Stony Brook University

The Roles of the L4-22K and L4-33K Proteins during Adenovirus Infection

A Dissertation Presented

by

Kai Wu

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Molecular Genetics and Microbiology

Stony Brook University

December 2012

Stony Brook University

The Graduate School

Kai Wu

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Dr. Patrick Hearing, Ph.D. - Advisor
Professor, Molecular Genetics and Microbiology

Dr. Carol Carter, Ph.D. - Chairperson of Defense
Professor, Molecular Genetics and Microbiology

Dr. Laurie Krug, Ph.D.
Assistant Professor, Molecular Genetics and Microbiology

Dr. Eckard Wimmer, Ph.D.
Distinguished Professor, Molecular Genetics and Microbiology

Dr. Rolf Sternglanz, Ph.D.
Distinguished Professor Emeritus, Biochemistry and Cell Biology

This dissertation is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

Abstract of the Dissertation

The Roles of the L4-22K and L4-33K Proteins during Adenovirus Infection

by

Kai Wu

Doctor of Philosophy

in

Molecular Genetics and Microbiology

Stony Brook University

2012

Adenoviruses are a common cause of upper respiratory tract infection, conjunctivitis, and gastroenteritis. Fortunately, most of these infections are self-limiting and rarely cause severe or lethal diseases. Based on its characteristics of safety, easy manipulation and production, adenovirus has been widely used as a gene therapy vector in preclinical settings.

Despite the intensive study of adenovirus in the past four decades, the mechanism for viral genome packaging is not fully understood. My dissertation focuses on two viral late gene products, L4-22K and L4-33K, to explore their functions during adenovirus infection, with an emphasis on their roles in viral genome packaging.

By making specific mutant viruses that are deficient in making L4-22K or L4-33K proteins, I have found both proteins are absolutely required to make progeny viruses in infected human cells. Different approaches were used to pinpoint the roles that L4-22K and L4-33K play during the virus life cycle. For L4-22K, my work has shown it plays important roles in virus assembly, viral genome packaging and viral infection-induced cell lysis, which consist of the crucial final steps of adenovirus infection to produce and spread the progeny viruses. It does so by switching the mode of infection from the early to late phase to assemble progeny viruses, initiating the viral genome packaging process to finalize virus production, and facilitating cell lysis to spread the viruses. This work establishes the essential roles of L4-22K during adenovirus

infection and provides insights for coordination of late events of adenovirus infection. For L4-33K, my results confirm its role as a virus-encoded alternative RNA splicing factor and identify its primary targets of regulation: proteins IIIa and pVI. Further analyses have shown that L4-33K regulates IIIa expression to package the viral genome into capsids, and may regulate adenovirus protease activity for virion maturation.

Table of Contents

List of Figures

List of Abbreviations

Ad	Adenovirus
ADP	Adenovirus death protein
AVP	Adenovirus protease
B1H	bacterial one-hybrid
bp	base pair
CAR	coxsackie and adenovirus receptor
ChIP	chromatin immunoprecipitation
Co-IP	co-immunoprecipitation
CPE	cytopathic effect
DBP	DNA binding protein
DE	downstream elements
DMEM	Dulbecco's modified Eagle medium
dsDNA	double-stranded DNA
EC	empty capsids
EM	electron microscopy
EMSA	electrophoretic mobility shift assay
FITC	fluorescein isothiocyanate
FFU	fluorescent focus unit
GFP	green fluorescent protein
HI	heavy intermediates
HV	helper virus
HDAd	helper-dependent adenoviral vector
hpi	hours post-infection
hpt	hours post-transfection

ITR	inverted terminal repeat
L1	late region 1
L2	late region 2
L3	late region 3
L4	late region 4
L5	late region 5
LI	light intermediates
MLP	major late promoter
MLTU	major late transcription unit
MV	mature virions
NFI	nuclear factor I
NPC	nuclear pore complex
nt	nucleotide
Oct1	octamer-binding transcription factor
ORF	open reading frame
PI	propidium iodide
PFU	plaque forming unit
PS	packaging sequence
pVIc	pVI C-terminal 11 aa peptide
qPCR	quantitive PCR
TF	transcription factor
UPE	upstream promoter elements
3-AT	3-amino-trizole
5-FOA	5-fluoro-orotic acid

Acknowledgments

My Ph.D. training in the Microbiology department means a lot to me. I want to first thank my advisor Dr. Patrick Hearing for the greatest mentorship I could expect for my Ph.D. training. His passion on science completely infected me at the very beginning and I have been very happy to work with him on the same bench in a fully engaged fashion for the past five years. I also want to thank his trust on me from the very beginning, which has given me the opportunity to build independence through the training. He also keeps his office door open whenever we need him. I had difficult times for the projects and I really appreciate his help and encouragement to get me back on track. With the backup projects moving well, I thank him very much for getting the reagents I need along the way. It's really awesome that he is able to get virtually everything I need for the experiments, and I thank him very much for that which helps me to get work done in time. I also want to thank him for introducing me to the Ad community, including giving presentation, manuscript preparation, meeting people in the field, learning the peer-review process, and how this business works. I'd very much like to be a virologist as he is.

I also want to thank my awesome dissertation committee members: Carol, Laurie, Eckard and Rolf, for their time spent with me, their encouragement, their guidance for my postdoc, and the opportunity to interact with them. I appreciate their time very much as I get to know how tight a professor's schedule could be. Their encouragement was essential to get me back on track. And I really appreciate the opportunity to further interact with each one of them.

I also want to thank Nancy and Senthil for the opportunities to do my rotations in their labs, where I learned not only different aspects of science, but also the styles of lab operation.

During the training, I got a lot of help from the neighboring labs, virtually every lab in the department and the CID. I want to thank them all for their reagents, instruments, protocols, and advices. I want to specially thank Eckard and Aniko for the opportunity to join their lab meetings and become an "affiliated member". I want to thank Wei-Xing, his postdoc Ji'An and student Zhixun, for their insights and help on my cell death and autophagy experiments. This is a really wonderful department to be in, and I want to specially thank our awesome Chair, Jorge, for his encouragement, his guidance for my postdoc application and organizing the student invited speakers, and his openness and willingness to help all graduate students. I also want to thank Susan and Guo-Wei for their helps on the EM and the deconvolution microscope, and Jason for his great help on flow cytometry.

I also want to thank many professors in the Ad community who shared important reagents for my study, including Drs. Gudrin Schiedner, Stefan Kochanek, Michael Imperiale, Arnold Levine, Peter van der Vliet, Carl Anderson, Walter Mangel, Daniel Engel, David Matthews, Christopher Wiethoff, Maxim Balakirev, Ann Tollefson, and William Wold.

I also want to thank the people in the Hearing lab. Thanks to Janet for her advice and input on my projects, as well as all the equipment from her lab. Thanks to Mena for the guidance to the 'packaging' world and her help to get me to the local life at Stony Brook. Suk-Young is a thinker and I learned a lot from our discussions, and thanks for many rides to get my "antique car" fixed. Thanks to Diana for the collaboration which helped me to finish on time. Speical thanks to Kasey for introducing me to many others in the life sciences. Thanks all for their helps, certainly also including Jeniffer, Varsha, and Ting.

I also want to thank Janet and Kate for their help to navigate me through the graduate school. Janet has always been there, make sure I'm on schedule and help me every way she could. Kate has been wonderful, especially for the time when I just arrived. She made my life much easier to adjust to the mode of the graduate school.

Thanks to all classmates of 2007 for their help to settle me down after my arrival and throughout the graduate school, and possibly in the future, including: Amy, DeAnna, Matt, Iris, Xiao, Will, Shawn, Annie, Molly and Chewie.

Thanks to all my friends at Stony Brook: Qinghong, Jie, Ying, Niu, Xiao, Ding, Ting, Li, Xiaoying, Zhixun, Chunling, Ying and Ping. Big and special thanks to Qinghong and Jie who gave me tremendous help from the very beginning! Thanks to Froy, Al, Denise, and Paola for the wonderful collaboration in GSO! Thanks to Jim and Lila for introducing me to the American democratic world.

I would not be able to survive the five years without my wife Fang. Without her, I would not be able to concentrate on my work. Beyond the lab, I would not be able to live without her. Thanks for everything! Thanks to my son Eli who introduced me to another wonderful world, and let me find the other aspects of myself, which I would not realize without him. Thanks to my mother-in-law for her crucial help during and after Eli's birth, this helped me to finish the graduate school on time. Thanks to my parents, my sister and my grandparents, who built me as who I am. I'm sorry that I couldn't come back more often than once per five years. Thanks a lot for understanding and the support. Please stay well and I'll come back.

Introduction

Adenovirus Background

Adenovirus (Ad) was discovered in 1953 by two groups in search of etiological agents for acute respiratory illness (24, 64). Ads infect a wide range of animals and over 50 different serotypes of human Ad have been identified. Human Ad infection can lead to respiratory tract, gastrointestinal, and ophthalmologic diseases. Most of these infections are self-limiting, but can be severe and even fatal in immunocompromised hosts and occasionally in healthy young children and adults (12). In the military, especially among recruits, Ads have been a very common cause for acute respiratory infections, which might be due to the crowded living conditions and stressful working environment. In 2011, a live attenuated vaccine of Ad serotypes 4 and 7 was approved by the FDA for immunizing military personnel (25). Given its relatively safe properties, easy manipulation and culture, and extensive study of its biology in the past half century, Ads have been used widely as a gene transfer vector in laboratory research and preclinical settings (28).

Ads are non-enveloped, icosahedral particles containing a linear, non-segmented, double-stranded DNA (dsDNA) viral genome. With a size of around 36,000 base pairs (bp), the genome encodes five early gene families (E1A, E1B, E2, E3, and E4), three intermediate gene families (E2 late, IVa2, and IX), and one major late transcription unit (MLTU) including five different groups of mRNAs (L1-L5), at various locations on both strands of the genome (Fig. 1A). In addition to protein coding regions, the Ad genome also harbors two inverted terminal repeats (ITRs) at each terminus of the genome, which serve as replication origins during viral DNA replication. Moreover, between the left ITR and the first viral protein coding region E1A, the *cis*-acting region referred to as the packaging sequence (PS) has been identified to play a critical role in mediating viral genome packaging into the Ad capsid.

Adenovirus Replication Cycle

In the course of Ad infection, the replication cycle is conventionally divided into early and late phases, separated by the onset of viral DNA replication. Early events commence upon Ad infection, which includes: (i) Absorption, mediated by the viral ligand fiber protein and coxsackie B and adenovirus receptor (CAR) expressed on the cell surface; (ii) Internalization, facilitated by the major capsid protein penton and integrin molecules on the cell surface, via receptor-mediated endocytosis; (iii) Endosome escape by the minor capsid protein VI and Ad

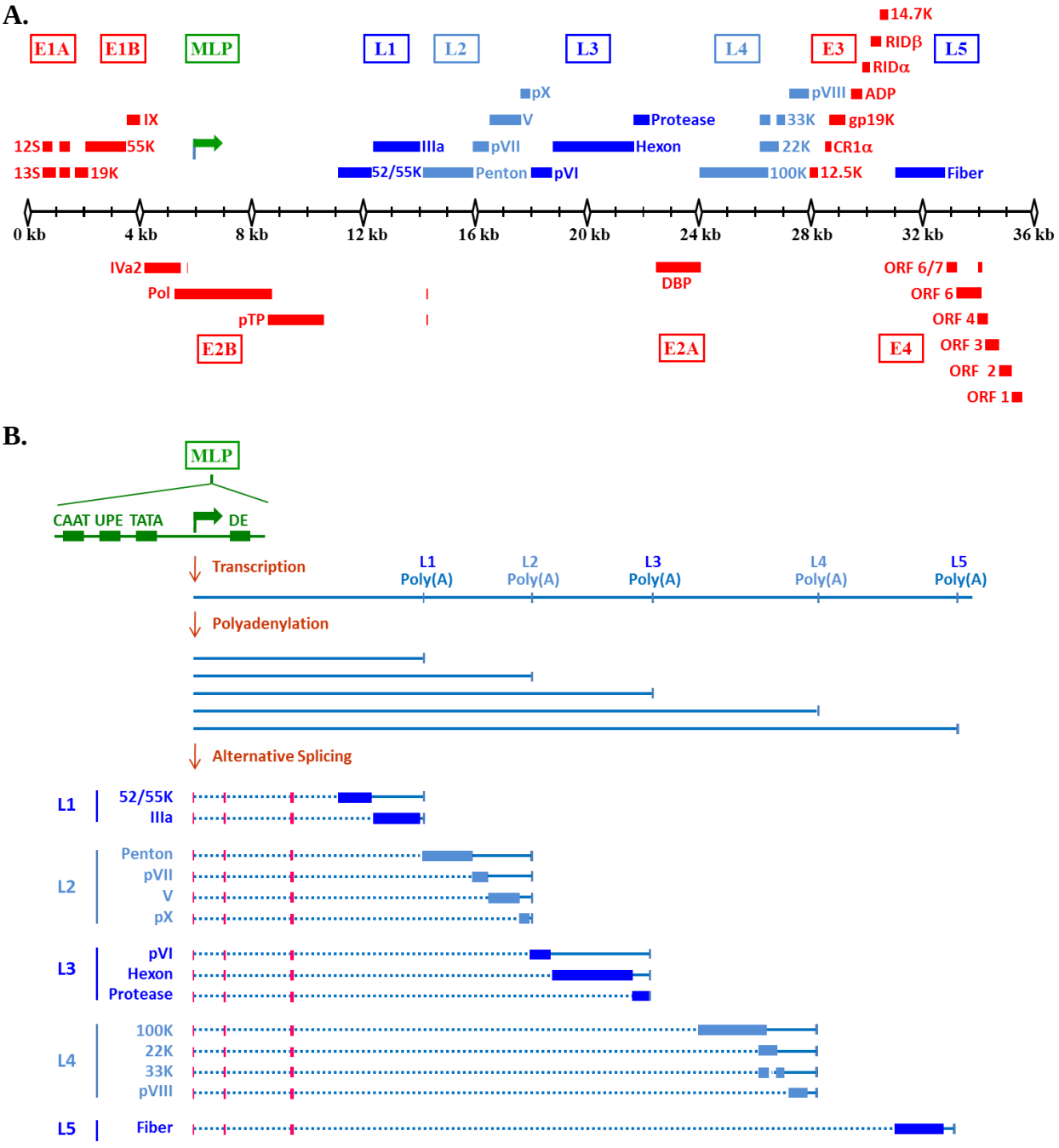


Figure 1. Adenovirus genome structure and viral late gene expression. A. Adenovirus genome with major late promoter (MLP; green), and open reading frames of early genes (red) and late genes (blue). **B.** Viral late gene expression. MLP structure is depicted in green, including the distal inverted CAAT box, upstream promoter element (UPE), TATA box, and downstream element (DE). pre-mRNA, RNA intermediates, and mRNAs of late genes are depicted in blue. Tripartite leader (TPL) is depicted in pink. The picture is drawn to scale.

protease (AVP); (iv) Movement of the partially uncoated virion to the nuclear pore complex (NPC) and transportation of the viral genome into the nucleus through the NPC; (v) Early gene expression. The immediate early gene product, E1A, is the first viral gene expressed and transactivates the Ad early genes (E1B, E2, E3, and E4). Ad early proteins counteract host responses to infection including the induction of apoptosis, DNA damage response, and aspects of the immune response, as well as set the stage for viral genome replication (5). Then, concomitant with viral DNA replication, the major late promoter (MLP) is fully activated and late gene products start to accumulate and assemble progeny virions with the viral genome packaged inside. Finally, activation of the AVP inside the immature virion cleaves precursor proteins and leads to the production of fully infectious mature virus particles.

Ad genome replication requires three virus-encoded proteins: precursor terminal protein (pTP), DNA polymerase (Pol) and DNA binding protein (DBP) (11). The replication is initiated by Pol-mediated transfer of dCMP to pTP, which serves as a protein primer. Serving as the origin of viral DNA replication, the ITR generally contains ~100 bp DNA sequence from each terminus of the viral genome and interacts with different viral and cellular proteins, such as pTP, Pol, nuclear factor I (NFI), and octamer-binding transcription factor (Oct1), to form the initiation complex. The transition from initiation to elongation is marked by a jump-back mechanism by the pTP-Pol complex. The elongation process requires DBP which unwinds dsDNA, binds ssDNA, protects ssDNA from nucleases, and increases the processivity of Pol.

In additional to viral genome amplification, the other consequence of the initiation of viral DNA replication is the full activation of the MLP (Fig. 1B). The MLP contains a core promoter consisting of a TATA box and transcription initiation element, two upstream activating elements (the distal inverted CAAT box and upstream promoter element [UPE]), and the downstream element (DE) (85). The MLP is activated by a number of cellular transcription factors and the viral protein, IVa2, which bind to the UPE and the DE of the MLP, respectively. Mutation in the UPE, the TATA box, the DE that abolishs IVa2 binding, or mutation of the IVa2 protein, however, only has a modest effect on Ad late gene expression, suggesting a redundancy of MLP elements (58, 85, 87). Even though the MLP is partially active before viral DNA replication, the elongation of transcription is not efficient and only L1-52/55K mRNA accumulates (2). Immediately following the initiation of viral DNA replication, there is a transient stage (a.k.a. early stage of the late phase of infection) when mRNAs from L1 and L4

are expressed (11). The role of early expression of the L1-52/55K protein is not understood, although there is a better understanding of the timing of such expression of the L4 gene. At this stage, the L4-22K protein is transcribed from a novel, internal promoter embedded in the open reading frame (ORF) of L4-100K (48). L4-22K activates the expression of Ad late genes at the transcriptional level by binding to the DE of the MLP (4, 51), and at the post-transcriptional level by an unknown mechanism (47).

The fully activated MLP gives rise to the approximately 28,000-nt pre-mRNA which becomes polyadenylated at one of the five possible polyadenylation sites, generating five families (L1-L5) of mRNAs with co-terminal 3'-ends (Fig. 1B). Following polyadenylation, the primary transcripts are spliced in such a way that each of the mature mRNA contains a common set of three short 5'-leader sequences, termed the tripartite leader (TPL). The TPL is then alternatively spliced to one of several 3' splice sites to form a total of more than 20 cytoplasmic mRNAs. The functions of the TPL include helping mRNA export and enhancing translation of viral late gene mRNAs (27, 38). In addition to the splicing mediated by cellular proteins, the viral protein L4-33K facilitates late gene expression by enhancing the alternative splicing of Ad late transcripts which contain weak 3' splice sites (72) and facilitating the early-to-late switch during infection (14).

With all of viral late gene proteins synthesized from the properly processed mRNAs, they translocate into the nucleus to assemble the capsid, package the viral genome, and cleave the precursor virion proteins to form the final mature viral particle (1). During the final stage of infection, the Ad death protein (ADP) induces cell lysis, which leads to the release of progeny viruses from the nucleus (70).

Adenovirus Viron and Genome Packaging

The mature Ad virion consists of major capsid proteins (hexon, penton, fiber), minor capsid proteins (IIIa, VI, VIII, IX), core proteins (V, VII, Mu, TP) that associate with the viral genome, and AVP (Fig. 2) (69). From a number of studies, it is believed that Ad empty capsids (EC) are first assembled in the nucleus, subsequently the viral genome is encapsidated to produce the heavy intermediates (HI), and finally AVP cleaves precursor proteins (IIIa, pVI, pVII, pVIII, pX, pTP) (Fig. 3) to produce the fully-infectious mature virions (MV) (Fig. 4A) (53). In addition to these three formats of Ad virions, light intermediates (LI) that only contain the left end of Ad genome have also been reported from the analysis of L1-52/55K and IIIa

temperature-sensitive mutants (10, 22). The EC band at a lighter density (1.29-1.30 g/cc) than the MV (1.34 g/cc) and HI (1.35 g/cc) in a CsCl density equilibrium gradient. MV contain proteolytically-processed forms of proteins VI, VII, VIII; precursor forms of these proteins (pVI, pVII, pVIII) are present in EC. In contrast to the MV, EC also contain the L1-52/55K protein (21). Core proteins V and VII, which are associated with the viral genome, are present in MV but are absent or present at lower levels in EC.

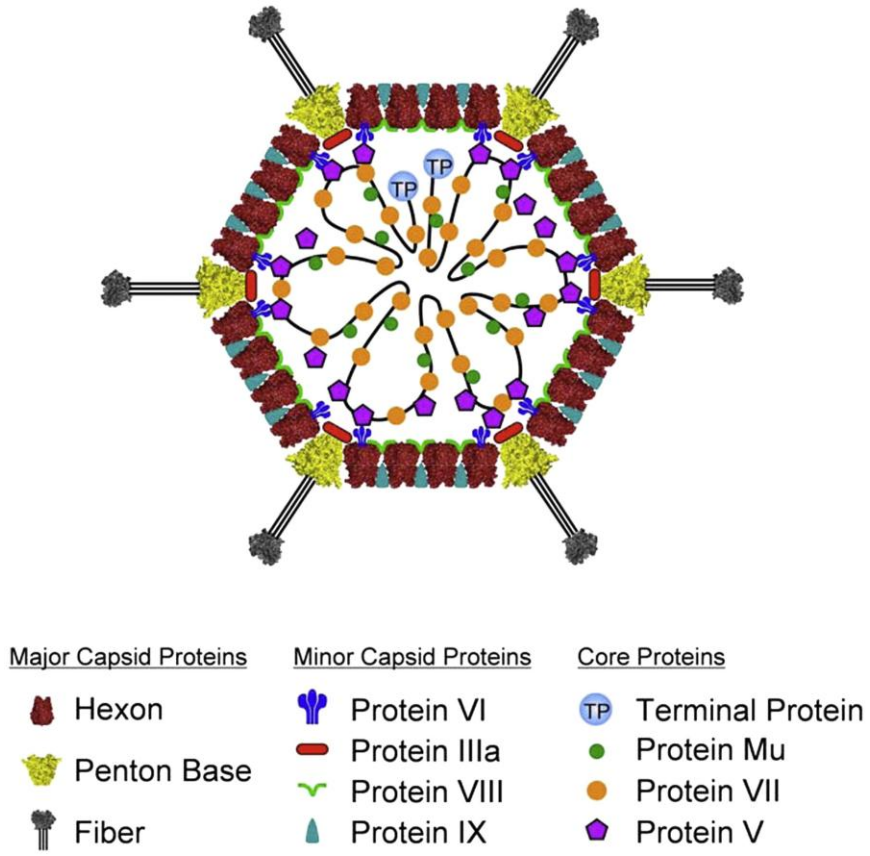


Figure 2. Schematic diagram of human adenovirus virion. Adenovirus structural proteins are depicted in the virion diagram. Adapted from Nemerow et al. Virology 2009.

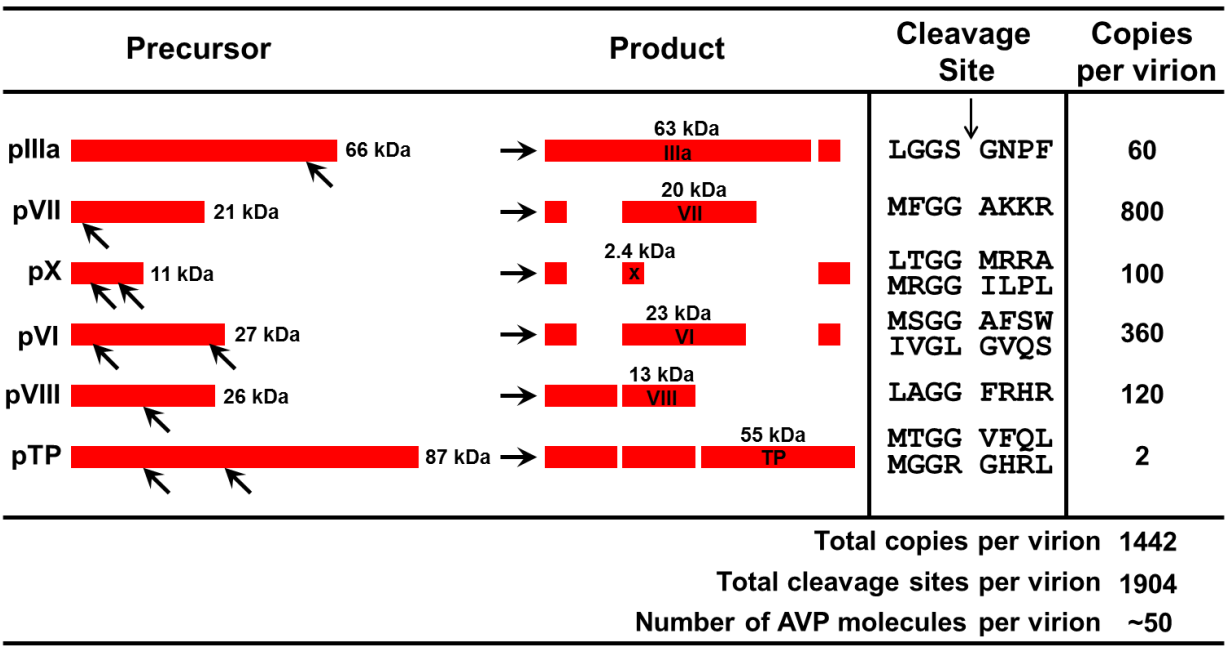


Figure 3. Cleavage events by adenovirus protease (AVP). The cleavage sites for each individual substrate of AVP, and the mature form of each substrate of AVP are depicted. The picture is adapted from Dr. Walter Mangel's presentation.

The Ad genome is packaged into the capsid utilizing the *cis*-acting PS located at the left end of the viral genome (Fig. 4B) (53). The PS overlaps with the E1A transcriptional enhancer region and seven functionally redundant AT-rich repeats, termed A1 to A7 repeats (Fig. 4B). Among these seven repeats, A1, A2, A5 and A6 are the most important for packaging activity, and share the consensus bipartite 5'-TTTG-N8-CG-3' motif, in which both bipartite parts and the 8 bp spacer are crucial in viral DNA packaging (Fig. 4B). Deletion of the PS results in only EC accumulation but no progeny virus production (23). The identification of the core consensus packaging repeats allowed us to use a simplified PS consisting of multimers of A repeats in the Ad genome, together with an E1-expressing cell line, to analyze the packaging process. To date, several cellular proteins (chicken ovalbumin upsteam promoter transcription factor [COUP-TF], Oct1, and CCAAT displacement protein [CDP]) and viral proteins (L1-52/55K, IVa2, L4-22K, and IIIa) have been shown to interact with the PS. However, results of mutation analysis in the PS have demonstrated that these cellular transcription factors are not required for Ad genome packaging *in vivo*, while IVa2, L1-52/55K and IIIa are essential for packaging from the analysis of respective mutant viruses where only EC are produced in their absence (10, 19, 50, 87). Electrophoretic mobility shift assays (EMSA) have shown that the IVa2 protein binds to the CG

motif of the PS, and another Ad late gene product, L4-22K, binds to the TTTG motif of the PS; binding of L4-22K to the PS is dependent on the IVa2 protein (Fig. 4C) (13, 54, 75, 83, 84, 86). Chromatin immunoprecipitation (ChIP) experiments have shown that the Ad proteins IVa2, L1-52/55K, and IIIa bind to the PS *in vivo* (41, 54, 61). The role of L4-22K *in vivo* has not been fully determined, although this protein binds to the PS *in vitro* in the presence of IVa2 (51).

The IIIa protein, one of the minor capsid proteins, was the first viral protein that was indicated to play a role in viral genome packaging since a IIIa temperature-sensitive mutant virus (*ts*112) only produces LI at the non-permissive temperature (10). Recent work from our laboratory demonstrated that IIIa plays an important role in the serotype specificity of Ad genome packaging through its interaction with L1-52/55K (41).

The L1-52/55K protein is unique from other late gene proteins in that it is expressed at the early point of infection, although the role of its early expression is not understood (2). Analysis of an L1-52/55K temperature-sensitive mutant (*ts*369) demonstrated that L1-52/55K is required for viral genome packaging, because only LI were produced in *ts*369-infected cells at the non-permissive temperature (22). In addition, this phosphorylated protein, shown as a 52/55 kD doublet on SDS-PAGE gel due to differential phosphorylation, is only present in EC and assembly intermediates, but not in MV, suggesting that L1-52/55K might also be a scaffold protein (21). The role of L1-52/55K in viral genome packaging is further confirmed by another L1-52/55K mutant virus *pm*8001, in which two stop codons were introduced at amino acid (aa) positions 18 and 21 of L1-52/55K into the Ad5 genome; *pm*8001 only produced EC (19). As an earlier study showed another viral protein IVa2 interacts with L1-52/55K by a yeast two-hybrid assay, as well as by co-immunoprecipitation (Co-IP) with Ad5 infection lysate, IVa2 was put on the stage to be involved in Ad genome packaging (20).

The IVa2 protein becomes detectable at an intermediate stage of Ad infection at the onset of DNA replication, and localizes to the nucleoplasm and nucleoli (39, 74). It was first reported to interact with DE of the MLP (74). But this interaction is functionally redundant and dispensable for MLP activation, as mutation in DE, abolishing the interaction with IVa2, only has a minor effect on late gene expression (85). Further characterization of the IVa2 protein has identified its N-terminal 50-136 aa as a nucleolar localization signal (NuLS) and its C-terminal 357-449 aa as a nuclear localization signal (NLS), both of which overlap its DNA binding domains (40). Based on the homology between the DE and the PS, Zhang and Imperiale first

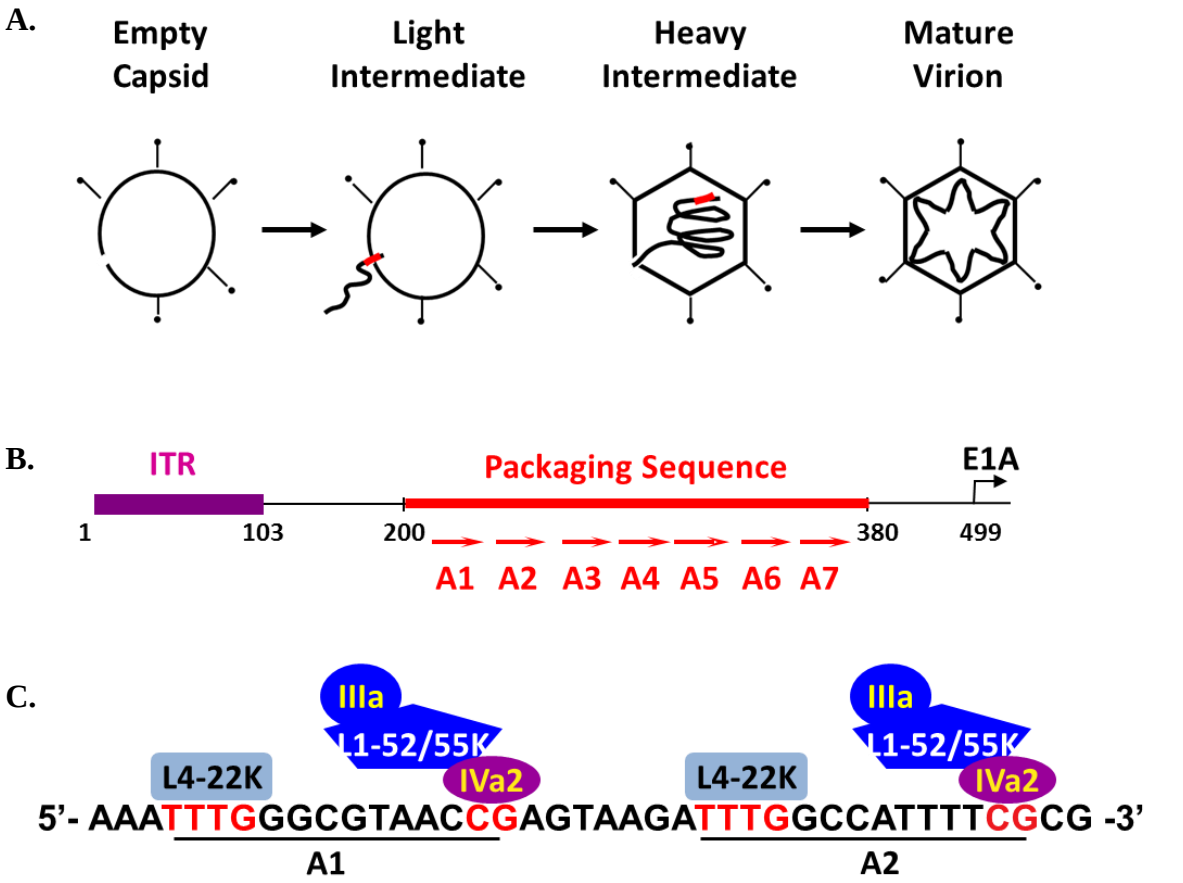


Figure 4. Adenovirus genome packaging. A. Adenovirus assembly intermediates. The red region is the packaging sequence (PS) located on the left end of genome. **B.** The left end of Ad5 genome. A1 to A7 repeats of the PS are shown. The inverted terminal repeats (ITRs) serve as the original of viral DNA replication. The nucleotide number of Ad5 is listed below. **C.** The interactions between the known packaging proteins (L4-22K, IVa2, L1-52/55K, and IIIa) and the PS are depicted.

showed that the IVa2 protein interacts with the PS (86), and it was shown later on that IVa2 is required for packaging of viral DNA based on the phenotype of only EC production with IVa2 mutant viruses: *pm*8002 and ΔIVa2 (50, 87). Results of mutation of the consensus motif of the PS revealed IVa2 binds directly to the CG bi-nucleotide motif (Fig. 4C) and mutation of CG gives rise to nonviable virus production *in vivo* (67). The IVa2 protein binds ATP (52) and is located at a single virion vertex (9). As mentioned above, L1-52/55K is also required for DNA packaging and interacts with IVa2. Further investigation by EMSA and ChIP assays showed that L1-52/55K does not bind the PS *in vitro*, however, it binds the PS *in vivo*, but not via IVa2 (54, 61).

These observations lead us to a working model of Ad DNA packaging (Fig. 4C) that the IVa2 and L4-22K proteins bind directly to the PS and recruit the L1-52/55K and IIIa proteins to promote viral DNA packaging; the IVa2 protein may function as a packaging motor by analogy to bacteriophage ATPases (18).

Adenovirus Late Region 4

The L4 region encodes four viral proteins: L4-100K, L4-22K, L4-33K and pVIII (Fig. 1A). L4-100K is a multifunctional protein, involved in host translation shutoff, facilitating translation of viral late mRNAs, hexon trimerization and hexon import into the nucleus. L4-100K selectively translates viral late mRNA while simultaneously suppressing cellular mRNA translation during the late phase of Ad infection (81). L4-100K binds to eIF4G, the scaffolding protein of the eIF4F complex, and the stronger binding displaces the protein kinase Mnk1 which normally binds to the same site and phosphorylates the cap binding protein eIF4E. As a result of this competitive displacement of Mnk1, eIF4E is dephosphorylated and not able to recognize the cap structure of mRNAs so that host cell translation is suppressed. Although Ad late mRNAs are also capped, the inhibition of their translation is compensated by the TPL which is present on all Ad late mRNAs (81). L4-100K binds to the TPL, which preferentially brings the translation initiation machinery into viral late mRNAs (81). L4-100K also serves as a chaperon for hexon trimerization and it is also indicated to help hexon nuclear import following its trimerization (7, 26, 46). pVIII is cleaved by AVP during virion maturation to form the mature form VIII, which is a minor capsid protein located in the inner surface of the Ad capsid (Fig. 2) (36, 62). An early study suggested that VIII plays an important role in virion structural stability (35).

The L4-22K protein was identified by our laboratory in 2006, and shares a common N-terminal 105 aa with L4-33K. It binds the PS and the DE *in vitro* in the presence of IVa2, and transfection into cells of an Ad5 infectious clone which contains a stop codon in L4-22K (pTG3602-22K^{-}) resulted in no infectious virus production, which indicates L4-22K plays an important role in viral genome packaging and/or virus assembly (22). By transfection analyses using the same Ad5 L4-22K mutant infectious clone, Morris and Leppard showed that L4-22K promotes Ad late gene expression at the post-transcriptional level by an unknown mechanism (47). In addition, Backstrom et al. showed L4-22K activates MLP transcription through its interaction with DE (4). Because these analyses indicated completely different roles for the L4-22K protein during Ad infection and due to the limitations of experiments involving transient

transfection assays, we sought to study L4-22K function(s) in the context of virus infection. In the L4-22K project, we generated two Ad5 L4-22K mutant viruses, Δ22K (a null mutant) and 22K^{-} (an amber mutant) using an L4-22K complementing cell line, to further explore the functions of L4-22K *in vivo*. We demonstrate that the L4-22K protein is important for appropriate temporal control of viral early and late gene expression. The effect of L4-22K in the activation of Ad late gene expression varied for different late genes. In addition, the L4-22K mutants only produced EC, but no MV, in non-complementing cells demonstrating that the L4-22K protein is required for genome packaging. ChIP assays demonstrated that L4-22K and IVa2 were dependent on each other for binding to the PS and that both of these proteins are required to recruit L1-52/55K to the PS *in vivo*. Finally, the L4-22K protein is required for expression of ADP at late times after infection and the induction of cellular cytopathic effect. We conclude that the Ad L4-22K protein is multifunctional and an integral component in crucial aspects of infection.

The L4-33K protein, which shares a common N-terminal 105 aa but has a distinct C-terminus from that of L4-22K, is an alternative RNA splicing factor (72). The L1 region represents an example of alternative RNA splicing where the 5' splice site could be joined to one of the two 3' splice sites, producing L1-52/55K or IIIa mRNAs (Fig. 1B). While the choice of L1-52/55K 3' splice site happens early during infection, it has been known that efficient usage of the IIIa 3' splice site depends on late gene synthesis (2). Using an *in vitro* splicing assay, Tomanen et al. showed L4-33K enhances splicing of viral late mRNAs that contain weak 3' splice sites, including IIIa, V, pVII and others (72). Furthermore, Tomanen Persson et al. recently showed that L4-33K could be phosphorylated by two protein kinases: DNA-dependent protein kinase (DNA-PK) and protein kinase A (PKA) (73). L4-33K interacts with the catalytic subunit of DNA-PK, and the phosphorylation by DNA-PK has an inhibitory role on L4-33K mediated alternative splicing of IIIa pre-mRNA. On the other hand, the phosphorylation of L4-33K by PKA stimulates IIIa splicing. These reports together suggest the complexity of the regulation of Ad late gene expression.

Besides being a virus-encoded alternative RNA splicing factor, earlier studies also suggested L4-33K plays a role in virus assembly (16, 17, 33). However, when Fessler and Young introduced an amber mutation at aa position 20 of L4-33K, they also disrupted the L4-22K protein, which shares the N-terminal 105 aa with L4-33K (16). Since L4-22K is required for

viral genome packaging (80), the assembly defect observed with that L4-33K mutant virus might be due to the deficiency in L4-22K, instead of L4-33K. In another two independent studies, Flint and colleagues made an L4-33K mutant infectious clone which truncated the C-terminal 47 aa of the L4-33K protein, without disrupting the adjacent E2 early (E2E) promoter and preserving the L4-22K reading frame; this L4-33K mutation was lethal for virus viability with a defect in virion production (17). In addition, Tikoo and colleagues showed L4-33K also plays a role in bovine Ad assembly (33). However, both of these latter studies did not conclusively examine viral late gene expression patterns. Given the known role of L4-33K as an alternative RNA splicing factor, the assembly defect observed in these two studies might reflect a deficiency in the expression of one or more late gene products whose expression is regulated by L4-33K, instead of a direct role for L4-33K in the virus assembly process. In order to determine the function(s) of the L4-33K protein in the context of viral infection, we characterized an Ad5 L4-33K mutant virus generated using a L4-33K complementing cell line. These results showed that the L4-33K mutant virus has comparable infectivity and viral genome replication with WT Ad5. We confirmed the role of L4-33K as an alternative RNA splicing factor in the context of viral infection, and identified the primary targets for L4-33K regulation during infection as proteins IIIa and pVI, with a lesser effect on fiber. We observed only EC production, but no MV, from L4-33K mutant virus-infected cells. However, the L4-33K protein neither preferentially binds to the PS nor influences the interaction of other packaging proteins with the PS *in vivo*. Finally, our results suggest that the regulation of IIIa expression by the L4-33K protein influences viral genome packaging and regulation of pVI expression may modulate AVP activity in the virus particle. These results demonstrate the multifunctional nature of the L4-33K protein and its involvement in several different and critical aspects of Ad infection.

Materials and Methods

Cells and viruses. A549 (human lung adenocarcinoma epithelial cell line; ATCC) and 293 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) containing penicillin and streptomycin, and supplemented with 10% bovine calf serum (Hyclone). N52.E6 cells (human amniocytes) (66) were maintained in DMEM containing penicillin and streptomycin, and supplemented with 10% FetalClone III serum (HyClone). N52.E6-Cre (a kind gift from Drs. Gudrin Schiedner and Stefan Kochanek, University Ulm) and 293-L1 (19) cells were maintained in DMEM containing penicillin and streptomycin, 200 μg/ml Geneticin, and supplemented with 10% FetalClone III or bovine calf serum, respectively. TetC4 cells were maintained in DMEM containing penicillin and streptomycin, 200 μg/ml Geneticin, 200 μg/ml hygromycin B, and supplemented with 10% FetalClone III serum (Hyclone). TetC4-IVa2 cells (50) were maintained in DMEM containing penicillin and streptomycin, 200 μg/ml Geneticin, 200 μg/ml hygromycin B, 10 μg/ml blasticidin S, 1 μg/ml doxycycline, and supplemented with 10% fetal bovine serum.

To make tetracycline-inducible cell lines (tet-off system) expressing L4-22K-ΔC, L4-33K, or IIIa, the N-terminal 173 aa of Ad5 L4-22K with a stop codon at the end, the coding regions for Ad5 L4-33K and Ad5 IIIa, were PCR-amplified and cloned into pUHD 10-3 (kindly provided by Dr. H. Bujard, Zentrum für Molekulare Biologie, Germany) under the tetracycline-response element and a minimal CMV promoter using *Eco*RI and *Xba*I sites, respectively. The resulting plasmids, pUHD-22K-ΔC, pUHD-33K, pUHD-IIIa, were independently cotransfected with pSuper-Blasticidin (Invitrogen) into the cell line TetC4 that constitutively expresses the tetracycline-controlled transactivator, tTA, in the background of N52.E6-Cre cells. The three cell lines (termed TetC4-22K-ΔC, TetC4-33K, and TetC4-IIIa) were selected and maintained in DMEM containing penicillin and streptomycin, 200 μg/ml Geneticin, 200 μg/ml hygromycin B, 10 μg/ml blasticidin S, 1 μg/ml doxycycline, and supplemented with 10% fetal bovine serum.

An Ad5-WT virion stock was generated by transfecting *Pac*I-linearized infectious clone pTG3602 (8) into N52.E6-Cre cells; a cell lysate was prepared and purified virions were produced using A549 cells. IVa2 mutants viruses *pm*8002 (87) and ΔIVa2 (51) were propagated using TetC4-IVa2 cells minus doxycycline in the culture medium. L1-52/55K mutant virus *pm*8001 (19) and an Ad5 virus containing loxP sites flanking the PS (termed Ad5-Ψ-loxP,

unpublished) were propagated in 293-L1 and N52.E6 cells, respectively. All virions were purified using CsCl equilibrium density gradient centrifugation, as described (50). For the study of L4-22K, the particle/PFU ratio of different virus particles was determined on TetC4-22K-ΔC cells and A549 cells were infected at an MOI of 12.5 PFU/cell; N52.E6, N52.E6-Cre, TetC4-IVa2, TetC4-22K-ΔC, and TetC4-33K cells were infected at an MOI of 4 PFU/cell. For the study of L4-33K, the particle/PFU ratio of different virus particles was determined on TetC4-33K cells and A549 cells were infected at an MOI of 5 PFU/cell; 293, TetC4, TetC4-33K, TetC4-IIIa cells were infected at an MOI of 1.5 PFU/cell. However, please note that the particle/cell ratios of Ad5-WT infection in A549 cells in both studies are very similar, because plaque formation in TetC4-33K cells is not as efficient as in TetC4-22K-ΔC cells.

The strategy used for making pTG3602-Δ22K was to delete the L4-33K intron so that the C-terminus of L4-22K would be deleted without interfering with L4-33K expression. First, we PCR amplified L4-33K cDNA from pcDNA3-L4-33K using primers 5-33K (CGCGGATCCACCATGGCACCCAAAAAGAAGCTGC) and Ad5 nt 27102-27071 (GAGAAAGGGCGCGAAACTAGTCCTTAAGAGTC). Second, we PCR amplified Ad5 27079-27590 fragment from pTG3602 using primers Ad5 nt 27079-27112 (AGGACTAGTTTCGCGCCCTTTCTCAAATTTAAGC) and Ad5 nt 27590-27575 (CGACTCGTCGTTGAGC). Third, we did a hybrid PCR using mixed L4-33K cDNA and Ad5 27079-27590 as the template and 5-33K and Ad5 nt 27590-27575 as primers to amplify the Δ22K sequence, which corresponds to Ad5 nt 26195-27590 with the L4-33K intron (nt 26511-26712) deleted. This fragment was cotransformed with *Spe*I-digested pTG3602 into *E*. coli BJ5183 cells for homologous recombination (8) and pTG3602-Δ22K was isolated and confirmed by nt sequencing. The construction of pTG3602-22K^{-} was described previously (51). To make Δ22K and 22K^{-} viruses, *Pac*I-linearized pTG3602-Δ22K and pTG3602-22K^{-} plasmids were transfected using NanoJuice (EMD Millipore) into TetC4-22K-ΔC cells which had been seeded on the plates without doxycycline for 2 days. Virus stocks were passaged twice on TetC4-22K-ΔC cells without doxycycline. The passage 2 lysate of each mutant virus was titered and used to generate purified virions.

To make the Ad5 L4-33K mutant virus, an amber mutation (TAG) was introduced at aa position 138 of L4-33K in the background of Ad5 infectious clone pTG3602 (8), termed pTG3602-33K^{-}. To make pTG3602-33K^{-}, we first PCR amplified Ad5 26011-26830 and Ad5

26801-27316 fragments from pTG3602, both of which contain an introduced amber mutation (TAG) at aa position 138 of L4-33K, using primers: Ad5 nt 26011-26026 (GCCTGCGTCATTACCC) and 33K-AvrII-1 (GCCCAAGAAATCCTAGGCGGCGGCAGC), 33K-AvrII-2 (GCTGCCGCCGCCTAGGATTTCTTGGGC) and Ad5 nt 27316-27301 (CGGATTCCGTTGACCC), respectively. Next, we did a hybrid PCR using the two fragments above as the template and Ad5 nt 26011-26026 and Ad5 nt 27316-27301 as primers to amplify the 33K^{-}sequence, which corresponds to Ad5 nt 26011-27316 with an introduced amber mutation (TAG) at aa position 138 of L4-33K. This final fragment was cotransformed with *Spe*I-digested pTG3602 into *E*. coli BJ5183 cells for homologous recombination and pTG3602-33K^{-} was isolated and confirmed by nt sequencing. Then, *PacI*-linearized pTG3602-33K^{-} plasmid was transfected using NanoJuice (EMD Millipore) into TetC4-33K cells which had been seeded in the culture medium without doxycycline for 24 h, and the transfected cells were overlaid for plaque assay at 48 h post-transfection. 10 days post-transfection, individual plaques were picked and amplified in TetC4-33K cells without doxycycline in the culture medium. The passage 2 lysate of the Ad5 L4-33K mutant virus was titered by plaque assay using TetC4-33K cells and used to generate purified virions. The purity of the mutant virus stock was verified by restriction endonuclease digestion (the amber codon insertion generates an AvrII site) and Southern blot analysis.

Virus growth and infectivity. For complementation assays, TetC4-22K-ΔC, TetC4-33K or TetC4-IIIa cells were seeded in the culture medium without doxycycline 24 h before infection by Ad-WT, Δ22K, 22K^{-} or 33K^{-} virus particles. Infected cell lysates were harvested at 48 hpi and titered by plaque assay on TetC4-22K-ΔC or TetC4-33K cells as described above.

For one-step viral growth curves, A549 cells were infected with Ad5-WT, Δ22K, 22K^{-} or 33K^{-} virus particles and infected cell lysates were harvest at 6, 12, 24 and 48 hpi for titration by plaque assay on TetC4-22K-ΔC or TetC4-33K cells as described above.

For fluorescent focus assays, A549 cells grown on glass coverslips were infected with 10^6 of Ad-WT, Δ22K, 22K^{-} or 33K^{-} virus particles in 24-well plates. Cells were fixed using methanol at 18 hpi and subjected to indirect immunofluorescence using a monoclonal antibody against DBP (mAb-B6, provided by Dr. Arnold Levine, Princeton Univ.) and secondary goat anti-mouse antibody conjugated with FITC (Invitrogen). Coverslips were examined under Zeiss Axiovert 200M digital deconvolution microscope and ten random fields of each coverslip were

counted for DBP-positive cells. The average number of DBP-positive cells (fluorescent focus unit [FFU]) per field was used to calculate total FFU in the well, and the particle/FFU ratio was calculated as: (particle number)/(total FFU in the well).

For plaque assays to determine the particle to PFU ratios, TetC4-22K-ΔC or TetC4-33K cells were seeded in the culture medium without doxycycline 24 h before infection with 1,000 Ad-WT, Δ22K, 22K^{-} or 33K^{-} virus particles; cells were overlaid at 1 hpi. Eight days post-infection, plaques were counted from each set of infected plates and the particle to PFU ratios were determined.

Viral DNA replication assay. A549 cells were infected with Ad5-WT, Δ22K, 22K^{-} or 33K^{-} viruses and harvested at 6, 12, 18 and 32 hpi for the L4-22K study, or 6, 12, 24 and 48 hpi for the L4-33K study. Cell pellets were washed once with PBS and resuspended in 500 μl isotonic NP-40 buffer (150 mM NaCl, 10 mM Tris [pH 7.5], 1.5 mM $MgCl_2$ and 0.6% NP-40). Cell lysates were kept on ice for 5-10 min, nuclei were pelleted by centrifugation at 2000 x g at 4 °C for 5 min, and then resuspended in 200 μl PBS for DNA extraction using DNeasy Blood and Tissue Kit (Qiagen) following the manufacturer's instruction. The final DNA sample was resuspended in TE and then subjected to quantitative PCR using Applied Biosystems 7500 Real-Time PCR System. The two sets of primers used for quantitative PCR are PACK-4/5 (Ad5 nt 44-63: TAATGAGGGGGTGGAGTTTG and Ad5 nt 280-261: GCGAAAATGGCCAAATGTTA), and GAPDH-5/6 (GAPDH-5: CCCCACACACATGCACTTACC and GAPDH-6: CCTAGTCCCAGGGCTTTGATT). Genome replication was calculated as the value of viral genome copy number divided by the GAPDH copy number. For comparison, the value of each virus at each time point was normalized to that of WT virus at 6 hpi, shown as the relative viral genome copy number per cell.

Northern blot analysis. A549 cells infected with Ad5-WT, Δ22K, 22K^{-} or 33K^{-} viruses were harvested at 12, 24, and 48 hpi for preparation of total cytoplasmic RNA using an RNeasy Mini Kit (Qiagen) following the manufacturer's instructions. 4 μg (for blots of L1, L2, L3-hexon, L5) or12 μg (for blots of E1A, L3-pVI, L3-AVP and L4) RNA of each sample was separated on a 1% formaldehyde-agarose gel and transferred to positively charged nylon membrane (GE Healthcare). The probes for detecting E1A and L1 to L5 mRNAs correspond to Ad5 nt 147–1334, 13026-13751, 16834-17452, 21573-22322, 26769-27590, and 31920-32465, respectively. The probe that specifically detects pVI mRNA corresponds to Ad5 nt 18003-18755. The probes

were amplified by PCR, purified, and labeled with ^{32}P dATP by random primer labeling using exo$^-$ Klenow DNA polymerase (NEB).

Reverse transcription PCR. 2 μg cytoplasmic RNA from Ad5-WT or 33K–-infected A549 cells were reverse transcribed using the oligo(dT) primer (NEB) and SuperScrip II Reverse Transcriptase (Invitrogen). 5% of the reverse transcription reaction was used as the template for the following 20-cycle PCR. The primer sets for detecting L4-22K and GAPDH are: Ad5 nt 9700-9719 (locates in TPL3 region) and Ad5 nt 26703-26682 (locates in L4-33K intron), GAPDH-1 (ACCCAGAAGACTGTGGATGG) and GAPDH-2 (TTCTAGACGGCAGGTCAGGT), respectively.

Western blot analysis. Mock- or virus-infected cells were lysed in 2x Laemmli sample buffer (0.15 M Tris pH 6.8, 4% SDS) and protein concentration was determined by a BCA Protein Assay Kit (Pierce). 30 or 40 μg of whole cell extract was loaded for SDS-PAGE and then transferred to nitrocellulose (Whatman) or PVDF (GE Healthcare) membranes. Membranes were blocked with 3% BSA in TBS for 1 h at room temperature and then probed with primary antibody for 1 h at room temperature or 4 °C overnight. Primary antibodies include: rabbit polyclonal E1A antibody (SC430; Santa Cruz Biotechnology) (1:500 dilution), mouse monoclonal DBP antibody (Arnold Levine, Princeton University) (1:1000 dilution), rabbit polyclonal DBP antibody (Peter van der Vliet, University of Utrecht) (1:2000 dilution), rabbit polyclonal IVa2 antibody (54) (1:1000 dilution), rabbit polyclonal L1-52/55K antibody (54) (1:1000 dilution), rabbit polyclonal IIIa antibody (41) (1:1000 dilution), rabbit polyclonal penton antibody (Carl Anderson, Brookhaven National Laboratory) (1:1000 dilution), rabbit polyclonal VII antibody (Daniel Engel, University of Virginia) (1:2000 dilution), rabbit polyclonal V antibody (David Matthews, University of Bristol) (1:1000 dilution), rabbit polyclonal VI antibody (Christopher Wiethoff, Loyola University Chicago) (1:5000 dilution), mouse monoclonal hexon antibody (MAB0850; Abnova) (1:5000 dilution), rabbit polyclonal hexon antibody (GTX44240; GeneTex) (1:1000 dilution), rabbit polyclonal L4-100K antibody (1:1000 dilution), rabbit polyclonal L4-22K antibody (#6601; against the distinct C-terminus of L4-22K (51))) (1:2000 dilution), rabbit polyclonal L4-33K antibody (#5091; against the distinct C-terminus of L4-33K) (1:2000 dilution), rabbit polyclonal ADP and pVIII antibody (Ann Tollefson and William Wold, St. Louis University) (1:400 dilution), rabbit polyclonal fiber antibody (Carl Anderson) (1:1000 dilution), rabbit polyclonal AVP antibody (Maxim Balakirev,

CEA-Grenoble) (1:500 dilution), rabbit polyclonal γ-Tubulin antibody (T5192; Sigma) (1:1000 dilution), and mouse monoclonal α-Tubulin antibody (T5192; Sigma) (1:10,000 dilution).

For virus particle analyses, ~2 $\times 10^8$ A549, N52.E6-Cre, N52.E6, TetC4-22K-ΔC and TetC4-33K cells were infected by Ad5-WT, Δ22K, 22K$^-$ or loxP viruses. 48 to 72 hpi, infected cell were harvested for preparation of virus particles by CsCl density gradient. Isolated MV or EC from CsCl gradient were subsequently analyzed by silver staining and western blot analysis as described (50).

Luciferase assay. 293 cells were transfected with 400 ng expression vector (pcDNA3, pcDNA3-L4, pcDNA3-L4-33K$^-$, pcDNA3-L4-22K$^-$, or pcDNA3-22K), and 400 ng reporter plasmid (pGL3-E1A, or pGL3-MLP), with 1 ng pRL-null as transfection control, using NanoJuice (EMD Millipore) following the manufacturer's instructions. Transfected cells were lysed at 24 h post-transfection and subjected to the luciferase assay using the dual-luciferase reporter assay system (Promega) following the manufacturer's instructions.

Chromatin immunoprecipitation. A549 cells were infected with Ad5-WT, Δ22K, 22K$^-$, 33K$^-$, *pm*8002 (a IVa2 mutant virus), ΔIVa2, or *pm*8001 (an L1-52/55K mutant virus) and TetC4-IVa2 or TetC4-22K-ΔC cells were infected by Ad5-WT, Δ22K, 22K$^-$, *pm*8002, or ΔIVa2. At 18 hpi, infected cells were cross-linked with formaldehyde, sonicated, and the supernatants from sonication were subjected to a pre-clear with protein A-agarose beads, and then overnight incubation with corresponding antibodies at 4 °C, as described (82). After overnight incubation with antibody, the manufacturer's instructions for Protein A Agarose/Salmon Sperm DNA (Millipore) were followed for washing and DNA precipitation processes. The final DNA sample was dissolved in 50 μl TE and subjected to quantitative PCR analysis using two sets of primers: the PACK-4/5 primer set described above to assess the copy number of the PS that had been specifically immunoprecipitated by the antibodies, and the E4-ORF6 primer set (Ad5 nt 33597-33578: TACCGGGAGGTGGTGAATTA and Ad5 nt 33428-33447: TTCAAAATCCCACAGTGCAA) to assess the copy number of an Ad5 right-end DNA fragment that had been nonspecifically immunoprecipitated or reflected incomplete sonication products. The capacity of each viral protein (L4-22K, L4-33K, IVa2, L1-52/55K, or IIIa) to bind the PS in virus-infected cells was calculated as the value of the PS copy number divided by the E4-ORF6 copy number. For comparison, the PACK/E4-ORF6 ratio of each antibody in a particular virus infection was normalized to ChIP using pre-immune serum. To confirm viral

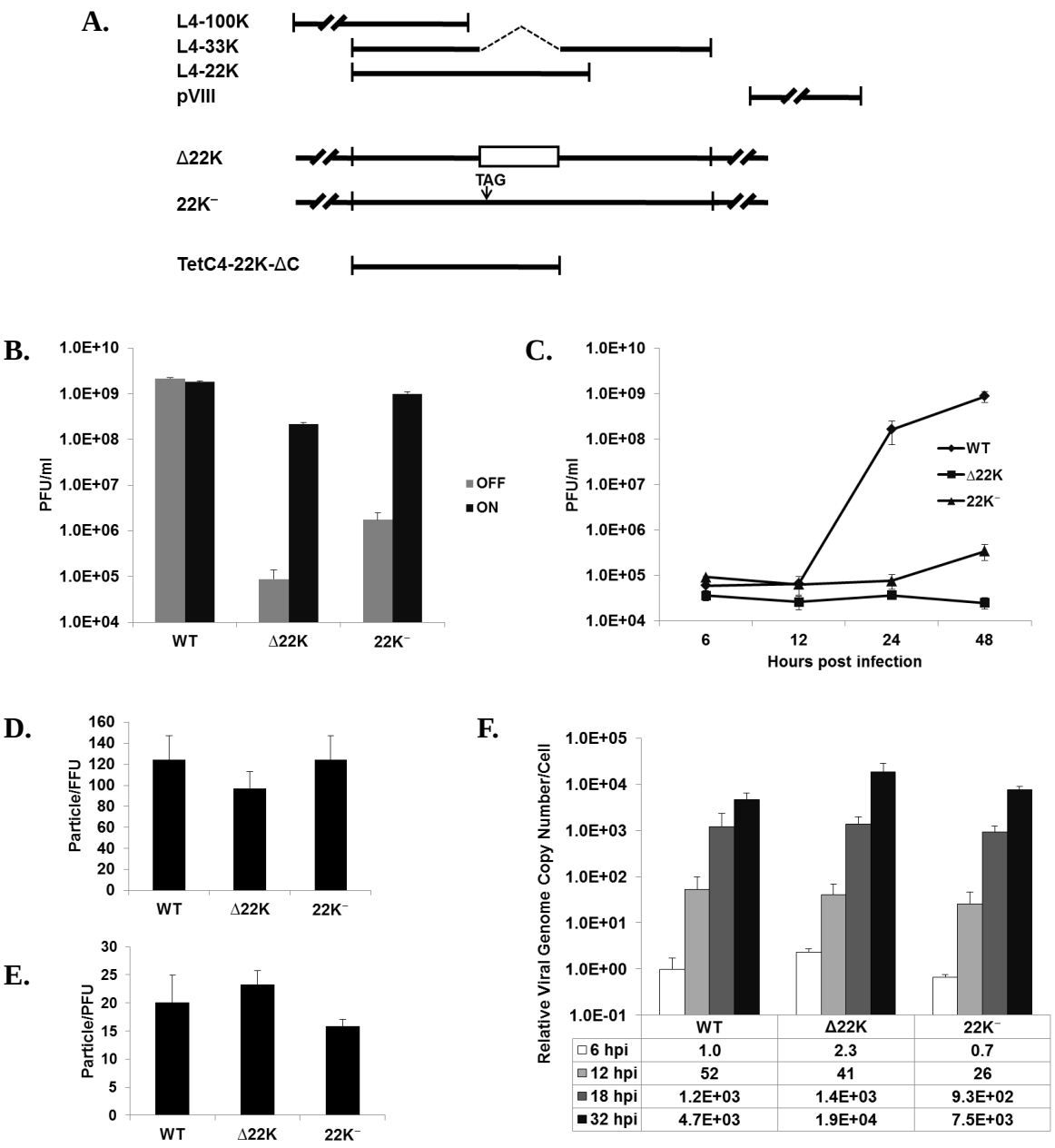
protein expression in ChIP experiments, 20 μl out of 600 μl supernatant after the pre-clearing process was boiled for 10 min in 0.1M DTT to reverse the cross-links, and analyzed by western blot as the input control.

Propidium iodide staining. Mock- or virus-infected A549 cells (including detached ones) were collected from 100 mm plates and washed with PBS once at 24, 48, or 72 hpi. A fraction of each set of cells were spun down and resuspended in culture medium with 1 μg/ml PI (Invitrogen). 10,000 cells of each sample were analyzed for cell viability, determined by PI exclusion using FACSCalibur flow cytometer (BD Biosciences).

Electron microscopy. 10 μl of virus particles purified from CsCl equilibrium centrifugation were applied to the shiny side of Formvar/Carbon film on 300 Square Mesh Nickel Grids (FCF300-Ni; Electron Microscopy Sciences), and fixed with 5% electron microscopy grade glutaraldehyde (16316-10; Electron Microscopy Sciences). After washes with TNE buffer (10 mM Tris, pH 7.5, 250 mM NaCl, 0.5 mM EDTA) and water, the grids were stained with NanoVan (2011; Nanoprobes) for 15 seconds. The grids with negatively stained virus particles were examined on an FEI Tecnai 12 BioTwin G02 microscope at an 80 kV, and digital images were acquired by AMT XR-60 charge-coupled-device digital camera system.

Analysis of L4-22K Functions

Generation, infectivity, and viral genome replication of L4-22K mutant viruses. In order to assess the role of L4-22K during Ad infection, two L4-22K mutant viruses were generated and characterized in this study: Δ22K (a null mutant) and 22K^{-} (an amber mutant) (Fig. 5A). A Δ22K infectious clone (pTG3602-Δ22K) was made by deleting the L4-33K intron leaving the L4-33K coding region intact but truncating the L4-22K protein at aa 105 (of 194). Deletion of the L4-33K intron shifts the reading frame of the remaining C-terminus of L4-22K and might generate an L4-22K variant which is predicted to be 5 kD larger than WT L4-22K. Since the L4-22K and L4-33K antibodies that we use throughout this study specifically recognize the distinct C-termini of L4-22K and L4-33K, respectively, we performed a western blot analysis of Δ22K-infected cell lysate using an antibody against the shared N-terminal region of L4-22K and L4-33K (75). However, we could not detect such an L4-22K variant because it would be masked by the L4-33K protein (Fig. 6). A 22K^{-} infectious clone (pTG3602-22K^{-}) which contains a stop codon (TAG) at aa 114 of L4-22K was previously constructed, without affecting the L4-33K ORF (51). Two different L4-22K mutant viruses were analyzed in these studies to verify the specificity of the results for the L4-22K mutations and because of uncertain effects that these different mutations may have on L4-33K expression. Our previous observation that transfection of pTG3602-22K^{-} resulted in no infectious virus production (51) suggested that an L4-22K complementing cell line would be needed to propagate the L4-22K mutant viruses. To reduce the chance of homologous recombination that would give rise to revertants, we made a tetracycline-inducible cell line expressing L4-22K-ΔC (1-173 aa of Ad5 L4-22K), termed TetC4-22K-ΔC (Fig. 5A). Earlier work from our laboratory had shown Ad17 L4-22K was able to complement the growth of pTG3602-22K^{-} in a complementation assay (79); Ad5 and Ad17 L4-22K proteins are well conserved except for an extra C-terminal 23 aa in Ad5 L4-22K. We reasoned that the extra C-terminal 23 aa of Ad5 L4-22K would not be required to complement Ad5 L4-22K mutant viruses, and that the removal of these sequences would eliminate or reduce the chance of generating revertants of the Δ22K and 22K^{-} mutant viruses, respectively. Southern blot analysis of DNA isolated from Δ22K- or 22K^{-}-infected A549 cells confirmed that both mutant virus stocks were pure (data not shown).



	WT	Δ22K	22K⁻
6 hpi	1.0	2.3	0.7
12 hpi	52	41	26
18 hpi	1.2E+03	1.4E+03	9.3E+02
32 hpi	4.7E+03	1.9E+04	7.5E+03

Figure 5. L4-22K mutant viruses. A. Schematic diagram of the Ad5 L4 region. Coding sequences for the L4-100K, L4-33K, L4-22K, and pVIII are shown. L4-22K mutants Δ22K and 22K⁻ are depicted along with the 22K-ΔC protein expressed in TetC4 cells. **B.** Complementation of Ad5-WT, Δ22K and 22K⁻ viruses in TetC4-22K-ΔC cells with (OFF) or without (ON) doxycycline (1 μg/ml) in the culture medium. Infected cell lysates were harvested at 48 hpi and subsequently titered on TetC4-22K-ΔC cells by plaque assay. **C**. Growth curves of Ad5-WT, Δ22K and 22K⁻ viruses in A549 cells. Infected A549 cell lysates were harvested at 6, 12, 24, 48 hpi and subsequently titered on TetC4-22K-ΔC cells by plaque assay. **D.** The particle/FFU ratios of Ad5-WT, Δ22K, and 22K⁻ viruses determined in A549 cells are presented. **E.** The particle/PFU ratios of Ad5-WT, Δ22K, and 22K⁻ viruses determined in TetC4-22K-ΔC cells are

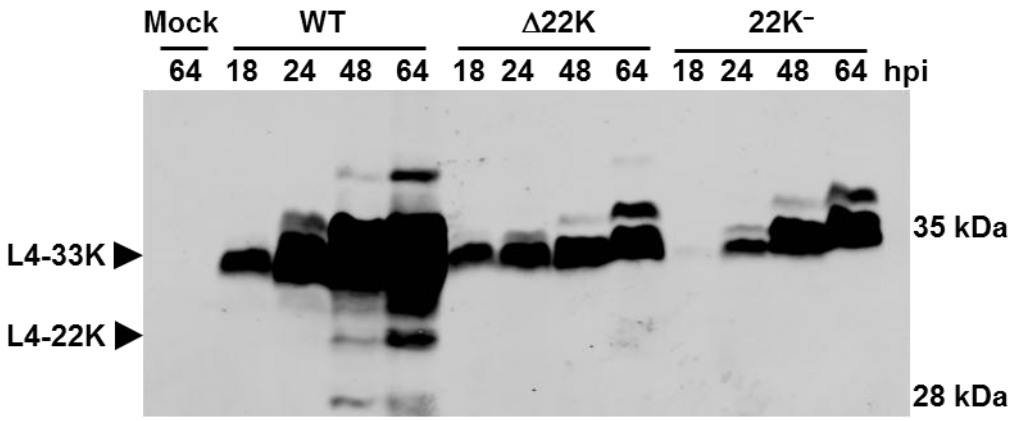


Figure 6. Relative abundance of L4-22K and L4-33K during Ad infection. Western blot analysis of L4-22K and L4-33K using whole cell extracts isolated from Ad5-WT-, Δ22K-, and 22K⁻-infected A549 cells at different times post-infection (indicated at the top, hpi). A mouse monoclonal antibody against the common N-terminus of L4-33K and L4-22K was used for the analysis. Protein designations are indicated on the left; protein molecular weight markers are indicated on the right.

To confirm the complementation of the L4-22K mutant viruses in TetC4-22K-ΔC cells, the titers of Ad5-WT, Δ22K and 22K⁻ viruses produced in TetC4-22K-ΔC cells with (OFF) or without (ON) doxycycline in the culture medium was measured. Indeed, TetC4-22K-ΔC cells efficiently complemented the growth of Δ22K and 22K⁻ mutant viruses (Fig. 5B). In addition, a one-step growth curve assay confirmed that mutation of L4-22K was lethal to virus viability in non-complementing A549 cells (Fig. 5C).

To test the infectivity of Δ22K and 22K⁻ viruses, virus particles were used in a plaque assay in the L4-22K complementing cell line, and in a fluorescent focus assay in A549 cells. The particle/PFU and particle/FFU ratios for each virus were calculated and the result showed no apparent differences between Ad5-WT and the two L4-22K mutant viruses (Fig. 5D and 5E). Next, we evaluated viral DNA replication (Fig. 5F). From the onset of infection to 18 hpi, Ad5-WT and both L4-22K mutant viruses exhibited similar input (6 hpi) and DNA replication levels (12, 18 hpi). At 32 hpi, however, the L4-22K mutant viruses replicated 2-4 fold higher than Ad5-WT. In contrast to these results, neither L4-22K mutant produced detectable infectious virus as

---

presented. **F.** Viral genome replication in Ad5-WT-, Δ22K-, or 22K⁻-infected cells. The genome copy number was normalized to GAPDH and this value was relative to that of WT 6 hpi for presentation.

measured in a growth curve (Fig. 5C). We conclude that both L4-22K mutant viruses had comparable infectivity and viral genome replication as Ad5-WT, and that the L4-22K protein is critical for the production of infectious progeny.

L4-22K mutant virus gene expression. To examine viral gene expression with the L4-22K mutant viruses, A549 cells were infected with Ad5-WT, Δ22K, or 22K^{-} viruses and total cell extracts were prepared at different time points and analyzed by western blot. With early gene expression, both L4-22K mutant viruses had higher E1A expression levels than Ad5-WT during the late phase of infection (Fig. 7 left, 24–64 hpi). In addition, DBP levels were also increased 2-4 fold at late times after infection, which corresponded to the higher viral genome replication with the L4-22K mutant viruses observed at 32 hpi (Fig. 5F). The analysis of intermediate (IVa2) and late (L1-L5) gene expression showed that there was a delay and decrease to various extents with the L4-22K mutant viruses (Fig. 7, left). Among the delayed and decreased gene products, major capsid proteins (hexon, penton, and fiber) and one minor capsid protein (IIIa) were modestly decreased, and IVa2 and L1-52/55K were delayed. Other minor capsid proteins (VI and VIII), core proteins (V and VII), and L4-33K exhibited a strong decrease compared with Ad5-WT. In contrast, L4-100K levels were quite similar with Ad5-WT and the L4-22K mutant viruses. Full-length L4-22K protein was not detected with the mutant viruses, as expected, and stable truncated products were not observed. Similar results were found following infection of N52.E6-Cre cells (human aminocytes) (Fig. 7 right), but the extent of delay and decrease for the late gene products was less apparent with the L4-22K mutant viruses when compared to Ad5-WT-infected A549 cells. In addition, the increases in E1A and DBP expression in L4-22K mutant virus infections were also observed in N52.E6-Cre cells during late phase of infection (Fig. 7, right). These results demonstrated that mutation of the L4-22K gene had pleiotropic effects on viral early and late gene expression; increased early gene expression and decreased late gene expression indicates that the L4-22K protein is important to drive the early-to-late phase switch during Ad infection.

To further explore the increased levels of E1A observed with L4-22K mutant viruses during the late phase of infection, northern blot analysis was performed to measure E1A mRNA levels. These results showed that E1A mRNA levels were strongly elevated in both Δ22K- and 22K^{-}-infected cells at 24 hpi compared to Ad5-WT-infected cells (Fig. 8). This increase suggests that the L4-22K protein plays a role in suppressing early gene expression.

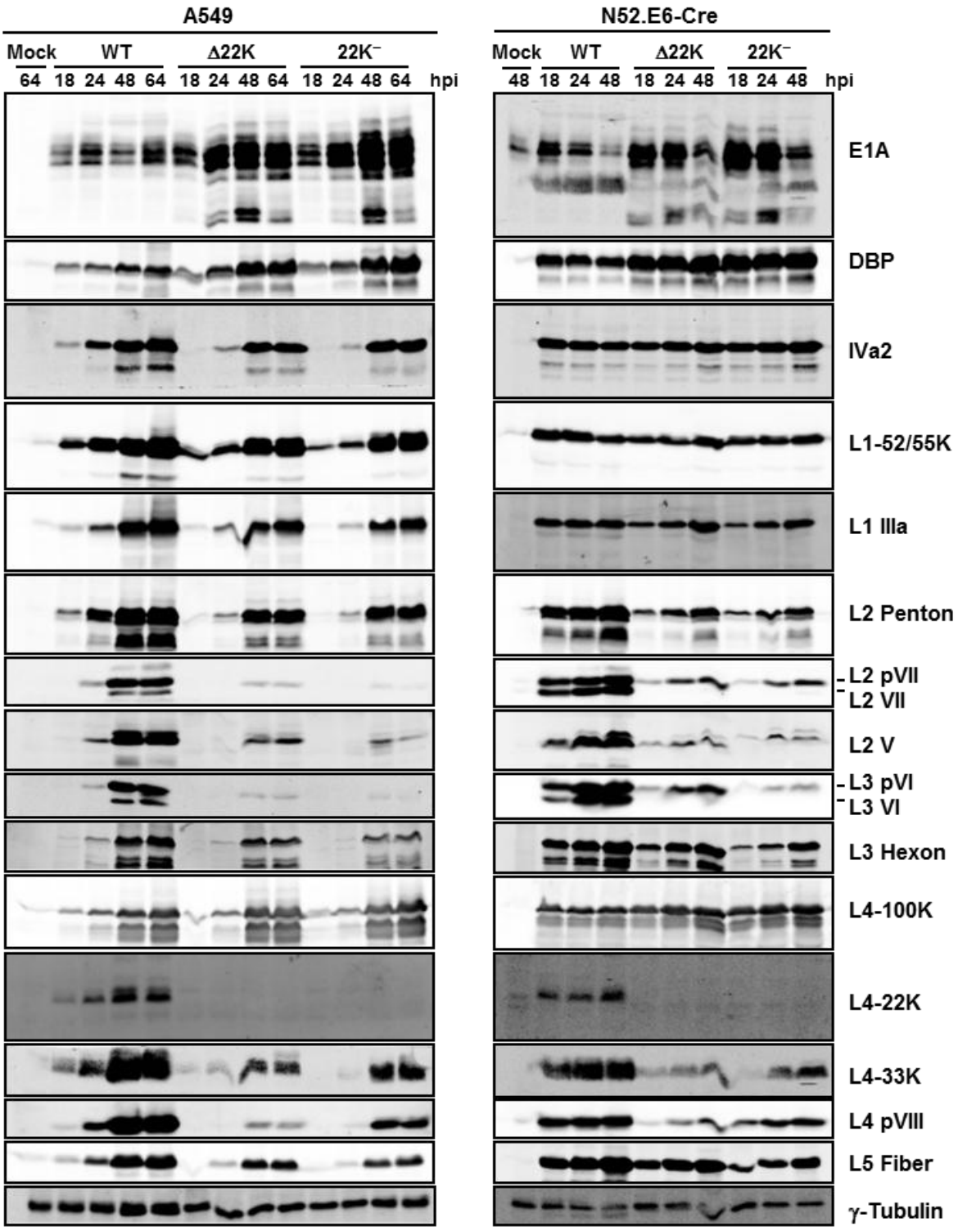


Figure 7. Western blot analysis of Ad early and late proteins. Representative early (E1A and DBP), intermediate (IVa2), and late (L1-L5) gene products were analyzed using whole cell extracts isolated from Ad5-WT-, Δ22K-, and 22K⁻-infected A549 cells (left) and N52.E6-Cre cells (right) at different times post-infection (indicated at the top, hpi). Protein designations are indicated on the right; Tubulin was analyzed as a loading control.

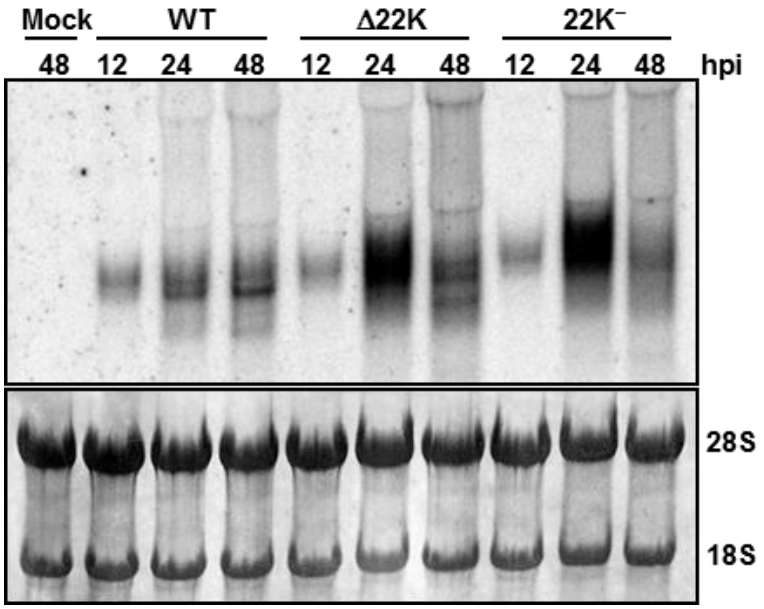


Figure 8. Northern blot analysis of E1A mRNA levels. Total cytoplasmic mRNAs isolated from Ad5-WT-, Δ22K-, and 22K⁻-infected A549 cells at 6, 12, 24, and 48 hpi were analyzed using a ^{32}P-labeled E1A probe (top panel). The blot was stained with methylene blue as a loading control (bottom panel). This experiment was conducted by Diana Orozco.

To investigate the mechanism by which L4-22K activates late gene expression while suppressing early gene expression, we performed luciferase reporter assays to analyze the role of L4-22K in early and late gene transcription. Consistent with a previous report, expression of L4-22K increased MLP transcription, while co-expression of IVa2 with L4-22K weakly boosted this increase (Fig. 9A). In addition, the transcription of the E1A promoter was also analyzed and the result showed there was a moderate reduction of E1A transcription in the presence of both L4-22K and IVa2, while L4-22K alone did not suppress E1A transcription (Fig. 9B). These results from luciferase assays suggest that L4-22K efficiently activates the MLP by itself, while it suppresses E1A transcription in the presence of IVa2. However, from western blot analysis of E1A expression in *pm*8002- (a IVa2 mutant virus) infected A549 cells, we observed that in the absence of full-length IVa2, E1A expression was comparable with that of Ad5-WT throughout the course of infection (Fig. 10), which suggests IVa2 is not necessary for suppressing E1A expression mediated by L4-22K during infection.

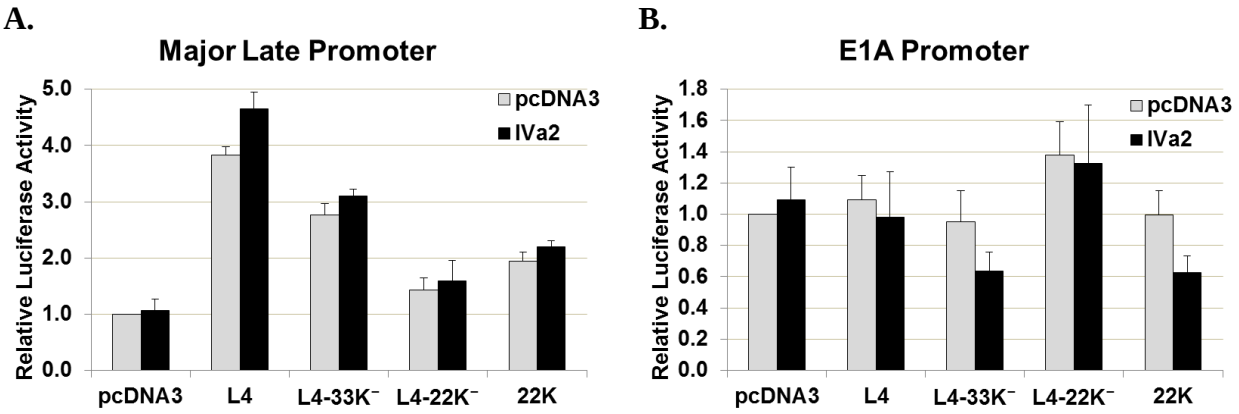


Figure 9. Transcriptional regulation of Ad early and late gene expression by L4-22K. Luciferase reporter plasmids of the MLP **(A)** or E1A promoter **(B)** were independently transfected with pcDNA3 (expression vector), L4 (expressing both L4-22K and L4-33K), L4-33K⁻ (expressing L4-22K only), L4-22K⁻ (expressing L4-33K only) or 22K, into 293 cells. 24 h post-transfection, luciferase assays were performed and the relative luciferase activity of the Firefly luciferase signal (test promoter) was normalized to the Renilla luciferase signal (transfection control).

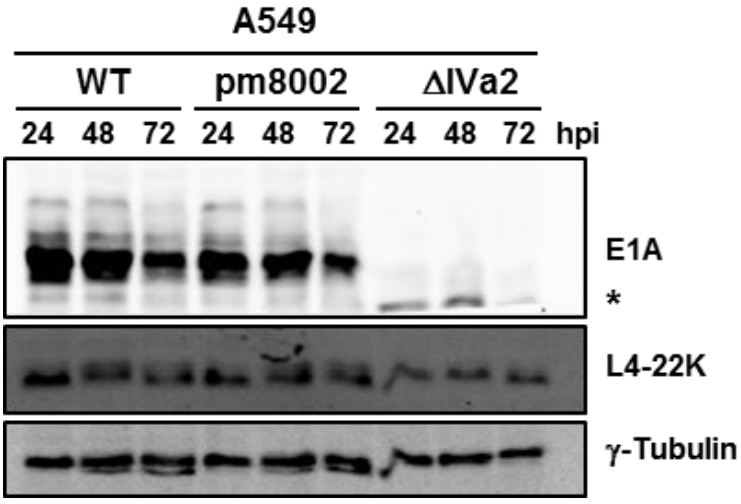


Figure 10. Western blot analysis of E1A expression in *pm*8002- and ΔIVa2-infected cells. E1A was analyzed using whole cell extracts isolated from Ad5-WT-, *pm*8002-, or ΔIVa2-infected A549 cells at different times post-infection (indicated at the top, hpi). Protein designations are indicated on the right; Tubulin was analyzed as a loading control. The asterisk indicates the truncated version of E1A expressed from ΔIVa2 virus due to its construction plan.

L4-22K mutant viruses only produce empty capsids. To determine if L4-22K is required for Ad genome packaging *in vivo*, A549 and N52.E6-Cre cells were infected with the L4-22K mutant viruses and virus particle production was analyzed using CsCl equilibrium density gradient centrifugation. In both A549 and N52.E6-Cre cells, only EC were observed from Δ22K- or 22K^{-}-infected cells which banded at a density of 1.29 g/cc (Fig. 11A). The yield of EC from both L4-22K mutant viruses was lower in A549 than in N52.E6-Cre cells which might reflect the more permissive nature of N52.E6-Cre cells compared to A549 cells since elevated levels of Ad5-WT were also observed. This phenotype of the L4-22K mutant viruses strongly indicates that the L4-22K protein is required for genome packaging. However, many of the late viral gene products were moderately or strongly delayed and decreased, including IVa2 and L4-33K that regulate late gene expression on transcriptional and post-transcriptional levels, respectively (Fig. 7) (1, 58, 72, 87). In order to examine the contribution of delayed IVa2 and reduced L4-33K gene expression in the L4-22K mutant virus phenotype, we infected IVa2- or L4-33K-expressing cell lines with the Δ22K mutant virus to ask if virion production was altered when IVa2 or L4-33K were supplemented *in trans*. These results showed that supplementation of IVa2 or L4-33K to Δ22K infection did not augment MV production (Fig. 11A). Curiously, Δ22K infection of the IVa2-expressing cell line did increase EC production, although we do not know if this represents complementation or a property of this particular cell subclone. Since EC were produced in L4-22K mutant virus infections that had not encapsidated the viral genome, we conclude that the L4-22K protein is required for viral DNA packaging.

To analyze whether the EC generated by Δ22K and 22K^{-} viruses have a similar protein composition to the control EC generated from Ad5-Ψ-loxP virus, we performed silver staining and western blot to analyze the viral proteins on these EC. Ad5-Ψ-loxP virus has loxP sites flanking the PS, so that upon infection in Cre-recombinase-expressing cells (N52.E6-Cre cells) only EC are produced due to the floxing of the PS. We also used the MV generated from Ad5-WT-infected cells as the MV control. These results revealed that EC generated from Δ22K or 22K^{-} mutant viruses in A549, N52.E6-Cre, TetC4-IVa2, or TetC4-33K cells had very similar protein composition to the EC of the grown in N52.E6-Cre cells (Figs. 11B and 11C). All of these EC lacked core protein V and the proteolytically-processed form of pVII, both characteristic of MV, and they all contained precursor forms of pVI and pVIII, as well as readily evident levels of L1-52/55K and L4-100K, which are present in a very low abundance in MV. In

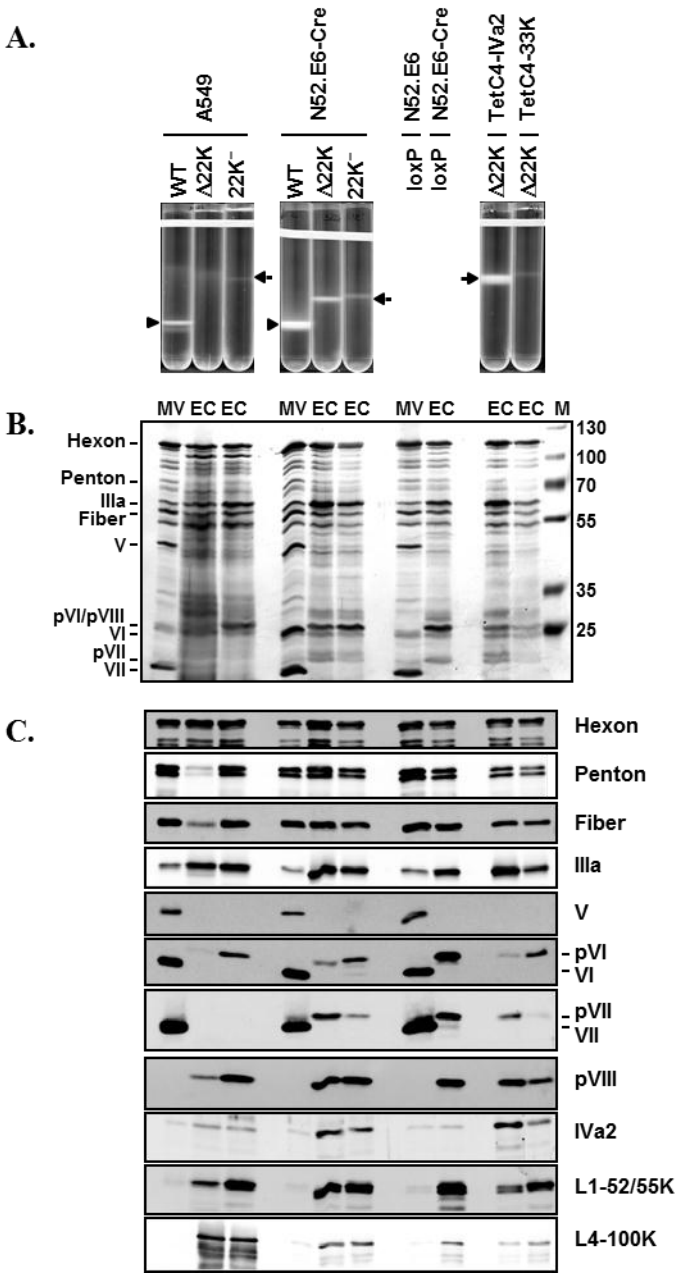


Figure 11. Virus particle production with L4-22K mutant viruses. A. CsCl density equilibrium gradient profiles of virus particle produced from Ad5-WT-, Δ22K-, 22K⁻-infected A549 and N52.E6-Cre cells, and Δ22K-infected TetC4-IVa2 and TetC4-33K cells. Arrows and arrowheads indicate mature virions (MV) and empty capsids (EC), respectively. **B.** Silver staining of protein components of MV and EC isolated from virus particles indicated in A; in addition MV and EC isolated from Ad5-ΨloxP-infected N52.E6 (MV) and N52.E6-Cre (EC) cells are shown. Protein designations are indicated on the left. M, molecular weight markers. **C.** Western blot analysis of protein components of MV and EC isolated from virus particles indicated in A. Protein designations are indicated on the right.

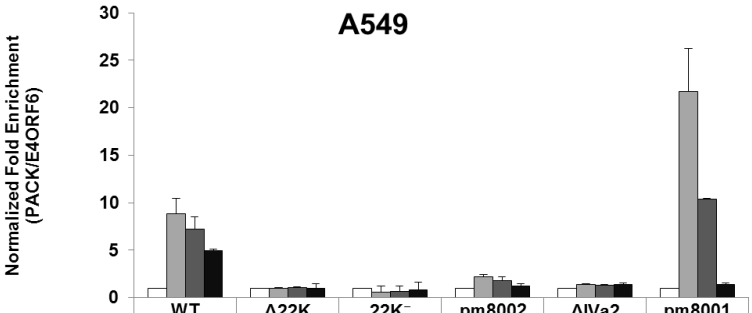
addition, the Δ22K or 22K^{-} EC had slightly elevated levels of IVa2 in comparison to MV or EC from Ad5-Ψ-loxP. Thus, the EC generated by the L4-22K mutant viruses are *bona fide* empty capsids.

The L4-22K protein binds to the PS *in vivo* only in the presence of IVa2. To explore packaging *in vivo*, we performed ChIP assays to decipher the relationship between the viral packaging proteins, IVa2, L4-22K, and L1-52/55K, and the PS in the genome. A549 cells infected with Ad5-WT, Δ22K, 22K^{-}, *pm*8002 (a IVa2 mutant virus) (87), ΔIVa2 (50), or *pm*8001 (an L1-52/55K mutant virus) (19) were analyzed by ChIP at 18 hpi using rabbit pre-immune serum, or rabbit serum directed against L4-22K, IVa2, or L1-52/55K. These results showed that the L4-22K protein bound to the PS of Ad5-WT *in vivo*, but this was not observed in Δ22K- or 22K^{-}-infected cells, as expected (Fig. 12A). In addition, L4-22K did not bind to the PS in *pm*8002- or ΔIVa2-infected cells, indicating that the IVa2 protein is required for L4-22K to bind to the PS *in vivo*, consistent with *in vitro* DNA-protein binding studies (51, 75, 83). Results from *pm*8001-infected cells demonstrated that in the absence of L1-52/55K, the interaction between L4-22K or IVa2 and the PS was not reduced. IVa2 bound to the PS in Ad5-WT- or *pm*8001-infected cells, but not in *pm*8002- or ΔIVa2-infected cells, consistent with previous reports (54, 61). However, IVa2 did not bind to the PS in Δ22K- or 22K^{-}-infected cells, implying that the L4-22K protein is essential for IVa2 to bind to the PS *in vivo*, in contrast to *in vitro* binding analyses (13, 54, 84). In addition, in the absence of L4-22K (Δ22K, 22K^{-}) or IVa2 (*pm*8002), the binding of L1-52/55K to the PS *in vivo* was completely abolished, indicating that both the L4-22K and IVa2 proteins are required for L1-52/55K to bind to the PS. The loss of L1-52/55K binding to the PS in the absence of the L4-22K protein may be due, at least in part, to decreased L1-52/55K expression in Δ22K- or 22K^{-}-infected cells (Fig. 7) and ChIP input control (Fig. 12D). This is not the case with *pm*8002 where WT levels of L1-52/55K were evident (Fig. 12D). Taken together, the ChIP experiments indicate that the IVa2 protein is essential for L4-22K and L1-52/55K to bind to the PS *in vivo*, the L4-22K protein is likely required for IVa2 and L1-52/55K to bind the PS *in vivo*, and the L1-52/55K protein is dispensable for IVa2 or L4-22K binding to the PS *in vivo*.

The L4-22K protein is required for IVa2 to bind to the PS *in vivo*. To examine the possibility that the loss of IVa2 binding to the PS in Δ22K- or 22K^{-}-infected cells was due to the decreased level of IVa2 expression, we infected TetC4-IVa2 cells with Ad5-WT, Δ22K, 22K^{-}, *pm*8002 (a IVa2 mutant virus), or ΔIVa2 for ChIP analyses; this provided similar levels of the IVa2 protein in the different virus infections (Fig. 12D). These results demonstrated that complementation of IVa2 protein levels in IVa2 mutant virus infections not only restored the binding of IVa2, but also the binding of the L4-22K and L1-52/55K proteins to the PS *in vivo* (Fig. 12B). However, the supplementation of IVa2 protein levels did not restore the binding of IVa2 or L1-52/55K to the PS in Δ22K- or 22K^{-}-infected cells (Fig. 12B), even though IVa2 protein levels were the same among the different virus-infected cells (Fig. 12D). These results ruled out the possibility that the decreased IVa2 protein expression was responsible for the loss of IVa2 binding to the PS in Δ22K- or 22K^{-}-infected cells, and confirmed that the L4-22K protein is required for IVa2 and L1-52/55K to bind to the PS *in vivo*. We note that there was a similar level of L1-52/55K expression among TetC4-IVa2 cells infected with the different viruses (Fig. 12D).

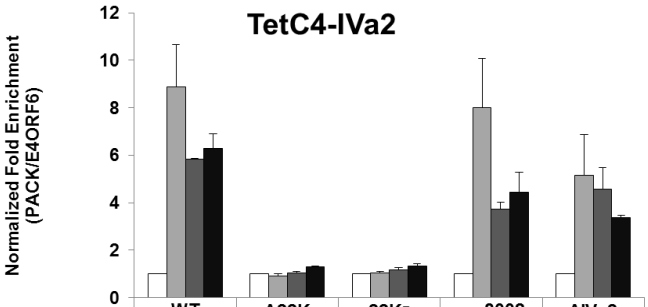
To further assure the relationship between L4-22K, IVa2, L1-52/55K and binding to the PS *in vivo*, the L4-22K expressing cell line, TetC4-22K-ΔC, was employed in ChIP analyses. Complementation of L4-22K fully restored the expression levels of IVa2 and L1-52/55K in cells infected with the L4-22K mutant viruses (Fig. 12D), and partially restored the binding of L4-22K-ΔC, IVa2, and L1-52/55K to the PS in Δ22K- or 22K^{-}-infected cells (Fig. 12C). The different extent of binding of IVa2 and L1-52/55K to the PS between Ad5-WT and two L4-22K mutant viruses may be due to the fact that the WT virus had additional full-length WT L4-22K expressed in addition to the L4-22K-ΔC provided by the cell line. Surprisingly, in *pm*8002-infected TetC4-22K-ΔC cells, where there was no full-length IVa2 expression (Fig. 12D), the binding of the IVa2, L4-22K, and L1-52/55K proteins to the PS was partially restored; this was not observed with the ΔIVa2 mutant virus (Fig. 12C). Western blot of the input samples showed that the 40 kD IVa2 isoform (57) was expressed in *pm*8002-infected TetC4-22K-ΔC cells, even though the full-length IVa2 protein was not evident, as expected (Fig. 12D). Since it is known that 40 kD IVa2 isoform binds to the PS (84), its presence may be responsible for the partially restored binding of the IVa2, L4-22K and L1-52/55K proteins to the PS in *pm*8002-infected TetC4-22K-ΔC cells.

A.



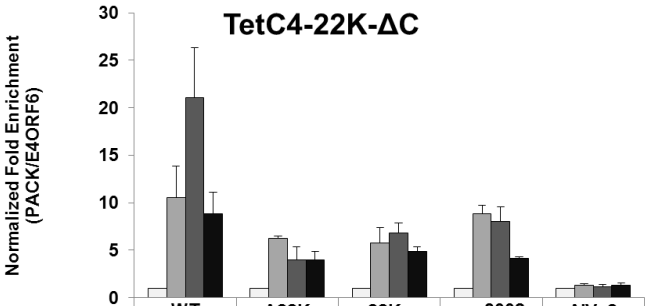
	WT	Δ22K	22K⁻	pm8002	ΔIVa2	pm8001
pre-immune	1.0	1.0	1.0	1.0	1.0	1.0
αL4-22K	8.9	1.0	0.6	2.2	1.4	21.7
αIVa2	7.2	1.1	0.6	1.8	1.3	10.4
αL1-52/55K	5.0	1.0	0.8	1.2	1.4	1.4

B.



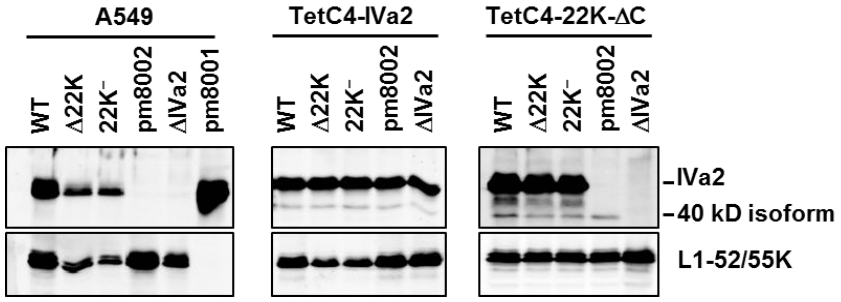
	WT	Δ22K	22K⁻	pm8002	ΔIVa2
pre-immune	1.0	1.0	1.0	1.0	1.0
αL4-22K	8.9	0.9	1.0	8.0	5.2
αIVa2	5.8	1.0	1.2	3.7	4.6
αL1-52/55K	6.3	1.3	1.3	4.5	3.4

C.



	WT	Δ22K	22K⁻	pm8002	ΔIVa2
pre-immune	1.0	1.0	1.0	1.0	1.0
αL4-22K	10.6	6.2	5.8	8.8	1.3
αIVa2	21.0	4.0	6.8	8.1	1.2
αL1-52/55K	8.8	3.9	4.9	4.1	1.3

D.



The L4-22K protein promotes Ad-induced cell lysis. During the course of these studies, we noted that there was a remarkable difference in the morphology of A549 cells infected with Ad5-WT and the L4-22K mutant viruses (Fig. 13A). With Ad5-WT-infected cells, apparent cytopathic effect was observed at 48 hpi with most of the cells rounded up and detached from the plate. In contrast, Δ22K- or 22K^{-}-infected cells had enlarged nuclei, but were still confluent and attached to the plate (Fig. 13A). In order to determine if the L4-22K protein influenced cell viability, PI staining was used to monitor cell death in Ad5-WT-, Δ22K-, or 22K^{-}-infected cells at different time points after infection. The results showed that Ad5-WT killed cells efficiently by 48–72 hpi, whereas both L4-22K mutant viruses showed very modest cell killing compared with mock-infected cells (Fig. 13B). This result indicated that the L4-22K protein contributes to cell killing during Ad infection. ADP is known to be responsible for Ad-induced cell lysis (70). We examined ADP expression levels in Ad5-WT-, Δ22K- or 22K^{-}-infected A549 cells. Western blot analysis showed that ADP expression levels were dramatically decreased in cells infected with the L4-22K mutant viruses compared to Ad5-WT (Fig. 13C). ADP is transcribed from the MLP during the late phase of infection (71), thus the L4-22K protein likely influences ADP synthesis in parallel with other Ad late genes. These data suggest that the L4-22K protein promotes Ad-induced cell lysis by influencing ADP expression.

Figure 12. ChIP assays for packaging proteins binding to the PS *in vivo*. ChIP assays for packaging protein binding to the PS *in vivo*. A549 **(A)**, TetC4-IVa2 **(B)**, and TetC4-22K-ΔC **(C)** cells were infected Ad5-WT, Δ22K, 22K^{-}, *pm*8002, ΔIVa2, or *pm*8001 and L4-22K-, IVa2-, and L1-52/55K-specific antibodies were used for ChIP to quantify binding to the PS. The readout is presented as normalized fold-enrichment by dividing the copy-number of PS pulled-down specifically by the antibody to the copy-number of E4-ORF6 fragment that was pulled down non-specifically. For each individual virus infection, the fold-enrichment value of each antibody was normalized to that of the pre-immune serum negative control. **D.** Western blot analysis of ChIP input samples for IVa2 and L1-52/55K protein expression levels.

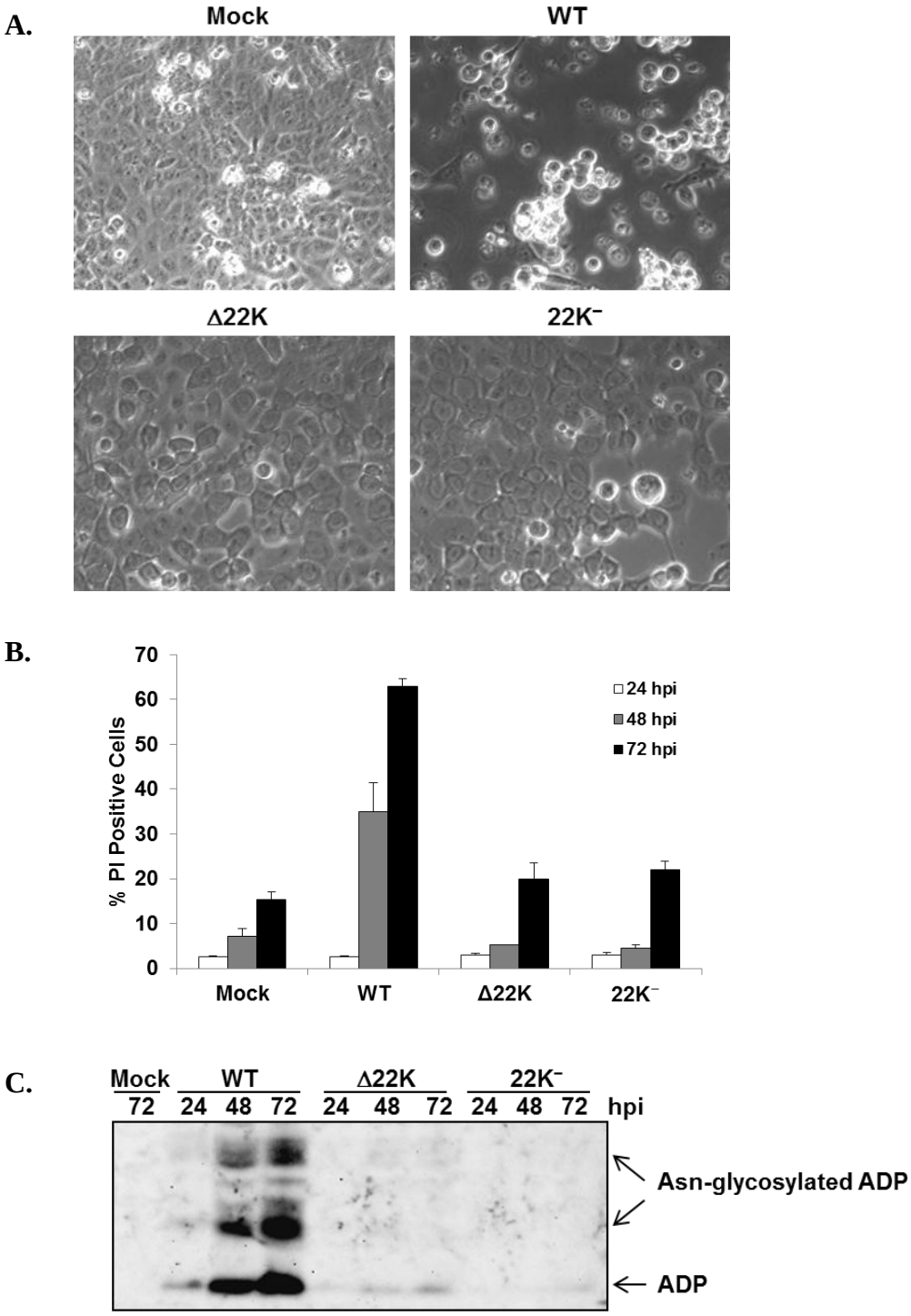


Figure 13. L4-22K promotes Ad-induced cell lysis by influencing ADP expression. A. Light microscopic images of Mock-, Ad5-WT-, Δ22K-, or 22K⁻-infected A549 cells at 48 hpi. **B.** Propidium iodide staining of Mock-, Ad5-WT-, Δ22K-, or 22K⁻-infected A549 cells at 24, 48, and 72 hpi. **C.** Western blot analysis of ADP expression levels in total cell extracts prepared from Mock-, Ad5-WT-, Δ22K-, or 22K⁻-infected A549 cells at 24, 48, and 72 hpi. Free and Asn-glycosylated ADP forms are indicated on the right.

Analysis of L4-33K Functions

Isolation of an Ad5 L4-33K mutant virus, infectivity, and viral genome replication. The Ad5 L4 region encodes four proteins: L4-100K, L4-33K, L4-22K, and pVIII (Fig. 14A). L4-33K and L4-22K share a common 105 aa N-terminus; the L4-33K mRNA is spliced to generate a unique L4-33K C-terminus. The E2E promoter is located on the opposite strand of the L4 region and overlaps with part of the L4-33K C-terminus (17). In order to make an Ad5 L4-33K mutant virus, an amber mutation (TAG) was introduced at aa position 138 of L4-33K in the background of the Ad5 infectious clone pTG3602 (termed pTG3602-33K$^-$) in which both the L4-22K reading frame and E2E promoter were kept intact (Fig. 14A). Consistent with previous reports that mutation of L4-33K is lethal (17, 33); pTG3602-33K$^-$ was not viable in non-complementing cells (data not shown). A tetracycline-inducible cell line expressing full-length Ad5 L4-33K, termed TetC4-33K, was isolated to complement the growth of the L4-33K mutant virus. Plaque-purified 33K$^-$ virus was amplified in TetC4-33K cells and mutant virus particles were purified by CsCl equilibrium centrifugation for use in all of the following experiments. Southern blot analysis of DNA isolated from 33K$^-$-infected A549 cells confirmed that the mutant virus stock was pure (data not shown). To confirm the complementation of the L4-33K mutant in TetC4-33K cells, the titers of Ad5-WT and 33K$^-$ viruses produced in TetC4-33K cells with (OFF) or without (ON) doxycycline in the culture medium was measured. These results showed that the TetC4-33K cell line complemented the growth of the 33K$^-$ mutant virus to a comparable level with Ad5-WT when L4-33K expression was turned on (Fig. 14B). In addition, a one-step growth curve assay confirmed that mutation of L4-33K was lethal to virus viability in non-complementing A549 cells (Fig. 14C). We note that turning on L4-33K expression slightly decreased the yield of Ad5-WT (Fig. 14B) and that extensive cell death was observed after turning on L4-33K expression for 72 h (data not shown). Thus, L4-33K expression has a toxic effect on cells that may relate to the disruption of normal splicing patterns of cellular transcripts in light of the role of L4-33K in regulating alternative RNA splicing (72).

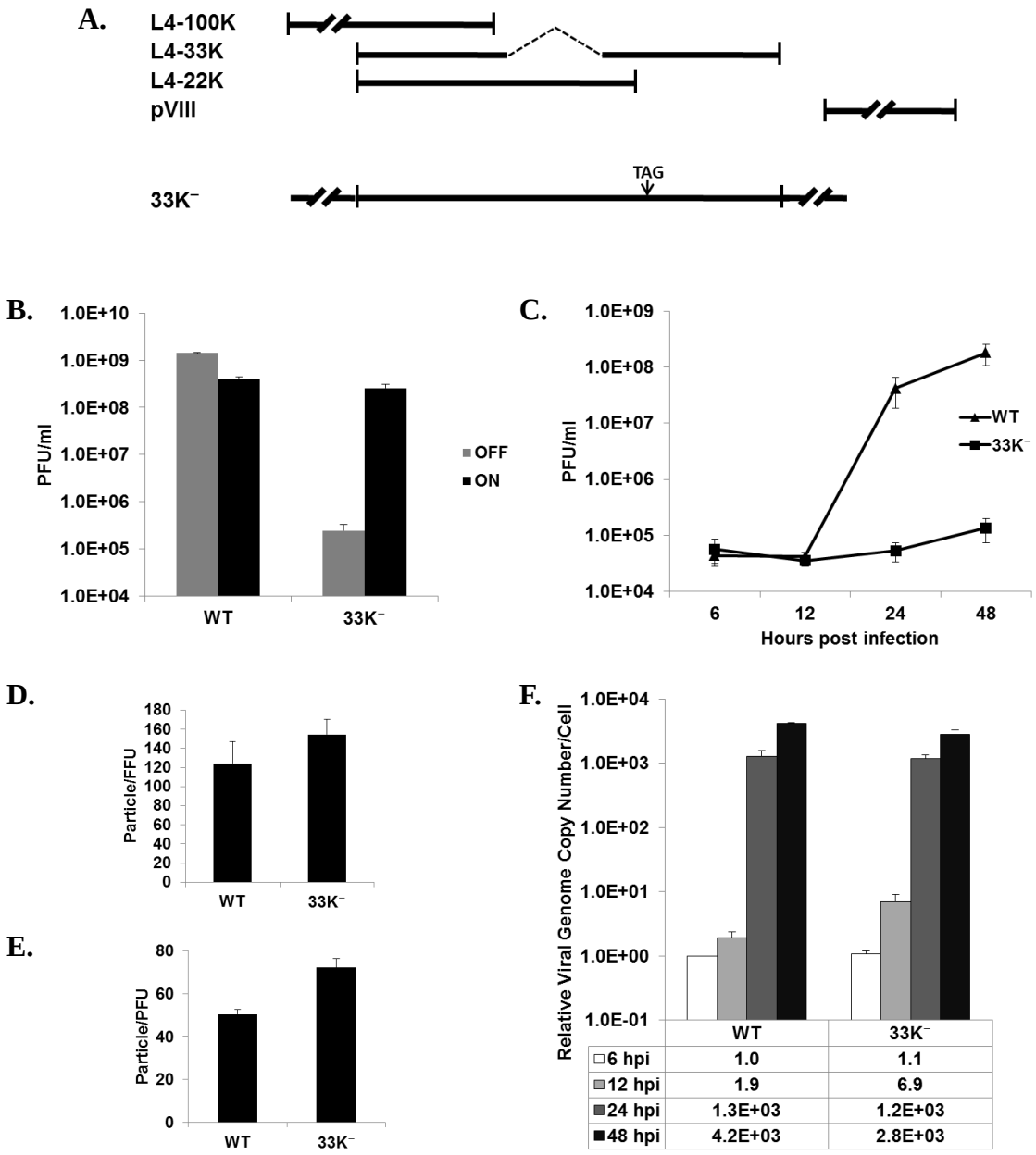


Figure 14. L4-33K mutant virus. A. Schematic diagram of the Ad5 L4 region. Coding sequences for the L4-100K, L4-33K, L4-22K, and pVIII are shown. L4-33K mutant 33K⁻ is depicted. **B.** Complementation of Ad5-WT and 33K⁻ viruses in TetC4-33K cells with (OFF) or without (ON) doxycycline (1 μg/ml) in the culture medium. Infected cell lysates were harvested at 48 hpi and subsequently titered on TetC4-33K cells by plaque assay. **C**. Growth curves of Ad5-WT and 33K⁻ viruses in A549 cells. Infected A549 cell lysates were harvested at 6, 12, 24, 48 hpi and subsequently titered on TetC4-33K cells by plaque assay. **D.** The particle/FFU ratios of Ad5-WT and 33K⁻ viruses determined in A549 cells are presented. **E.** The particle/PFU ratios of Ad5-WT and 33K⁻ viruses determined in TetC4-33K cells are presented. **F.** Viral genome replication in Ad5-WT- and 33K⁻- infected cells. The genome copy number was normalized to GAPDH and this value was relative to that of WT 6 hpi for presentation.

In order to investigate at which step(s) of the Ad life cycle L4-33K plays an essential role(s), the infectivity of 33K$^-$ virus particles was determined using a fluorescent focus assay in A549 cells using antibody against DBP and plaque assay in TetC4-33K cells in comparison to Ad5-WT. The results from both assays demonstrated that the 33K$^-$ virus had similar particle to FFU and PFU ratios to those of Ad5-WT (Figs. 14D and 14E), suggesting the 33K$^-$ mutant virus behaved comparably to Ad5-WT from the onset of infection to viral genome replication. Next, viral genome replication of Ad5-WT and 33K$^-$ was measured in infected A549 cells (Fig. 14F). These results exhibited comparable replication kinetics between the two viruses, although 33K$^-$ had a slightly lower replication than Ad5-WT at 48 hpi. From these data, we conclude that 33K$^-$ virus has similar infectivity and viral genome replication to Ad5-WT and that the growth defect of the 33K$^-$ mutant virus is likely due to later events after viral genome replication.

L4-33K mutant virus gene expression. Given the facts that 33K$^-$ virus has comparable infectivity and genome replication with Ad5-WT and that the L4-33K protein is known as an alternative RNA splicing factor (72), we performed northern blot and western blot analyses to evaluate viral gene expression patterns. Consistent with previous results using an *in vitro* splicing assay (72), northern blot results confirmed that IIIa mRNA levels were strongly reduced with the 33K$^-$ mutant virus (Fig. 15A, L1 IIIa). L3 pVI mRNA levels also were strongly reduced in 33K$^-$-infected cells, and L2 V, L2 pVII, L3 hexon, and L5 fiber mRNA levels also were diminished, but to a lesser extent (Fig. 15A). In contrast, notable increases in L1-52/55K, L4-100K, and L4-22K mRNAs were observed in 33K$^-$-infected A549 cells (Figs. 15A and 15B). L4-22K mRNAs were not detectable by northern blot so were evaluated by semi-quantitative RT-PCR (Fig. 15B)

To further explore viral gene expression at the protein level, we performed western blot analyses of viral early and late proteins using Ad5-WT- or 33K$^-$-infected A549 and TetC4 whole cell extracts (WCE) and nuclear extracts (NE) (Fig. 15C). These results largely recapitulated the northern blot assays and demonstrated that mutation of L4-33K led to increased levels of L1-52/55K, L4-100K, and L4-22K proteins, and decreased levels of IIIa, fiber, and pVI. The levels of V, pVII, and hexon remained almost the same as with Ad5-WT (Fig. 15C). The most notable effects in these analyses with the 33K$^-$ virus was a strong reduction in pVI and strong increase in L4-22K. Comparable results were obtained using nuclear extracts from infected A549 cells

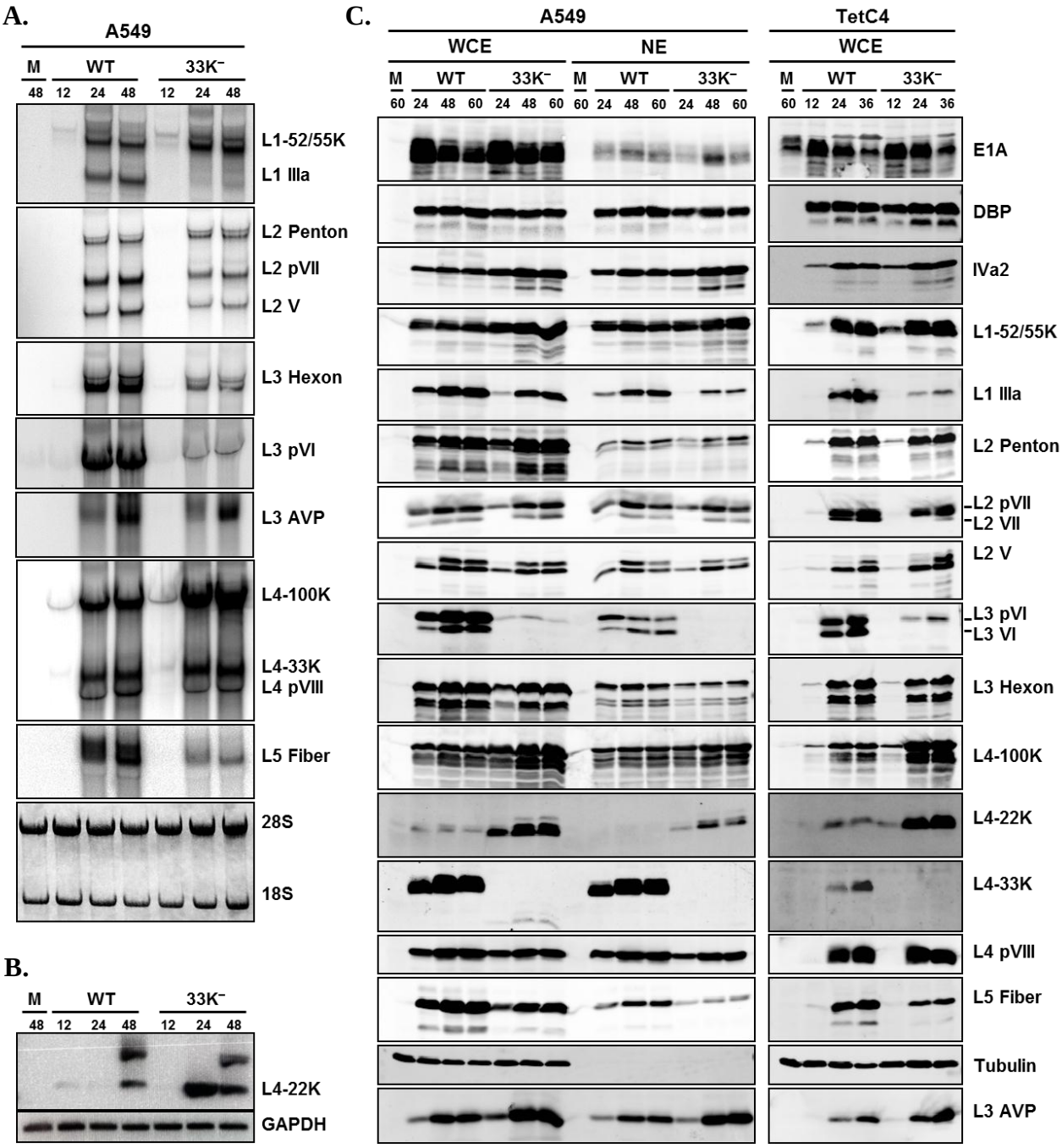


Figure 15. L4-33K mutant virus gene expression. A. Nothern blot analysis of Ad late gene mRNAs in Ad5-WT- or 33K⁻-infected A549 cells. Total cytoplasmic mRNAs isolated from the infected cells at 12, 24, and 48 hpi were analyzed using ^{32}P-labeled probes for each late region. The blot was stained with methylene blue as a loading control. **B.** RT-PCR analysis of L4-22K mRNA level in Ad5-WT- or 33K⁻-infected A549 cells. **C.** Western blot analysis of Ad early and late proteins in A549 and TetC4 cells. Representative early (E1A and DBP), intermediate (IVa2), and late (L1-L5) gene products were analyzed using whole cell extracts (WCE) and nuclear extracts (NE) isolated from Ad5-WT- and 33K⁻-infected A549 cells (left), and WCE isolated

indicating no apparent changes in nuclear import of Ad proteins with the 33K^{-} virus. Notably, early and intermediate gene products E1A, DBP, and IVa2 were not affected in 33K^{-}-infected cells; this also confirmed that mutation of L4-33K did not disrupt E2E transcription. In addition, there was no full-length L4-33K expression with 33K^{-} virus, but trace amounts of an apparently truncated version were evident which assured the integrity of the L4-33K mutant virus stock. The reduction of most viral late proteins (Fig. 15C) was not as evident as the corresponding reduction of mRNA levels (Fig. 15A) in 33K^{-}-infected cells. This discrepancy might be due to the increased level of L4-100K, which is known to increase translation of viral late gene transcripts that contain a tripartite leader (TPL) (81).

In conclusion, comprehensive northern blot and western blot analyses demonstrate that the L4-33K protein regulates viral late gene expression and that L4-33K controls the proper transition of gene expression during the late phase of infection.

The L4-33K mutant virus only produces empty capsids. We analyzed virus production in 33K^{-}-infected cells to determine whether mutation of L4-33K has any effect on virus assembly, viral genome packaging, or other late events of the Ad life cycle. Only EC, but not MV, were produced in 33K^{-}-infected A549, 293, and TetC4 cells (Fig. 16A). This result demonstrates that instead of having a virion assembly defect, mutation of L4-33K impaired the viral genome packaging process. The morphology of EC produced from 33K^{-}-infected cells was assessed by electron microscopy (Fig. 16B, 16C and 16D right). The 33K^{-} EC exhibited higher electron density compared to MV produced by Ad5-WT (Fig. 16D, left), which was very similar to the EC produced from a virus lacking the PS (loxP) (Fig. 16D, middle).

---

from Ad5-WT- and 33K^{-}-infected TetC4 cells (right) at different time (h) post-infection (indicated at the top). Protein designations are indicated on the right; Tubulin was analyzed as a loading control. The northern blot and RT-PCR were conducted by Diana Orozco.

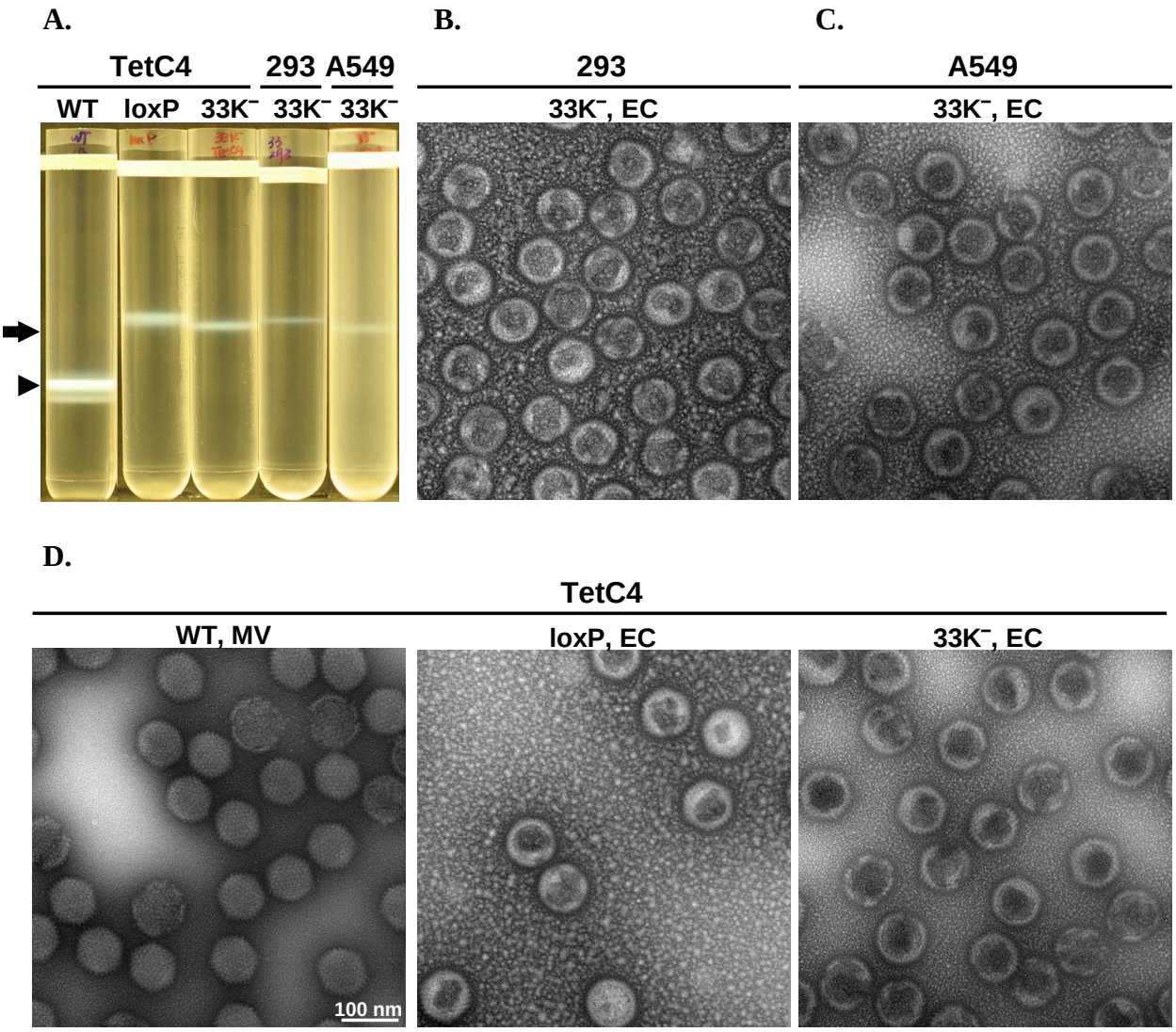
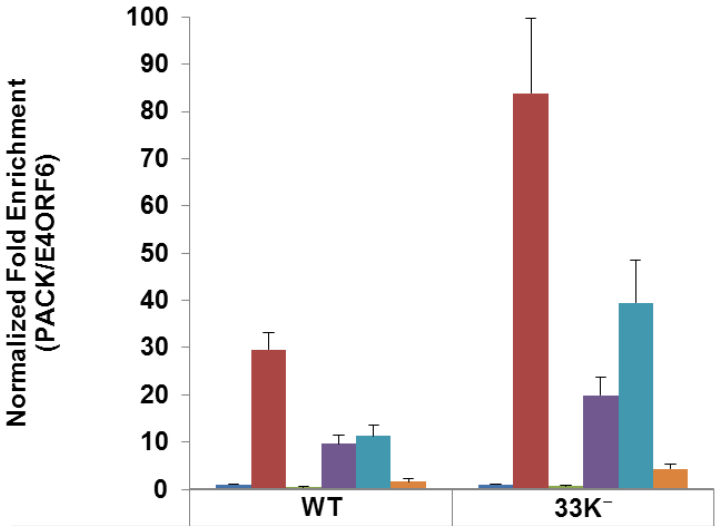


Figure 16. Virus particle production with L4-33K mutant virus. A. CsCl density equilibrium gradient profiles of virus particles produced from Ad5-WT-, Ad5-Ψ-loxP-, or 33K⁻-infected TetC4 (Cre recombinase-positive), 33K⁻-infected 293 cells, and 33K⁻-infected A549 cells. **B.** Electron microscopy picture of EC produced from 33K⁻-infected 293 cells. **C.** Electron microscopy picture of EC produced from 33K⁻-infected A549 cells. **D.** Electron microscopy picture of EC produced from Ad5-Ψ-loxP-, or 33K⁻-infected TetC4 cells, and MV from Ad5-WT-infected TetC4 cells. All of MV and EC were purified from the gradient in A. Arrows and arrowheads indicate MV and EC, respectively.

The L4-33K protein does not preferentially bind to the PS in vivo. Given the viral genome packaging defect observed in 33K^{-}-infected cells, we performed ChIP assays to determine if the L4-33K protein is part of the genome packaging machinery that interacts with the PS during infection. The results from Ad5-WT-infected A549 cells showed that L4-33K did not preferentially bind to the PS *in vivo* (Fig. 17, αL4-33K vs. pre-immune and WT vs. 33K^{-}). Next, we examined the binding of other viral packaging proteins to the PS in Ad5-WT or 33K^{-}-infected A549 cells. These results showed that mutation of L4-33K did not interfere with the binding of viral packaging proteins IVa2, L4-22K, L1-52/55K, or IIIa to the PS *in vivo* (Fig. 17, WT vs. 33K^{-}). Based on these results, we conclude that L4-33K is not likely to be directly involved in the viral genome packaging process.



	WT	33K⁻
pre-immune	1.0	1.0
αL4-22K	29.6	83.8
αL4-33K	0.4	0.7
αIVa2	9.7	20.0
αL1-52/55K	11.2	39.5
αIIIa	1.7	4.3

Figure 17. ChIP assays for packaging protein binding to the PS *in vivo*. ChIP assays for packaging protein binding to the PS *in vivo*. A549 were infected Ad5-WT or 33K^{-} and L4-22K-, L4-33K-, IVa2-, L1-52/55K-, and IIIa-specific antibodies were used for ChIP to quantify binding to the PS. The results are presented as normalized fold-enrichment by dividing the copy-number of PS pulled-down specifically by the antibody-antigen complex to the copy-number of E4-ORF6 fragment that was pulled down non-specifically by the antibody. For each virus infection, the fold-enrichment value of each antibody was normalized to that of the pre-immune serum negative control.

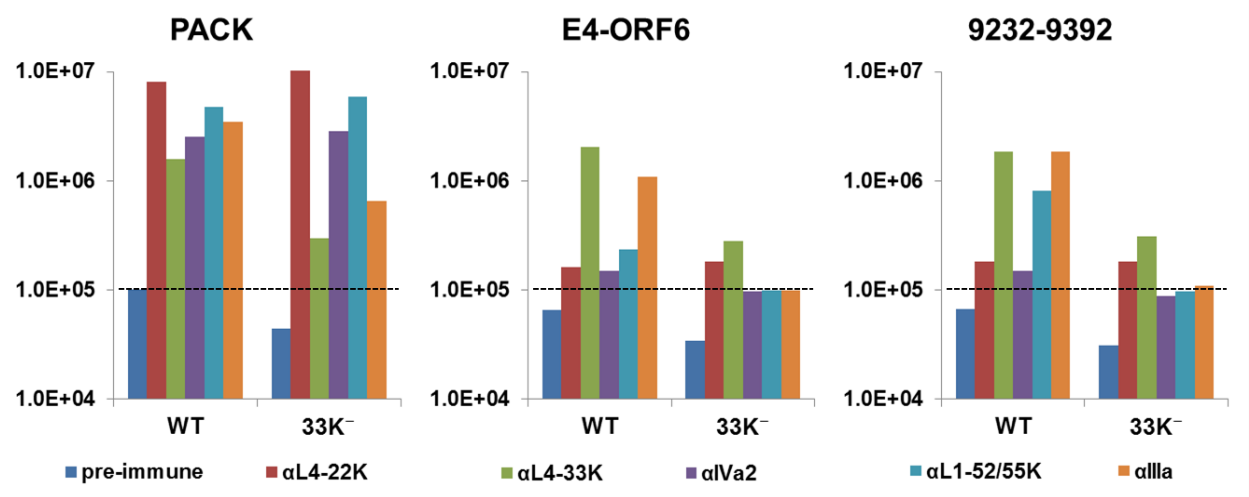


Figure 18. Absolute quantification of the PS (PACK), the E4-ORF6, and Ad5 9232-9392 DNA regions that were pulled down by each individual antibody from Ad5-WT- or 33K⁻-infected A549 cells in the ChIP assays performed in the Figure 17.

Although the normalized PACK/E4-ORF6 value of L4-33K bindings to the PS is lower than the negative control (pre-immune) (Fig. 17), we noticed that the absolute quantification of the DNA fragments pulled down by L4-33K antibody remained consistently high level among the PS, E4-ORF6, and Ad5 9232-9392 regions (Fig. 18). This result suggests that L4-33K might be associated with Ad genome across a broad region, but not preferentially with the PS.

IIIa protein level is critical for viral genome packaging. As IIIa is one of the primary targets regulated by L4-33K (Fig. 15) and it is involved in viral genome packaging (41), we examined whether the viral genome packaging defect with the 33K$^-$ mutant virus is due to the reduction of IIIa protein expression. A tetracycline-inducible cell line was generated that inducibly expresses the Ad5 IIIa protein to determine whether supplementation of additional IIIa would rescue the 33K$^-$ packaging defect. As shown in Figs. 19A and 19B, supplementation of IIIa protein rescued growth of the 33K$^-$ mutant virus by 1 log, although a severe defect in infectious virus production was still evident compared with Ad5-WT. We analyzed virus production in the IIIa cell line infected with 33K$^-$ and found that in addition to EC, heavy intermediate virus particles (HI) also were produced (Fig. 19C; density 1.35 g/cc). This phenotype demonstrated that supplementation of IIIa protein in 33K$^-$-infected cells did partially rescue the defect in viral genome packaging; however, the maturation of these virus particles was impaired since HI were produced. These

results demonstrate that IIIa protein levels and the regulation of IIIa gene expression by L4-33K are important for viral genome packaging.

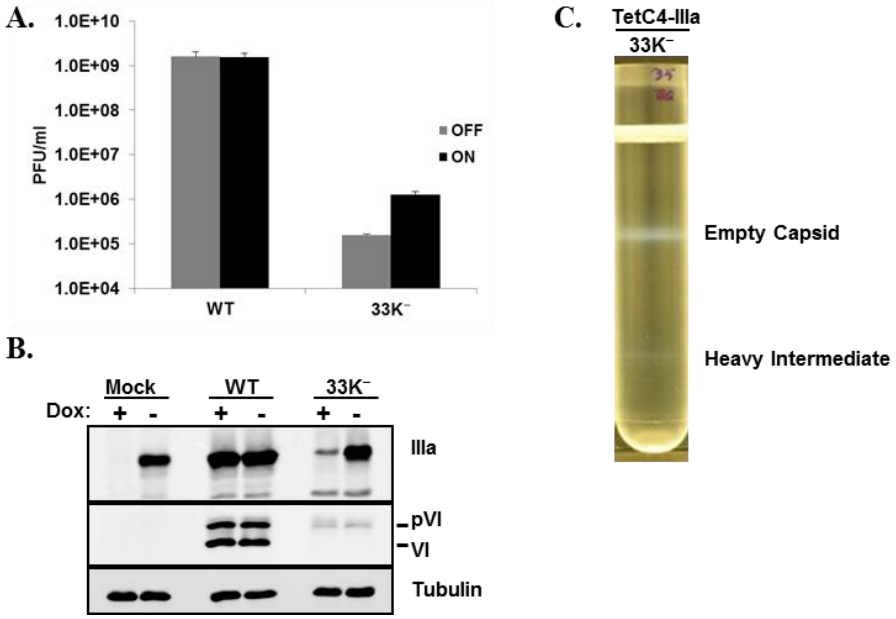


Figure 19. IIIa protein expression partly restores 33K⁻ viral genome packaging. A. Complementation of Ad5-WT and 33K⁻ viruses in TetC4-IIIa cells with (OFF) or without (ON) doxycycline (1 μg/ml) in the culture medium. Infected cell lysates were harvested at 48 hpi and subsequently titered on TetC4-33K cells by plaque assay. **B.** Western blot analysis of the infected cell lysates described in A. **C.** Virus particle production from 33K⁻-infected TetC4-IIIa cells without doxycycline in the culture medium.

Discussion

L4-22K functions. Our lab previously demonstrated that L4-22K is an essential viral protein and that it binds to the TTTG motif of the PS *in vitro* in the presence of IVa2, suggesting a role in viral genome packaging (51). Morris and Leppard have shown that the L4-22K protein is important for temporal control of viral gene expression by influencing intermediate and late gene expression (47). We generated and characterized two L4-22K mutant viruses, Δ22K and $22K^{-}$, to further explore L4-22K functions during Ad infection (Figs. 5A and 5B). Our results showed that L4-22K is indeed important for temporal control of viral gene expression, not only by activating late gene expression, but also by suppressing early gene expression. We also showed that the L4-22K protein binds to the PS *in vivo* and is essential to recruit two other packaging proteins, IVa2 and L1-52/55K, to this region. The elimination of L4-22K only gave rise to the production of EC, but no MV, which confirms that the L4-22K protein is required for Ad genome packaging. Finally, L4-22K contributes to Ad-induced cell death by influencing expression of ADP. Thus, the Ad L4-22K protein is multifunctional and an integral component of crucial aspects of infection.

The L4-22K mutant viruses behaved similarly as Ad5-WT through the onset of infection to viral DNA replication (Figs. 4D-4F). Alterations were observed with the L4-22K mutant viruses in both early and late gene expression as the late phase progressed consistent with the findings by Morris and Leppard (47). L4-22K protein has been suggested to influence late gene expression on the levels of gene transcription and/or RNA processing (4, 47). The overall delay and reduction of late gene products observed with both L4-22K mutant viruses may relate to its role in activating MLP transcription (Fig. 9A) (4). However, the extent of reduction of different late gene products was quite different which supports a role for L4-22K in the post-transcriptional processing of certain viral mRNAs (47). Interestingly, both L4-22K mutant viruses showed apparently increased E1A expression and mRNA levels late after infection (Figs. 7 and 8) and moderately increased DBP expression at this time (Fig. 7). This may explain the modestly higher genome replication level that was observed with both L4-22K mutant viruses at late times after infection (Fig. 5F). These results also suggest that L4-22K suppresses early gene expression during the late phase of infection. A role for an Ad late *trans*-acting component in the down-regulation of viral early gene expression was previously proposed (15).The increased E1A

protein level that we observed correlated well with E1A mRNA levels suggesting that L4-22K suppresses early gene expression. The Ad PS overlaps with the E1A enhancer region (23) suggesting that the binding of packaging proteins to this region might suppress E1A transcription during the late phase of infection. However, this model is open to question since the elimination of IVa2 expression with *pm*8002 (a IVa2 mutant virus), and consequently the binding of both IVa2 and L4-22K to the PS (Fig. 12A), did not enhance E1A expression at 24 hpi (87) or at later times of infection (Fig. 10) in A549 cells. But it is still plausible that independent of its binding to the PS, L4-22K indirectly suppresses E1A expression by affecting the activity of an E1A transcription factor(s) or that L4-22K influences E1A expression post-transcriptionally. The significance of suppressing E1A expression during the late phase of infection might be to redistribute the cellular transcriptional/translational machinery from early genes to late genes. Overall, our results and those published by others cited above indicate that the L4-22K protein plays an important role in regulating the temporal control of viral gene expression during the course of Ad infection.

The L4-22K mutant viruses only produced EC indicating that L4-22K is essential for Ad genome packaging (Fig. 11). But other aspects of mutant virus infection might contribute to this phenotype. First, the IVa2 and L4-33K proteins regulate Ad late gene expression at transcriptional and post-transcriptional levels (3, 14, 58, 72) and the L4-22K mutant viruses exhibited decreased expression of both of these proteins, particularly with L4-33K (Fig. 7). Supplementation of IVa2 or L4-33K using complementing cell lines, however, did not rescue the ability of the L4-22K mutants to produce MV (Fig. 11). Second, the strongly reduced levels of V, pVII/VII and pVIII (Fig. 7) might also underlie the L4-22K mutant phenotype with respect to MV production. However, protein VII is not required for Ad genome packaging (Ostapchuk and Hearing, unpublished data) and a mutant in protein VIII did not display a packaging defect (35), but this possibility is not fully resolved. Third, pVI promotes hexon import into the nucleus (77) and the reduction of pVI in L4-22K mutant virus infections (Fig. 7) may contribute to capsid formation. However, the yield of EC from Δ22K-infected N52.E6-Cre cells (Fig. 11) indicates that pVI is not limiting in the absence of the L4-22K protein. While we cannot exclude the possibility of combinatorial effects of reduction of different Ad late proteins, the binding of the L4-22K protein to the PS *in vitro* (51, 75, 83) and *in vivo* (Fig. 12) and the lack of MV production with both L4-22K mutant viruses (Fig. 11) lead us to conclude that the L4-22K

protein is required for Ad genome packaging. This conclusion is strengthened by the analyses of viruses carrying mutations in the PS where a precise correlation was observed between the binding of the L4-22K protein to PS *in vitro* and virus viability *in vivo* (54).

Our results demonstrate that the L4-22K and IVa2 proteins are dependent upon each other for binding to the PS *in vivo* and that both of these proteins are required to recruit the L1-52/55K protein to this region of the Ad genome (Fig. 12). These results are consistent with previous studies that showed that IVa2 is required for L4-22K to bind to the PS *in vitro* (51, 75, 83). However, *in vitro* binding studies also have shown that purified recombinant IVa2 can bind to the PS alone (13, 54, 84). A possible explanation for this discrepancy is that in virus-infected cells the IVa2 protein needs L4-22K to remove other factors that might anchor on or close to IVa2 binding site and impede IVa2 DNA binding. Perhaps more likely is the observation that the IVa2 and L4-22K proteins binds to the PS cooperatively (83); we anticipate that weak DNA-protein interactions can be identified using *in vitro* binding studies but that the infected cell nucleus *in vivo* provides a more rigorous environment where only more stable DNA-protein interactions are maintained. It has been reported that L1-52/55K binds to the PS in *pm*8002-infected cells in the absence of IVa2 (61). However, we could not detect such binding in our experimental setting (Fig. 12A). Different L1-52/55K antibodies and cells were used is these two studies. These differences also may reflect the use of semi-quantitative PCR (61) versus quantitative PCR (Fig. 12) in these analyses. We observed partially-restored binding of L4-22K, IVa2, and L1-52/55K with the PS in *pm*8002-infected TetC4-22K-ΔC cells (Fig. 12C). Western blot analysis confirmed the existence of 40 kD IVa2 isoform in *pm*8002-infected TetC4-22K-ΔC cells, but not in *pm*8002-infected A549 cells at 18 hpi (Fig. 12D). In this context, the 40 kD IVa2 isoform appears to have partially restored the binding of L4-22K, IVa2, and L1-52/55K to the PS *in vivo* (Fig. 12C).

The L4-22K protein contributes to expression of ADP (Fig. 13) which in turn contributes to the lysis of Ad-infected cells (70). Even though the ADP coding region is located in the E3 region, ADP is transcribed from the MLP and not expressed in a great abundance until the late phase of infection (71). The link between Ad genome packaging and augmented ADP expression to induce cell lysis, and the involvement of the L4-22K protein in both of these processes, makes perfect sense with respect to the temporal coordination of the Ad infection cycle. A number of recent studies also have shown that Ad infection induces autophagy which also contributes to the

killing of infected cells (29, 30, 63). It will be interesting to determine if the autophagy process is altered by mutation of L4-22K, and if ADP, L4-22K, or other viral genes influenced by L4-22K are responsible for induction of autophagy during Ad infection.

L4-33K functions. The Ad L4-33K protein has been linked with two disparate functions: one as a virus-encoded alternative RNA splicing factor that regulates Ad late mRNA splicing patterns and a second as a protein involved in virus assembly (16, 17, 33, 72). To further explore the function(s) of L4-33K in the context of Ad infection, we made and characterized an Ad5 L4-33K mutant virus using a newly-developed complementing cell line (Figs. 14A and 14B). Characterization of this mutant virus demonstrates that mutation of L4-33K is lethal to Ad5 viability. Our results demonstrate that mutation of L4-33K does not block virus assembly, but impairs viral genome packaging and virion maturation. Further, our results confirm *in vivo* the essential role of L4-33K as an Ad late mRNA splicing regulator in the context of virus infection.

As mutation of L4-33K is lethal to virus growth (Fig. 14C), we determined at which step(s) L4-33K may play an essential role(s). The L4-33K mutant virus had comparable infectivity and viral genome replication to Ad5-WT (Figs. 14D-14F), which demonstrates there are no apparent differences between Ad5-WT and the mutant virus from the onset of infection through viral genome replication. These results suggested that the growth defect arose from a late event during infection. Northern blot analysis of cytoplasmic viral mRNAs (Fig.15A) confirmed the previous report that L4-33K functions as a virus-encoded alternative RNA splicing factor (72). Consistent with this report, we found that IIIa, V, pVII, and hexon mRNA levels were regulated by L4-33K, although the magnitude of the effect on V and pVII mRNAs levels *in vivo* was modest compared to that observed *in vitro*. Multiple factors may influence cytoplasmic mRNA levels *in vivo* including nuclear transport and mRNA stability that may relate to these differences. In addition, we also found that pVI and fiber mRNAs were reduced in L4-33K mutant virus-infected cells whereas increased levels of L1-52/55K, L4-100K, and L4-22K mRNAs were observed in the absence of L4-33K (Figs. 15A and 15B). The previous report showed that IIIa mRNA accumulation does not occur at the expense of L1-52/55K mRNA accumulation using an *in vivo* splicing assay by co-transfecting an L4-33K expression vector with a vector to express an L1/L2 region splicing substrate (72). These subtle differences likely reflect the different types of assays utilized, but do not affect the clear interpretation that L4-33K

is a highly important for the appropriate processing of certain Ad late mRNAs. Our analysis of the full panel of Ad late gene transcripts in the context of mutant virus infection demonstrates that there is a clear switch of viral late gene mRNA accumulation mediated by L4-33K: the early set of late gene mRNAs (L1-52/55K, L4-100K, and L4-22K) first accumulate at the relatively early points of the late phase (Figs. 15A and 15B, 12 hpi) consistent with a prior report (34). Subsequently, Ad late mRNAs are produced corresponding to all five late gene families (L1–L5), some of whose expression depends completely (IIIa) or in part (V, pVII, hexon, pVI, and fiber) upon the expression of L4-33K. Because this switch in mRNA expression patterns did not occur efficiently in the absence of L4-33K, we observed increased mRNA levels of L1-52/55K, L4-100K, and L4-22K. It is possible that other viral proteins whose levels are altered in 33K^{-}-infected cells (IIIa, V, pVII, pVI, hexon, fiber, L1-52/55K, or L4-100K) might contribute to the switch of viral late gene mRNA accumulation, but this possibility is discounted because those gene products have no known function on gene transcription, RNA processing, or mRNA export. However, since the L4-22K protein activates Ad late gene expression at the transcriptional and post-transcriptional levels (4, 47, 80), we cannot rule out the possibility that increased L4-22K (Fig. 15B and 15C) may contribute to the increased mRNA levels of L1-52/55K and L4-100K. Interestingly, viral mRNAs that are increased in 33K^{-}-infected cells (L1-52/55K, L4-100K, and L4-22K) are located in the L1 and L4 regions, which corresponds to an earlier publication that L1 and L4 transcripts accumulate in the early time points of the late phase of Ad infection (34). These results lead us to conclude that L4-33K temporally controls the switch from the early set of viral late mRNAs to the late ones during progression of the late phase of Ad infection.

Western blot results largely, but not completely, recapitulated changes in Ad late mRNAs levels observed by northern blot analysis (Fig. 15). For example, the strong decrease in IIIa cytoplasmic mRNA levels observed in 33K^{-}-infected cells resulted in only a moderate decrease in IIIa protein levels. One explanation for this discrepancy is that the increased level of L4-100K observed in 33K^{-}-infected cells might compensate for the reduced viral mRNA level, since an important function of L4-100K is to enhance translation of viral late gene transcripts through its interaction with both the TPL in the mRNA and the translation initiation complex (81). However, in contrast to IIIa, decreases in pVI and fiber mRNA levels resulted in concomitant decreases in protein levels making this interpretation uncertain since all of the Ad late gene mRNAs contain

the TPL sequences. The exact basis for differences in viral late mRNA levels versus proteins levels in unknown at this time. Since pVI has been reported to augment hexon protein transport into nucleus (77), the apparent reduction in pVI levels in 33K^{-}-infected cells could be a cause of the previous reported assembly defect of L4-33K mutants (16, 17, 33). However, our results showed similar nuclear accumulation of hexon in Ad5-WT- and 33K^{-}-infected cells (Fig.15C) discounting this interpretation of the prior results.

Three previous reports concluded that L4-33K plays a role in virus assembly (16, 17, 33). Our results clarify the interpretations of results obtained with different L4-33K mutants. Only EC, and no MV, were produced in L4-33K mutant virus-infected A549, 293, and TetC4 cells (Fig. 16). This demonstrates there is no virus assembly defect, but rather a viral genome packaging defect in the absence of L4-33K. The accumulation of EC also confirms that most viral late gene products accumulate properly in the nucleus as shown by western blot (Fig. 15C, NE), except for decreased levels of IIIa, pVI, and fiber. The viral genome packaging defect led us to investigate how L4-33K might be involved in this process. ChIP analyses showed that the L4-33K protein itself does not preferentially binds to the PS *in vivo*, and its mutation does not impair the interaction of other viral packaging proteins (IVa2, L4-22K, L1-52/55K, and IIIa) with the PS (Fig. 17). In fact, we observed stronger binding of these packaging proteins with the PS in L4-33K mutant virus infected-cells (Fig. 17). The reason for this stronger binding may be due to the increased levels of L4-22K and L1-52/55K in 33K^{-}-infected cells (Fig. 15C). As L4-22K and IVa2 are dependent on each other and are essential to recruit L1-52/55K to bind to the PS *in vivo* (Fig. 12), the increased levels of L4-22K and L1-52/55K in 33K^{-}-infected cells may enhance the formation of the packaging complex, even though the total pool of IIIa is reduced. However, the discrepancy between the strong binding of the known viral packaging proteins with the PS and no effective genome packaging suggests either that we lack one or more other important components of the viral genome packaging machinery in 33K^{-}-infected cells or the ChIP approach does not reveal the complete nature of the packaging process.

To further track the cause of genome packaging defect with the L4-33K mutant virus, a cell line was made that inducibly expresses the Ad5 IIIa protein in order to supplement this protein in 33K^{-}-infected cells. Our results showed additional expression of the Ad5 IIIa protein in cells infected with the L4-33K mutant virus partly restored the packaging process since

infectious virus production was increased (Fig. 19A). However, the analysis of virion production only showed the appearance of HI particles (Fig. 19C). The HI production explains why the 33K$^-$ titer obtained following induction of IIIa expression in the cell line was not fully restored (Fig. 19A), since HI are not fully infectious. We also note that HI production was inefficient, even though the IIIa protein was provided at comparable levels to that observed in Ad5-WT-infected cells (Fig. 19B). This observation suggests that in addition to IIIa as a crucial component for viral genome packaging, other altered viral or cellular proteins in the context of 33K$^-$ infection might be accountable for the inefficient genome packaging process. The accumulation of HI suggests AVP activity was impaired since AVP is required for final virus particle maturation (43). However, the level of AVP expression in 33K$^-$-infected cells was slightly higher than that observed in Ad5-WT-infected cells (Fig. 15C). We believe that the strongly decreased levels of pVI observed in 33K$^-$-infected cells likely contributes to a decrease in AVP activity since a pVI C-terminal 11 aa peptide (pVIc) is a very important co-factor to stimulate AVP activity (43). We cannot exclude the possible role of pVI in viral genome packaging, although such a role has not been suggested previously.

In summary, our results demonstrate that the phenotype of an Ad5 L4-33K mutant virus is complex. The L4-33K protein regulates the accumulation of selective Ad late gene mRNAs and is involved in the proper transition of gene expression during the late phase of infection. The L4-33K protein also plays a role in adenovirus morphogenesis by regulating the levels of protein IIIa which is required for genome packaging and protein pVI which is required to activate the adenovirus protease during virion maturation.

Future Directions

Mechanism by which L4-22K activates late gene expression. It has been reported that L4-22K activates the expression of Ad late genes at the transcriptional level by binding to the DE of the MLP (4, 51), and at the post-transcriptional level by an unknown mechanism (47). We confirmed the transcriptional role of L4-22K in activating the MLP by luciferase reporter assay (Fig. 9A). Regarding the post-transcriptional role of L4-22K in activating the MLP, because L4-33K protein level is strongly decreased in the L4-22K mutants-infected cells (Fig. 7) and L4-33K is an alternative RNA splicing factor, we need to exclude the possibility that L4-33K might contribute to the phenotype of the post-transcriptional regulation which we observed with the L4-22K mutant viruses. This issue will be addressed by infecting TetC4-33K cells with the L4-22K mutant viruses and comparing the late gene expression profile between Ad5-WT and the L4-22K mutants. If the extent of reduction of late gene products is still quite different, it will further support the role of L4-22K in the post-transcriptional regulation of late gene expression. If the reduction is very similar across the panel of late gene products, it will demonstrate that L4-22K does not regulate late genes at the post-transcriptional level. However, we cannot exclude L4-33K being a target of the post-transcriptional regulation by L4-22K because it would be provided by the cell line. Second, after excluding the possible involvement of L4-33K, we can do northern blot analysis, or quantitative PCR, of nuclear pre-mRNAs and cytoplasmic mRNAs of viral late genes to directly confirm the post-transcriptional regulation by L4-22K. This experiment will also shed light on which steps of RNA processing that L4-22K might play: polyadenylation, splicing, mRNA export, and/or mRNA stability.

Mechanism by which L4-22K suppresses E1A expression. In addition to the phenotype we observed from the L4-22K mutants, we also noticed a moderate reduction of E1A transcription when L4-22K and IVa2 are both present (Fig. 9B). Because the Ad PS overlaps with the E1A transcriptional enhancer, we speculated that the binding of L4-22K to the PS is important for L4-22K to suppress E1A transcription. However, this model is open to question since the elimination of IVa2 expression with *pm*8002 (a IVa2 mutant virus), and consequently the binding of both IVa2 and L4-22K to the PS (Fig. 12A), did not enhance E1A expression at 24 hpi (87) or at later times of infection (Fig. 10) in A549 cells. This result also challenges the requirement of both L4-22K and IVa2 to suppress E1A transcription (Fig. 9B). But it is still plausible that independent of its binding to the PS, L4-22K indirectly suppresses E1A expression

by affecting the activity of an E1A transcription factor(s) or that L4-22K regulates E1A expression post-transcriptionally. It will be interesting to investigate E1A expression profile in the TetC4-22K-ΔC cells with or without doxycycline, in order to test the post-transcriptional role that L4-22K might have on E1A expression.

Interacting partners of L4-22K. To further decipher the mechanisms by which L4-22K executes its functions, it will be necessary to find out its interacting partners which would be involved in the gene regulation and/or viral genome packaging. We have tried a couple of different conditions for Co-IP in order to identify viral protein(s) that might interact with L4-22K. So far the results are negative. The difficulty of finding novel viral protein binding partners for L4-22K might come from two aspects: (i) the interaction between L4-22K and other viral proteins might be transient and/or weak; (ii) the expression level of L4-22K during viral infection is low. For the second point, it is well demonstrated in Fig. 6 where both L4-33K and L4-22K were analyzed by western blotting using a mouse monoclonal antibody against the common N-terminus of L4-33K and L4-22K. The L4-22K protein level is far less that L4-33K, which has made it difficult to work with. Alternatively, we can tag L4-22K, purify its partners by tandem affinity purification, and identify possible partners by mass spectrometry.

The possible role of pVI/VI in Ad morphogenesis. The functions of pVI/VI include: (i) virion capsid protein VI counteracts Daxx-mediated suppression of E1A transcription (68); (ii) VI is required for efficient endosome escape of Ad (76); (iii) VI is important for Ad to transport to the nucleus by the microtubule-dependent trafficking (78); and, (iv) pVIc is important for AVP activity (42). Because IIIa and pVI are the primary regulation targets of L4-33K (Fig. 15), and supplementation of IIIa to 33K$^-$ virus only partially rescued the viral genome packaging defect of 33K$^-$ virus and produced HI (Fig. 19), two questions arose regarding the possible role of pVI/VI in Ad morphogenesis: (i) Does pVI/VI participate in the viral genome packaging process? And, (ii) Does the reduction of pVI contribute to the HI production when additional IIIa was provided to 33K$^-$ virus? One could transfect pVI expression plasmid to the IIIa cell line and analyze virion production from 33K$^-$ virus. It would also be great if a cell line expressing both IIIa and pVI could be obtained and used for the second question. Another angle to address this question is to make a pVI-null mutant virus using a pVI-expressing cell line. This approach will avoid the function of pVI/VI at early phase of infection since the mutant virus particles contain VI from the complementing cell line and we could analyze the virion production to address its possible

role in viral genome packaging. If pVI is indeed important for viral genome packaging, we would see a packaging defect from the mutant virus. If pVI is not involved in the packaging process, we would anticipate HI production due to the role of pVIc as the co-factor for AVP.

References

1. **Akusjarvi, G.** 2008. Temporal regulation of adenovirus major late alternative RNA splicing. Front Biosci **13:**5006-5015.
2. **Akusjarvi, G., and H. Persson.** 1981. Controls of RNA splicing and termination in the major late adenovirus transcription unit. Nature **292:**420-426.
3. **Ali, H., G. LeRoy, G. Bridge, and S. J. Flint.** 2007. The adenovirus L4 33-kilodalton protein binds to intragenic sequences of the major late promoter required for late phase-specific stimulation of transcription. J Virol **81:**1327-1338.
4. **Backstrom, E., K. B. Kaufmann, X. Lan, and G. Akusjarvi.** 2010. Adenovirus L4-22K stimulates major late transcription by a mechanism requiring the intragenic late-specific transcription factor-binding site. Virus research **151:**220-228.
5. **Berk, A. J.** 2005. Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. Oncogene **24:**7673-7685.
6. **Bett, A. J., L. Prevec, and F. L. Graham.** 1993. Packaging capacity and stability of human adenovirus type 5 vectors. J Virol **67:**5911-5921.
7. **Cepko, C. L., and P. A. Sharp.** 1982. Assembly of adenovirus major capsid protein is mediated by a nonvirion protein. Cell **31:**407-415.
8. **Chartier, C., E. Degryse, M. Gantzer, A. Dieterle, A. Pavirani, and M. Mehtali.** 1996. Efficient generation of recombinant adenovirus vectors by homologous recombination in Escherichia coli. J Virol **70:**4805-4810.
9. **Christensen, J. B., S. A. Byrd, A. K. Walker, J. R. Strahler, P. C. Andrews, and M. J. Imperiale.** 2008. Presence of the adenovirus IVa2 protein at a single vertex of the mature virion. J Virol **82:**9086-9093.
10. **D'Halluin, J. C., M. Milleville, P. A. Boulanger, and G. R. Martin.** 1978. Temperature-sensitive mutant of adenovirus type 2 blocked in virion assembly: accumulation of light intermediate particles. J Virol **26:**344-356.
11. **de Jong, R. N., P. C. van der Vliet, and A. B. Brenkman.** 2003. Adenovirus DNA replication: protein priming, jumping back and the role of the DNA binding protein DBP. Current topics in microbiology and immunology **272:**187-211.
12. **Echavarria, M.** 2008. Adenoviruses in immunocompromised hosts. Clin Microbiol Rev **21:**704-715.
13. **Ewing, S. G., S. A. Byrd, J. B. Christensen, R. E. Tyler, and M. J. Imperiale.** 2007. Ternary complex formation on the adenovirus packaging sequence by the IVa2 and L4 22-kilodalton proteins. J Virol **81:**12450-12457.
14. **Farley, D. C., J. L. Brown, and K. N. Leppard.** 2004. Activation of the early-late switch in adenovirus type 5 major late transcription unit expression by L4 gene products. J Virol **78:**1782-1791.
15. **Fessler, S. P., and C. S. Young.** 1998. Control of adenovirus early gene expression during the late phase of infection. J Virol **72:**4049-4056.
16. **Fessler, S. P., and C. S. Young.** 1999. The role of the L4 33K gene in adenovirus infection. Virology **263:**507-516.
17. **Finnen, R. L., J. F. Biddle, and J. Flint.** 2001. Truncation of the human adenovirus type 5 L4 33-kDa protein: evidence for an essential role of the carboxy-terminus in the viral infectious cycle. Virology **289:**388-399.

18. **Guo, P., and T. J. Lee.** 2007. Viral nanomotors for packaging of dsDNA and dsRNA. Molecular microbiology **64:**886-903.
19. **Gustin, K. E., and M. J. Imperiale.** 1998. Encapsidation of viral DNA requires the adenovirus L1 52/55-kilodalton protein. J Virol **72:**7860-7870.
20. **Gustin, K. E., P. Lutz, and M. J. Imperiale.** 1996. Interaction of the adenovirus L1 52/55-kilodalton protein with the IVa2 gene product during infection. J Virol **70:**6463-6467.
21. **Hasson, T. B., D. A. Ornelles, and T. Shenk.** 1992. Adenovirus L1 52- and 55-kilodalton proteins are present within assembling virions and colocalize with nuclear structures distinct from replication centers. J Virol **66:**6133-6142.
22. **Hasson, T. B., P. D. Soloway, D. A. Ornelles, W. Doerfler, and T. Shenk.** 1989. Adenovirus L1 52- and 55-kilodalton proteins are required for assembly of virions. J Virol **63:**3612-3621.
23. **Hearing, P., and T. Shenk.** 1983. The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. Cell **33:**695-703.
24. **Hilleman, M. R., and J. H. Werner.** 1954. Recovery of new agent from patients with acute respiratory illness. Proc Soc Exp Biol Med **85:**183-188.
25. **Hoke, C. H., Jr., A. Hawksworth, and C. E. Snyder, Jr.** 2012. Initial assessment of impact of adenovirus type 4 and type 7 vaccine on febrile respiratory illness and virus transmission in military basic trainees, March 2012. Msmr **19:**2-4.
26. **Hong, S. S., E. Szolajska, G. Schoehn, L. Franqueville, S. Myhre, L. Lindholm, R. W. Ruigrok, P. Boulanger, and J. Chroboczek.** 2005. The 100K-chaperone protein from adenovirus serotype 2 (Subgroup C) assists in trimerization and nuclear localization of hexons from subgroups C and B adenoviruses. Journal of molecular biology **352:**125-138.
27. **Huang, W., and S. J. Flint.** 1998. The tripartite leader sequence of subgroup C adenovirus major late mRNAs can increase the efficiency of mRNA export. J Virol **72:**225-235.
28. **Imperiale, M. J., and S. Kochanek.** 2004. Adenovirus vectors: biology, design, and production. Curr Top Microbiol Immunol **273:**335-357.
29. **Ito, H., H. Aoki, F. Kuhnel, Y. Kondo, S. Kubicka, T. Wirth, E. Iwado, A. Iwamaru, K. Fujiwara, K. R. Hess, F. F. Lang, R. Sawaya, and S. Kondo.** 2006. Autophagic cell death of malignant glioma cells induced by a conditionally replicating adenovirus. Journal of the National Cancer Institute **98:**625-636.
30. **Jiang, H., E. J. White, C. I. Rios-Vicil, J. Xu, C. Gomez-Manzano, and J. Fueyo.** 2011. Human adenovirus type 5 induces cell lysis through autophagy and autophagy-triggered caspase activity. J Virol **85:**4720-4729.
31. **Kochanek, S.** 1999. High-capacity adenoviral vectors for gene transfer and somatic gene therapy. Hum Gene Ther **10:**2451-2459.
32. **Kochanek, S., P. R. Clemens, K. Mitani, H. H. Chen, S. Chan, and C. T. Caskey.** 1996. A new adenoviral vector: Replacement of all viral coding sequences with 28 kb of DNA independently expressing both full-length dystrophin and beta-galactosidase. Proc Natl Acad Sci U S A **93:**5731-5736.
33. **Kulshreshtha, V., L. A. Babiuk, and S. K. Tikoo.** 2004. Role of bovine adenovirus-3 33K protein in viral replication. Virology **323:**59-69.

34. **Larsson, S., C. Svensson, and G. Akusjarvi.** 1992. Control of adenovirus major late gene expression at multiple levels. Journal of molecular biology **225:**287-298.
35. **Liu, G. Q., L. E. Babiss, F. C. Volkert, C. S. Young, and H. S. Ginsberg.** 1985. A thermolabile mutant of adenovirus 5 resulting from a substitution mutation in the protein VIII gene. J Virol **53:**920-925.
36. **Liu, H., L. Jin, S. B. Koh, I. Atanasov, S. Schein, L. Wu, and Z. H. Zhou.** 2010. Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. Science **329:**1038-1043.
37. **Liu, Y., C. Wang, S. Mueller, A. V. Paul, E. Wimmer, and P. Jiang.** 2010. Direct interaction between two viral proteins, the nonstructural protein 2C and the capsid protein VP3, is required for enterovirus morphogenesis. PLoS pathogens **6:**e1001066.
38. **Logan, J., and T. Shenk.** 1984. Adenovirus tripartite leader sequence enhances translation of mRNAs late after infection. Proceedings of the National Academy of Sciences of the United States of America **81:**3655-3659.
39. **Lutz, P., and C. Kedinger.** 1996. Properties of the adenovirus IVa2 gene product, an effector of late-phase-dependent activation of the major late promoter. J Virol **70:**1396-1405.
40. **Lutz, P., F. Puvion-Dutilleul, Y. Lutz, and C. Kedinger.** 1996. Nucleoplasmic and nucleolar distribution of the adenovirus IVa2 gene product. J Virol **70:**3449-3460.
41. **Ma, H. C., and P. Hearing.** 2011. Adenovirus structural protein IIIa is involved in the serotype specificity of viral DNA packaging. J Virol **85:**7849-7855.
42. **Mangel, W. F., W. J. McGrath, D. L. Toledo, and C. W. Anderson.** 1993. Viral DNA and a viral peptide can act as cofactors of adenovirus virion proteinase activity. Nature **361:**274-275.
43. **Mangel, W. F., D. L. Toledo, J. Ding, R. M. Sweet, and W. J. McGrath.** 1997. Temporal and spatial control of the adenovirus proteinase by both a peptide and the viral DNA. Trends Biochem Sci **22:**393-398.
44. **Meng, X., M. H. Brodsky, and S. A. Wolfe.** 2005. A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. Nat Biotechnol **23:**988-994.
45. **Mohr, I., and Y. Gluzman.** 1996. A herpesvirus genetic element which affects translation in the absence of the viral GADD34 function. The EMBO journal **15:**4759-4766.
46. **Morin, N., and P. Boulanger.** 1986. Hexon trimerization occurring in an assembly-defective, 100K temperature-sensitive mutant of adenovirus 2. Virology **152:**11-31.
47. **Morris, S. J., and K. N. Leppard.** 2009. Adenovirus serotype 5 L4-22K and L4-33K proteins have distinct functions in regulating late gene expression. J Virol **83:**3049-3058.
48. **Morris, S. J., G. E. Scott, and K. N. Leppard.** 2010. Adenovirus late-phase infection is controlled by a novel L4 promoter. J Virol **84:**7096-7104.
49. **Noyes, M. B., X. Meng, A. Wakabayashi, S. Sinha, M. H. Brodsky, and S. A. Wolfe.** 2008. A systematic characterization of factors that regulate Drosophila segmentation via a bacterial one-hybrid system. Nucleic Acids Res **36:**2547-2560.
50. **Ostapchuk, P., M. Almond, and P. Hearing.** 2011. Characterization of Empty adenovirus particles assembled in the absence of a functional adenovirus IVa2 protein. J Virol **85:**5524-5531.

51. **Ostapchuk, P., M. E. Anderson, S. Chandrasekhar, and P. Hearing.** 2006. The L4 22-kilodalton protein plays a role in packaging of the adenovirus genome. J Virol **80:**6973-6981.
52. **Ostapchuk, P., and P. Hearing.** 2008. Adenovirus IVa2 protein binds ATP. J Virol **82:**10290-10294.
53. **Ostapchuk, P., and P. Hearing.** 2005. Control of adenovirus packaging. J Cell Biochem **96:**25-35.
54. **Ostapchuk, P., J. Yang, E. Auffarth, and P. Hearing.** 2005. Functional interaction of the adenovirus IVa2 protein with adenovirus type 5 packaging sequences. J Virol **79:**2831-2838.
55. **Palmer, D. J., and P. Ng.** 2005. Helper-dependent adenoviral vectors for gene therapy. Hum Gene Ther **16:**1-16.
56. **Palmer, D. J., and P. Ng.** 2008. Methods for the production of helper-dependent adenoviral vectors. Methods Mol Biol **433:**33-53.
57. **Pardo-Mateos, A., and C. S. Young.** 2004. A 40 kDa isoform of the type 5 adenovirus IVa2 protein is sufficient for virus viability. Virology **324:**151-164.
58. **Pardo-Mateos, A., and C. S. Young.** 2004. Adenovirus IVa2 protein plays an important role in transcription from the major late promoter in vivo. Virology **327:**50-59.
59. **Parks, R. J., L. Chen, M. Anton, U. Sankar, M. A. Rudnicki, and F. L. Graham.** 1996. A helper-dependent adenovirus vector system: removal of helper virus by Cre-mediated excision of the viral packaging signal. Proc Natl Acad Sci U S A **93:**13565-13570.
60. **Parks, R. J., and F. L. Graham.** 1997. A helper-dependent system for adenovirus vector production helps define a lower limit for efficient DNA packaging. J Virol **71:**3293-3298.
61. **Perez-Romero, P., R. E. Tyler, J. R. Abend, M. Dus, and M. J. Imperiale.** 2005. Analysis of the interaction of the adenovirus L1 52/55-kilodalton and IVa2 proteins with the packaging sequence in vivo and in vitro. J Virol **79:**2366-2374.
62. **Reddy, V. S., S. K. Natchiar, P. L. Stewart, and G. R. Nemerow.** 2010. Crystal structure of human adenovirus at 3.5 A resolution. Science **329:**1071-1075.
63. **Rodriguez-Rocha, H., J. G. Gomez-Gutierrez, A. Garcia-Garcia, X. M. Rao, L. Chen, K. M. McMasters, and H. S. Zhou.** 2011. Adenoviruses induce autophagy to promote virus replication and oncolysis. Virology **416:**9-15.
64. **Rowe, W. P., R. J. Huebner, L. K. Gilmore, R. H. Parrott, and T. G. Ward.** 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. Proc Soc Exp Biol Med **84:**570-573.
65. **Schiedner, G., S. Hertel, M. Johnston, V. Biermann, V. Dries, and S. Kochanek.** 2002. Variables affecting in vivo performance of high-capacity adenovirus vectors. J Virol **76:**1600-1609.
66. **Schiedner, G., S. Hertel, and S. Kochanek.** 2000. Efficient transformation of primary human amniocytes by E1 functions of Ad5: generation of new cell lines for adenoviral vector production. Human gene therapy **11:**2105-2116.
67. **Schmid, S. I., and P. Hearing.** 1997. Bipartite structure and functional independence of adenovirus type 5 packaging elements. J Virol **71:**3375-3384.
68. **Schreiner, S., R. Martinez, P. Groitl, F. Rayne, R. Vaillant, P. Wimmer, G. Bossis, T. Sternsdorf, L. Marcinowski, Z. Ruzsics, T. Dobner, and H. Wodrich.** 2012.

Transcriptional activation of the adenoviral genome is mediated by capsid protein VI. PLoS pathogens **8:**e1002549.

69. **Smith, J. G., C. M. Wiethoff, P. L. Stewart, and G. R. Nemerow.** 2010. Adenovirus. Curr Top Microbiol Immunol **343:**195-224.
70. **Tollefson, A. E., A. Scaria, T. W. Hermiston, J. S. Ryerse, L. J. Wold, and W. S. Wold.** 1996. The adenovirus death protein (E3-11.6K) is required at very late stages of infection for efficient cell lysis and release of adenovirus from infected cells. J Virol **70:**2296-2306.
71. **Tollefson, A. E., A. Scaria, S. K. Saha, and W. S. Wold.** 1992. The 11,600-MW protein encoded by region E3 of adenovirus is expressed early but is greatly amplified at late stages of infection. J Virol **66:**3633-3642.
72. **Tormanen, H., E. Backstrom, A. Carlsson, and G. Akusjarvi.** 2006. L4-33K, an adenovirus-encoded alternative RNA splicing factor. J Biol Chem **281:**36510-36517.
73. **Tormanen Persson, H., A. K. Aksaas, A. K. Kvissel, T. Punga, A. Engstrom, B. S. Skalhegg, and G. Akusjarvi.** 2012. Two cellular protein kinases, DNA-PK and PKA, phosphorylate the adenoviral L4-33K protein and have opposite effects on L1 alternative RNA splicing. PloS one **7:**e31871.
74. **Tribouley, C., P. Lutz, A. Staub, and C. Kedinger.** 1994. The product of the adenovirus intermediate gene IVa2 is a transcriptional activator of the major late promoter. J Virol **68:**4450-4457.
75. **Tyler, R. E., S. G. Ewing, and M. J. Imperiale.** 2007. Formation of a multiple protein complex on the adenovirus packaging sequence by the IVa2 protein. J Virol **81:**3447-3454.
76. **Wiethoff, C. M., H. Wodrich, L. Gerace, and G. R. Nemerow.** 2005. Adenovirus protein VI mediates membrane disruption following capsid disassembly. J Virol **79:**1992-2000.
77. **Wodrich, H., T. Guan, G. Cingolani, D. Von Seggern, G. Nemerow, and L. Gerace.** 2003. Switch from capsid protein import to adenovirus assembly by cleavage of nuclear transport signals. The EMBO journal **22:**6245-6255.
78. **Wodrich, H., D. Henaff, B. Jammart, C. Segura-Morales, S. Seelmeir, O. Coux, Z. Ruzsics, C. M. Wiethoff, and E. J. Kremer.** 2010. A capsid-encoded PPxY-motif facilitates adenovirus entry. PLoS pathogens **6:**e1000808.
79. **Wohl, B. P., and P. Hearing.** 2008. Role for the L1-52/55K protein in the serotype specificity of adenovirus DNA packaging. J Virol **82:**5089-5092.
80. **Wu, K., D. Orozco, and P. Hearing.** 2012. The Adenovirus L4-22K Protein Is Multifunctional and Is an Integral Component of Crucial Aspects of Infection. J Virol **86:**10474-10483.
81. **Xi, Q., R. Cuesta, and R. J. Schneider.** 2004. Tethering of eIF4G to adenoviral mRNAs by viral 100k protein drives ribosome shunting. Genes & development **18:**1997-2009.
82. **Yang, J., and P. Hearing.** 2007. Chromatin immunoprecipitation to study the binding of proteins to the adenovirus genome in vivo. Methods Mol Med **131:**113-121.
83. **Yang, T. C., and N. K. Maluf.** 2012. Cooperative heteroassembly of the adenoviral L4-22K and IVa2 proteins onto the viral packaging sequence DNA. Biochemistry **51:**1357-1368.

84. **Yang, T. C., Q. Yang, and N. K. Maluf.** 2009. Interaction of the adenoviral IVa2 protein with a truncated viral DNA packaging sequence. Biophysical chemistry **140:**78-90.
85. **Young, C. S.** 2003. The structure and function of the adenovirus major late promoter. Current topics in microbiology and immunology **272:**213-249.
86. **Zhang, W., and M. J. Imperiale.** 2000. Interaction of the adenovirus IVa2 protein with viral packaging sequences. J Virol **74:**2687-2693.
87. **Zhang, W., and M. J. Imperiale.** 2003. Requirement of the adenovirus IVa2 protein for virus assembly. J Virol **77:**3586-3594.

Appendix

Specific Aim

Modify Helper-Dependent Adenoviral Vector System with an Altered-Specificity IVa2 Protein and an Altered-Specificity Genome Packaging Sequence.

Introduction

Besides basic research on the Ad life cycle, exploring Ad as a gene therapy vector has been under development for decades. The great advantages of Ad to be used as a gene therapy vector are: (i) wide tropism to efficiently transduce a variety of cell types; (ii) infection independent of cell cycle; (iii) high levels of gene expression; and, (iv) low chance of integrating into cellular chromosome. From the original so-called first generation Ad vectors to the second generation, one or more early regions were deleted and replaced with foreign genes instead. However, the majority of the Ad genome, especially most of the structural genes, was still present in the gene therapy vector, which would directly induce cytotoxic effects and an adaptive immune response following transduction. Subsequently, cells containing the vector would be cleared by inflammatory or immune responses, resulting in transient expression of foreign genes (31, 55).

Most recently, the helper-dependent Ad vectors (HDAds, also called gutted, gutless, high-capacity) have been developed without any viral protein coding region, but only bearing ~500 bp *cis*-acting Ad sequences for vector DNA replication (ITRs) and packaging (PS) (32, 59). In addition to the general advantages that early versions of Ad vectors have, HDAds allow simultaneous gene transfer of several expression cassettes for multiple-gene defects, large or tissue-specific promoters, and some genes in their natural genomic context. Furthermore, due to the absence of viral gene expression, HDAds are able to direct high-level, long-term transgene expression without chronic toxicity. Because the size limit of the Ad genome that can be packaged into virus requires between ~27.7 kb (60) and ~38 kb (6), due to the different length of foreign genes, stuffer DNA is used in HDAds which usually consists of non-coding mammalian DNA and minimal repeat sequences (65).

In general, to grow HDAds, the genome is released from a plasmid using an appropriate restriction enzyme and then transfected into recombinase Cre-expressing 293 cells. Subsequently, an E1-deleted helper virus (HV), with loxP sites flanking the PS, is used to infect those cells in order to complement the growth of HDAds in *trans* (Fig. A1). At the same time, the activity of Cre recombinase, would excise the PS from the HV genome, referred to as floxing, so that the HV genome would not be packagable and only HDAds could be produced (Fig. A1) (56). However, HV contamination still happens because of the efficiency of Cre recombinase activity, even though high levels of Cre-expressing cells have been used (56). Currently, the best quality of HDAds that can be obtained following CsCl purification contains 0.02-0.01% HV contamination (56).

Based on our knowledge on Ad genome packaging process that the viral protein IVa2 binds specifically to the CG motif of the PS (Fig. 4C), the ultimate goal of this project is to find a pair of altered-specificity IVa2 (AS-IVa2) and altered-specificity PS (AS-PS). In the ideal scenario, AS-IVa2 will only recognize AS-PS, but not the wild-type PS (WT-PS) any more. We will independently introduce the AS-IVa2 and AS-PS to replace WT-IVa2 and WT-PS in the HV

and gene therapy vector, respectively (Fig. A2). On the other hand, the HV still contains WT-PS. As a result, during the production of HDAd, even though HV contamination still happens because of the efficiency of Cre recombinase activity, the HV genome that contains the WT-IVa2 will not be recognized by the AS-IVa2, so that it will not be packaged during the HDAd production (Fig. A2). Through this approach, we will be able to further decrease the HV contamination problem during HDAd production.

In order to find the pair of AS-PS and AS-IVa2 for eliminating HV contamination, three steps are required for the success of finding the perfect pair (Fig. A3). The first step is to identify the AS-PS. The best AS-PS will bind to WT-IVa2 poorly and therefore causes Ad growth defect. After getting the AS-PS, the second step involves a library screening to find IVa2 mutants that would restore its binding to the AS-PS and consequently rescue the growth defects caused by the AS-PS. With some IVa2 mutants in hand, the third step is to confirm the specificity of these IVa2 mutants to find the AS-IVa2 that will specifically bind to the AS-PS but not the WT-PS any more. The ultimate criteria is that the AS-IVa2 will rescue the growth defect of the AS-PS virus, but cause a growth defect for the WT-PS virus.

For the first step, point mutants of the CG motif in A1 and A2 repeats (A1/2) were analyzed *in vitro* and *in vivo* in order to identify the AS-PS for subsequent screening. For the second step, we employed three different approaches to find mutant IVa2 that restores its binding to AS-PS, including the bacterial one-hybrid (B1H) system, the *in vivo* selection system, and the serial blind passaging approach.

The B1H system was reported by Wolfe and his colleagues to perform high-throughput identification of DNA binding sequences for specific transcription factors (TFs) (44, 49). In this system, also called omega B1H system, TFs are cloned into an expression vector and fused with the omega subunit of bacterial RNA polymerase on the N-terminus, while randomized DNA sequences are cloned upstream of reporter genes (yeast *URA3* and *HIS3*) in a reporter plasmid (Fig. A4A). After co-transformation into an *E.coli* selective strain (omega subunit and homologs of the yeast *URA3* and *HIS3* knocked out), reporter genes would be expressed if the TF can bind the DNA binding sequence on the reporter plasmid based on the recruitment of other RNA polymerase components by the omega subunit fused with the TF. 3-amino-trizole (3-AT) and 5-fluoro-orotic acid (5-FOA) are used for positive- and counter-selection, respectively. 3-AT is a competitive inhibitor of HIS3, providing selection for an active promoter; while 5-FOA would be converted to a toxic compound by the uracil biosynthesis pathway that consequently selects against an active promoter. In addition, the omega-Zif12 B1H system was also developed by Wolfe's laboratory, which contains fingers 1 and 2 (F1F2) of the Zif268 protein and their DNA binding sequences (f1f2) in expression and reporter vectors, respectively (Fig. A4B). This modified system is aimed to improve weak interaction between protein and DNA with the help of F1F2 binding to f1f2 on the reporter plasmid. Of note, different lengths of spacer between the DNA binding site and -35 box of the promoter was reported to be crucial to activate reporter gene expression, which was explained to properly expose the DNA binding site to transcription factors and the core *E. coli* promoter (49).

The *in vivo* selection is the most powerful tool to obtain different kinds of mutant viruses. However, due to the large size of the Ad infectious clone, it is very difficult to clone a random mutated IVa2 library in the context of the whole Ad genome. Instead, we want to split the Ad genome into two pieces: the left end with AS-PS on a transfer plasmid and the right end digested from ΔIVa2 viral DNA (Fig. A6). A random mutated IVa2 library will be cloned under CMV promoter in the transfer vector pE1A-dl194-811-AS-PS, which contains AS-PS and has most of

E1A deleted. Then the left end released from the transfer vector and the right end from the ΔIVa2 viral DNA will be co-transfected into an E1-expressing cell line. Due to the overlapping region, homologous recombination will happen in the transfected cells and produce Ad5 viral DNA containing the AS-PS and mutant IVa2 protein (Fig. A6). If the mutant IVa2 protein recognizes the AS-PS, the virus will grow in the transfected cells and could be subsequently purified using plaque purification method. Titration and further characterization will then be applied to the mutant virus in order to confirm the AS-IVa2.

The serial blind passaging approach is a straightforward but powerful way to obtain the suppressor mutants which would rescue the defect of a previous mutation for both RNA and DNA viruses (37, 45). From the first step of this project, we will identify a defective virus which bares mutations in the CG motif. By serial blind passaging of this defective virus, we hope to find a second-site mutant which will rescue the growth defect of the parental virus. We hope the second-site mutation locates in the IVa2 region, although it does not necessarily have to be. Since green fluorescent protein (GFP) is built in the genome of the parental virus, it will be very convenient to monitor the virus growth throughout the serial blind passaging process.

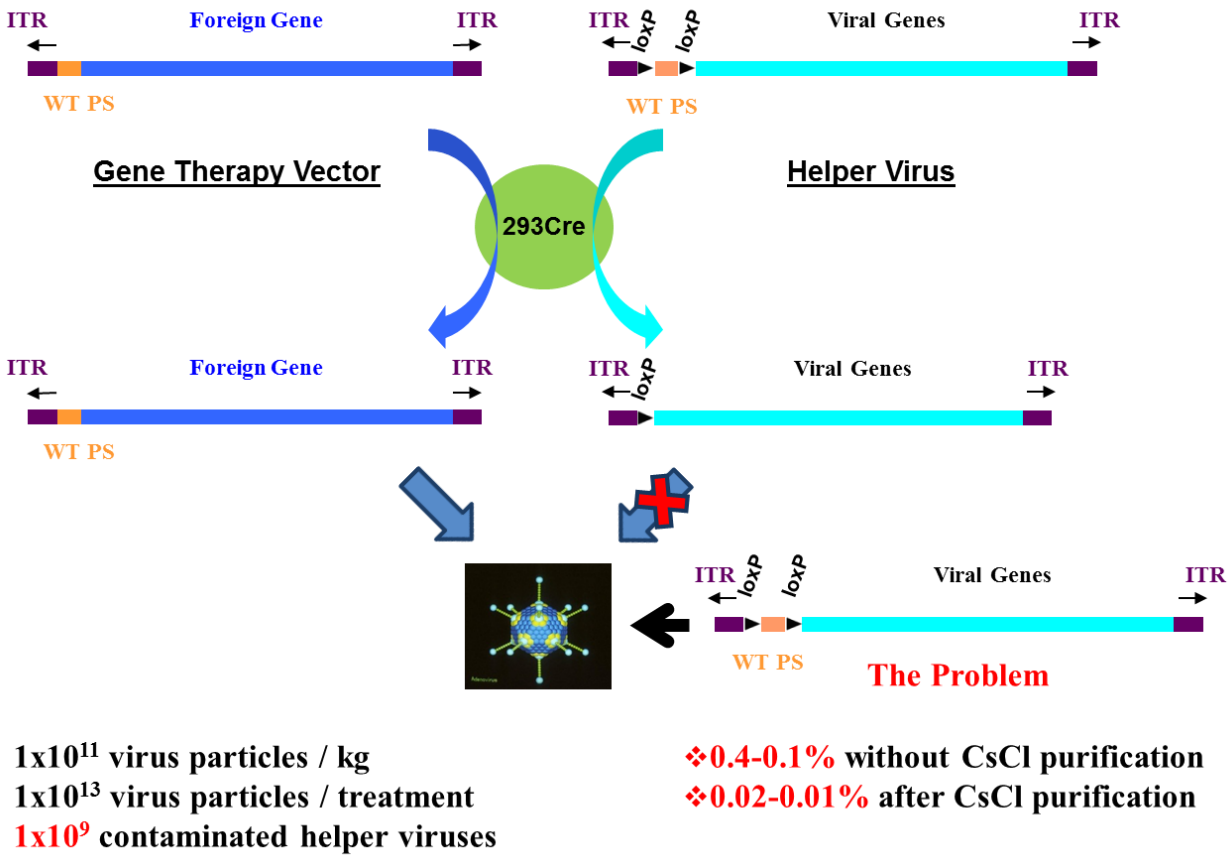


Figure A1. Helper-dependent adenovirus vector system. ITR: inverted terminal repeat. PS: packaging sequence. Modified from Palmer and Ng. Hum Gene Ther. 2005

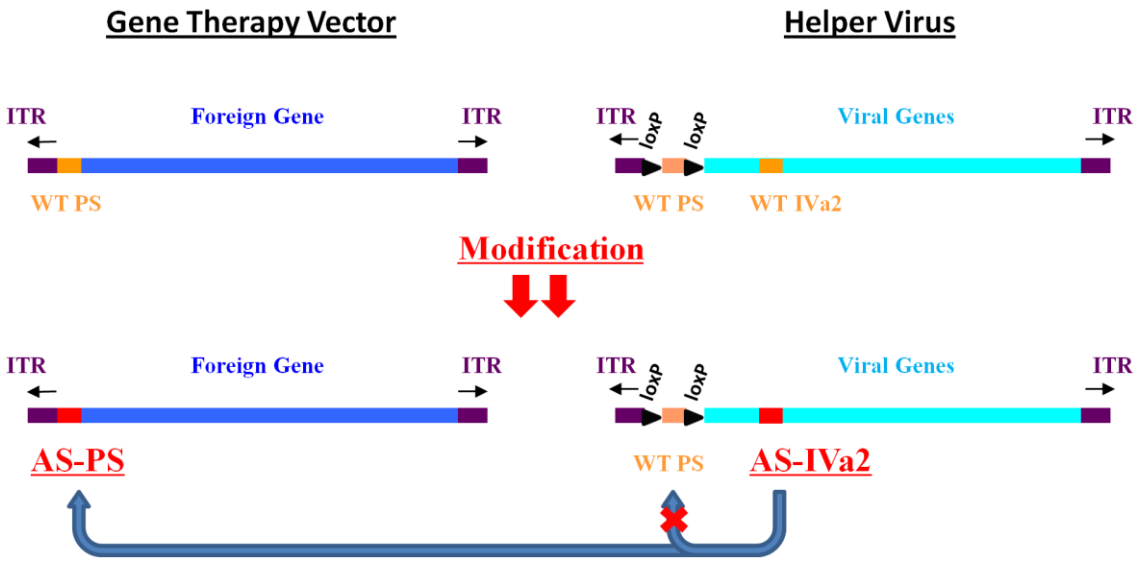


Figure A2. Schematic diagram of the altered-specificity modification to the HDAd system. The genomes of the gene therapy vector and helper virus are shown in the upper panel with wild-type packaging sequence (WT PS) and wild-type IVa2 protein (WT IVa2) in orange color. The altered-specificity IVa2 (AS-IVa2) and altered-specificity packaging sequence (AS-PS) are introduced into the HDAd system in the lower panel in red color. Because AS-IVa2 does not bind to the WT PS, helper virus genome will not be able to be packaged, even it's not cleaved by Cre recombinase. ITR: inverted terminal repeat.

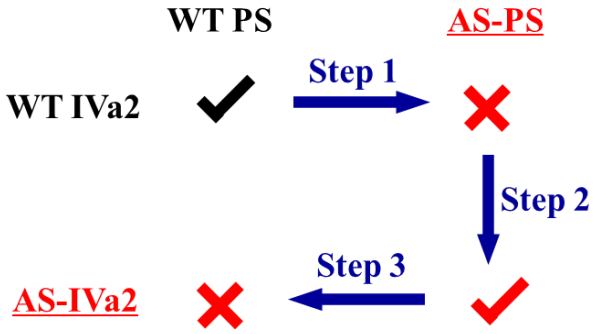


Figure A3. Work flows to find the pair of AS-PS and AS-IVa2.

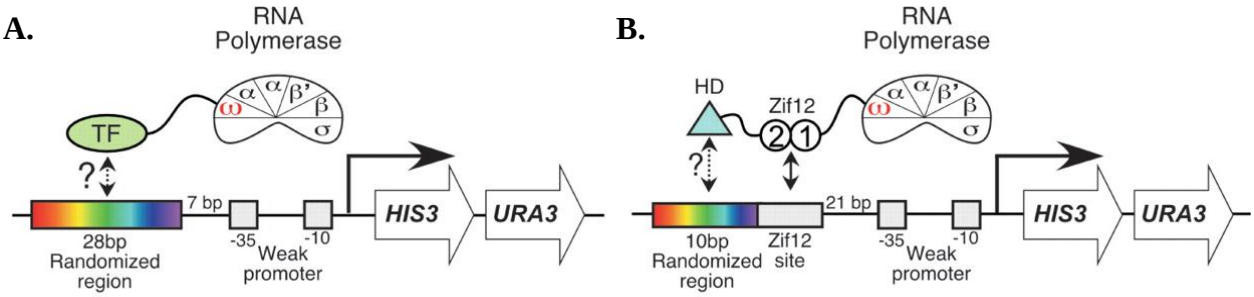


Figure A4. Overview of the B1H system. A. The omega-based B1H system. The DNA binding site upstream of the promoter allows the TF to bind and subsequently to bring other bacterial RNA polymerase components through the omega subunit fused with the TF. The consequent reporter gene expression will be selected on selective plates. **B.** The omega-Zif12-based B1H system. Fingers 1 and 2 (F1F2) of the Zif268 protein are inserted between the TF and omega subunit. The interaction of F1F2 and its DNA binding site (f1f2) on the reporter plasmid helps the weak interaction of a TF with its DNA binding site. This figure is modified from ref. 49.

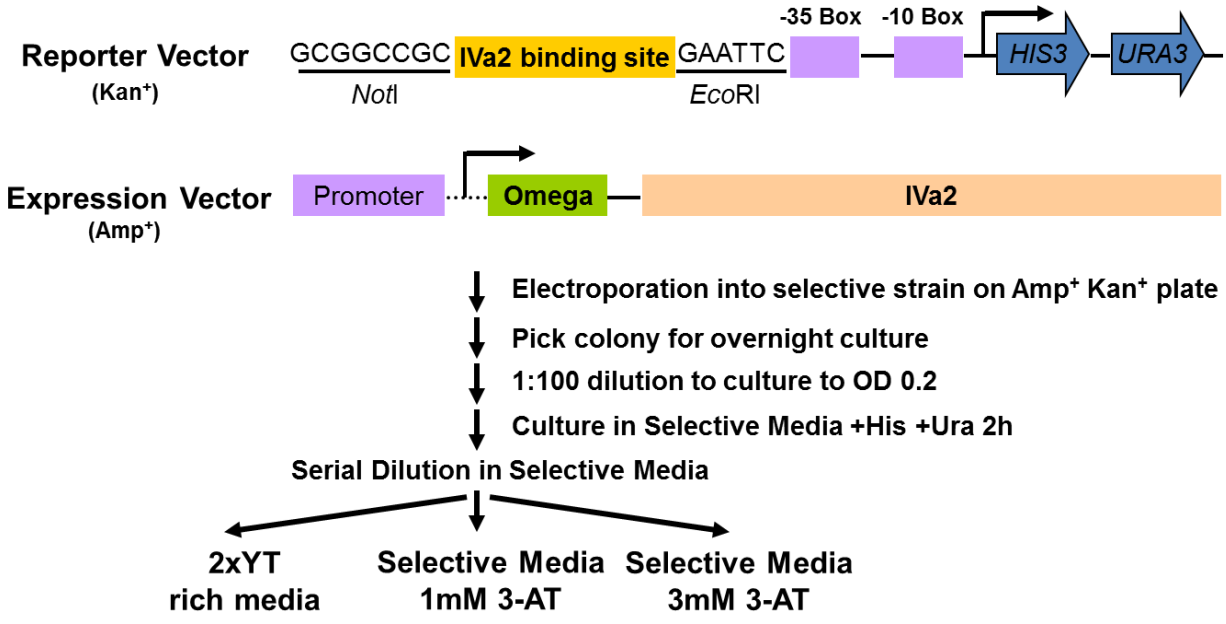


Figure A5. B1H selection protocol.

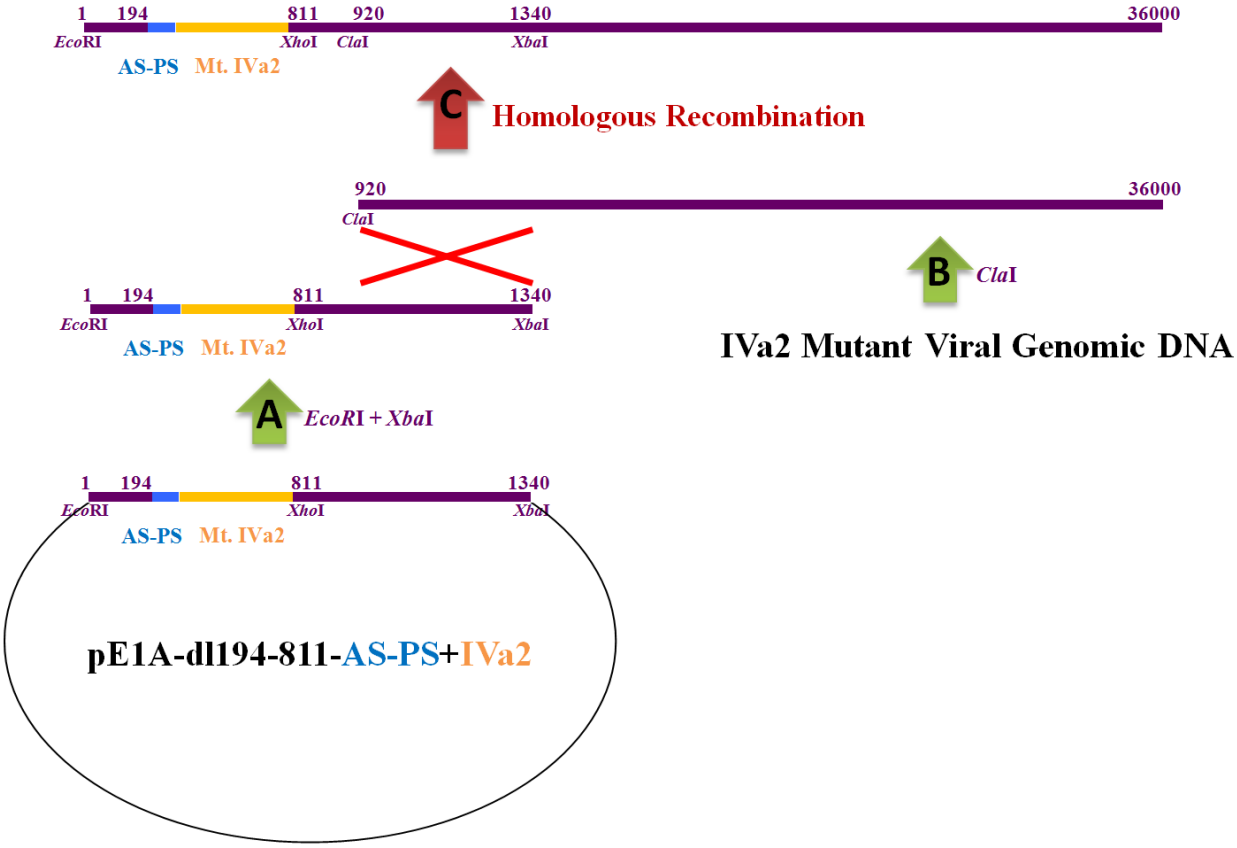


Figure A6. The *in vivo* selection system. A. The left end of the Ad5 genome is released from the transfer vector, containing AS-PS and mutant IVa2 cloned from the random PCR mutated IVa2 library. **B.** IVa2 mutant (ΔIVa2) viral genomic DNA is purified from the virions and digested with *Cla*I to release the right end of its genome. **C.** Due to ~400 bp overlapping region, homologous recombination between the left and the right ends will generate an intact viral genome with AS-PS and mutant IVa2, which will produce progeny viruses if the mutant IVa2 recognizes the AS-PS. Mt.: mutant.

Results

Step 1: Find the AS-PS.

In order to evaluate the effect of mutations in the CG motif on IVa2 binding *in vitro*, WT and point mutations of the CG motif in oligonucleotides containing A1/2 (termed C-A, C-G, C-T, G-A, G-C, and G-T, respectively) were synthesized and analyzed for the binding ability with WT IVa2 by EMSA (Figs. A7A and A7B). 2 ng of purified baculovirus-expressed WT IVa2 protein was incubated with ^{32}P-labeled WT and six point mutated A1/2 oligonucleotides in the presence of three different concentrations of cold non-specific competitor poly(dI:dC). The EMSA result (Fig. A7B) was relatively quantified (Fig. A7C), and showed that all six mutant PS had defects on IVa2 binding, while C-T had the least defect and G-T had the most defect.

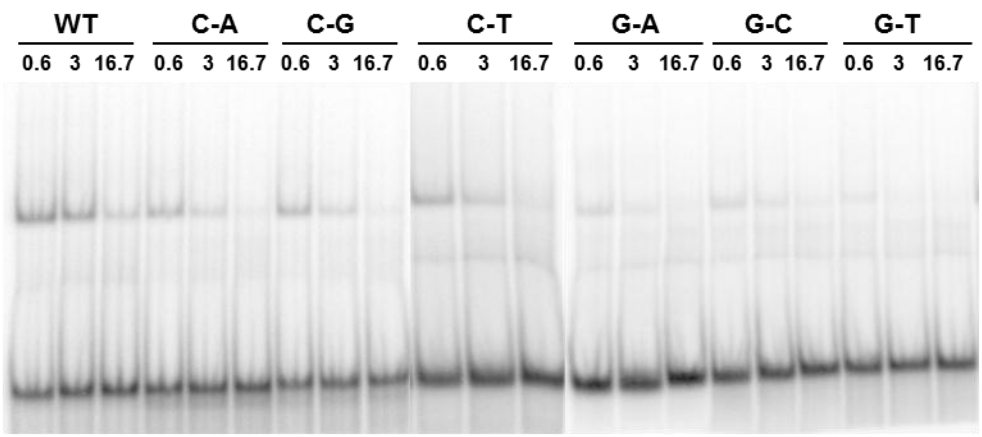
In order to test the effect of point mutations of the CG motif on virus growth *in vivo*, I introduced the same six mutations of the CG motif in the context of A1/2 dimer into the Ad genome and titered different mutant viruses to pick the most defective mutation as the AS-PS for the following experiments. Transfection of different infectious clones into the E1-expressing N52.E6-Cre cells yielded slight cytopathic effect (CPE) at ~14 days post transfection. The cells were harvested and lysed by 3 cycles of freeze-thaw process to release virus. In this way, I obtained my virus stocks of the different mutants for the second passage (P2). Viruses in P2 lysates were then titered in N52.E6-Cre cells by measuring GFP signal in infected cells using flow cytometry. The results showed that G-C had about 4-log decrease of infectivity, while other mutants presented comparable infectivity as the virus containing WT A1/2 dimer (Fig. A7D). Based on this result, we picked G-C as the AS-PS for the following screening.

Figure A7. Look for the AS-PS by mutagenesis of the CG motif *in vitro* and *in vivo*. A. Schematic diagram of the mutations (red) made in the CG motif of A1 and A2 (A1/2). **B.** EMSA using purified IVa2 protein from baculovirus-infected insect cells and the mutants of A1/2. Different concentrations of poly(dI-dC) were used to prevent non-specific binding of proteins to the probe. **C.** Relative quantification of the binding between IVa2 and the mutants of A1/2 from the EMSA in B. **D.** Growth of viruses that have WT or mutant A1/2 dimers. WT or mutant A1/2 dimers was introduced to an Ad5 infectious clone (pGS66) where Ad5 PS has been replaced; in addition, E1A and E1B were also substituted by GFP under the control of the CMV promoter. *Swa*I-linearized viral genomes were independently transfected into N52.E6-Cre cells (expressing Ad5 E1A and E1B), and the lysates of the second passage (P2) were titered in N52.E6-Cre cells by flow cytomertry. The purified IVa2 protein is a gift from Dr. Philomena Ostapchuk.

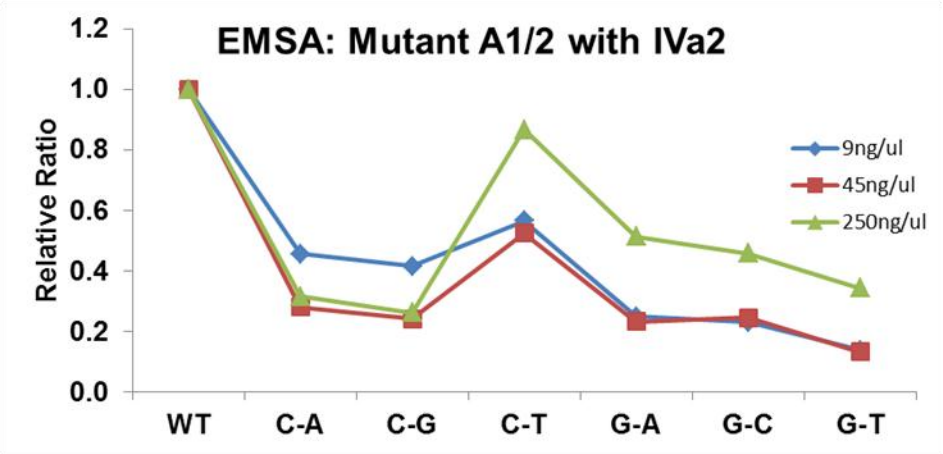
A.

```
WT A1/2: AGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCG
    C-A: AGTAAATTTGGGCGTAACAGAGTAAGATTTGGCCATTTTAGCG
    C-G: AGTAAATTTGGGCGTAACGGAGTAAGATTTGGCCATTTTGGCG
    C-T: AGTAAATTTGGGCGTAACTGAGTAAGATTTGGCCATTTTTGCG
    G-A: AGTAAATTTGGGCGTAACCAAGTAAGATTTGGCCATTTTCACG
    G-C: AGTAAATTTGGGCGTAACCCAGTAAGATTTGGCCATTTTCCCG
    G-T: AGTAAATTTGGGCGTAACCTAGTAAGATTTGGCCATTTTCTCG
```

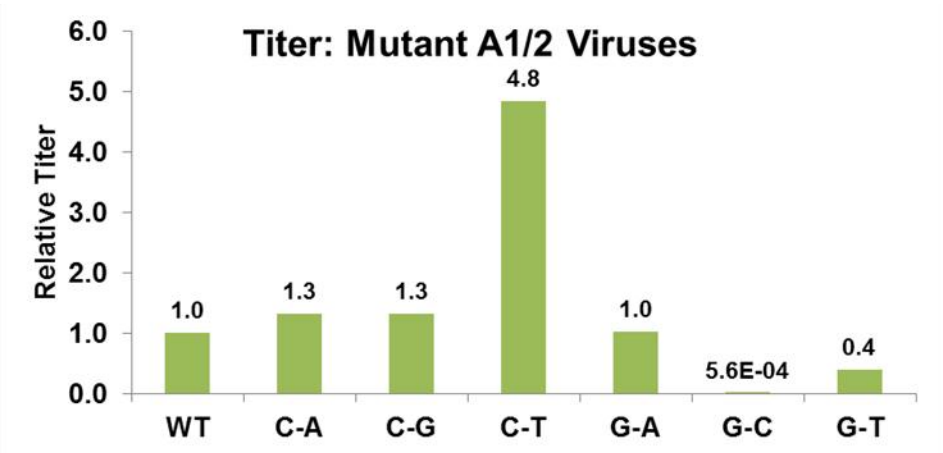
B.



C.



D.



Step 2: Find the mutant IVa2 which rescues the defect of the AS-PS virus.

Approach 1: B1H system

The different ways of optimizing the B1H system is summarized in Fig. A8.

A. System Control: To make sure the B1H system works in my hands, positive and negative system controls were analyzed according to the protocol from Dr. Scot Wolfe's lab (Fig. A9). Zif268, a zinc-finger DNA binding protein fused with omega subunit (Omega-Zif268), was used as a positive control with its binding site in the reporter plasmid (Zif268 Binding Site) and they worked perfectly well up to 10mM 3-AT (data not shown).

B. Experimental Control: With this system working, the interaction of WT IVa2 and WT-PS (the experimental positive control) needs to be determined in the B1H system before screening for AS-IVa2 using AS-PS.

   1) A1+A2 repeats (A1/2) series: A1/2 with different lengths of random spacers, were cloned into reporter vectors, respectively. These A1/2 reporter plasmids were transformed into selective strain individually or together with the Omega-IVa2 expression plasmid. Unfortunately, it turned out that the A1/2 series had a greater level of self-activity and it did not respond to Omega-IVa2 for activation of the B1H reporter genes at all (Fig. A10). In order to eliminate self-activity of the reporter plasmid, a shorter IVa2 binding sequence is desired.

   2) Shortened A1 series: Because IVa2 binds to the CG motif and the two nucleotides next to CG motif are conserved among the most important A1, A2, A5 and A6 repeats, I put the core motif 'CGAG' of the IVa2 binding site as the center, together with different lengths of flanking region, into the reporter vector (Fig. A11A). This series of shortened oligonucleotides (35, 44, 53, 80 and 85) showed that 44 and 80 had nearly no self-activity, while 35, 53, and 85 had high self-activity. When co-transformed with the Omega-IVa2 expression vector, 44 and 80 presented minor activation response while others survived better due to self-activity (Fig. A11C). In addition, this series of oligonucleotides were also tested for their binding capacity with WT IVa2 protein *in vitro*. The EMSA result showed the trend of IVa2 binding ability: $85 \approx 80 > 53 >> 44 >> 35$, while there was nearly no shift in 35 and 44 (Fig. A11B). Based on both self-activity and the *in vitro* binding capacity to IVa2, 80 was picked for the next step to improve its interaction with IVa2 in the B1H system in order to get better reporter gene activation.

   3) 80 Series: Different lengths of spacers, adapted from the negative control 'empty' reporter vector, were added between the 80 sequence and -35 box of the core promoter (Fig. A5). Self-activity and the activation response of the 80 series with Omega-IVa2 were measured (Fig. A12). The results revealed that this series of 80 sequences had a little self-activation, and obtained 1-2 log more growth rate on 1 mM 3-AT plate when Omega-IVa2 was added. However, this trend was gone under 3 mM 3-AT selection pressure. This is a little better than the original 80 sequence, but more activation is absolutely desired.

4) 80 dimer series: The 80 dimer sequence with different lengths of spacers were synthesized and analyzed in the system (data not shown). Disappointingly, the results showed a similar pattern of activation to the 80 series without any further improvement.

5) Omega-IVa2 vs. IVa2-Omega: IVa2-Omega fusion protein, with the same linker as Omega-IVa2 has, was constructed in the expression vector and tested with the 80 series (data not shown) or 80 dimer series (Fig. A13) of reporter vectors. However, no better activation was obtained either.

6) Flipped 80 Series: Although we know that IVa2 binds to the CG motif of A repeats, we do not know whether it binds to the plus or minus strand. I cloned the flipped 80 sequence with different lengths of spacers into the reporter vector and subsequently tested the self-activity as well as the activation with Omage-IVa2 or IVa2-Omega (Fig. A14). However, this series didn't present better activation response.

7) Leucine zipper motif as the linker between Omega and IVa2: Considering the weak interaction between IVa2 and its DNA binding sequence, we constructed Omega-Zipper-IVa2 and IVa2-Zipper-Omega with the leucine zipper motif as the linker in the middle, with the hope that homodimer of IVa2 might improve the interaction and activation. The two constructs were tested with the 80 dimer series, but no better activation was observed either (Fig. A15).

8) Omega-Zif12 B1H system: With a limited activation response from WT IVa2 and its binding site in the original B1H system, I turned to the omega-Zif12 B1H system, which introduces finger 1 and 2 (F1F2) of Zif268 protein and its DNA binding sequence (f1f2) into the expression and reporter vectors, respectively. This system is aimed to improve weak interaction between protein and DNA with the help of F1F2 binding to its binding site (f1f2) on the reporter plasmid (Fig. A4B). Omega-F1F2-IVa2 protein and its corresponding reporter plasmid with or without spacers were constructed and analyzed in the same *E.coli* selective strain as used above. As shown in Fig. A16, the positive control Omega-F1F2-Oct1 survived very well up to 10 mM 3-AT concentration (data not shown) due to interaction of both Oct1 and F1F2 with their corresponding binding sites. When Omega-F1F2-IVa2 was added with the same Oct1-f1f2 binding sequence as used for the positive control above, the effect of F1F2 and its binding site was measured as the F1F2 background, which gave 5-log bacteria growth in 1 mM 3-AT. However, when Omega-F1F2-IVa2 was added with its own 80 binding site with or without the 5nt spacer, the results showed the same activation response as the F1F2 background; so, no activation from IVa2 and its binding site was obtained (Fig. A16).

9) Flexible Glycine linker between Finger 1+2 and IVa2: Since there is only one Threonine between Fingers 1+2 and IVa2 in the construct of Omega-F1F2-IVa2, a 15-aa flexible linker (GGGGS repeats) was cloned in between and then tested in the Omega-Zif12 system. Disappointingly again, there was no better activation for my experimental positive control (WT IVa2 and its binding site) (Fig. A17).

10) Specifity of the B1H reporter gene expression: Since we have known from the *in vitro* and *in vivo* analyses that the G-C mutation is the most defective one for IVa2 binding (Fig. A7), G-C or G-T mutations were introduced to the 80+6 sequence, and they were tested in the original B1H system in order to test whether the activation of reporter gene expression is specific to the interaction between IVa2 and its binding site (Fig. A18). The data showed that only in the case of IVa2-Omega with the 80+6 G-C, there was a slight decrease of reporter gene activation. However, the issue of the small window of activation between the positive and negative controls makes it impossible to do any convincing library screen. These results closed the door of this approach.

Approach 2: *In vivo* selection system

In order to access if the system is feasible, we started with WT IVa2 and WT A1/2 dimer or G-C A1/2 dimer as positive and negative controls, respectively (Fig. A6). For the right-end Ad genome fragment, we purified ΔIVa2 virus particles produced from TetC4-IVa2 cells, extracted the viral genome from the purified particles, digested it with *Cla*I to release the left-end genome containing the PS, and purified the right end fragment (~30 kb) from sucrose gradient. For the transfer plasmid with the left-end genome (pE1A-dl194-811), we cloned WT IVa2 into the E1A position, together with WT-PS (positive control) or AS-PS (negative control). The purified left end that was released from the transfer plasmid by digestion of *Eco*RI and *Xba*I, was co-transfected with the right-end ΔIVa2 genome into 293 cells. 2 days post-transfection, the transfected cells were overlaid for a plaque assay. We got 20 plaques on the positive control plate (WT IVa2 + WT-PS), and none on the negative control plate (WT IVa2 + AS-PS) (Fig. A19). However, first, the window between positive and negative controls is not great. Second, the 20 plaques from the positive control could not be amplified by passaging the viruses in 293 cells, which later turned to be due to the insufficient IVa2 expression level expressed from the CMV promoter than the natural Ad5-WT infection (data not shown), even though we could not explain why these plaques grew in the first place. These results closed the door of this approach.

Approach 3: Serial blind passage of the G-C virus

Inspired from the suppressor mutant works done in RNA and DNA viruses (37, 45), I did serial blind passaging of the G-C (AS-PS) virus. At P2, there were barely any GFP-positive cells after infection by the G-C virus, compared with the WT virus (Fig. A7). However, at the seventh passage (P7), the infectivity of the G-C virus was very comparable with the WT virus (Fig. A20A). Viral DNA was extracted from the infection lysate, amplified by PCR, and sequenced for the following regions: the PS, IVa2, L4-22K, L1-52/55K, and IIIa. The sequencing results showed that both the WT and G-C viruses still contains the respective PS, but in a trimer format, instead of its original dimer format (Fig. A20B). Strikingly, instead of finding a mutation in the IVa2 region, we found L4-22K had a mutation at its aa position 107: lysine to glutamic acid (K107E). To test whether this mutation is sufficient to rescue the defect of the G-C virus, we built this mutation into the Ad5 infectious clone which contains the WT A1/2 dimer (termed WT K107E) or G-C A1/2 dimer (G-C K107E) as the PS, and tested their growth. The result showed

that the G-C K107E virus grew to a comparable level with the WT virus, increasing the titer of 4-log from its parental G-C virus, demonstrating that the L4-22K-K107E is sufficient to rescue the growth defect of the G-C virus (Fig. A20C). However, the L4-22K-K107E does not suppress the growth of the WT virus, which means the L4-22K-K107E does not present the altered-specificity property.

The outcome of this approach is very interesting: (i) the G-C virus, which contains four nucleotide changes (G to C) in the context of A1/2 dimer, did not simply revert these four nucleotides back to the WT version; and, (ii) the suppressor mutation locates in the L4-22K region, instead of the IVa2 region, even though we mutated the IVa2 binding site (CG motif). Along with the second point, we need to investigate the mechanism by which the L4-22K-K107E rescues the G-C mutation. Comparison of the full panel of viral gene expression will be needed to explore the difference between the WT and the G-C K107E viruses. In addition, EMSA should be performed using the nuclear extract of the WT- or the G-C K107E-infected cells with the WT or G-C A1/2, in order to evaluate the interaction of IVa2, L4-22K and the PS. We think it is very important to learn the mechanism and it will give us new information about the Ad genome packaging process, because the outcome was not we designed or expected.

Furthermore, even though the L4-22K-K107E does not present the altered-specificity property, it will be very interesting to test whether mutating the lysine107 to other aa will make it have the property while maintaining the recognition to the G-C A1/2.

A.

```
A1/2:       CTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGC
A1/2+2:     CTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGC GA
A1/2+4:     CTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGC GATC
A1/2+6:     CTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGC GATCGA
A1/2+8:     CTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGC GATCGATC
A1/2+10:    CTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGC GATCGATCGA
```

B.

```
A1/2:       CTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCGGC
35:                            AACCGAGTAAG
44:                          TAACCGAGTA
53:                        GTAACCGAGTAA
80:                      GGCGTAACCGAG
```

C.

```
80:     GGCGTAACCGAG
80+2:   GGCGTAACCGAG GT
80+4:   GGCGTAACCGAG GTGC
80+6:   GGCGTAACCGAG GTGCAC
80+8:   GGCGTAACCGAG GTGCACTC
80+10:  GGCGTAACCGAG GTGCACTCCG
```

D.

```
80di:     GGCGTAACCGAG GGCGTAACCGAG
80di+2:   GGCGTAACCGAG GGCGTAACCGAG GT
80di+4:   GGCGTAACCGAG GGCGTAACCGAG GTGC
80di+6:   GGCGTAACCGAG GGCGTAACCGAG GTGCAC
80di+8:   GGCGTAACCGAG GGCGTAACCGAG GTGCACTC
80di+10:  GGCGTAACCGAG GGCGTAACCGAG GTGCACTCCG
```

E.

```
f80:     CTCGGTTACGCC
f80+2:   CTCGGTTACGCC GT
f80+4:   CTCGGTTACGCC GTGC
f80+6:   CTCGGTTACGCC GTGCAC
f80+8:   CTCGGTTACGCC GTGCACTC
f80+10:  CTCGGTTACGCC GTGCACTCCG
```

F.

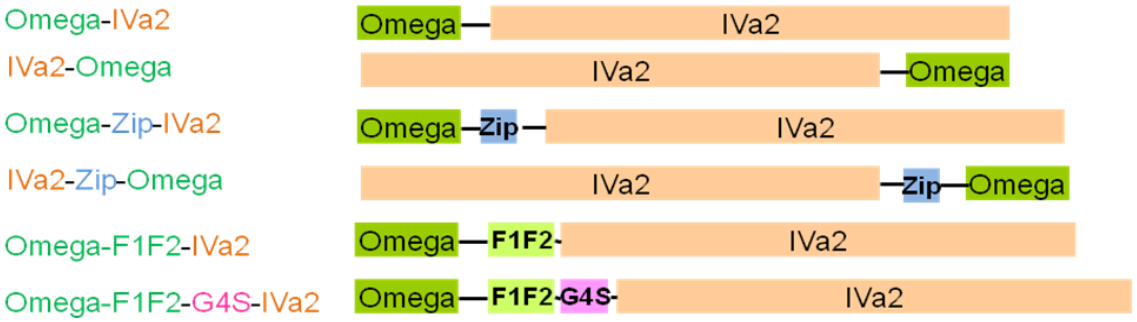


Figure A8. Summary of approaches taken for optimizing the B1H system. Schematic of different A1/2 sequences and Omega-IVa2 expression constructs. **A.** Sequences of the A1/2 series of oligonucleotides with different lengths of random spacers; **B.** Shortened A1 series; **C.** 80 series; **D.** 80 dimer series; **E.** flipped 80 series. **F.** Schematic of different IVa2 expression constructs. Zip: leucine zipper motif; G4S: 15 aa of GGGGS repetitive linker.

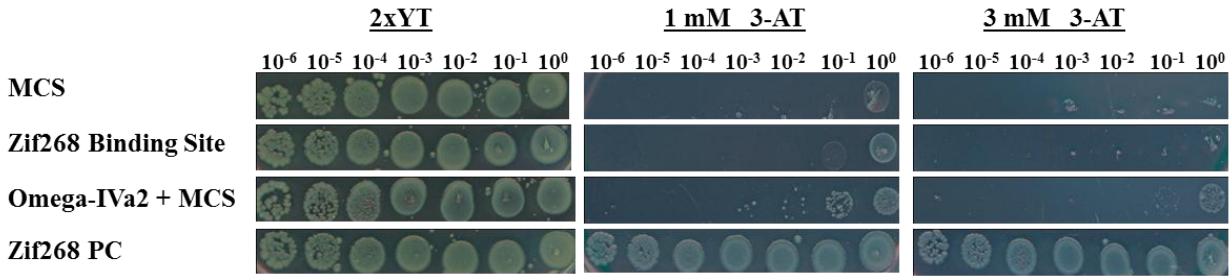


Figure A9. B1H system controls. Both MCS (the empty reporter plasmid) and Zif268 binding site (the reporter plasmid that contains the Zif268 binding site) are negative controls of the B1H system. Zif268 PC (Omega-Zif268 from the expression vector, and Zif268 binding site in the reporter plasmid) is the system positive control. Omega-IVa2+MCS is to test the background of Omega-IVa2, if any.

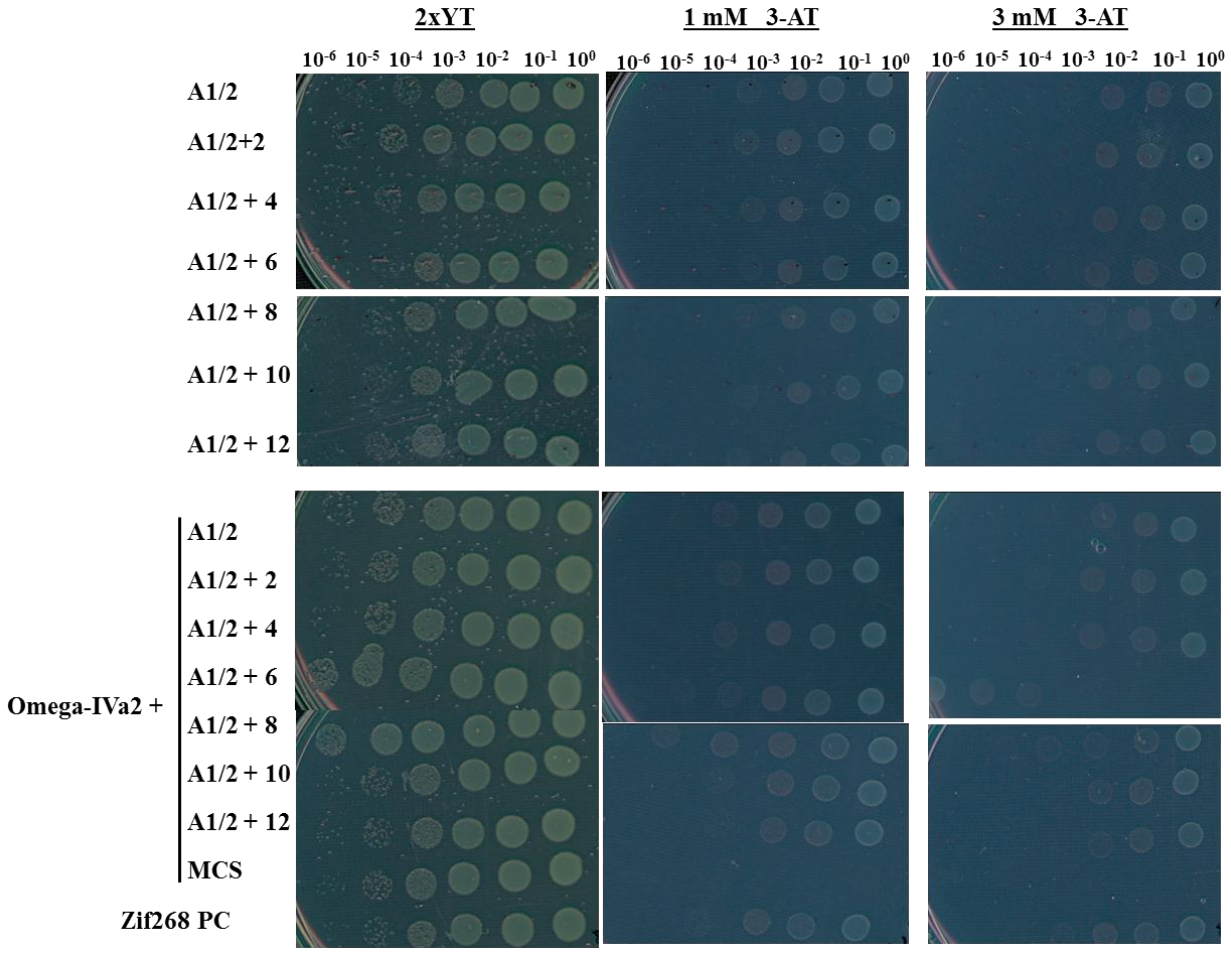
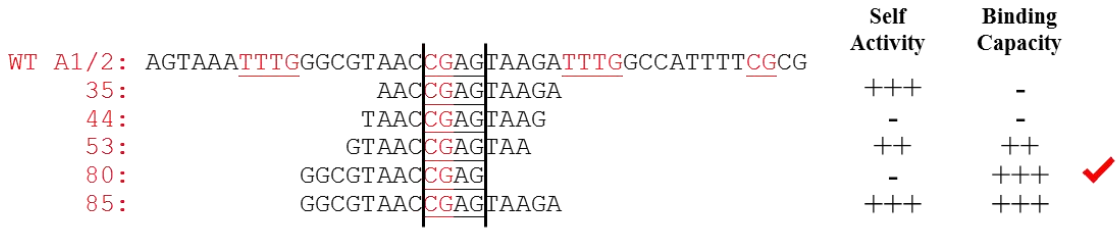
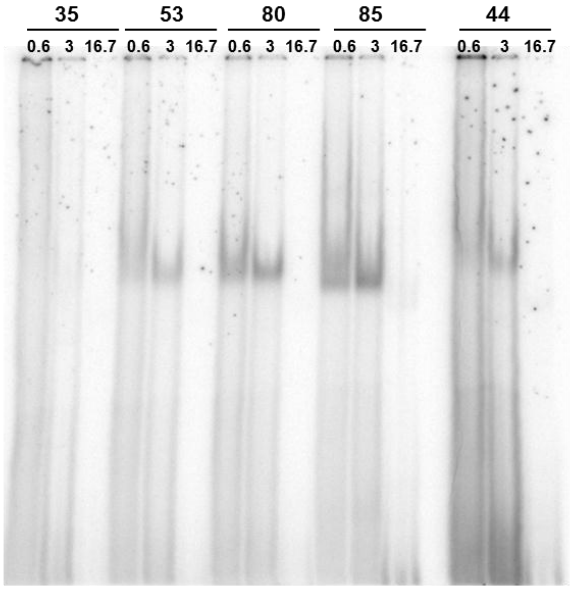


Figure A10. A1/2 series. A1/2 and A1/2 with different length of spacers were cloned into the reporter plasmid. Their self-activity and interaction with Omega-IVa2 were tested. Compared with the negative control (Omega-IVa2+MCS) and the positive control (Omega-Zif268+Zif268 binding site) in Fig. A9, it is clear that the A1/2 series has a greater level of self-activity and it does not respond to Omega-IVa2 for activation of the B1H reporter genes at all.

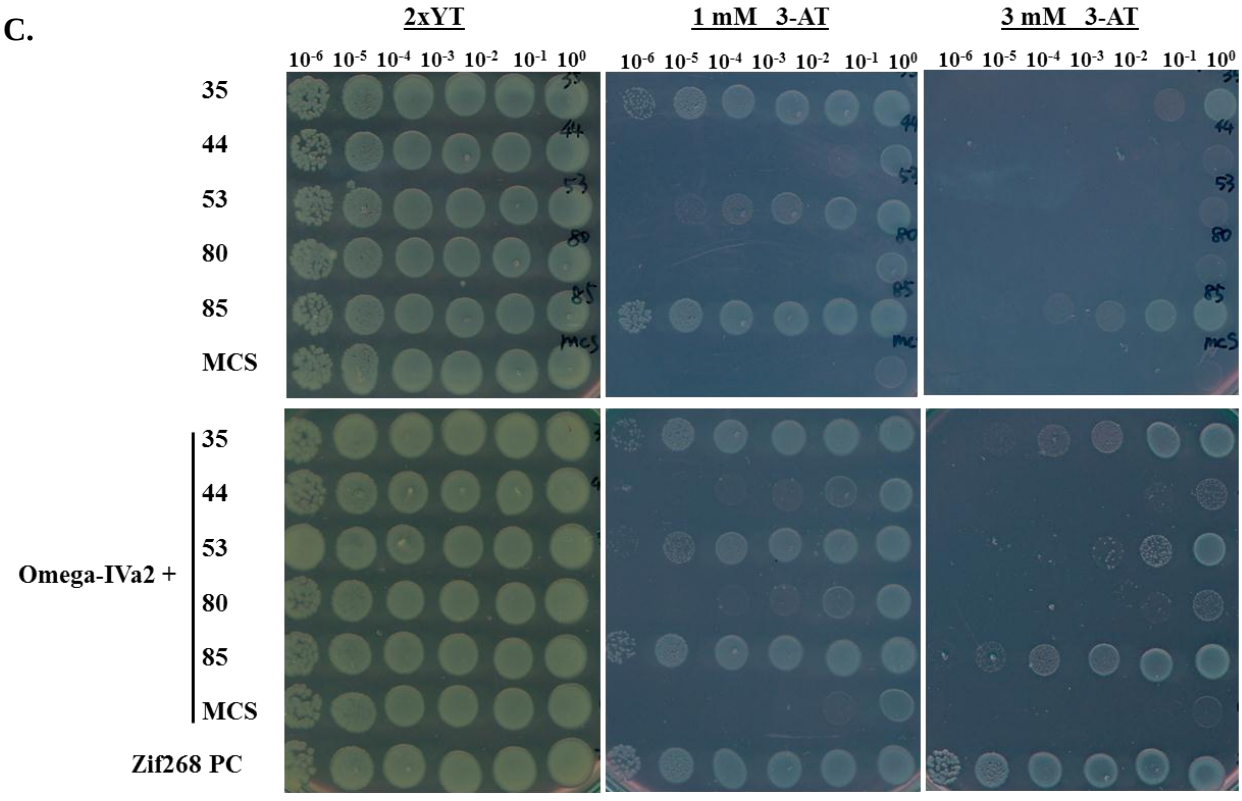
A.


Self Activity
Binding Capacity
WT A1/2: AGTAAATTTGGGCGTAACCGAGTAAGATTTGGCCATTTTCGCG
35: AACCGAGTAAGA +++ -
44: TAACCGAGTAAG - -
53: GTAACCGAGTAA ++ ++
80: GGCGTAACCGAG - +++
85: GGCGTAACCGAGTAAGA +++ +++

B.


35
53
80
85
44
0.6 3 16.7 0.6 3 16.7 0.6 3 16.7 0.6 3 16.7 0.6 3 16.7

C.


2xYT
1 mM 3-AT
3 mM 3-AT
10-6 10-5 10-4 10-3 10-2 10-1 100
35
44
53
80
85
MCS
Omega-IVa2 +
35
44
53
80
85
MCS
Zif268 PC

80+2: GGCGTAACCGAG GT
80+4: GGCGTAACCGAG GTGC
80+6: GGCGTAACCGAG GTGCAC
80+8: GGCGTAACCGAG GTGCACTC
80+10: GGCGTAACCGAG GTGCACTCCG
80+12: GGCGTAACCGAG GTGCACTCCGGA

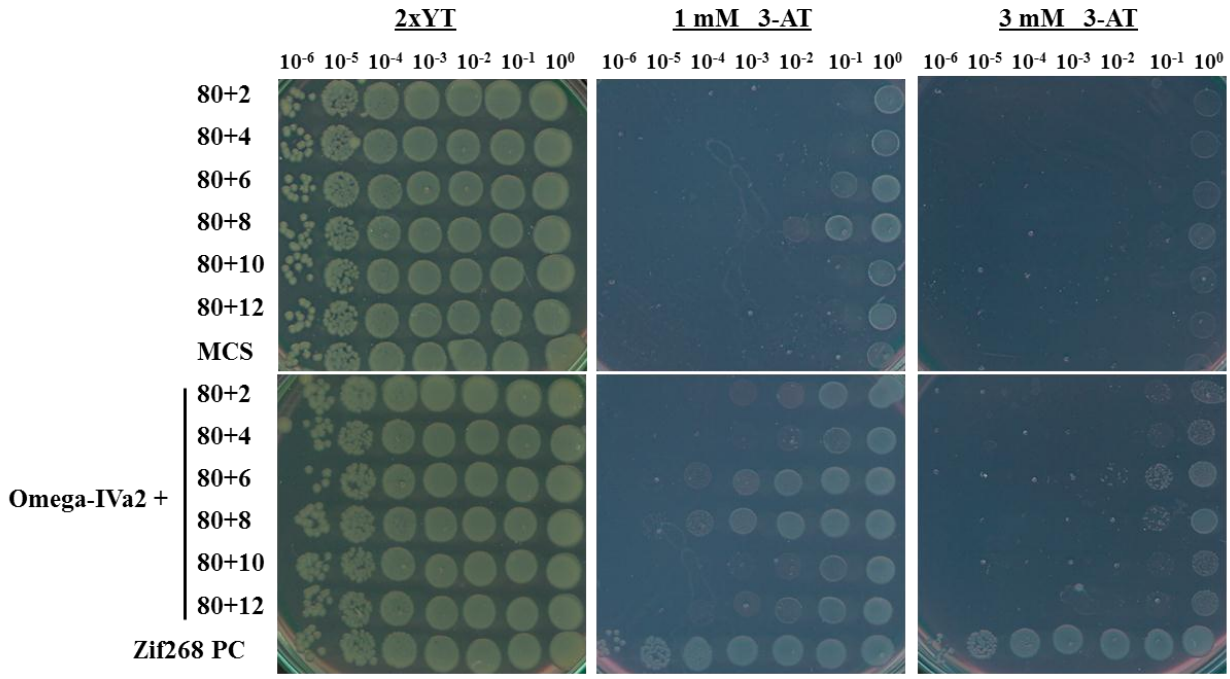


Figure A12. The 80 series. The 80 sequence with different length of spacers were cloned into the reporter plasmid. Their self-activity and interaction with Omega-IVa2 were tested.

Figure A11. Shortened A1 Series. In order to reduce the self-activity, the CGAG motif with different lengths of flanking region **(A)** were tested for binding with IVa2 by EMSA **(B)** and cloned into the reporter plasmid to test their self-activity and response to Omega-IVa2 **(C)**. The result showed that the 80 sequence has the best combination of low self-activity and strong binding to IVa2, even though its response to Omega-IVa2 is still weak.

80di: GGCGTAACCGAG GGCGTAACCGAG
80di+2: GGCGTAACCGAG GGCGTAACCGAG GT
80di+4: GGCGTAACCGAG GGCGTAACCGAG GTGC
80di+6: GGCGTAACCGAG GGCGTAACCGAG GTGCAC
80di+8: GGCGTAACCGAG GGCGTAACCGAG GTGCACTC
80di+10: GGCGTAACCGAG GGCGTAACCGAG GTGCACTCCG
80di+12: GGCGTAACCGAG GGCGTAACCGAG GTGCACTCCGGA

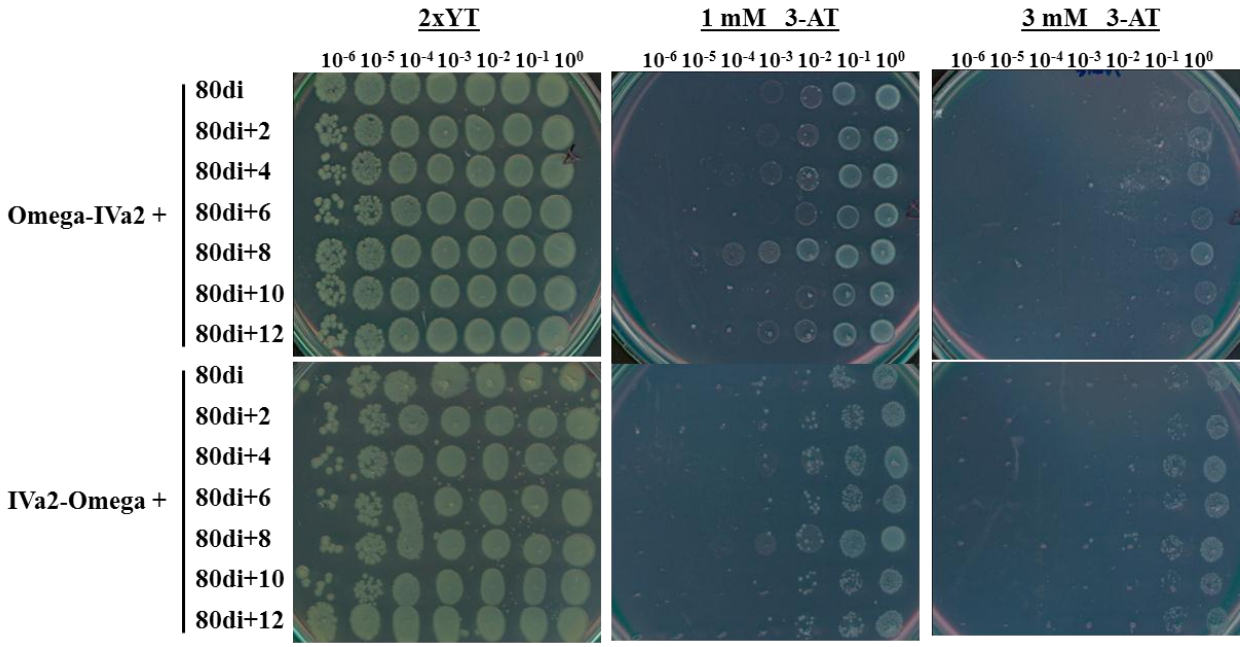


Figure A13. The 80 dimer series with Omega-IVa2 or IVa2-Omega. The results showed that for the 80 dimer series, there is also a good correlation of higher self-activity and stronger response to Omega-IVa2, which suggest the stronger response is not specific to the interaction of IVa2 and the binding sequence. It also showed in this particular setting that putting IVa2 at the N-terminus of the fusion protein does not help to gain better activation of B1H reporter genes.

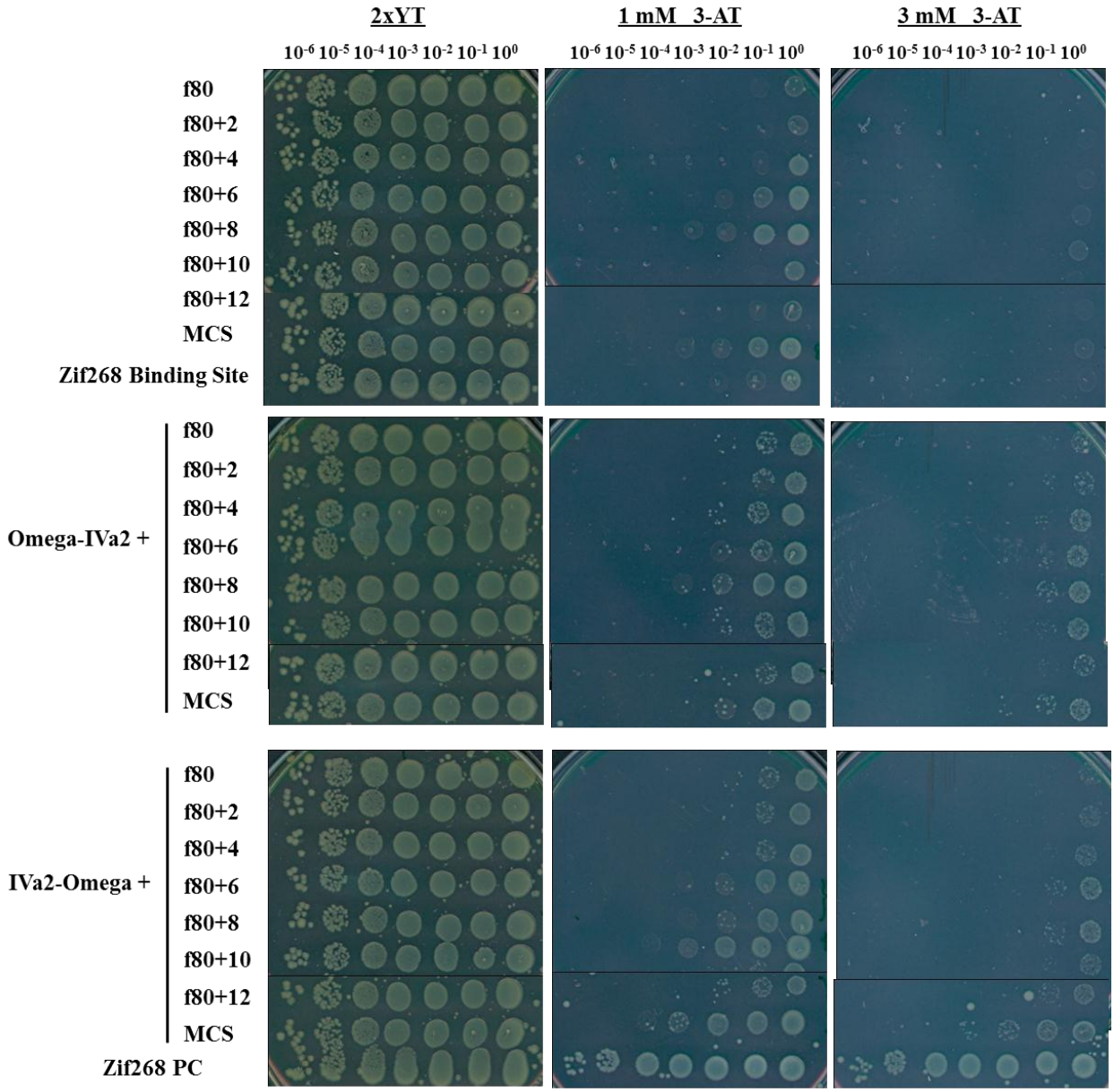


Figure A14. The flipped 80 series with Omega-IVa2 or IVa2-Omega. The results showed that most of these combinations did not present a better activation. The best candidate pair from these combinations is f80+10 with IVa2-Omega. But the window between self-activity and reporter gene activation is too narrow to conduct a convincing library screen.

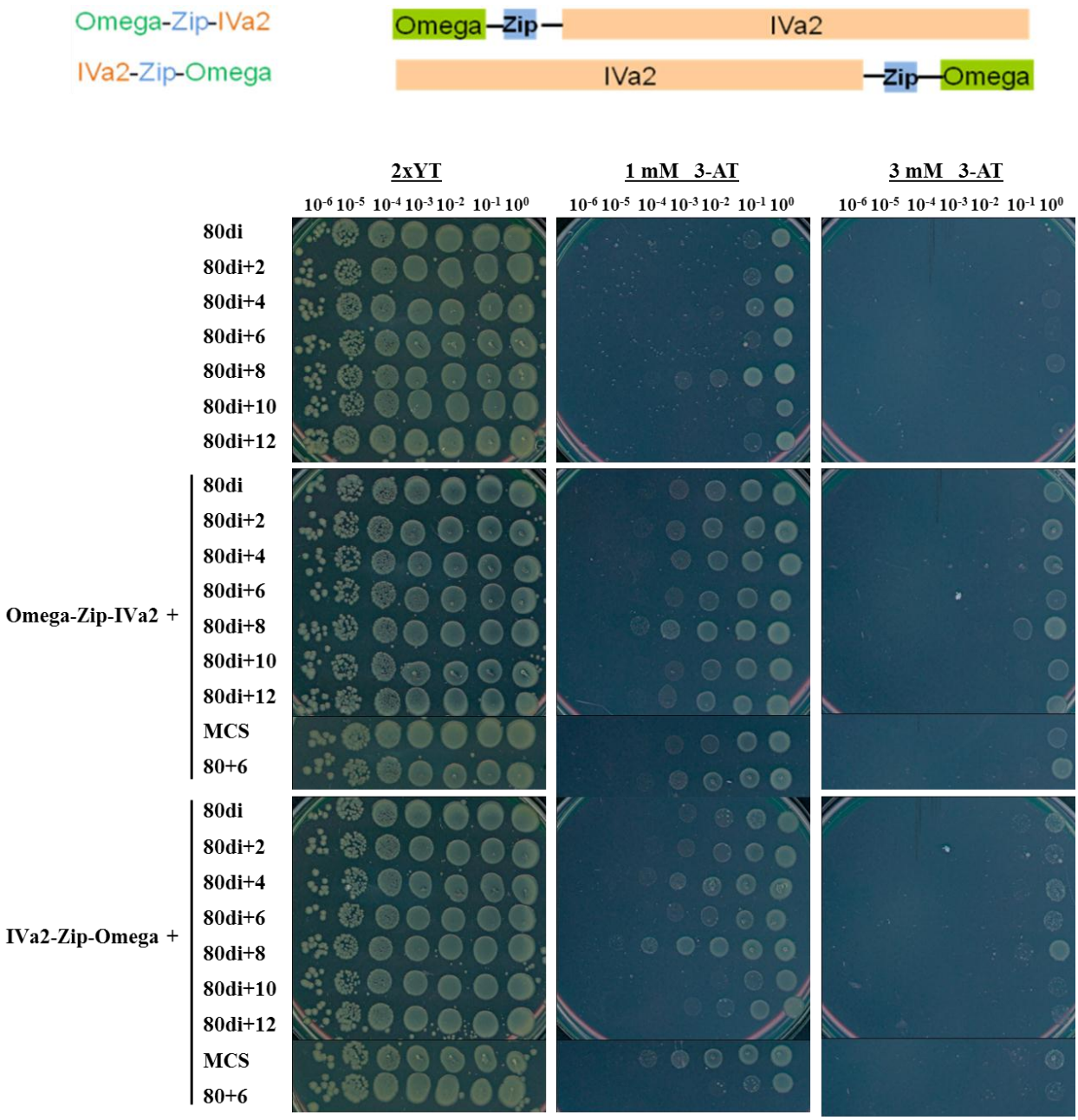


Figure A15. Leucine zipper motif as the linker between Omega and IVa2. Considering the weak interaction between IVa2 and its DNA binding sequence, we constructed Omega-Zipper-IVa2 and IVa2-Zipper-Omega (top panel) with the leucine zipper motif as the linker in the middle, with the hope that homodimerization of IVa2 by the leucine zipper might improve the interaction with DNA and stimulate activation. The two constructs were tested with the 80 dimer series, but no better activation was observed either.

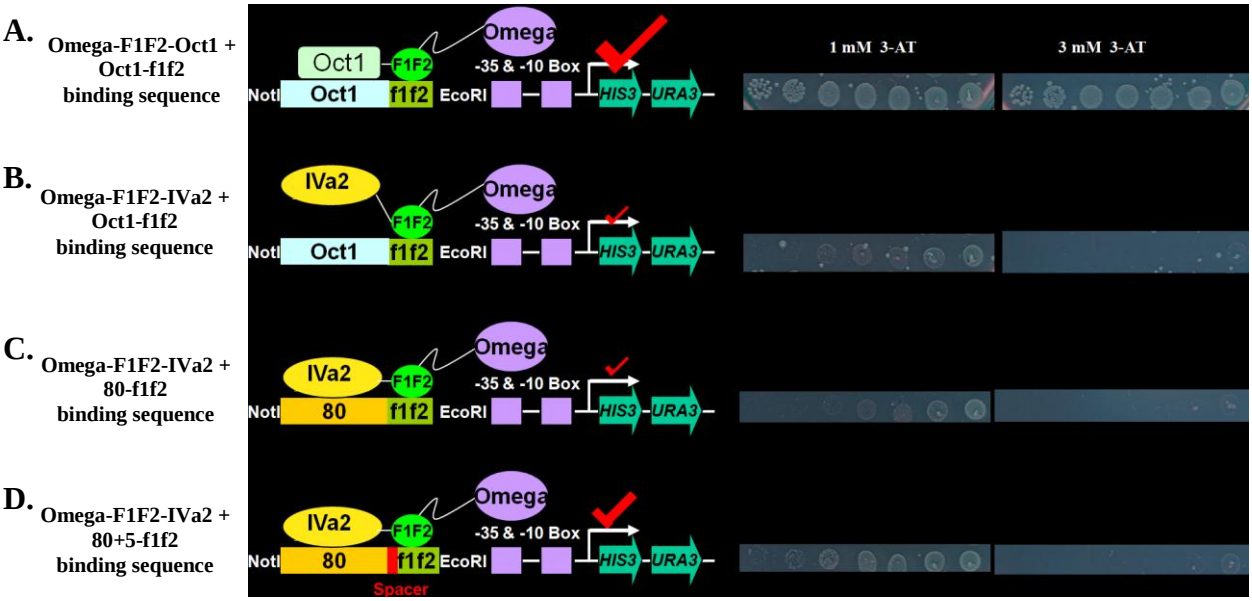


Figure A16. The Omega-Zif12 B1H system controls. With a limited activation response from WT IVa2 and its binding site in the original B1H system, I turned to the omega-Zif12 B1H system, which introduces finger 1 and 2 (F1F2) of the Zif268 protein and its DNA binding sequence (f1f2) into expression and reporter vectors, respectively. This system is aimed to improve weak interactions between protein and DNA with the help of F1F2 binding to f1f2 on the reporter plasmid. The system positive control **(A)** and finger 1 and 2 control **(B)** was determined with general B1H protocol. IVa2 protein was also tested in this system together with the reporter vectors with **(C)** or without **(D)** a 5 nt spacer.

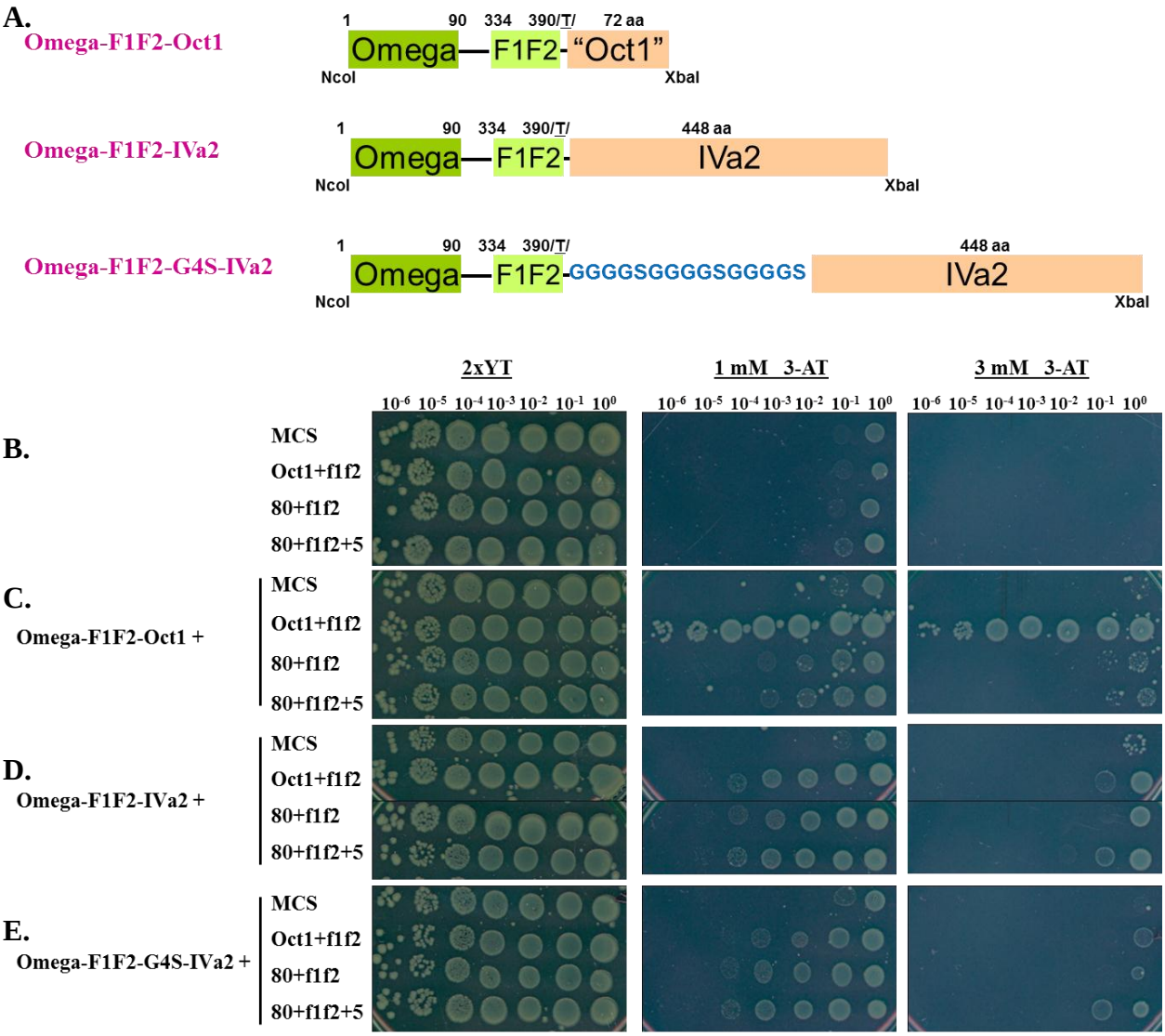


Figure A17. The Omega-Zif12 B1H system using the 80+f1f2 series and the flexible glycine linker. A. Diagrams of different expression vectors. **B.** Self-activity of the reporter plasmids. Omega-F1F2-IVa2 protein and its corresponding reporter plasmid with (80+f1f2+5) or without (80+f1f2) a 5-nt spacer were constructed and analyzed. **C.** The positive control Omega-F1F2-Oct1 survived very well due to interaction of both Oct1 and F1F2 with their corresponding binding sites. **D.** When Omega-F1F2-IVa2 was added with the same Oct1-f1f2 binding sequence as used for the positive control above, the effect of F1F2 and its binding site was measured as the F1F2 background, which gave a 5-log bacteria growth in 1 mM 3-AT. However, when Omega-F1F2-IVa2 was added with its own 80 binding site with or without the 5-nt spacer, the results showed the same activation response as F1F2 background. **E.** Since there is only one threonine between Finger 1+2 and IVa2 in the construct of Omega-F12-IVa2, a 15-aa flexible linker (GGGGS repeats) was cloned in between for testing. However, there was no better activation obtained. In conclusion, for the interaction of IVa2 and its binding site, this system did not give rise to better activation than the original B1H system.

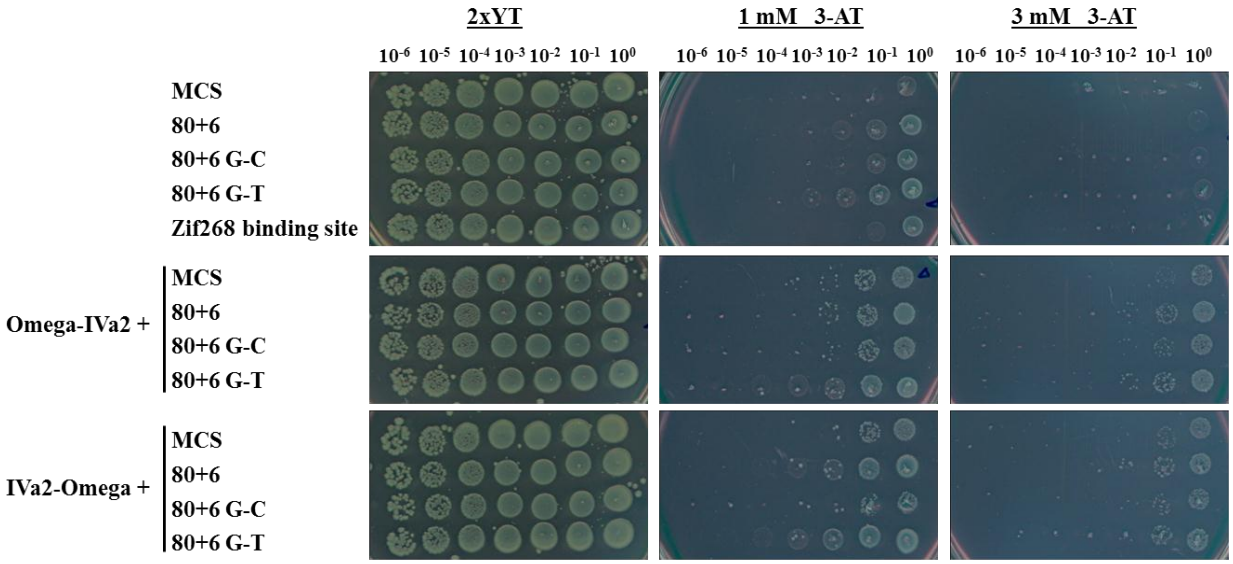
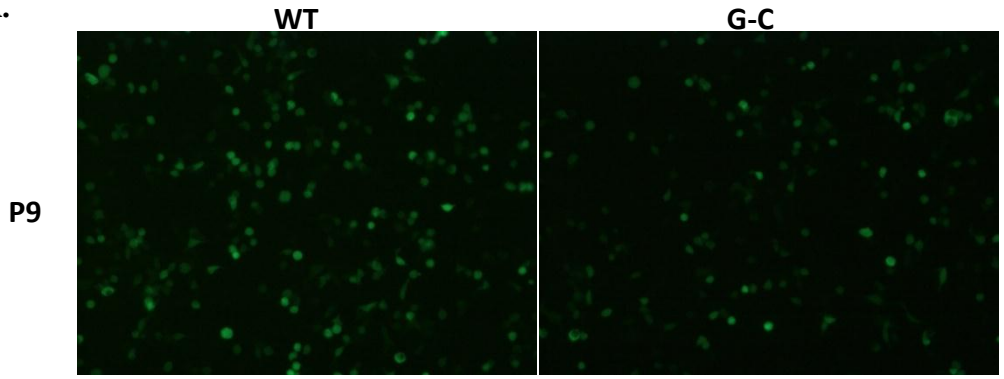


Figure A18. Specifity of the B1H reporter gene expression. Since we have known from the *in vitro* and *in vivo* analyses that the G-C mutation is the most defective one for IVa2 binding (Fig. A7), G-C or G-T mutations were introduced to the 80+6 sequence, and they were tested in the original B1H system in order to test whether the activation of reporter gene expression is specific to the interaction between IVa2 and its binding site. The data showed that only in the IVa2-Omega case, there was a slight decrease of reporter gene activation in 80+6 G-C. However, the issue of the small window of activation between the positive and negative controls make it impossible to do any convincing library screen.

Figure A19. The *in vivo* selection system controls. First, the window between positive (20 plaques per 60 mm plate transfection) and negative controls (no plaques) is not great. Second, the 20 plaques from the positive control could not be amplified by passaging the virus in 293 cells, which later turned to be due to the insufficient IVa2 expression level expressed from the CMV promoter than the natural Ad5-WT infection (data not shown), even though we could not explain why these plaques grew in the first place. These results closed the door of this approach.

A.



B.

Virus	A1/2	IVa2	L4-22K	L1-52/55K	IIIa
WT	WT A1/2 trimer	WT	WT	WT	WT
G-C	G-C A1/2 trimer	WT	K107E	WT	WT

C.

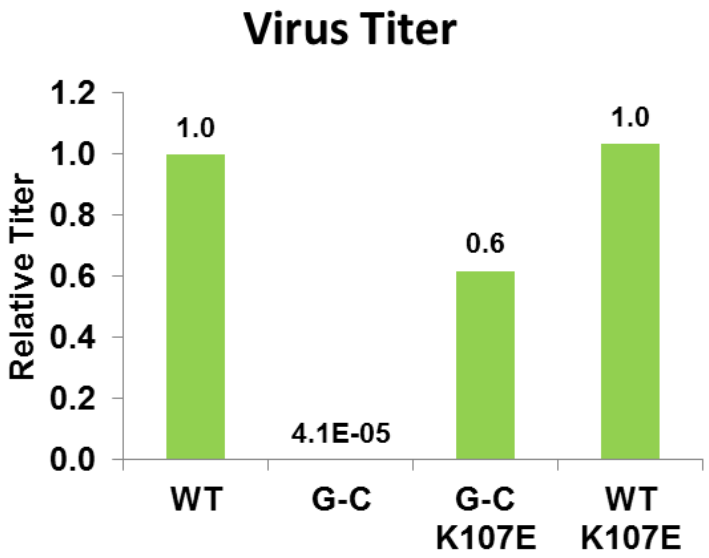


Figure A20. An L4-22K suppressor mutant rescues the packaging defect due to the G-C mutation in the IVa2 binding site. A. GFP-positive N52.E6-Cre cells at 18 hpi by WT or G-C viruses that have been passaged in N52.E6-Cre cells for nine passages. **B.** Summary of DNA sequences of A1/2 region, and the protein coding sequences of L4-22K, IVa2, L1-52/55K and IIIa from the passage 9 (P9) lysates of WT or G-C viruses. **C.** The titers of the indicated viruses which had been transfected by corresponding infectious clones for ten days. The lysates (passage 1; P1) of each transfection were used to infect N52.E6-Cre cells for titration of the growth of each individual virus by measuring GFP signal using flow cytometry.