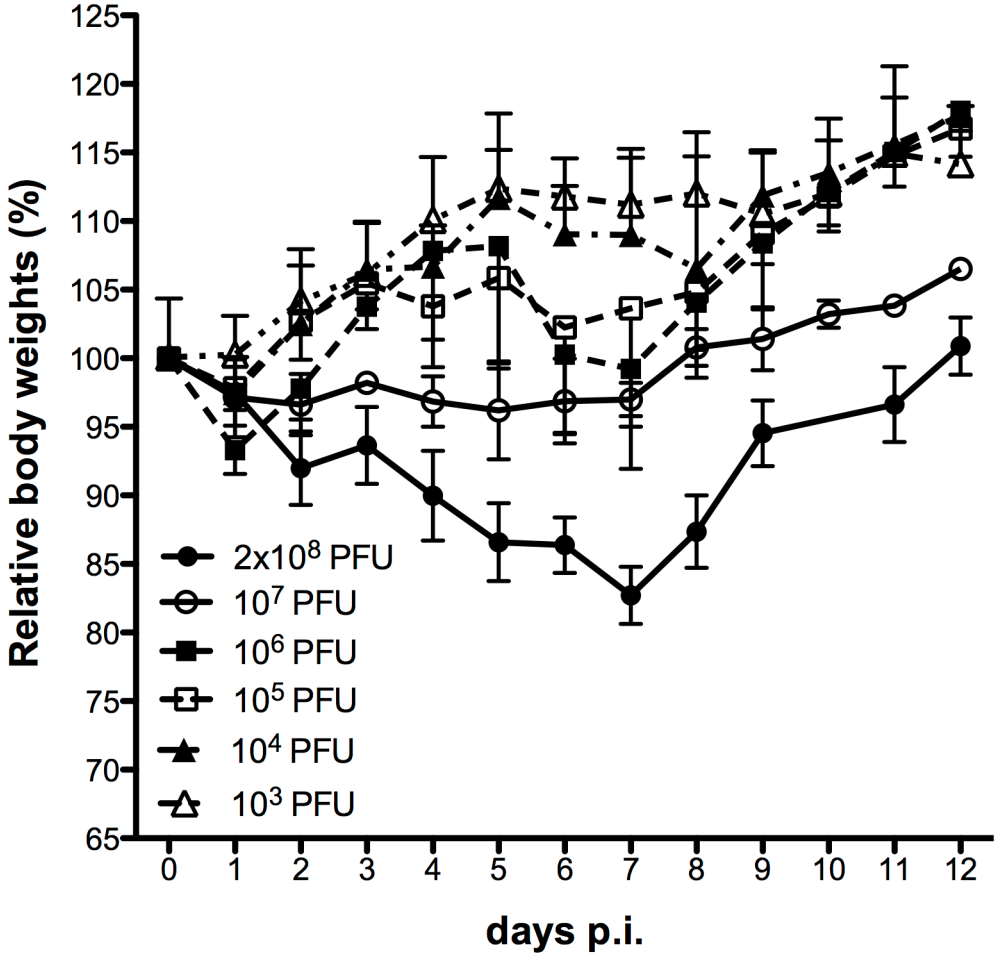


Figure 6.3 Dynamics of mouse body weights after L1^{sdmax} infection.
Five-to-six weeks old male BALB/c mice were infected with different doses of L1^{sdmax} virus intranasally. Their body weights were continuously measured for the following 12 days.

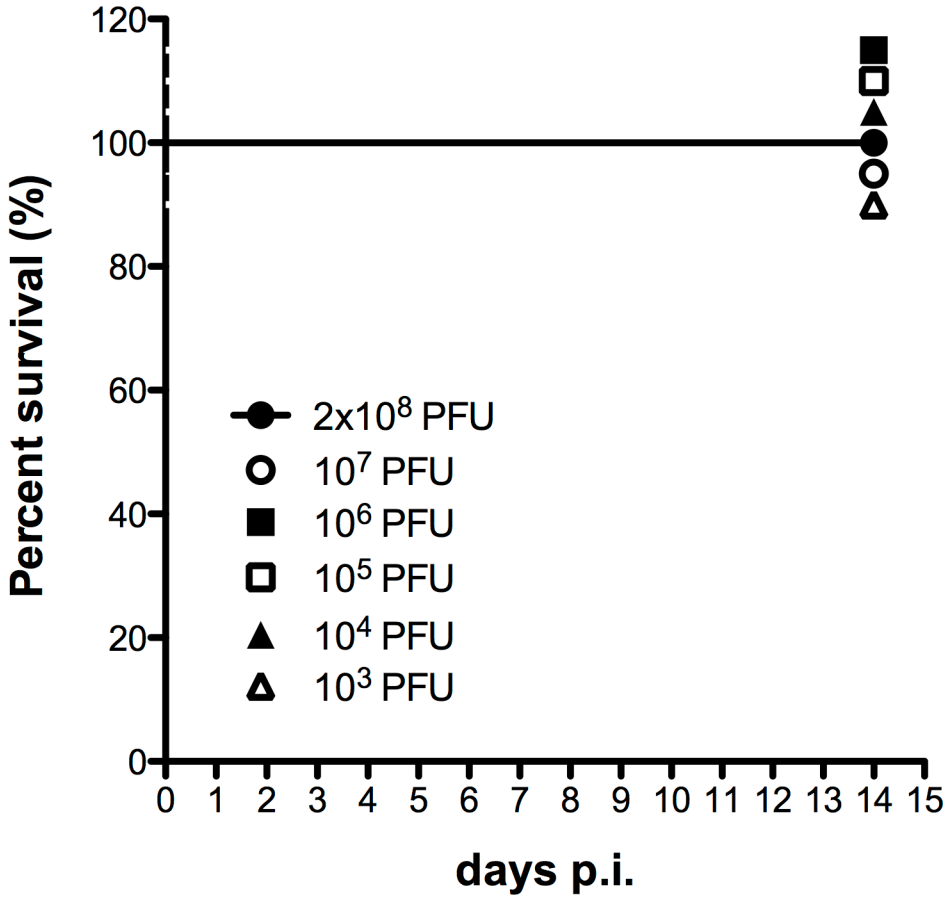


Figure 6.4 Survival rates of L1^{sdmax}-infected animals.
None of the L1^{sdmax}-infected animals succumbed to virus infection or lost more than 25% of their initial weights.

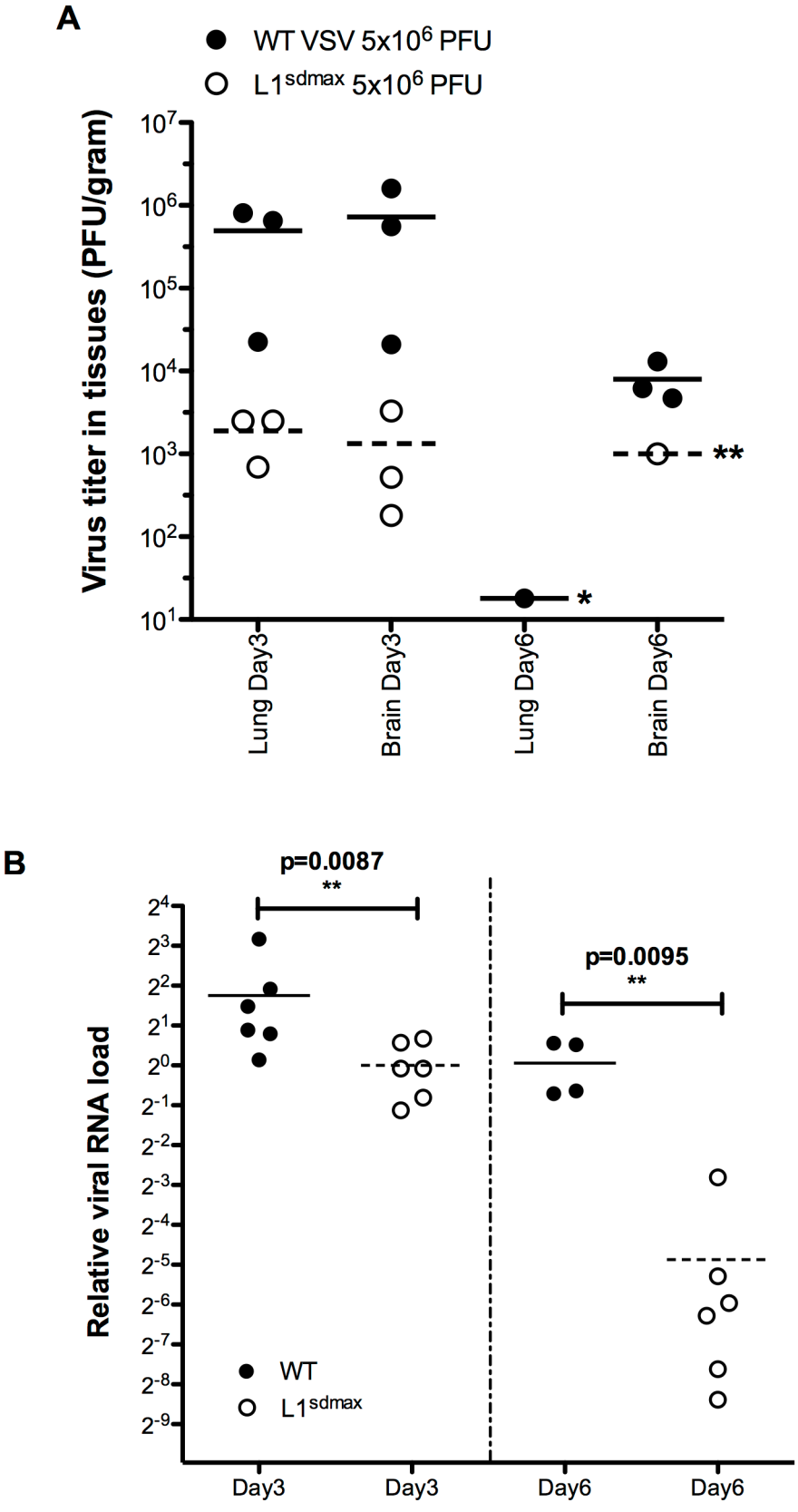

A
WT VSV 5x10^6 PFU
L1^sdmax 5x10^6 PFU
Virus titer in tissues (PFU/gram)
10^7
10^6
10^5
10^4
10^3
10^2
10^1
**
*
Lung Day3
Brain Day3
Lung Day6
Brain Day6
B
p=0.0087
**
p=0.0095
**
Relative viral RNA load
2^4
2^3
2^2
2^1
2^0
2^-1
2^-2
2^-3
2^-4
2^-5
2^-6
2^-7
2^-8
2^-9
WT
L1^sdmax
Day3
Day3
Day6
Day6

Figure 6.5 Virus burden in WT/L1^{sdmax}-infected organs.
(A) Five-to-six weeks old male BALB/c mice were intranasally infected with $5x10^6$ PFU viruses. At 3 and 6 dpi, 3 mice of each group were euthanized and their organs were harvested and homogenized. The virus titers in the homogenates were subsequently determined by plaque forming assay in Vero cells. * at 6 dpi the infectious viruses were not detectable (< 30 PFU/gram) in all but one WT-infected lung; ** at 6 dpi the presence of infectious virus could only be detected in one L1^{sdmax}-infected brain (N=3) (B) The viral RNA load in mice was measured using two-step quantitative PCR. Total RNA was extracted from infected brain tissues and reverse transcribed with random hexamers, followed by a real-time PCR directed against the VSV N gene. All the reactions were normalized to the mRNA level of the mouse β-actin gene. The asterisks indicated a significant statistical difference (**$p<0.01$, Two tailed Mann-Whitney U test).

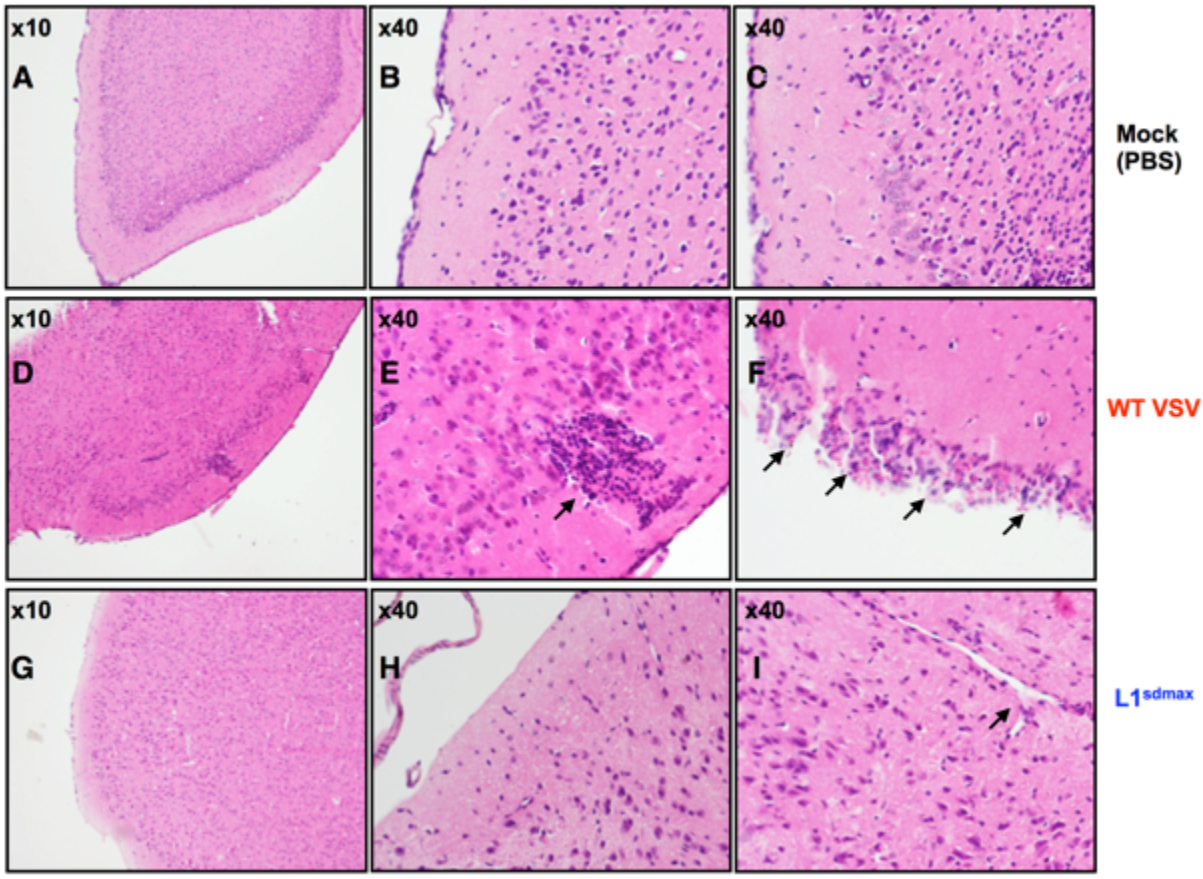


Figure 6.6 Brain histopathology of WT/L1^{sdmax}-infected mice.
Mice were infected with 5x10^{6} PFU WT (D-F) or L1^{sdmax} (G-I) virus intranasally. Horizontal brain sections containing cerebral cortex were stained for H&E at 5 dpi. Representative brain histology was shown for each group. Compared to PBS inoculated brains (A-C) and L1^{sdmax}-infected group (G-I), WT-infected brains (D-F) appeared to have severe inflammation. Arrow in (E) indicates the presence of multifocal perivascular cuffing and mononuclear cell infiltration in brain parenchyma. Arrows in (F) indicate a meningeal inflammation, characterized by thickened meninges, heavy infiltration of inflammatory cells and meningeal-based hemorrhage.

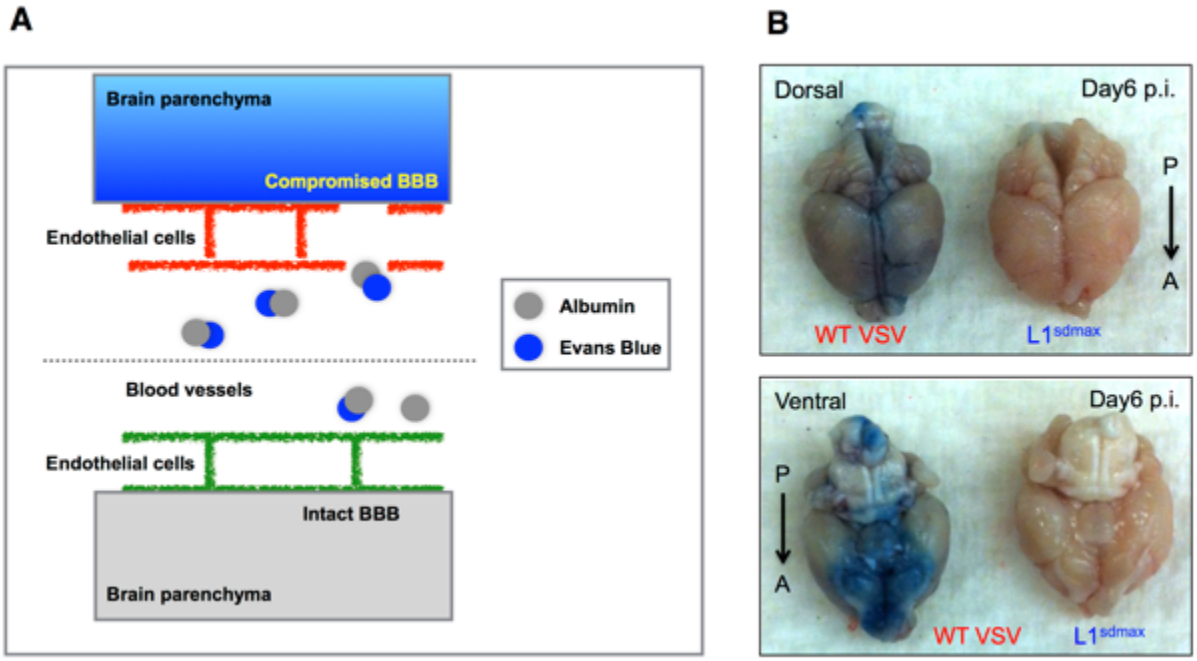


Figure 6.7 Evans blue assay to evaluate the integrity of the blood brain barrier.
(A) a schematic diagram to illustrate the mechanism of the Evans blue assay. As a dye, Evans blue has a high affinity towards albumin, a large molecule that normally can not cross the blood brain barrier(20). The parenchyma tissue is stained only when the blood brain barrier is compromised. (B) Mice were infected with viruses at a dose of $5x10^6$ PFU intranasally. At 6 dpi, mice were injected with Evans blue in PBS intraperitoneally and the dye was allowed to circulate for 1 hr. Then the mice were euthanized, perfused with saline and dissected. A severely compromised blood brain barrier was observed on day 6 after WT infection whereas $L1^{sdmax}$-infected brains showed no sign of BBB damage. Experiments were carried out only when WT-infected animals exhibited severe weight loss because VSV pathogenesis developed with variable kinetics.

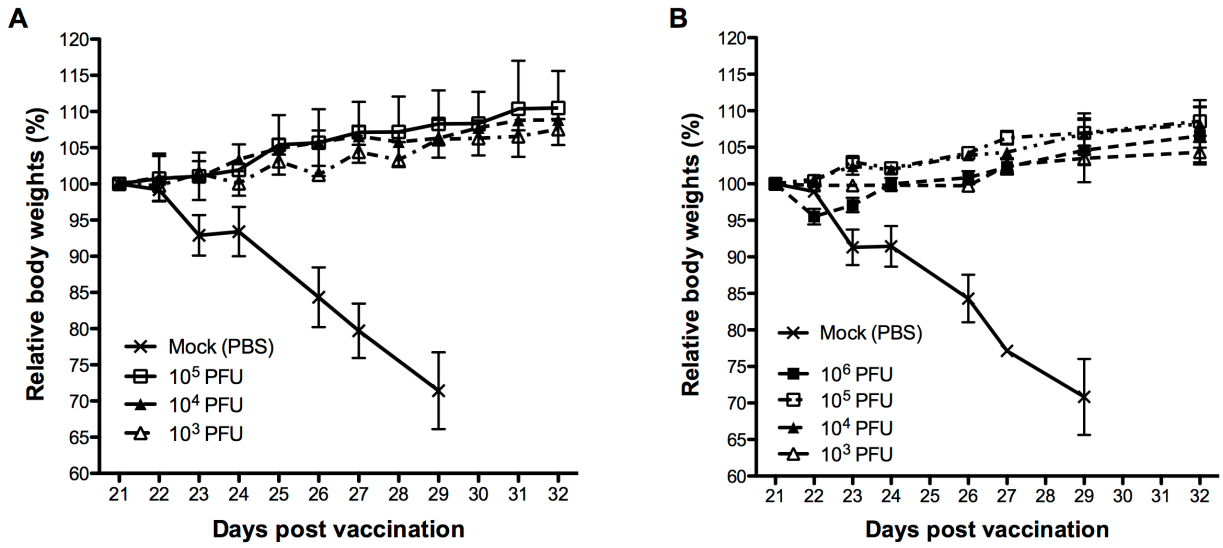


Figure 6.8 Immunogenicity of the L1^{sdmax} virus.
Dynamics of immunized mice body weights after a lethal challenge. Five-to-six weeks old male BALB/c mice were immunized with WT (A) or L1^{sdmax} (B) virus intranasally at different doses on day 0. After three weeks, all vaccinated mice that survived from the primary infection as well as three age-matched virus-naive mice were challenged with a lethal dose of WT virus (10^7 PFU per animal). Their body weights were closely monitored until day 32.

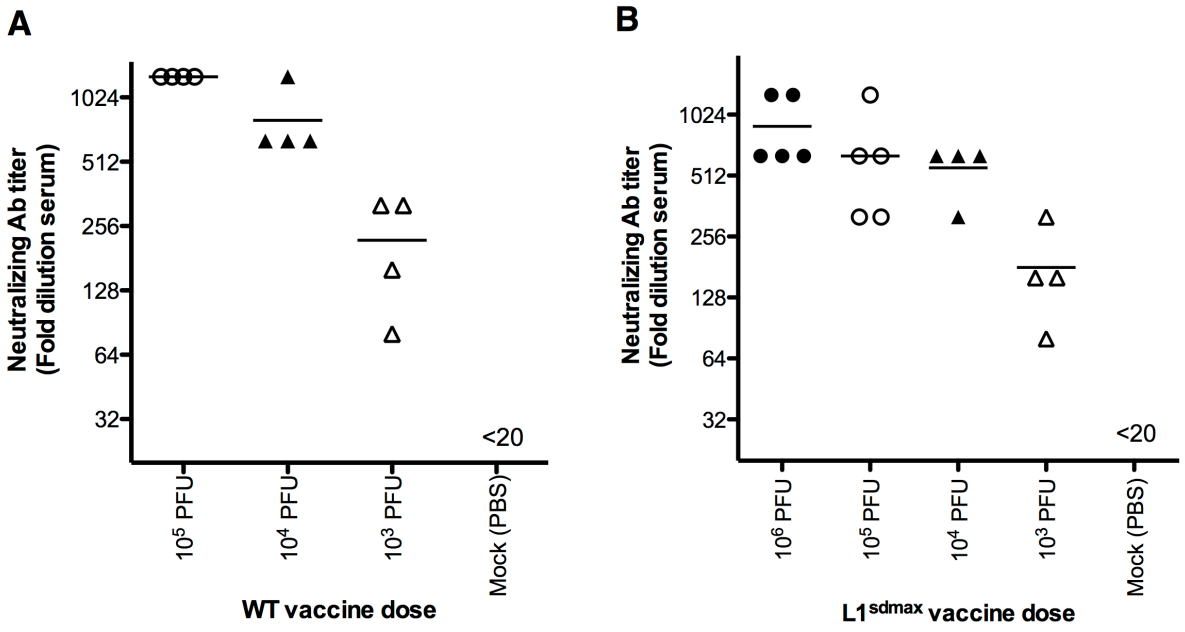


Figure 6.9 Virus-specific neutralizing antibody (nAb) titers in WT- (A) and L1^{sdmax}-(B) immunized mice at 21 dpi. Neutralizing antibody titers were measured as described in Materials and Methods.

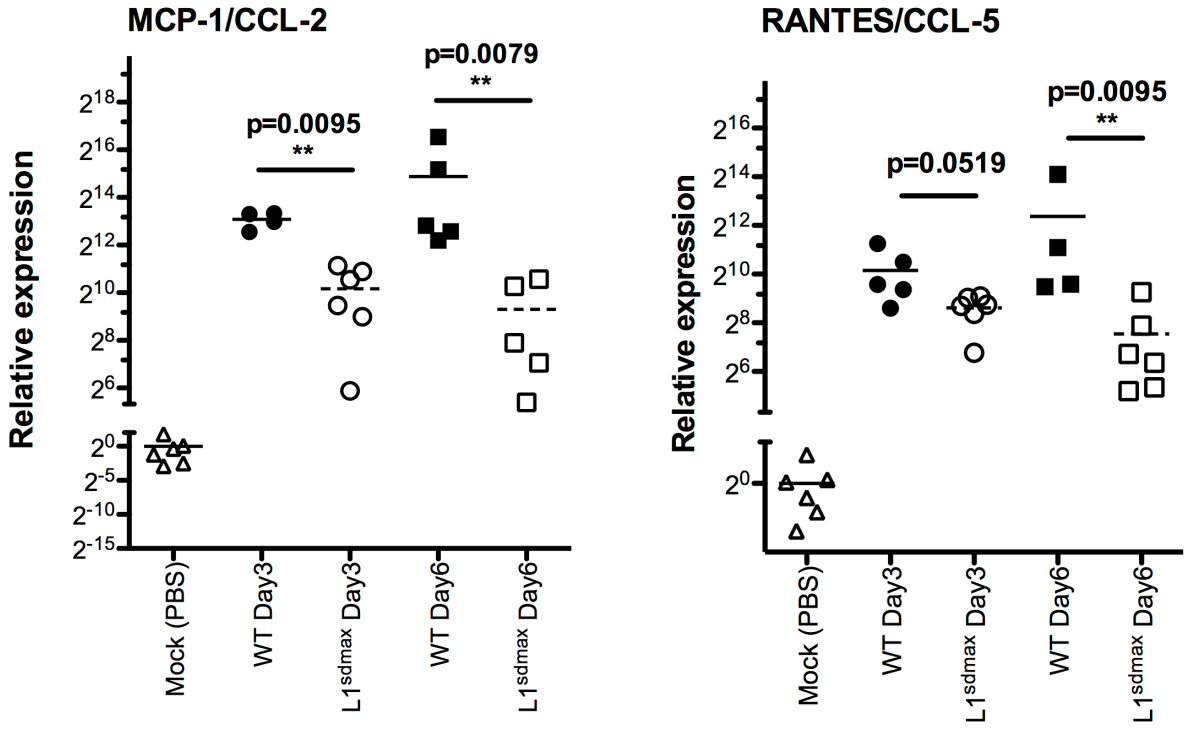


Figure 6.10 Transcriptional induction of chemokines in mice after virus infection.
Chemokine (MCP-1 and RANTES) expressions were measured by quantitative RT-PCR and normalized to the transcription level of the mouse β-actin gene. Data was collected from 4-6 mice per group. Statistical significance was demarcated by asterisk (**p<0.01, Two tailed Mann-Whitney U test).

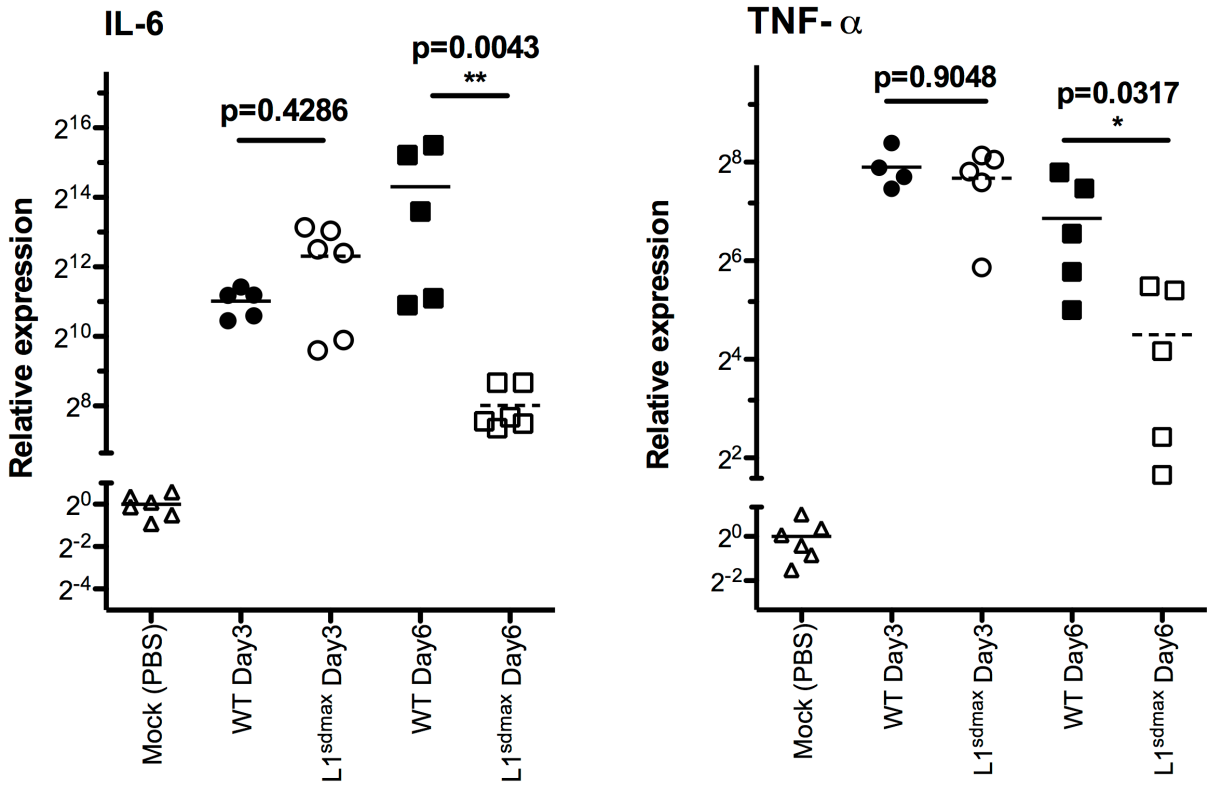


Figure 6.11 Transcriptional induction of pro-inflammatory cytokines in mice.
Pro-inflammatory cytokine induction (IL-6 and TNF-α) was measured in PBS/WT/L1^{sdmax} infected mice. Statistical significance was demarcated by asterisk (*p<0.05, **p<0.01, Two tailed Mann Whitney U test).

	Forward primer	Reverse primer
VSV N primer	5' ATGACAAATGGTTGCCTTTGTATCTACTT 3'	5' ACGACCTTCTGGCACAAGAGGT 3'
Mouse MCP-1 primer	5' TTAAAAACCTGGATCGGAACCAA 3'	5' GCATTAGCTTCAGATTTACGGGT 3'
Mouse RANTES primer	5' GCTGCTTTGCCTACCTCTCC 3'	5' TCGAGTGACAAACACGACTGC 3'
Mouse IL-6 primer	5' TGGGAAATCGTGGAAATGAG 3'	5' CTCTGAAGGACTCTGGCTTTG 3'
Mouse TNF-α primer	5' GCCTCTTCTCATTCCTGCTTGT 3'	5' GATGATCTGAGTGTGAGGGTCTG 3'
Mouse β-actin primer	5' AAGGCCAACCGTGAAAAGAT 3'	5' GTGGTACGACCAGAGGCATAC 3'

Table 6.1 The primers used for quantitative RT-PCR.

Chapter 7 Characterization of the G^{min} virus in cell culture and in the mouse

7.1 Introduction

The utility of codon pair bias-based recoding ("min" strategy) for the development of live attenuated vaccines has been iteratively demonstrated. The two poliovirus variants $P1^{min-XY}$ and $P1^{min-Z}$, whose P1 regions (encoding the viral capsid) contained more underrepresented codon pairs than the corresponding WT sequence, were highly attenuated both *in vitro* and *in vivo*(87). In a similar manner, a recombinant influenza PR8 variant $(NA+HA)^{min}$, which contained a total of 618 silent mutations, also exhibited remarkable attenuation phenotype in an animal model(93). In a collaborative effort in attenuating respiratory syncytial virus (RSV), we found that the recombinant virus Min B, which contained the deoptimized versions of viral G and F genes (encoding the two viral surface proteins), provided the greatest restrictive effects (92). The last but not the least, a recombinant dengue virus E^{hmin} that harbored a deoptimized envelope protein E appeared to be highly attenuated in newborn mice (109). Interestingly, in the four studies aforementioned, codon pair deoptimization of the viral genes that encode structural proteins (for instance, polio capsid protein, influenza HA and NA proteins, respiratory syncytial virus G and F proteins and dengue virus E protein) all led to a desirable balance between viral virulence and immunogenicity, suggesting that deoptimizing a structural protein in VSV in a similar manner could probably lead to a profound and more predictable attenuation phenotype.

Unlike the L polymerase, the VSV glycoprotein has no enzymatic activity and is not involved in genome replication. It has been well known that VSV can be pseudo-typed by the glycoproteins from closely related virus species, which indicates that the VSV G gene should be devoid of RNA sequence signatures that are essential for virus replication(144).

In this chapter I designed, constructed and characterized a VSV variant G^{min} that contained a deoptimized version of the glycoprotein gene G. Although considerably less viral glycoproteins were present in G^{min}-infected cells, the virus itself exhibited a robust growth phenotype in cell culture and a neurovirulence in mice that was comparable to the parental WT strain.

7.2 Results

7.2.1 Design, construction and recovery of recombinant G^{min} virus

A deoptimized/"min" version of the VSV glycoprotein gene G (nt 3211-4613, numbers referring to the complete genome of VSV Indiana serotype) was designed and chemically synthesized as described before. This artificial sequence was designated as G^{min}.

The G^{min} sequence had a significant enrichment of underrepresented codon pairs, which led to a codon pair score as low as -0.49. In total, 360 synonymous nucleotides substitutions were introduced into this region, accounting for 22.0% of the full-length sequence(Fig7.1). The recombinant virus G^{min} was recovered using vaccinia virus-mediated reverse genetics system and plaque purified.

7.2.2 Plaque-forming phenotype of the G^{min} virus in Vero cells

The plaque forming assay was carried out in Vero cells to compare the plaque morphology of G^{min} virus with its parental strain. Briefly, Vero cells were infected with serially diluted viruses, overlaid with 0.25% agar and incubated for overnight or prolonged period. Interestingly, G^{min} virus exhibited an indistinguishable plaque phenotype, from WT VSV(Fig7.2A).

7.2.3 Multi-step growth kinetics in BHK21 cells

BHK21 cells were infected with either WT or G^{min} virus at a MOI of 0.001. The supernatants were harvested at the indicated time points and further analyzed in Vero cells to determine virus titers. Surprisingly, G^{min} grows equally well as WT. At 48 hpi, when the full CPE occurred, G^{min} virus reached a maximum titer, which was close to 10^9 PFU/ml (Fig7.2B).

7.2.4 Western Blot analysis and Immunocytochemistry to measure viral protein synthesis

Two protein assays, immunoblot and immunocytochemistry, were carried out to examine viral protein synthesis in G^{min}-infected cells. For Western Blot analysis, BHK21 cells were infected with WT or G^{min} virus at a MOI of 5 and the whole cell lysates were collected at 3 hpi.

As expected, significantly less glycoprotein (~5 fold less, quantitation of band intensities by ImageJ) was synthesized in G^{min}-infected cells relative to the WT-infected ones (Fig7.3A).

For immunocytochemistry, HeLa cells were infected at a MOI of 5, followed by a co-staining of VSV N and G proteins at 3 hpi. Consistent with the finding in Western Blot, less glycoprotein was detected in G^{min}-infected cells while the synthesis of viral N protein was not affected at all (Fig7.3B). Taken together, these results indicate that in cell culture codon pair deoptimization of VSV glycoprotein gene G leads to a significantly damped protein synthesis, even though virus replication does not seem to be negatively affected.

7.2.5 Quantitative RT-PCR to measure viral RNA synthesis in infected cells

During VSV infection, at each intergenic junction, approximately 30% of the viral RNA polymerases fail to initiate the transcription of the downstream gene (transcription attenuation)(40). Thereby, theoretically the ratio of viral N and G mRNA in VSV-infected cells should be close to 100:35 (100x0.7x0.7x0.7) while the ratio of N and M should be about 100:49 (100x0.7x0.7)(Fig7.4A).

To test whether the reduction of G gene expression during G^{min} virus infection occurred at the transcription level, I quantitated multiple viral gene transcripts in the cell extracts by real-time RT-PCR and estimated their ratios. Interestingly, much less G mRNA, but not M mRNA, was identified in G^{min}-infected cells in comparison to the ones infected with WT virus (Fig7.4B).

7.2.6 Characterization of the G^{min} virus in mice

The virulence of the G^{min} virus was examined in five-to-six weeks old male BALB/c mice. Strikingly, a similar mortality rate was observed in G^{min}-infected groups when the virus was given intranasally. The mice did not develop any neuronal disorder symptoms until 5 dpi, but afterwards they lost weights rapidly and many of them succumbed to infection. The LD_{50} of G^{min} virus was approximately $5x10^5$ PFU (Fig7.5), suggesting that the G^{min} virus was almost as virulent as the WT VSV in mice.

7.3 Discussion

The finding that the G^{min} virus is not attenuated struck us as a surprise, especially considering our successful track record in developing live attenuated vaccine strains by deoptimizing the viral genes that encode structural proteins. The VSV glycoprotein is the only structural protein embedded on the virion surface, which undoubtedly plays a critical role in the activities involved with receptor binding and morphogenesis. Therefore, if G gene expression was reduced as expected, the G^{min} virus would be highly attenuated because of a predictable incompetence of host entry or budding, or both. It could take longer for the accumulation of glycoproteins in the cytoplasm, a prerequisite of the initiation of virus budding; it could also alter the kinetics of virus entry, if glycoprotein was present on the virus surface at a lower density. However, the results described aforementioned suggest that both the growth phenotype and neurovirulence of the G^{min} virus are indistinguishable from its parental strain, even though less glycoprotein is synthesized.

Recently it was reported that rabies virus (RV) glycoprotein was not a dominant factor for virus pathogenicity(145). A recombinant RV that carried a codon deoptimized version of the glycoprotein gene, expressed 2-3 fold less glycoprotein in tissue culture but demonstrated a WT-like mortality rate in experimental animals. Both VSV and RV belong to *Rhabdoviridae*, suggesting that viruses in this family may possess a high tolerance for the shortage of the glycoproteins.

It is of significance to note that in mice the disease onset is slightly delayed, probably because of a transient reduction of viral glycoproteins. Depending on the infection dose, the neuronal disorder symptoms usually appeared between 5 dpi and 10 dpi when the WT virus is given intranasally. During the G^{min} virus infection, however, the symptoms that were caused by virus infection was not observed until 7 dpi, which probably suggests that relative to WT VSV, G^{min} virus is marginally less competent in replication and dissemination *in vivo*.

As shown in Fig7.4, the quantitative RT-PCR assay clearly showed that much less G gene mRNA was available in G^{min}-infected cells, which likely partially contributed to the low level of viral G protein. The untranslated regulatory sequences that flank the G gene ORF were never altered during the recoding, hence the decrease of G gene mRNA was not likely due to an inadequate transcription.

Although the detailed molecular mechanism still remains unknown, we speculate that codon pair deoptimization might affect several mechanisms simultaneously, which interfere with gene expression in a context-dependent manner. For instance, it might be difficult for the ribosome to read through a string of "rare" codon pairs, a predicament leading to less protein production per mRNA. In addition, it is known that the dinucleotide frequency of CpG and UpA in eukaryotic RNA viruses, as well as in mammalian mRNA in general, is lower than expected, a phenomenon that must have been advantageous to viruses. Generating rare codon pairs by rearranging existing synonymous codons inadvertently leads to an enrichment of UpA and CpG dinucleotides in mRNA sequences (with the increase mapping to between codons, since codon use remains unchanged in the recoded sequence). It could be true that the accumulation of rare codon pairs (or dinucleotides) results in an increased instability of the corresponding mRNA.

7.4 Materials and methods

Synthesis of the deoptimized G gene and reverse genetics of the G^{min} virus

The same algorithm described above was employed to design the nucleotide sequence for the G^{min} variant. The entire recoded segment was chemically synthesized by GenScript (NJ, USA) and cloned into the parental VSV antigenomic plasmid pVSV1+ to replace the corresponding WT region. The recombinant virus was recovered using the well-established vaccinia virus-mediated reverse genetics system, followed by plaque purification and sequencing confirmation.

Plaque forming assay and multi-step growth kinetics

Plaque forming assay was conducted in Vero cells as described above. The plaques were developed at either 18 hpi or 48 hpi.

For multi-step growth kinetics, BHK21 cells were infected at a MOI of 0.001. The supernatants were collected at the indicated time points and the virus titers were determined by plaque forming assay in Vero cells.

Western Blot analysis to measure viral protein synthesis in infected cells

BHK21 cells were infected with either WT or G^{min} virus at a MOI of 5. Whole cell lysates were harvested at the indicated time points and probed with anti-VSV G antibody [P5D4] (purchased from Abcam, Inc).

Immunocytochemistry

HeLa cells were infected with either WT or G^{min} virus at a MOI of 5. At 3 hpi, cells were washed with PBS for 4 times, fixed with 3.7% PFA for 30min and then permeabilized with 0.2% saponin/2%BSA/PBS for 15min. For VSV N protein staining, cells were stained with monoclonal antibody [10G4] (purchased from KeraFAST) at a dilution of 1:300 followed by a 1:500 dilution of a secondary anti-mouse antibody conjugated to AlexaFluor-594 (Invitrogen Molecular Probes). For VSV-G protein detection, cells were stained with goat polyclonal antibody to VSV-G tag (FITC labeled, purchased from Abcam, Inc).

Quantitative RT-PCR to measure the mRNA level of multiple viral genes

BHK21 cells were infected with viruses at a MOI of 3. At the indicated time points, total RNA was extracted and reverse transcribed with Oligo $(dT)_{20}$. Quantitative PCR was then carried out with the primers that specifically target viral genes.

Virus infection and LD_{50} determination in mice

Five-to-six weeks old male BALB/c mice were anesthetized with ketamine/xylazine and inoculated intranasally with different doses of G^{min} virus in a total volume of 50 µl. Animal body weights and the symptoms of neurological disorder were monitored on a daily basis for 14 days post infection. The virus LD_{50} was calculated based on the Reed-Muench method.

7.5 Figures and tables

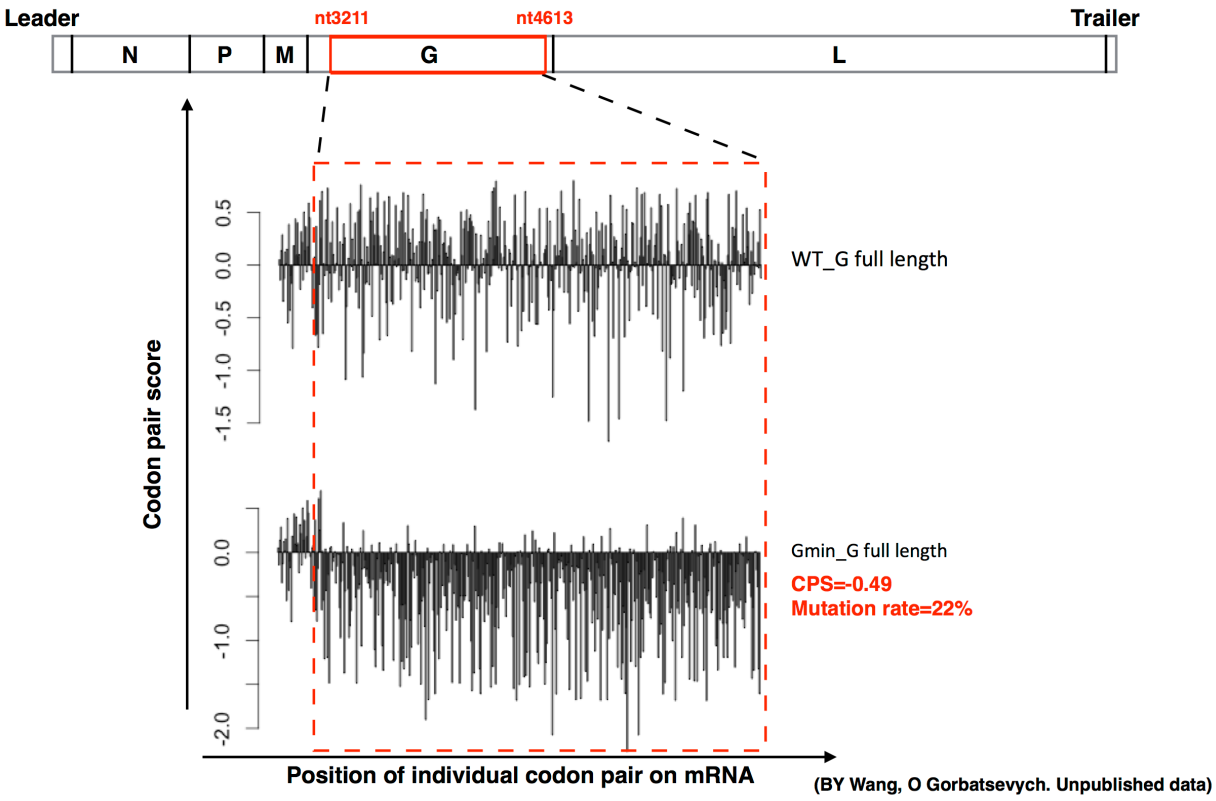


Figure 7.1 The distribution of codon pair bias in WT and G^{min} sequences.
The location of the recoded region in the G ORF was highlighted. The position of each codon pair was marked on the X-axis and the codon pair score (the bar graph) represented the frequency of individual codon pair in the host genome. The non-biased codon pair was set to 0.0. More or less frequently used codon pairs were scaled up or down accordingly.

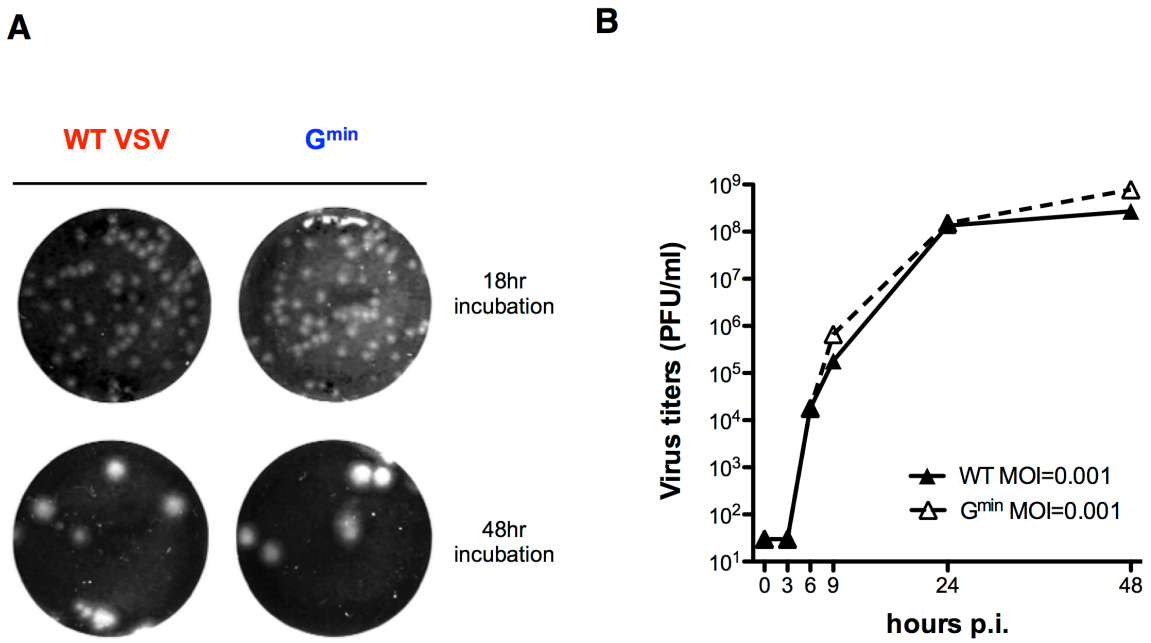


Figure 7.2 Growth phenotype of the G^{min} virus in tissue culture.
(A) Plaque morphology of the G^{min} virus in Vero cells. Plaques were developed at both 18 hpi and 48 hpi. (B) Multi-step growth kinetics of WT VSV and G^{min} virus in BHK21 cells. At the indicated time points, cell supernatants were collected and virus titers were determined by plaque forming assays in Vero cells.

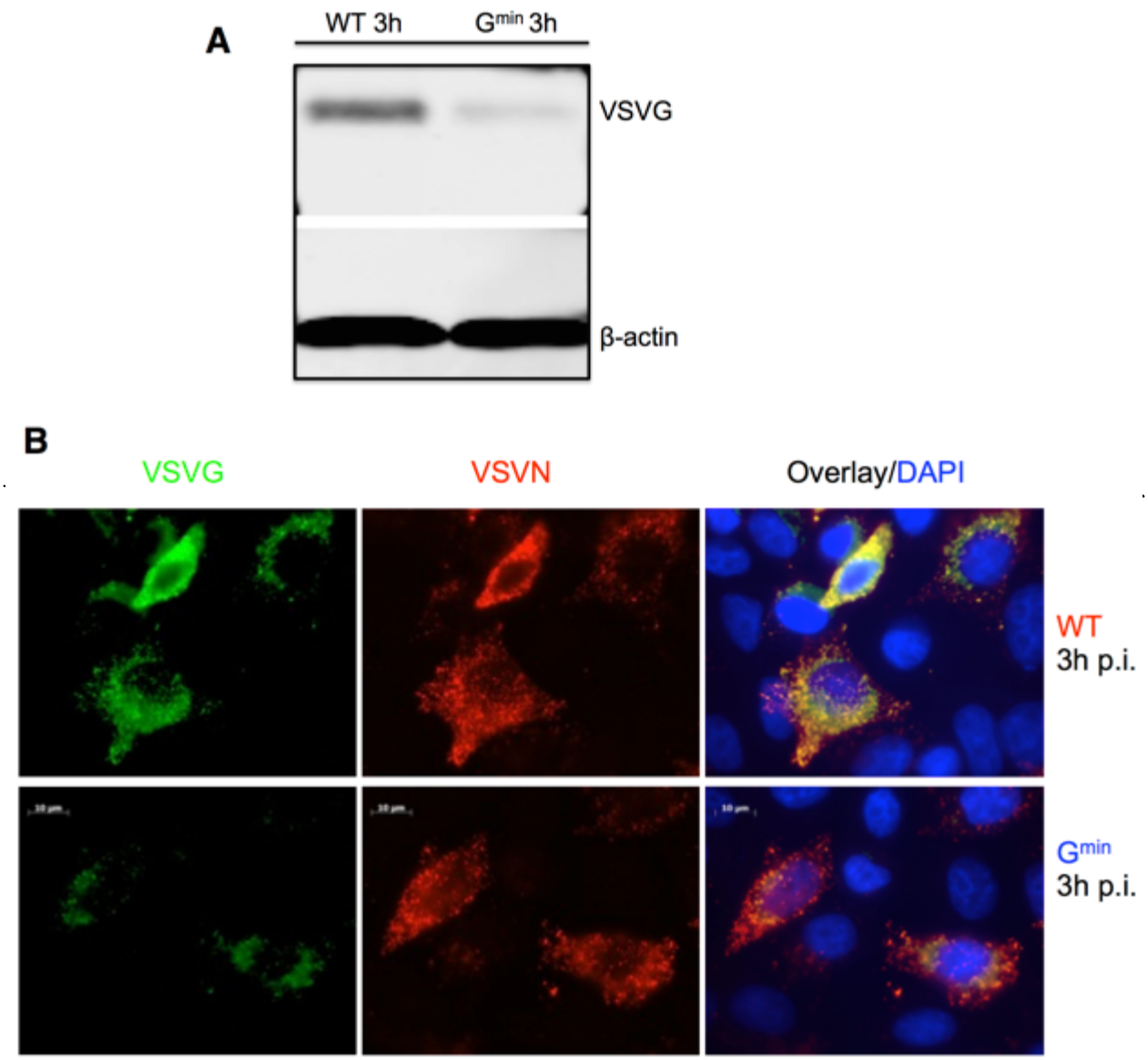


Figure 7.3 Viral protein synthesis in G^{min}-infected cells.
(A) Western Blot analysis to determine the expression level of viral glycoprotein at 3 hpi in infected cells. (B) Immunocytochemistry to measure the expression of viral N and G proteins in infected cells.

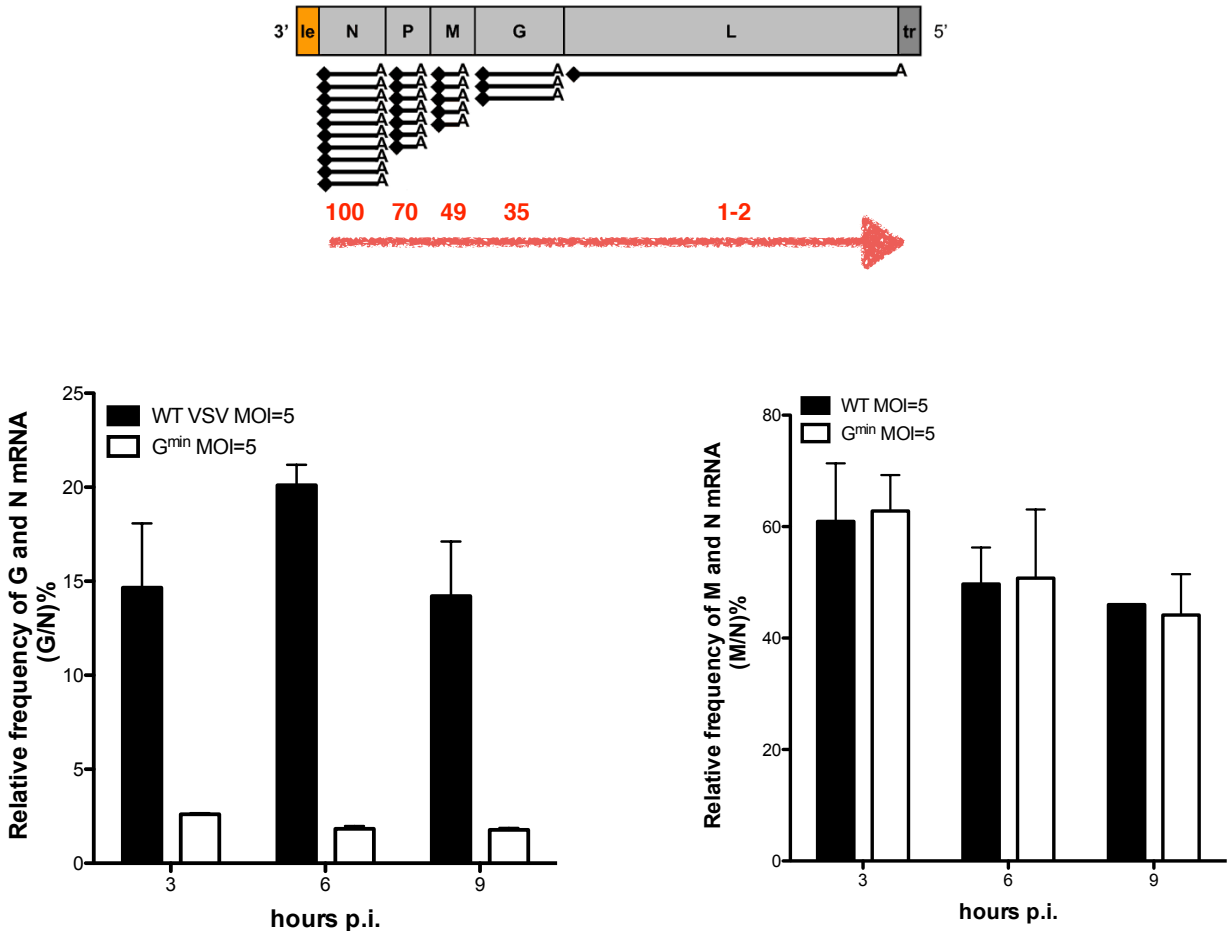


Figure 7.4 Viral RNA synthesis in G^{min}-infected cells.
BHK21 cells were infected with WT or G^{min} virus at a MOI of 3. At the indicated time points, total RNA was extracted and quantitative RT-PCR was performed to analyze the transcription of multiple viral genes.

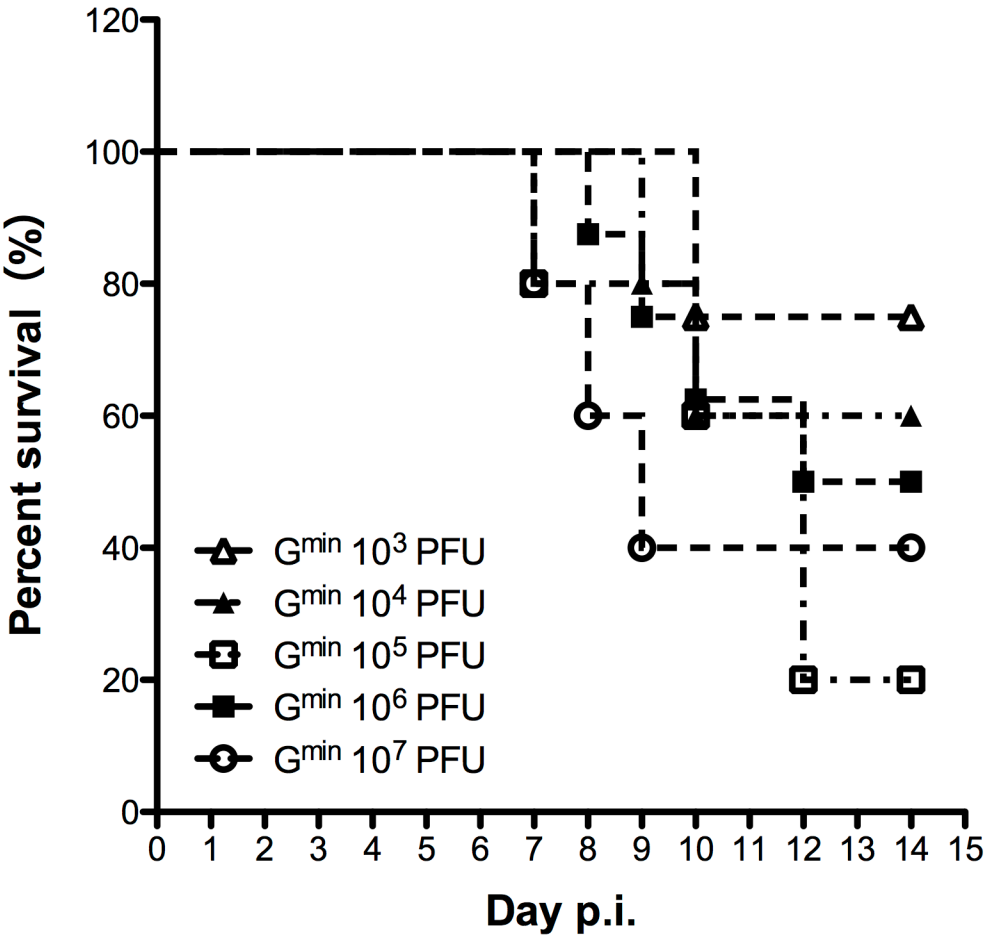


Figure 7.5 Survival rates of G^{min}-infected animals.
Five-to-six male BALB/c mice were infected with different doses of G^{min} virus intranasally. Their body weights and symptoms of neuronal disorder were closely monitored for 14 days after infection.

Chapter 8 Summary and Future directions

Summary

Vesicular stomatitis virus, along with a wide range of mammalian pathogens, belongs to the family of rhabdovirus in the order of *Mononegavirales*. As a well-known animal virus that usually causes vesicular disease in livestock, VSV has served as a model organism for decades, to probe the molecular virology of other notorious human pathogens. It also appears to be an attractive viral platform for the delivery of foreign antigens and anti-cancer regimens, showing great potential in a variety of clinical applications.

However, even though VSV is considered to be "innocuous" in general, this virus is not always as meek as a lamb in human beings. In fact, amongst the human infection cases that are rare and mostly self-restricted, a single case of brain disease in a 3-year-old child has been related to VSV exposure. Also, mounting evidence indicates that WT VSV is neurotoxic in experimental animals, which undoubtedly poses serious safety concerns for the direct use of WT VSV in clinical practices.

Numerous efforts have been invested during the past few years to improve the safety profile of VSV backbone, most of which involved a limited number of point mutations mapped in viral genome. However, RNA viruses, including VSV, exhibit a much higher spontaneous mutation rate than any other organisms, a natural obstacle that the conventional vaccine strategies can hardly overcome.

To address the unmet medical needs aforementioned, my research focuses on the development of a novel, live attenuated VSV vaccine candidate. Specifically, to achieve a desirable balance between safety and efficacy, I adopted a unique large-scale genome recoding strategy based on codon pair bias (CPB), which is a highly conserved but under-appreciated genomic phenomenon. The results described here characterize three newly designed and synthesized versions of VSV, $L1^{sdmax}$, $L1^{min}$ and G^{min} virus, using two distinct recoding strategies "sdmax" and "min". Nucleotide changes were made in the regions of the viral L and G genes by shuffling the positions of synonymous codons without changing either codon usage or amino acid sequence. Specifically for the $L1^{sdmax}$ design, a total of 858 silent mutations were introduced into the 5'-terminal segment of the VSV polymerase gene L, creating hundreds of

overrepresented codon pairs. In contrast, for the $L1^{min}$ design, 623 synonymous nucleotide substitutions were put into the same L1 region, resulting in a much lower codon pair bias score relative to the corresponding WT sequence. Similarly, the G^{min} design contains 323 silent mutations, accounting for 22% of the full-length G gene sequence.

As expected, the $L1^{min}$ design led to a non-viable virus, most likely due to a poor L gene expression during the reverse genetics. Strikingly, the $L1^{sdmax}$ design resulted in a modest attenuation in cell culture, as judged by a smaller plaque size, delayed release of infectious viral progeny and decreased yields of viral products. The attenuation of the $L1^{sdmax}$ virus was more pronounced in a mouse model, which was evidenced by a low virulence, considerably less virus burden as well as diminished histopathological changes in infected tissues. Moreover, the $L1^{sdmax}$ virus demonstrated a superior immunogenicity that elicited a potent humoral response and fully protected vaccinated mice from a lethal WT challenge. On the other hand, the G^{min} design was a disappointment in terms of vaccine development. This variant replicated just as well as WT VSV in the cell culture and showed neurovirulence in mice, even though it is less competent in producing viral glycoproteins.

The detailed molecular mechanism underlying virus attenuation still remains enigmatic and codon pair bias seems to affect multiple aspects of gene expression in a context-dependent manner. For instance, deoptimization of the viral glycoprotein gene G results in a reduction of the G mRNA, providing compelling evidence that nucleotide sequence containing more under-represented codon pairs might be favored for *in vivo* degradation. As for the $L1^{sdmax}$ design, the substantial codon shuffling that disrupted the natural distribution pattern of "good" or "bad" codon pairs in the L gene ORF might perturb the coordination between ribosome translocation and other events related to translation, such as co-translational folding.

Taken together, in this study, recombinant VSVs with diverse codon pair biases were rationally designed, constructed and analyzed in both cell culture and a rodent model. One such recombinant virus, $L1^{sdmax}$, showed promise as an effective vaccine candidate as it demonstrated a favorable safety profile but still retained the WT-like immunogenicity, which is reminiscent of the encouraging results of poliovirus genome recoding. In addition, $L1^{sdmax}$ virus has the great potential for scaled-up production at an affordable cost, given the fact that this virus can easily grow up to a high titer in cell culture. My study not only expands the existing approaches to

attenuate VSV for the prophylactic treatment of vesicular disease, but may also serve as a paradigm for the vaccine development of other negative strand RNA viruses.

Future directions

Studies are currently underway in the lab and other collaborative groups, to understand the molecular mechanism underlying codon pair bias-mediated virus attenuation. The multiple virus constructs characterized in this study, could be subjected to ribosome profiling assay to determine whether the codon pairs or other closely related genetic patterns (e.g. CpG/UpA dinucleotides) play a role in regulating the ribosome translocation rate along a messenger RNA. Detailed analysis of RNA stability would also be conducted to test the hypothesis that increasing underrepresented codon pairs in a given ORF destabilizes the RNA transcript.

As mentioned above, according to the working model of codon pair bias, the alteration of gene expression is presumably an accumulating effect, resulting from the numerous synonymous substitutions. This unique feature of virus attenuation caused by codon pair bias is critical for the development of vaccine strain against RNA viruses, whose spontaneous mutation rates are much higher than other microorganisms. Compared to the conventional attenuation strategy that usually depends on a few point mutations, one of the anticipated advantages of my approach is the stability of the viral genetic material, which could make reversion nearly impossible. In fact, I continuously passaged the $L1^{sdmax}$ virus in BHK21 cells for 5 times at low MOI and did not observe any sign of phenotypic reversion. Nevertheless, the long-term passaging across multiple cell types *in vitro* is required to validate the finding. Also the genetic stability of this virus variant *in vivo* remained to be further investigated.

A striking observation in this study is that $L1^{sdmax}$ virus underwent a limited replication and dissemination that caused little morbidity, even when they were given to animals at an extremely high dose. The unnoticeable virus infection itself is still capable of eliciting a robust immune response. Nonetheless, the current construct may have to further attenuated, probably combined with certain existing attenuation strategy, to avoid all proliferation in the central nervous system.

The recent Ebola outbreak in West Africa and the sporadic cases in the Western Hemisphere have elicited international concern. One of the most promising vaccine candidates against Ebola infection, VSV-ZEBOV, is a VSV chimera that expresses Ebola genetic information. It will be of interest to further evaluate the safety and efficacy of $L1^{sdmax}$ as a novel platform to deliver Ebola antigens. In addition, certain human paramyxoviruses, including human respiratory syncytial virus (hRSV) and human metapneumovirus (hMPV), are also

potential viral pathogens, against which the $L1^{sdmax}$-based vaccine vehicles could be developed. As these viruses account for the majority of respiratory diseases in infants/elderly and their natural infection is via mucosal surface, the use of $L1^{sdmx}$-based antigen delivery is predicted to confer an efficacious mucosal immunity without causing any neurological complications.

References

1. **Brown F, Bishop DH, Crick J, Francki RI, Holland JJ, Hull R, Johnson K, Martelli G, Murphy FA, Obijeski JF, Peters D, Pringle CR, Reichmann ME, Schneider LG, Shope RE, Simpson DI, Summers DF, Wagner RR.** 1979. Rhabdoviridae. Report of the Rhabdovirus Study Group, International Committee on Taxonomy of Viruses. Intervirology **12:**1-7.
2. **Lichty BD, Power AT, Stojdl DF, Bell JC.** 2004. Vesicular stomatitis virus: re-inventing the bullet. Trends in molecular medicine **10:**210-216.
3. **Mead DG, Ramberg FB, Besselsen DG, Mare CJ.** 2000. Transmission of vesicular stomatitis virus from infected to noninfected black flies co-feeding on nonviremic deer mice. Science **287:**485-487.
4. **Stallknecht DE.** 2000. VSV-NJ on Ossabaw Island, Georgia. The truth is out there. Annals of the New York Academy of Sciences **916:**431-436.
5. **Quiroz E, Moreno N, Peralta PH, Tesh RB.** 1988. A human case of encephalitis associated with vesicular stomatitis virus (Indiana serotype) infection. The American journal of tropical medicine and hygiene **39:**312-314.
6. **Roberts A, Buonocore L, Price R, Forman J, Rose JK.** 1999. Attenuated vesicular stomatitis viruses as vaccine vectors. Journal of virology **73:**3723-3732.
7. **Sabin AB, Olitsky PK.** 1937. Influence of Host Factors on Neuroinvasiveness of Vesicular Stomatitis Virus : Ii. Effect of Age on the Invasion of the Peripheral and Central Nervous Systems by Virus Injected into the Leg Muscles or the Eye. The Journal of experimental medicine **66:**35-57.
8. **Sabin AB, Olitsky PK.** 1937. Influence of Host Factors on Neuroinvasiveness of Vesicular Stomatitis Virus : I. Effect of Age on the Invasion of the Brain by Virus Instilled in the Nose. The Journal of experimental medicine **66:**15-34.
9. **Sabin AB, Olitsky PK.** 1938. Influence of Host Factors on Neuroinvasiveness of Vesicular Stomatitis Virus : Iv. Variations in Neuroinvasiveness in Difierent Species. The Journal of experimental medicine **67:**229-249.
10. **Sabin AB, Olitsky PK.** 1938. Influence of Host Factors on Neuroinvasiveness of Vesicular Stomatitis Virus : Iii. Effect of Age and Pathway of Infection on the Character and Localization of Lesions in the Central Nervous System. The Journal of experimental medicine **67:**201-228.
11. **Plakhov IV, Arlund EE, Aoki C, Reiss CS.** 1995. The earliest events in vesicular stomatitis virus infection of the murine olfactory neuroepithelium and entry of the central nervous system. Virology **209:**257-262.
12. **Cornish TE, Stallknecht DE, Brown CC, Seal BS, Howerth EW.** 2001. Pathogenesis of experimental vesicular stomatitis virus (New Jersey serotype) infection in the deer mouse (Peromyscus maniculatus). Veterinary pathology **38:**396-406.

13. **Johnson JE, Nasar F, Coleman JW, Price RE, Javadian A, Draper K, Lee M, Reilly PA, Clarke DK, Hendry RM, Udem SA.** 2007. Neurovirulence properties of recombinant vesicular stomatitis virus vectors in non-human primates. Virology **360:**36-49.

14. **Schellekens H, Smiers-de Vreede E, de Reus A, Dijkema R.** 1984. Antiviral activity of interferon in rats and the effect of immune suppression. The Journal of general virology **65 (Pt 2):**391-396.

15. **Shinozaki K, Ebert O, Suriawinata A, Thung SN, Woo SL.** 2005. Prophylactic alpha interferon treatment increases the therapeutic index of oncolytic vesicular stomatitis virus virotherapy for advanced hepatocellular carcinoma in immune-competent rats. Journal of virology **79:**13705-13713.

16. **Fenner FJ.** 1968. The pathogenesis of viral infections: the influence of age on resistance to viral infections. Academic Press, Inc., New York **vol. 2.**

17. **Collins PL, Melero JA.** 2011. Progress in understanding and controlling respiratory syncytial virus: Still crazy after all these years. Virus Res **162:**80-99.

18. **van Bleek GM, Osterhaus AD, de Swart RL.** 2011. RSV 2010: Recent advances in research on respiratory syncytial virus and other pneumoviruses. Vaccine **29:**7285-7291.

19. **Roberts A, Kretzschmar E, Perkins AS, Forman J, Price R, Buonocore L, Kawaoka Y, Rose JK.** 1998. Vaccination with a recombinant vesicular stomatitis virus expressing an influenza virus hemagglutinin provides complete protection from influenza virus challenge. Journal of virology **72:**4704-4711.

20. !!! INVALID CITATION !!!

21. **Schlereth B, Rose JK, Buonocore L, ter Meulen V, Niewiesk S.** 2000. Successful vaccine-induced seroconversion by single-dose immunization in the presence of measles virus-specific maternal antibodies. Journal of virology **74:**4652-4657.

22. **Reuter JD, Vivas-Gonzalez BE, Gomez D, Wilson JH, Brandsma JL, Greenstone HL, Rose JK, Roberts A.** 2002. Intranasal vaccination with a recombinant vesicular stomatitis virus expressing cottontail rabbit papillomavirus L1 protein provides complete protection against papillomavirus-induced disease. Journal of virology **76:**8900-8909.

23. **Mire CE, Versteeg KM, Cross RW, Agans KN, Fenton KA, Whitt MA, Geisbert TW.** 2013. Single injection recombinant vesicular stomatitis virus vaccines protect ferrets against lethal Nipah virus disease. Virology journal **10:**353.

24. **Mire CE, Miller AD, Carville A, Westmoreland SV, Geisbert JB, Mansfield KG, Feldmann H, Hensley LE, Geisbert TW.** 2012. Recombinant vesicular stomatitis virus vaccine vectors expressing filovirus glycoproteins lack neurovirulence in nonhuman primates. PLoS neglected tropical diseases **6:**e1567.

25. **Mire CE, Geisbert JB, Agans KN, Satterfield BA, Versteeg KM, Fritz EA, Feldmann H, Hensley LE, Geisbert TW.** 2014. Durability of a vesicular stomatitis virus-based marburg virus vaccine in nonhuman primates. Plos One **9:**e94355.

26. **Kapadia SU, Rose JK, Lamirande E, Vogel L, Subbarao K, Roberts A.** 2005. Long-term protection from SARS coronavirus infection conferred by a single immunization with an attenuated VSV-based vaccine. Virology **340:**174-182.

27. **Kahn JS, Roberts A, Weibel C, Buonocore L, Rose JK.** 2001. Replication-competent or attenuated, nonpropagating vesicular stomatitis viruses expressing respiratory syncytial virus (RSV) antigens protect mice against RSV challenge. Journal of virology **75:**11079-11087.

28. **Geisbert TW, Geisbert JB, Leung A, Daddario-DiCaprio KM, Hensley LE, Grolla A, Feldmann H.** 2009. Single-injection vaccine protects nonhuman primates against infection with marburg virus and three species of ebola virus. Journal of virology **83:**7296-7304.

29. **Geisbert TW, Feldmann H.** 2011. Recombinant vesicular stomatitis virus-based vaccines against Ebola and Marburg virus infections. The Journal of infectious diseases **204 Suppl 3:**S1075-1081.

30. **Geisbert TW, Daddario-Dicaprio KM, Lewis MG, Geisbert JB, Grolla A, Leung A, Paragas J, Matthias L, Smith MA, Jones SM, Hensley LE, Feldmann H, Jahrling PB.** 2008. Vesicular stomatitis virus-based ebola vaccine is well-tolerated and protects immunocompromised nonhuman primates. PLoS pathogens **4:**e1000225.

31. **Garbutt M, Liebscher R, Wahl-Jensen V, Jones S, Moller P, Wagner R, Volchkov V, Klenk HD, Feldmann H, Stroher U.** 2004. Properties of replication-competent vesicular stomatitis virus vectors expressing glycoproteins of filoviruses and arenaviruses. Journal of virology **78:**5458-5465.

32. **Cobleigh MA, Wei X, Robek MD.** 2013. A vesicular stomatitis virus-based therapeutic vaccine generates a functional CD8 T cell response to hepatitis B virus in transgenic mice. Journal of virology **87:**2969-2973.

33. **Cobleigh MA, Buonocore L, Uprichard SL, Rose JK, Robek MD.** 2010. A vesicular stomatitis virus-based hepatitis B virus vaccine vector provides protection against challenge in a single dose. Journal of virology **84:**7513-7522.

34. **Stojdl DF, Lichty BD, tenOever BR, Paterson JM, Power AT, Knowles S, Marius R, Reynard J, Poliquin L, Atkins H, Brown EG, Durbin RK, Durbin JE, Hiscott J, Bell JC.** 2003. VSV strains with defects in their ability to shutdown innate immunity are potent systemic anti-cancer agents. Cancer cell **4:**263-275.

35. **Stojdl DF, Lichty B, Knowles S, Marius R, Atkins H, Sonenberg N, Bell JC.** 2000. Exploiting tumor-specific defects in the interferon pathway with a previously unknown oncolytic virus. Nature medicine **6:**821-825.

36. **van den Pol AN, Davis JN.** 2013. Highly attenuated recombinant vesicular stomatitis virus VSV-12'GFP displays immunogenic and oncolytic activity. Journal of virology **87:**1019-1034.

37. **Ozduman K, Wollmann G, Piepmeier JM, van den Pol AN.** 2008. Systemic vesicular stomatitis virus selectively destroys multifocal glioma and metastatic carcinoma in brain. The Journal of neuroscience : the official journal of the Society for Neuroscience **28:**1882-1893.

38. **Fang XK, Zhang SK, Sun XD, Li JJ, Sun T.** 2012. Evaluation of attenuated VSVs with mutated M or/and G proteins as vaccine vectors. Vaccine **30:**1313-1321.

39. **Hastie E, Grdzelishvili VZ.** 2012. Vesicular stomatitis virus as a flexible platform for oncolytic virotherapy against cancer. The Journal of general virology **93:**2529-2545.

40. **Barr JN, Whelan SP, Wertz GW.** 2002. Transcriptional control of the RNA-dependent RNA polymerase of vesicular stomatitis virus. Biochimica et biophysica acta **1577:**337-353.

41. **Sleat DE, Banerjee AK.** 1993. Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. Journal of virology **67:**1334-1339.

42. **Rahmeh AA, Li J, Kranzusch PJ, Whelan SP.** 2009. Ribose 2'-O methylation of the vesicular stomatitis virus mRNA cap precedes and facilitates subsequent guanine-N-7 methylation by the large polymerase protein. Journal of virology **83:**11043-11050.

43. **Ogino T, Banerjee AK.** 2007. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. Molecular cell **25:**85-97.

44. **Li J, Rahmeh A, Morelli M, Whelan SP.** 2008. A conserved motif in region v of the large polymerase proteins of nonsegmented negative-sense RNA viruses that is essential for mRNA capping. Journal of virology **82:**775-784.

45. **Li J, Fontaine-Rodriguez EC, Whelan SP.** 2005. Amino acid residues within conserved domain VI of the vesicular stomatitis virus large polymerase protein essential for mRNA cap methyltransferase activity. Journal of virology **79:**13373-13384.

46. **Grdzelishvili VZ, Smallwood S, Tower D, Hall RL, Hunt DM, Moyer SA.** 2005. A single amino acid change in the L-polymerase protein of vesicular stomatitis virus completely abolishes viral mRNA cap methylation. Journal of virology **79:**7327-7337.

47. **Emerson SU, Wagner RR.** 1973. L protein requirement for in vitro RNA synthesis by vesicular stomatitis virus. Journal of virology **12:**1325-1335.

48. **Resa-Infante P, Jorba N, Coloma R, Ortin J.** 2011. The influenza virus RNA synthesis machine: advances in its structure and function. RNA biology **8:**207-215.

49. **Das K, Aramini JM, Ma LC, Krug RM, Arnold E.** 2010. Structures of influenza A proteins and insights into antiviral drug targets. Nature structural & molecular biology **17:**530-538.

50. **Poch O, Blumberg BM, Bougueleret L, Tordo N.** 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. The Journal of general virology **71 (Pt 5):**1153-1162.

51. **Rahmeh AA, Schenk AD, Danek EI, Kranzusch PJ, Liang B, Walz T, Whelan SP.** 2010. Molecular architecture of the vesicular stomatitis virus RNA polymerase. Proceedings of the National Academy of Sciences of the United States of America **107:**20075-20080.

52. **Li J, Rahmeh A, Brusic V, Whelan SP.** 2009. Opposing effects of inhibiting cap addition and cap methylation on polyadenylation during vesicular stomatitis virus mRNA synthesis. Journal of virology **83:**1930-1940.

53. **Leyrat C, Yabukarski F, Tarbouriech N, Ribeiro EA, Jr., Jensen MR, Blackledge M, Ruigrok RW, Jamin M.** 2011. Structure of the vesicular stomatitis virus N(0)-P complex. PLoS pathogens **7:**e1002248.

54. **Leyrat C, Schneider R, Ribeiro EA, Jr., Yabukarski F, Yao M, Gerard FC, Jensen MR, Ruigrok RW, Blackledge M, Jamin M.** 2012. Ensemble structure of the modular and flexible full-length vesicular stomatitis virus phosphoprotein. Journal of molecular biology **423:**182-197.

55. **Green TJ, Macpherson S, Qiu S, Lebowitz J, Wertz GW, Luo M.** 2000. Study of the assembly of vesicular stomatitis virus N protein: role of the P protein. Journal of virology **74:**9515-9524.

56. **Green TJ, Luo M.** 2009. Structure of the vesicular stomatitis virus nucleocapsid in complex with the nucleocapsid-binding domain of the small polymerase cofactor, P. Proceedings of the National Academy of Sciences of the United States of America **106:**11713-11718.

57. **Mire CE, White JM, Whitt MA.** 2010. A spatio-temporal analysis of matrix protein and nucleocapsid trafficking during vesicular stomatitis virus uncoating. PLoS pathogens **6:**e1000994.

58. **von Kobbe C, van Deursen JM, Rodrigues JP, Sitterlin D, Bachi A, Wu X, Wilm M, Carmo-Fonseca M, Izaurralde E.** 2000. Vesicular stomatitis virus matrix protein inhibits host cell gene expression by targeting the nucleoporin Nup98. Molecular cell **6:**1243-1252.

59. **Rajani KR, Pettit Kneller EL, McKenzie MO, Horita DA, Chou JW, Lyles DS.** 2012. Complexes of vesicular stomatitis virus matrix protein with host Rae1 and Nup98 involved in inhibition of host transcription. PLoS pathogens **8:**e1002929.

60. **Quan B, Seo HS, Blobel G, Ren Y.** 2014. Vesiculoviral matrix (M) protein occupies nucleic acid binding site at nucleoporin pair (Rae1 * Nup98). Proceedings of the National Academy of Sciences of the United States of America **111:**9127-9132.

61. **Jayakar HR, Murti KG, Whitt MA.** 2000. Mutations in the PPPY motif of vesicular stomatitis virus matrix protein reduce virus budding by inhibiting a late step in virion release. Journal of virology **74:**9818-9827.

62. **Irie T, Carnero E, Okumura A, Garcia-Sastre A, Harty RN.** 2007. Modifications of the PSAP region of the matrix protein lead to attenuation of vesicular stomatitis virus in vitro and in vivo. The Journal of general virology **88:**2559-2567.

63. **Obiang L, Raux H, Ouldali M, Blondel D, Gaudin Y.** 2012. Phenotypes of vesicular stomatitis virus mutants with mutations in the PSAP motif of the matrix protein. The Journal of general virology **93:**857-865.

64. **Irie T, Licata JM, Harty RN.** 2005. Functional characterization of Ebola virus L-domains using VSV recombinants. Virology **336:**291-298.

65. **Compton T, Ivanov IE, Gottlieb T, Rindler M, Adesnik M, Sabatini DD.** 1989. A sorting signal for the basolateral delivery of the vesicular stomatitis virus (VSV) G protein lies in its luminal domain: analysis of the targeting of VSV G-influenza hemagglutinin chimeras. Proceedings of the National Academy of Sciences of the United States of America **86:**4112-4116.

66. **Finkelshtein D, Werman A, Novick D, Barak S, Rubinstein M.** 2013. LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. Proceedings of the National Academy of Sciences of the United States of America **110:**7306-7311.

67. **Burns JC, Friedmann T, Driever W, Burrascano M, Yee JK.** 1993. Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. Proceedings of the National Academy of Sciences of the United States of America **90:**8033-8037.

68. **Cureton DK, Massol RH, Saffarian S, Kirchhausen TL, Whelan SP.** 2009. Vesicular stomatitis virus enters cells through vesicles incompletely coated with clathrin that depend upon actin for internalization. PLoS pathogens **5:**e1000394.

69. **Cureton DK, Massol RH, Whelan SP, Kirchhausen T.** 2010. The length of vesicular stomatitis virus particles dictates a need for actin assembly during clathrin-dependent endocytosis. PLoS pathogens **6:**e1001127.

70. **Mire CE, Dube D, Delos SE, White JM, Whitt MA.** 2009. Glycoprotein-dependent acidification of vesicular stomatitis virus enhances release of matrix protein. Journal of virology **83:**12139-12150.

71. **Heinrich BS, Cureton DK, Rahmeh AA, Whelan SP.** 2010. Protein expression redirects vesicular stomatitis virus RNA synthesis to cytoplasmic inclusions. PLoS pathogens **6:**e1000958.

72. **Jayakar HR, Jeetendra E, Whitt MA.** 2004. Rhabdovirus assembly and budding. Virus Res **106:**117-132.

73. **Robison CS, Whitt MA.** 2000. The membrane-proximal stem region of vesicular stomatitis virus G protein confers efficient virus assembly. Journal of virology **74:**2239-2246.

74. **Chen BJ, Lamb RA.** 2008. Mechanisms for enveloped virus budding: can some viruses do without an ESCRT? Virology **372:**221-232.

75. **Ingham RJ, Gish G, Pawson T.** 2004. The Nedd4 family of E3 ubiquitin ligases: functional diversity within a common modular architecture. Oncogene **23:**1972-1984.

76. **Han Z, Lu J, Liu Y, Davis B, Lee MS, Olson MA, Ruthel G, Freedman BD, Schnell MJ, Wrobel JE, Reitz AB, Harty RN.** 2014. Small-molecule probes targeting the viral PPxY-host Nedd4 interface block egress of a broad range of RNA viruses. Journal of virology **88:**7294-7306.

77. **Moerdyk-Schauwecker M, Hwang SI, Grdzelishvili VZ.** 2014. Cellular proteins associated with the interior and exterior of vesicular stomatitis virus virions. Plos One **9:**e104688.

78. **Moerdyk-Schauwecker M, Hwang SI, Grdzelishvili VZ.** 2009. Analysis of virion associated host proteins in vesicular stomatitis virus using a proteomics approach. Virology journal **6:**166.

79. **Plotkin JB, Robins H, Levine AJ.** 2004. Tissue-specific codon usage and the expression of human genes. Proceedings of the National Academy of Sciences of the United States of America **101:**12588-12591.

80. **Gustafsson C, Govindarajan S, Minshull J.** 2004. Codon bias and heterologous protein expression. Trends in biotechnology **22:**346-353.

81. **Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E.** 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.

82. **Burns CC, Shaw J, Campagnoli R, Jorba J, Vincent A, Quay J, Kew O.** 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.

83. **Gutman GA, Hatfield GW.** 1989. Nonrandom utilization of codon pairs in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America **86:**3699-3703.

84. **Boycheva S, Chkodrov G, Ivanov I.** 2003. Codon pairs in the genome of Escherichia coli. Bioinformatics **19:**987-998.

85. **Wang FP, Li H.** 2009. Codon-pair usage and genome evolution. Gene **433:**8-15.

86. **Moura G, Pinheiro M, Arrais J, Gomes AC, Carreto L, Freitas A, Oliveira JL, Santos MAS.** 2007. Large Scale Comparative Codon-Pair Context Analysis Unveils General Rules that Fine-Tune Evolution of mRNA Primary Structure. Plos One **2**.

87. **Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S.** 2008. Virus attenuation by genome-scale changes in codon pair bias. Science **320:**1784-1787.

88. **Cello J, Paul AV, Wimmer E.** 2002. Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. Science **297:**1016-1018.

89. **Wimmer E.** 2006. The test-tube synthesis of a chemical called poliovirus. The simple synthesis of a virus has far-reaching societal implications. EMBO reports **7 Spec No:**S3-9.

90. **Wimmer E, Mueller S, Tumpey TM, Taubenberger JK.** 2009. Synthetic viruses: a new opportunity to understand and prevent viral disease. Nature biotechnology **27:**1163-1172.

91. **Mueller S, Coleman JR, Wimmer E.** 2009. Putting Synthesis into Biology: A Viral View of Genetic Engineering through De Novo Gene and Genome Synthesis. Chem Biol **16:**337-347.

92. **Le Nouen C, Brock LG, Luongo C, McCarty T, Yang L, Mehedi M, Wimmer E, Mueller S, Collins PL, Buchholz UJ, DiNapoli JM.** 2014. Attenuation of human

respiratory syncytial virus by genome-scale codon-pair deoptimization. Proceedings of the National Academy of Sciences of the United States of America **111:**13169-13174.

93. **Yang C, Skiena S, Futcher B, Mueller S, Wimmer E.** 2013. Deliberate reduction of hemagglutinin and neuraminidase expression of influenza virus leads to an ultraprotective live vaccine in mice. Proceedings of the National Academy of Sciences of the United States of America **110:**9481-9486.

94. **Sitaraman V, Hearing P, Ward CB, Gnatenko DV, Wimmer E, Mueller S, Skiena S, Bahou WF.** 2011. Computationally designed adeno-associated virus (AAV) Rep 78 is efficiently maintained within an adenovirus vector. Proceedings of the National Academy of Sciences of the United States of America **108:**14294-14299.

95. **Mueller S, Coleman JR, Papamichail D, Ward CB, Nimnual A, Futcher B, Skiena S, Wimmer E.** 2010. Live attenuated influenza virus vaccines by computer-aided rational design. Nature biotechnology **28:**723-726.

96. **Combe M, Sanjuan R.** 2014. Variation in RNA virus mutation rates across host cells. PLoS pathogens **10:**e1003855.

97. **Song YT, Liu Y, Ward CB, Mueller S, Futcher B, Skiena S, Paul AV, Wimmer E.** 2012. Identification of two functionally redundant RNA elements in the coding sequence of poliovirus using computer-generated design. Proceedings of the National Academy of Sciences of the United States of America **109:**14301-14307.

98. **Lawson ND, Stillman EA, Whitt MA, Rose JK.** 1995. Recombinant vesicular stomatitis viruses from DNA. Proceedings of the National Academy of Sciences of the United States of America **92:**4477-4481.

99. **Whelan SP, Ball LA, Barr JN, Wertz GT.** 1995. Efficient recovery of infectious vesicular stomatitis virus entirely from cDNA clones. Proceedings of the National Academy of Sciences of the United States of America **92:**8388-8392.

100. **Harty RN, Brown ME, Hayes FP, Wright NT, Schnell MJ.** 2001. Vaccinia virus-free recovery of vesicular stomatitis virus. Journal of molecular microbiology and biotechnology **3:**513-517.

101. **Cyrklaff M, Risco C, Fernandez JJ, Jimenez MV, Esteban M, Baumeister W, Carrascosa JL.** 2005. Cryo-electron tomography of vaccinia virus. Proceedings of the National Academy of Sciences of the United States of America **102:**2772-2777.

102. **Hodges EN, Connor JH.** 2013. Translational control by negative-strand RNA viruses: methods for the study of a crucial virus/host interaction. Methods **59:**180-187.

103. 2004. Biology of negative strand RNA viruses : the power of reverse genetics. Springer, Berlin New York.

104. **Rosales R, Sutter G, Moss B.** 1994. A Cellular Factor Is Required for Transcription of Vaccinia Viral Intermediate-Stage Genes. Proceedings of the National Academy of Sciences of the United States of America **91:**3794-3798.

105. **Pekosz A, He B, Lamb RA.** 1999. Reverse genetics of negative-strand RNA viruses: closing the circle. Proceedings of the National Academy of Sciences of the United States of America **96:**8804-8806.

106. **Neumann G, Whitt MA, Kawaoka Y.** 2002. A decade after the generation of a negative-sense RNA virus from cloned cDNA - what have we learned? The Journal of general virology **83:**2635-2662.

107. **Fuerst TR, Niles EG, Studier FW, Moss B.** 1986. Eukaryotic Transient-Expression System Based on Recombinant Vaccinia Virus That Synthesizes Bacteriophage-T7 Rna-Polymerase. Proceedings of the National Academy of Sciences of the United States of America **83:**8122-8126.

108. **Tekes G, Rahmeh AA, Whelan SP.** 2011. A freeze frame view of vesicular stomatitis virus transcription defines a minimal length of RNA for 5' processing. PLoS pathogens **7:**e1002073.

109. **Shen SH, Stauft CB, Gorbatsevych O, Song Y, Ward CB, Yurovsky A, Mueller S, Futcher B, Wimmer E.** 2015. Large-scale recoding of an arbovirus genome to rebalance its insect versus mammalian preference. Proceedings of the National Academy of Sciences of the United States of America **112:**4749-4754.

110. **Kranzusch PJ, Whelan SP.** 2012. Architecture and regulation of negative-strand viral enzymatic machinery. RNA biology **9:**941-948.

111. **Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, Pan T, Dahan O, Furman I, Pilpel Y.** 2010. An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell **141:**344-354.

112. **Buggele WA, Horvath CM.** 2013. MicroRNA Profiling of Sendai Virus-Infected A549 Cells Identifies miR-203 as an Interferon-Inducible Regulator of IFIT1/ISG56. Journal of virology **87:**9260-9270.

113. **Valle M.** 1971. Factors affecting plaque assay of animal viruses--with special reference to vesicular stomatitis and vaccinia virus. Acta pathologica et microbiologica Scandinavica. Section B: Microbiology and immunology **219:**Suppl 219:211-269.

114. **El Asmi F, Maroui MA, Dutrieux J, Blondel D, Nisole S, Chelbi-Alix MK.** 2014. Implication of PMLIV in Both Intrinsic and Innate Immunity. PLoS pathogens **10**.

115. **Thompson KA, Yin J.** 2010. Population dynamics of an RNA virus and its defective interfering particles in passage cultures. Virology journal **7:**257.

116. **Irwin B, Heck JD, Hatfield GW.** 1995. Codon pair utilization biases influence translational elongation step times. The Journal of biological chemistry **270:**22801-22806.

117. **Angov E.** 2011. Codon usage: nature's roadmap to expression and folding of proteins. Biotechnology journal **6:**650-659.

118. **Ciryam P, Morimoto RI, Vendruscolo M, Dobson CM, O'Brien EP.** 2013. In vivo translation rates can substantially delay the cotranslational folding of the Escherichia coli cytosolic proteome. Proceedings of the National Academy of Sciences of the United States of America **110:**E132-140.

119. **Komar AA.** 2009. A pause for thought along the co-translational folding pathway. Trends in biochemical sciences **34:**16-24.

120. **Komar AA, Lesnik T, Reiss C.** 1999. Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. FEBS letters **462:**387-391.

121. **Pechmann S, Frydman J.** 2013. Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. Nature structural & molecular biology **20:**237-243.

122. **Zhang G, Hubalewska M, Ignatova Z.** 2009. Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. Nature structural & molecular biology **16:**274-280.

123. **Zhou M, Guo J, Cha J, Chae M, Chen S, Barral JM, Sachs MS, Liu Y.** 2013. Non-optimal codon usage affects expression, structure and function of clock protein FRQ. Nature **495:**111-115.

124. **Canter DM, Perrault J.** 1996. Stabilization of vesicular stomatitis virus L polymerase protein by P protein binding: a small deletion in the C-terminal domain of L abrogates binding. Virology **219:**376-386.

125. **Connor JH, McKenzie MO, Parks GD, Lyles DS.** 2007. Antiviral activity and RNA polymerase degradation following Hsp90 inhibition in a range of negative strand viruses. Virology **362:**109-119.

126. **Rima BK, McFerran NV.** 1997. Dinucleotide and stop codon frequencies in single-stranded RNA viruses. The Journal of general virology **78 (Pt 11):**2859-2870.

127. **Rothberg PG, Wimmer E.** 1981. Mononucleotide and dinucleotide frequencies, and codon usage in poliovirion RNA. Nucleic acids research **9:**6221-6229.

128. **Atkinson NJ, Witteveldt J, Evans DJ, Simmonds P.** 2014. The influence of CpG and UpA dinucleotide frequencies on RNA virus replication and characterization of the innate cellular pathways underlying virus attenuation and enhanced replication. Nucleic acids research.

129. **Habjan M, Penski N, Spiegel M, Weber F.** 2008. T7 RNA polymerase-dependent and -independent systems for cDNA-based rescue of Rift Valley fever virus. The Journal of general virology **89:**2157-2166.

130. **Al-Saif M, Khabar KS.** 2012. UU/UA dinucleotide frequency reduction in coding regions results in increased mRNA stability and protein expression. Molecular therapy : the journal of the American Society of Gene Therapy **20:**954-959.

131. **Washenberger CL, Han JQ, Kechris KJ, Jha BK, Silverman RH, Barton DJ.** 2007. Hepatitis C virus RNA: dinucleotide frequencies and cleavage by RNase L. Virus Res **130:**85-95.

132. **Doma MK, Parker R.** 2006. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature **440:**561-564.

133. **Morin B, Kranzusch PJ, Rahmeh AA, Whelan SP.** 2013. The polymerase of negative-stranded RNA viruses. Current opinion in virology **3:**103-110.

134. **Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM.** 2007. A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science **315:**525-528.

135. **Sarangi U, Paithankar KR, Kumar JU, Subramaniam V, Sreedhar AS.** 2012. 17AAG Treatment Accelerates Doxorubicin Induced Cellular Senescence: Hsp90 Interferes with Enforced Senescence of Tumor Cells. Drug target insights **6:**19-39.

136. **Hou YJ, Banerjee R, Thomas B, Nathan C, Garcia-Sastre A, Ding AH, Uccellini MB.** 2013. SARM Is Required for Neuronal Injury and Cytokine Production in Response to Central Nervous System Viral Infection. J Immunol **191:**875-883.

137. **Garman RH.** 2011. Histology of the central nervous system. Toxicologic pathology **39:**22-35.

138. **Bi Z, Barna M, Komatsu T, Reiss CS.** 1995. Vesicular stomatitis virus infection of the central nervous system activates both innate and acquired immunity. Journal of virology **69:**6466-6472.

139. **Komatsu T, Ireland DD, Chung N, Dore A, Yoder M, Reiss CS.** 1999. Regulation of the BBB during viral encephalitis: roles of IL-12 and NOS. Nitric oxide : biology and chemistry / official journal of the Nitric Oxide Society **3:**327-339.

140. **Manaenko A, Chen H, Kammer J, Zhang JH, Tang JP.** 2011. Comparison Evans Blue injection routes: Intravenous versus intraperitoneal, for measurement of blood-brain barrier in a mice hemorrhage model. J Neurosci Meth **195:**206-210.

141. **Griffin DE.** 2003. Immune responses to RNA-virus infections of the CNS. Nature reviews. Immunology **3:**493-502.

142. **Schafer A, Brooke CB, Whitmore AC, Johnston RE.** 2011. The role of the blood-brain barrier during Venezuelan equine encephalitis virus infection. Journal of virology **85:**10682-10690.

143. **Reed L MM.** 1938. A simple method for estimating fifty percent endpoints. Am J Hyg **27:**493-497.

144. **Cronin J, Zhang XY, Reiser J.** 2005. Altering the tropism of lentiviral vectors through pseudotyping. Current gene therapy **5:**387-398.

145. **Wirblich C, Schnell MJ.** 2011. Rabies virus (RV) glycoprotein expression levels are not critical for pathogenicity of RV. Journal of virology **85:**697-704.