

Adeno-associated Virus and the Development of Adeno-associated Virus Vectors: A Historical Perspective

Barrie J. Carter*

Targeted Genetics Corporation, Seattle, WA 98101, USA

*To whom correspondence and reprint requests should be addressed at Targeted Genetics Corp., 1100 Olive Way, Suite 100, Seattle, WA 98101, USA.

Available online 18 October 2004

Contents

Introduction	981
Discovery of AAV	982
AAV Molecular Biology	982
AAV Latency	983
AAV Genetics and Molecular Clones	983
The First AAV Vectors	985
AAV Latency Revisited	986
Additional Revisitation of History	987
References	988

INTRODUCTION

Adeno-associated virus (AAV) was discovered almost 40 years ago [1–3] and the first AAV vectors were described 20 years ago [4–6], which led to initiation of the first human clinical trial of an AAV vector some 8 years ago, in cystic fibrosis patients [7]. These events also delineate three periods in the history of AAV vector development. In the first 20 years, the basic elements of the AAV genome structure, growth cycle, and latency were described. In the early 1980s, infectious molecular clones of AAV serotype 2 were generated and the AAV2 DNA sequence was determined. This led to the initial genetic analysis of AAV and the development of a *trans*-complementation system that enabled generation of the first AAV vectors by the mid-1980s.

In the next 10 years until 1995, important developments included refinement of the basic design of AAV vectors, the first *in vivo* experiments, and preparation of the first AAV vector to enter clinical trials. Also, the AAV Rep proteins were identified and shown to have a remarkable series of biochemical properties, including binding and site-specific nicking at the AAV DNA termini and some potential for cytostasis. We learned also that, in cultured cells, wild-type AAV integrates with high efficiency and specificity at a particular site on human chromosome 19 in an interaction mediated by binding of the AAV Rep protein to the AAV DNA termini and a

site within chromosome 19. The immediate implication for AAV vectors, which generally are deleted of AAV rep, was that they would not integrate with great efficiency at the same site and this led to reinvestigation of AAV vector persistence. The cytostatic property of the Rep proteins initially presented an obstacle to generation of efficient packaging systems for AAV vectors using stable producer cell lines, as had been accomplished for retroviral vectors.

In the most recent 10 years, there has been an explosion of interest in AAV vectors and many additional important developments. First, the *trans*-complementation transfection systems have been improved for AAV production and the cytostatic properties of Rep protein have been overcome to generate producer cell systems that do not require transfection as part of the vector production process. Second, there have been many insights into the biology of AAV, including analysis of additional serotypes and trafficking of vectors in mammalian cells. Third, AAV vector genomes have been shown to persist in host cells *in vivo* mainly as duplex, circular episomes. Finally, accumulating evidence suggests that AAV vectors show a very encouraging safety profile in clinical trials and there is a rapid expansion of AAV vectors entering clinical trials for various clinical applications.

In this article, I will discuss mainly the first 20 years of AAV to describe the earliest work on AAV and AAV vectors. This was a time when only a small handful of

investigators were involved and it generally may not be familiar to the vast number of individuals who now use AAV vectors. I will comment on the second 20 years of AAV briefly and only to the extent that it helps connect the early years to the present. The greatly expanded accomplishments of this period, as well as some of the earlier work, are summarized in a recent review [8]. I will present a somewhat personalized view and I have included a number of anecdotes that help reflect the novelty of the early days of the field.

Much of the earliest work on AAV, particularly the molecular biology aspects, reflected the work on bacteriophage molecular biology in the 1960s. Several of us, such as Ken Berns and myself, had been thoroughly steeped in the lessons about circular and linear phage genomes, circular permutations, terminal redundancies, and repeats, and this colored much of the thinking of early studies on AAV. I first learned about AAV as a postdoc in Lionel Crawford's laboratory at the Imperial Cancer Research Fund (ICRF), in London, in 1969 when I read a fascinating paper from Jim Rose and Ken Berns at the U.S. National Institutes of Health (NIH). Soon thereafter, Jim wrote a letter to Lionel Crawford to recruit a molecular biologist, and in October 1970, I arrived at NIH to work on AAV in Jim's laboratory in the National Institute of Allergy and Infectious Diseases (NIAID).

DISCOVERY OF AAV

Adeno-associated viruses were discovered almost 40 years ago in the laboratories of Bob Atchison at Pittsburgh [1] and Wallace Rowe at NIH [2]. Both groups recognized that small particles, which were initially thought to be subunit contaminants of adenovirus preparations, were actually a genus of the parvoviruses, now known as dependoviruses. AAV is replication defective and is dependent upon helper functions provided by adenovirus or herpesviruses for productive replication in mammalian cells. Immediately after its discovery, AAV was also isolated from humans, but serological studies, especially those led by Neil Blacklow [3], indicated that it was not associated with causing any disease even though it appears to travel in human populations with the helper virus, particularly adenovirus.

The absence of pathogenicity due to AAV infection probably explains why it did not generate much interest in the virology community at large. However, in 1969, in the almost prehistoric days prior to molecular cloning of genes or DNA sequencing, AAV offered some unique advantages as an experimental system. Because it had a DNA genome of only 5 kb and plus and minus strands that were packaged into individual particles but could be separated, AAV represented an identifiable genetic system in which gene expression might be studied in human cells. Also, it was a defective virus and required adenovirus as a helper and mutants of adenovirus were just

beginning to be identified. Thus, even though the field of mammalian cell genetics was rudimentary, it seemed possible that adenovirus mutants could be used to study AAV in the context of host background mutations. This provided a prospect of connecting biochemistry and genetics and was what originally attracted me to AAV.

Following the discovery of AAV, three parallel lines of investigation to understand its molecular biology, latency, and genetics converged to inform the design and generation of the first AAV vectors [4–6]. However, the subsequent expanded development and use of AAV vectors has taught us as much about AAV biology as the study of AAV directly has taught us about vector development.

AAV MOLECULAR BIOLOGY

AAV vectors reflect two major unique properties of AAV DNA. They contain a single strand of DNA, the transgene cassette, bounded by the AAV genome inverted terminal repeat (ITR) sequences. The analysis of AAV DNA to identify its unique structure is a fascinating example of deductive reasoning from structural analysis in the days before DNA sequencing was readily available.

The first unique feature of AAV DNA was alluded to by Lionel Crawford [9], who performed a careful physicochemical comparison of bacteriophage Φ X174, the autonomous parvovirus minute virus of mice (MVM), and AAV. All three viruses were similar in particle size, composition, and buoyant density in the ultracentrifuge but DNA extracted from Φ X174 and MVM was single-stranded with a molecular weight of about 1.5×10^6 , whereas that from AAV appeared to be double-stranded with a molecular weight of about 3×10^6 . To resolve this conundrum, Lionel suggested that individual AAV particles must contain a single DNA strand of either plus or minus complementary sense that reassociate to form duplex DNA upon extraction from particles. At NIAID, Jim Rose and Ken Berns [10] used a modification of the classic Meselson and Stahl density-shift experiment from a decade earlier to devise an elegant test of Crawford's hypothesis. They isolated DNA from AAV that had been grown in normal culture medium or in medium in which thymidine residues were substituted by BUdR to give DNA with a heavier buoyant density in CsCl gradients. Duplex DNA from the individual virus stocks banded at either light or heavy density as expected. However, if the virus stocks were mixed before DNA isolation, the resulting duplex DNA molecules banded at the intermediate density. Thus, the experiment showed that there were equal numbers of particles having a single-strand DNA genome of either minus or plus sense [10]. This was the paper that I read while in Crawford's laboratory at the ICRF.

The density substitution experiment also led to a preparative separation of the AAV DNA minus and plus strands. An asymmetry of T residues between the two

strands causes denatured BU-substituted minus strands to band at a higher density than plus strands in CsCl. This was useful for defining the template strand for AAV mRNA (the minus strand) and for determining the polarity of the genome [11], but more importantly, it led to the observation of inverted repeats in the AAV DNA termini. In 1971 at NIAID, together with the late George Khoury, I was analyzing AAV RNA transcription by assaying AAV RNA-DNA hybrids with S1 nuclease and hydroxylapatite using a system George had set up with Malcolm Martin to analyze SV40 RNA. However, we made the much more interesting observation that separated minus or plus single strands of AAV DNA bound and eluted from hydroxylapatite as if they were double stranded, but only a small region of the strand was responsible for this behavior [12]. We guessed that the AAV DNA termini contained an inverted repeat because this would give the same sequence at both ends of any duplex replicating form and might explain how replication could be initiated from either end ultimately to package separate strands into individual particles. This was quickly proven by direct visualization in the electron microscope, which showed that AAV DNA single strands could form hydrogen-bonded single-strand circles and must therefore contain an inverted terminal repeat [13].

Meanwhile, Ken Berns and Tom Kelly, at Johns Hopkins University in Baltimore, used limited exonuclease degradation of duplex AAV DNA followed by electron microscopy and observed the formation of duplex circular molecules [14]. This showed that there was a direct terminal repeat that could mediate formation of duplex circular molecules. The presence of both inverted and direct repeat terminal repeat sequences indicated that there must be a terminal palindromic sequence that would have both types of properties [15]. Terminal palindromes would be expected to form hairpin structures and additional structural complexities were shown in a series of studies in my laboratory (by then at the National Institute of Diabetes, Digestive, and Kidney Diseases, NIH) and that of Berns. Using the newly emerging restriction endonucleases [16] we observed two types of duplex termini that formed at each end of AAV DNA, and experiments with David Denhardt on 5' ³²P-end-labeling of AAV duplex DNA showed that one type was a linear duplex and the other had a nonlinear cruciform or rabbit-ear structure [17].

The structure of the termini as hairpins was further substantiated when Steve Straus and Jim Rose at NIAID made the novel observation that duplex replicating forms of AAV DNA were concatemeric head-to-head and tail-to-tail structures that contained hairpin sequences at one or the other terminus and that AAV DNA replicated from the terminus that could form a self-priming hairpin [18]. Additional evidence came from Bill Hauswirth, in the Berns lab [19], who showed, using isotope labeling kinetics and restriction cleavage, that the termini are

the origins of replication, and from Luis de la Maza in my lab [20], who showed that internally deleted AAV DI (defective-interfering) genomes have an exactly analogous hairpin terminus at one genome end. Finally, direct sequencing of the AAV2 termini in 1980 in the Berns laboratory showed that they do indeed contain an ITR that is a 145-nucleotide sequence of which the first 125 nucleotides form a T-shaped, base-paired hairpin structure [21]. The direct sequencing also revealed the inversion of the internal region of the hairpin ("flip" and "flop") that occurs as a result of hairpin transfer at each round of replication, according to a model proposed earlier by Cavalier-Smith to explain how linear eukaryotic chromosome termini might be duplicated [22]. It was nearly 10 years later, after our identification of the AAV Rep proteins [23] and their initial purification and *in vitro* assay in the Muzyzcka laboratory [24], that the remarkable biochemistry of the Rep protein interaction with the ITRs began to be elucidated.

AAV LATENCY

A program to screen cell lines for vaccine production at NIH resulted in observations that infection of some batches of African green monkey kidney cells and one or two primary cultures of human embryonic kidney cells with adenovirus resulted in rescue of infectious AAV. This led David Hoggan at NIAID in 1972 to test the idea that AAV might be able to establish a latent state in host cells. He took a human cell line, Detroit 6, that was originally established from a nasopharyngeal carcinoma and infected it with AAV at a high multiplicity in the absence of any helper virus. After growth through at least 10 cell passages infectious AAV was no longer detectable from cell lysates but superinfection of the cultures with adenovirus resulted in rescue of infectious AAV [25]. This was the first indication that AAV may persist in a cell in the absence of helper when conditions are not permissive for productive replication. Further analysis, by Cheung *et al.* [26], of a cell line carrying latent AAV cloned from the originally infected Detroit 6 cell culture, showed that a relatively low number of AAV genomes were integrated into the host cell chromosome mostly as head-to-tail tandem repeats. The idea that AAV could establish a latent state was an early impetus to thinking about using AAV as a gene delivery vector.

AAV GENETICS AND MOLECULAR CLONES

By 1980 we had a general summary of the structure of AAV DNA, a definitive proof of the palindromic arrangement of the ITRs, and an overall understanding that AAV DNA replication involved self-priming from the ITR hairpins and amplification via head-to-head and tail-to-tail concatemers that were precursors of the progeny single strands. Maureen Myers showed that these progeny

single-stranded genomes are packaged into preformed capsids [27]. Largely through the efforts of Cathy Laughlin [28] and subsequently Carol Marcus-Sekura [29], we had established the basic transcription map of AAV, including the three overlapping mRNA families in both spliced and unspliced versions that were generated from the three transcription promoters p5, p19, and p40. However, we knew much less about the genetics of AAV except that the three AAV capsid proteins, VP1, VP2, and VP3, had overlapping sequences and were coded by the smaller 2.3-kb mRNA transcribed off the p40 promoter in the right half of the genome [30]. We did not know what the left half of the genome coded for via the p5- and p19-promoted transcripts, but suspected that this might be a replication function that was predicted based on the properties of AAV DI genomes [31]. At this time, virtually all the molecular biology on AAV was being undertaken with AAV serotype 2 and this remained the case until into the 1990s. We now know that expanding the AAV work to a variety of serotypes offers a greater range of biology and sophistication in development of vectors. However, it was probably important that the early work focused on AAV2 because, with only a handful of labs working on it, the cumulative results were more readily coherent.

By 1980, and in the general context of the early development of gene delivery for therapeutic applications, there were clear reasons to develop AAV vectors because the parental virus was not a pathogen and because it appeared to have a mechanism for persistence. These are still the primary reasons for using AAV vectors even though the knowledge about the persistence mechanism has taken several interesting turns. We also knew then that AAV has a very high burst size, well in excess of 100,000 particles per cell, and could be produced readily at high titers, particularly in suspension culture. At the time we simply thought that AAV was easy to grow. We now recognize that this was important in ultimately being able to achieve AAV vector production at high specific productivity (yield per cell) that helps to minimize the biomass (number of cells) needed for a given total yield. This is of critical importance for the eventual commercial production of AAV vectors as therapeutic products. However, in 1980, several obstacles still stood in the way and important steps remained to be taken before AAV vectors could be developed.

Analysis of adenovirus mutants in the 1980s did identify the adenovirus genes that are necessary and sufficient to establish the helper function for AAV replication as a subset of adenovirus early genes, including E1, E2A, E4orf6, and the VA1 RNA gene [32]. Subsequently, this has been of importance in developing AAV vector production systems [8]. However, establishing a direct genetic analysis of AAV was difficult because it replicated only in the presence of a lytic helper, so there was no easy plaque- or focus-forming assay that permitted selection of

mutants. Molecular clones of AAV were needed, but in the early days of the guidelines established by the NIH Recombinant DNA Advisory Committee cloning of AAV DNA into plasmids was not straightforward. Aside from the less well developed technology at the time, there was a regulatory issue that cloning of an intact human virus in bacterial plasmids required a high level of containment and even though AAV was not a human pathogen there was ambiguity about the containment that would be required. However, an ad hoc working group of about 12 of us virologists was convened by Ken Berns at the ASM meeting in 1980 in Miami Beach. At that meeting it was recommended, and agreed on by the NIH RAC soon after, that recombinant DNA viruses including AAV, adenovirus, and herpesvirus generally could be safely studied in BL2 containment. This permitted us to proceed with cloning of AAV into plasmids.

Two approaches were taken to yield molecular clones of AAV2. In Gainesville, Florida, Nic Muzyczka and Ken Berns, with their newly arrived graduate student Jude Samulski, constructed molecular clones in the plasmid pBR322 by taking AAV2 duplexes and adding GC tails followed by insertion at the *Pst*I site in the plasmid and successfully derived molecular clones [33]. However, AAV2 has three *Pst*I sites within its sequence so such clones would not enable direct excision of an intact AAV DNA duplex. At NIAID, Cathy Laughlin, Jon-Duri Tratschin, and I adopted a different strategy by adding *Bgl*III linkers and cloning into a derivative of pBR322 that had been engineered to contain a *Bgl*III site (which is not present in AAV2 DNA) so that we would be able to excise intact duplex AAV2 genomes from the plasmid [34]. Purified AAV DNA had been shown to be infectious [35] but we all expected that we would need to excise intact AAV DNA from the plasmid. However, this need was superseded by a fortuitous, but unforeseen, property of the AAV molecular clones.

Transfection of the AAV plasmids into mammalian cells that were infected with adenovirus resulted in rescue and replication of infectious AAV genomes free of plasmid DNA sequence and production of infectious AAV particles [33,34]. Thus, the infectivity of these clones essentially recapitulated the rescue of latent AAV from the Detroit 6 clones of David Hoggan. However, our understanding of this rescue of latent AAV was fairly rudimentary and it was not clearly appreciated beforehand that the plasmid clones of AAV would be so readily infectious. Remarkably, the efficiency of rescue and replication was sufficiently high that excision of the intact duplex prior to transfection had little impact on the yield of infectious AAV [34]. The availability of the infectious molecular clones opened the way to understanding AAV genetics and the development of AAV vectors.

The presence of the ITRs in AAV DNA presented some difficulties in generating the initial plasmid clones. In

adding *Bgl*II linkers, we had to resort to generating two half-genome molecules and then reassembling a single intact clone because we had been unable to recover intact clones when we tried adding *Bgl*II linkers to both ends of intact AAV duplexes in a single step [34]. This probably reflected several aspects of AAV termini, including the terminal flip-flop, which renders 50% of the duplex ITRs nonlinear, and that the cleavage of AAV DNA is staggered so that the 5' terminal sequence is TTG-, TT-, or G-. Consequently the proportion of termini that were linear and blunt-ended for optimal ligation of a linker was probably very low. Nonetheless, once the intact molecular clones were assembled they were quite stable, including the ITR regions.

Interestingly, the GC-tailing approach gave infectious clones also, even though almost none of the clones initially isolated had intact ITRs [33]. Generally, varying portions of the ITRs were deleted or were unstable, perhaps because of the combination of the GC tails and the GC stacks in the internal region of the ITR. However, examination of the progeny AAV DNA genomes that were rescued by replication from these plasmids showed that the deleted regions of an ITR had been repaired, which indicated that during AAV replication the ITRs must interact and undergo a process of gene conversion [36]. Provided that there was a complete ITR represented in aggregate by the ITR regions at the two ends, infectivity was maintained. Subsequently, clones having up to the first 120 nucleotides of the ITR deleted at one or the other end proved to retain infectivity [37].

Introduction of mutations within the AAV sequence of the molecular clones enabled genetic analysis of AAV and interpretation of this analysis was aided by the completion of the entire DNA sequence of AAV2 in 1983 by the Berns lab [38] using the Maxam-Gilbert chemical sequencing procedure. Thus, by 1984, the initial genetic analysis [39,40] quickly showed that mutations in the ITRs had a *cis*-acting Ori⁻ phenotype consistent with their function as replication origins. The DNA sequence showed two large open reading frames in the right and left halves of the genome and the genetic analysis demonstrated that each coded for a *trans*-acting function. The right half gene had a cap⁻ phenotype and thus coded for the AAV capsid proteins as previously demonstrated biochemically. The left half gene had a rep⁻ phenotype that was AAV DNA replication negative and thus coded for the previously predicted replication function. Interestingly, another phenotype (lip⁻ or inf⁻) was revealed by mutations now known to lie in the amino-terminal region of VP1 and resulted in a failure to assemble stable infectious particles. These studies delineated the basic approach to making AAV vectors. Since the ITRs are *cis*-acting they must be retained in AAV vector genomes, but the *trans*-acting Rep and Cap functions can be provided by

complementation. Thus, most of the AAV genome, except the ITRs, can be deleted and the maximum payload for an AAV vector is therefore about 4.5 kb if all of the unique sequence is deleted.

THE FIRST AAV VECTORS

The first AAV vectors, described in 1984, were generated using a two-plasmid (vector plasmid and complementing plasmid) transfection system in either human KB cells or 293 cells that were also infected with adenovirus [4,5]. The vectors were generated as crude cell lysates in which the contaminating infectious adenovirus was removed by heat inactivation. In the vector plasmids, generally most of the capsid gene was deleted but a functional rep gene was retained and therefore the vectors that were generated had a rep⁺cap⁻ phenotype. This largely reflected the availability of convenient restriction endonuclease sites. The plasmids that were used to provide complementing Rep or Cap functions either contained the entire wild-type sequence or had extensive overlapping homology with the AAV sequences in the vector. Consequently, all the early AAV vector particles, regardless of their rep and cap phenotypes, were contaminated with wild-type AAV derived directly from the complementing plasmid or by recombination between the complementing plasmid and the vector plasmid. There were a number of implications of the presence of active rep genes in the early vector preps as we shall consider below.

We generated an AAV2 vector, AVHiCAT, which had the bacterial gene CAT (chloramphenicol acetyl transferase) inserted at a position immediately in front of the intron, at a convenient single *Hind*III site, so that it was expressed from the AAV p40 promoter [4]. This vector still expressed a functional rep gene but was deleted for about 40% of the AAV genome (between map units 52 and 91) by removal of a *Hinc*II fragment that removed most of the cap gene and thus was rep⁺cap⁻. We also constructed a second vector, AVBcCAT, that retained the cap gene but had the CAT cassette inserted between the p19 and the p40 promoters by substitution of a region of rep so that CAT was expressed from the p19 promoter. Thus, AVBcCAT was a rep⁻cap⁺ vector. The AVHiCAT vector genome was slightly smaller than wt AAV, whereas the AVBcCAT genome was about 3% larger than wt AAV, but both could be packaged into AAV particles by complementation with pAV2, a plasmid containing wild-type AAV2. Infection of these vectors onto fresh 293 cells led to expression of CAT enzyme activity in a short-term transient assay, although it was necessary to add fresh adenovirus to detect CAT expression. The titer of the packaged vectors was low and at the time we assumed that the effect of the adenovirus was to catalyze conversion of the vector genomes to a duplex molecule and also to activate the p19 promoter and the p40 promoter in the rep⁺ vectors [4]. These experiments

clearly showed that AAV vectors could be packaged into AAV particles and used to mediate foreign transgene after infection of cells.

In a similar way, and also using heat-inactivated crude cell lysates, Hermonat and Muzyczka [5] constructed an AAV vector, dl52-91/neo, with the bacterial gene *neo* as a selective marker expressed from an SV40 early region promoter inserted in place of the cap gene 52-91 deletion. Complementation of this vector with a plasmid containing a wild-type infectious AAV was not successful, apparently because wt AAV was preferentially packaged, so a plasmid with a 1-kb fragment of inserted extraneous DNA was used in the hope that this complementing genome would be too large to be packaged. However, at least 20% of their vector preparations were wild-type AAV particles. Nevertheless, these vectors were used to infect several cell lines, including human Detroit 6 and KB cells and thymidine kinase-deficient mouse L cells, and generated colonies that were resistant to the antibiotic Geneticin when the cultures were placed under selective pressure. The dl52-91/neo vector had a genome that was 4.7 kb, approximately the same size as the wt AAV2 genome. An analogous vector, dl52-92/dhfr, which contained the mammalian cell selective marker, dihydrofolate reductase (dhfr) cDNA rather than *neo*, and that had a genome about 6% larger than wt, was also capable of being packaged into particles but gave much lower yields [5]. This experiment went a step further and showed that in the transduced cell colonies derived under selective pressure, the vector was presumably integrated because no low-molecular-weight genomes were observed unless the cells were superinfected with adenovirus, which led to amplified vector genomes of the expected size. Thus, it was concluded that the vector had integrated into chromosomal sites and that rescue was possible because it was rep⁺.

We also generated AAV-*neo* vectors, analogous to AVHiCAT, that were rep⁺cap⁻ and in which the *neo* gene was expressed off the p40 promoter, but we added one more modification by purification of the vector particles on CsCl, in addition to heat inactivation, to help remove the adenovirus particles [6]. Also, this AAV-*neo* vector had a genome of 3.4 kb so it banded in CsCl at a density intermediate between that of adenovirus and wild-type AAV and this helped to reduce, but not eliminate, the amount of wt AAV contamination. This vector gave very similar results to the dl52-91/neo vector in that it yielded Geneticin-resistant colonies in HeLa or 293 cells at a frequency of up to several percentages, and the vector genome was integrated in host cell chromosomes but could be rescued by superinfection with adenovirus or adenovirus plus wt AAV.

Consequently, by the mid-1980s there was definitive proof-of-concept that AAV particles having recombinant genomes could be generated and could transduce mam-

malian cells to express foreign transgenes from either AAV or non-AAV transcription promoters. For the selective marker *neo*, the vector genome was apparently integrated into the cell chromosome but was rescuable by superinfection with adenovirus. However, most of these vectors were rep⁺ and in any case were contaminated with appreciable amounts of wt AAV particles. This led to some ambiguity regarding the mechanism of the integration and rescue of vector genomes and several years later led to a transient, but erroneous, impression regarding the efficiency and specificity of integration by AAV vectors. Also, in other studies of AAV genetics we were beginning to appreciate the properties of Rep in negative and positive regulation of expression of AAV genes and perhaps other cellular genes [41,42].

The presence of appreciable amounts of the wt AAV in the early vector preparations was clearly undesirable. However, the genetic studies had already shown that only the ITR sequences were required *in cis* and that Rep and Cap could both be provided *in trans*. Peri Senapathy and I analyzed generation of infectious AAV by recombination after transfecting adenovirus-infected cells with pairs of plasmids containing two subgenomic regions of AAV and showed that the degree of recombination was proportional to the amount of overlapping homology between the two subgenomic regions [43]. Thus, an alternate approach to generation of AAV vectors with much less wt AAV was to ensure that overlapping homology was minimized, and subsequently AAV vectors began to be generated in two-plasmid transfection systems based on this principle [44,45]. Nevertheless, there was still a need for a producer cell line expressing Rep and Cap, analogous to producer cell lines that were being developed for retrovectors. However, one of the negative (cytostatic) functions of Rep led to difficulty in generating cell lines that constitutively express it, and this delayed generation of producer cell lines for AAV vectors until some 10 years later when Jim Trempe [46] produced cells with an inducible rep gene. Then, Phil Johnson's lab generated cell lines stably containing functional rep and cap genes [47] by taking advantage of an earlier observation by Ella Mendelson [48] that rep⁺ AAV gene cassettes could stably express Rep in HeLa cells, but not in 293 cells, presumably because the p5 promoter was not active in the former cells.

AAV LATENCY REVISITED

Further studies of AAV latency by Cathy Laughlin using wt AAV2 in KB cells [49] showed results analogous to those of Cheung *et al.* [26] seen earlier in the Detroit 6 cells, that the AAV was integrated as head-to-tail concatemers. Similarly, McLaughlin *et al.* [50] using the rep⁺ dl52-91/neo vector (or a rep⁻ version that was contaminated with wt AAV), showed that Geneticin-resistant

clones derived from transduction of Detroit 6 or KB cells had the vector genomes integrated as head-to-tail cassettes. Both groups noted that the persistent head-to-tail concatemers could not have been generated by the productive mode of AAV DNA replication, characterized by Straus and Rose, which gives head-head and tail-tail concatemers. A novel experiment then led to a remarkable insight into integration, but also generated mistaken impressions regarding integration or persistence of AAV vectors.

Rob Kotin and Ken Berns revisited the state of the integrated AAV in the original Detroit 6 clones of Dave Hoggan and succeeded in determining the sequence of the junctions between the ITRs and the host cell DNA [51]. These flanking sequences were then used to screen large number of clones of KB, Detroit 6, or HeLa cells containing either integrated wt AAV or the dl52-92/neo vector that had been generated independently, including those from the Laughlin [49] and Muzyczka [50] labs, as well as additional clones from Jude Samulski at Pittsburgh. Remarkably, most of the integrations had occurred in a defined region (now called AAVS1) on the q arm of human chromosome 19 [52]. The authors of this study [50] were careful to note that the site specificity of the vector integration may have been mediated by the contaminating wt AAV in the vector preps or because they were rep⁺ vectors. Despite this caveat, these studies led to some incorrect impressions by many others that AAV vectors integrate efficiently and site-specifically at AAVS1. However, this was an important caveat and we now know that rep⁻ vectors in the absence of wt AAV integrate with only very low efficiency, if at all, and certainly without specificity for AAVS1 [53]. Indeed, wt AAV in cultured cells integrates at the AAVS1 site but the efficiency and site specificity of the process require the AAV rep gene because it is mediated by binding of the Rep68/78 protein to the ITR and to sequences at AAVS1 [54]. It still remains to be determined if this phenomenon occurs with AAV naturally in its human or animal host. However, we now know from experiments in animals that the more common mode of AAV vector persistence *in vivo* is as head-to-tail concatemers [8] and that in nondividing cells in the absence of selection most of these concatemers remain episomal [55].

The episomal persistence of rep⁻ vectors is interesting for two additional reasons, viewed from a historical perspective. The head-to-tail arrangement indicates that they are not generated by the productive mode of AAV DNA replication. Also, the early studies of integrated wt AAV or *neo*-expressing vectors (probably contaminated with wt AAV) showed that the integrated genomes were predominantly head-to-tail concatemers. Thus, in the absence of a helper virus, the AAV ITRs seem to drive concatemerization and circularization to persist episodically but the presence of a functional rep gene (or

Rep68/78 protein) can drive these structures to integrate into AAVS1.

ADDITIONAL REVISITATION OF HISTORY

The misunderstandings of AAV vector integration were not the only misconceptions that were generated about AAV vectors, particularly in the early 1990s, as many additional investigators began studies of AAV vectors. We had noted in our earliest AAV vector study that conversion of the incoming single-stranded vector genome may be limiting for expression and thus added adenovirus to observe expression of the CAT enzyme activity. In the mid-1990s, several groups revisited this and looked for ways to enhance the process and this led some to suggest that AAV did not transduce cells unless the cells were actively replicating. This was thought to reflect the need for actively dividing cells to convert the single-strand genome to a transcribable duplex template. However, it turns out that this is yet another example showing that gene delivery vectorology must be an *in vivo* science and that *in vitro* studies on cultured cells may be non-predictive and misleading. The first *in vivo* studies of AAV vectors were not reported until 1993 [56], but from those and a multitude of additional studies it is now abundantly clear that the most preferred substrates for transduction by AAV vectors *in vivo* are terminally differentiated or nondividing cells [8]. Apparently, in such cells, the rep⁻ AAV vector is afforded time to traffic to the nucleus, uncoat, and form duplex concatemeric episomes. In a rapidly dividing cell this process may occur too slowly to avoid dilution of the vector genome. Thus, in dividing cells a rep⁺ vector may be able to integrate, presumably at AAVS1, and thus be retained through multiple cell divisions as was observed in the first AAV vector studies.

The enormously increased interest in AAV vectors in the past decade has led to greatly expanded developments, including use of dual vectors to increase the packaging limit, self-complementary vectors to overcome the single- to double-strand conversion, and the use of capsids of alternate serotypes, or modifications of the capsid sequence, to enhance interaction with various cellular receptors and cellular trafficking. This is facilitating a more sophisticated approach to enhance targeting of specific cell types or organs for the development of AAV vectors as therapeutics. In addition, there has been a vast improvement in methods for producing AAV vectors and, for instance, cumbersome CsCl gradients are now a thing of the past.

Overall, I believe that AAV vectorology has now reached a satisfying stage. During the evolution of many fields after the initial developments, often there is an enormous expansion of observations that in the

short term lead to some degree of inconsistency and confusion. This is then followed by a third phase in which sufficient new information is obtained that a more coherent pattern emerges and the field can thus advance rapidly, but in a more rational fashion. I believe that we have reached this third phase with AAV vectors.

RECEIVED FOR PUBLICATION AUGUST 18, 2004; ACCEPTED SEPTEMBER 14, 2004.

REFERENCES

- Atchison, R. W., Casto, B. C., and Hammon, M. (1965). Adenovirus-associated defective virus particles. *Science* **49**: 754–756.
- Hoggan, M. D., Blacklow, N. R., and Rowe, W. P. (1966). Studies of small DNA viruses found in various adenovirus preparations: physical, biological and immunological characteristics. *Proc. Natl. Acad. Sci. USA* **55**: 1467–1472.
- Blacklow, N. R. (1988). Adeno-associated viruses of humans. In *Parvoviruses and Human Disease* (J. R. Pattison, Ed.), pp. 165–174. CRC Press, Boca Raton, FL.
- Tratschin, J. D., West, M. H. P., Sandbank, T., and Carter, B. J. (1984). A human parvovirus, adeno-associated virus as a eukaryotic vector: transient expression and encapsidation of the prokaryotic gene for chloramphenicol acetyl transferase. *Mol. Cell. Biol.* **4**: 2072–2081.
- Hermonat, P. L., and Muzyczka, N. (1984). Use of adeno-associated virus as a mammalian cDNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells. *Proc. Natl. Acad. Sci. USA* **81**: 6466–6470.
- Tratschin, J. D., Miller, I. L., Smith, M. G., and Carter, B. J. (1985). Adeno-associated virus vectors for high frequency integration, expression and rescue of genes in mammalian cells. *Mol. Cell. Biol.* **5**: 3251–3260.
- Flotte, T. R., et al. (1996). A phase I study of an adeno-associated virus CFTR gene vector in adult CF patients with mild lung disease. *Hum. Gene Ther.* **7**: 1145–1159.
- Carter, B. J., Burstein, H. B., and Peluso, R. W. (2004). AAV vectors for gene therapy. In *Gene and Cell Therapy: Therapeutic Mechanisms and Strategies*, 2nd ed. (N. Templeton-Smith, Ed.), pp. 53–101. Dekker, New York.
- Crawford, L. V., Follett, E. A. C., Burdon, M. G., and McGeoch, D. G. (1969). The DNA of a minute virus of mice. *J. Gen. Virol.* **4**: 37–48.
- Rose, J. A., Berns, K. I., Hoggan, M. D., and Kocot, F. J. (1969). Evidence for a single-stranded adenovirus-associated virus genome: formation of a DNA density hybrid upon release of viral DNA. *Proc. Natl. Acad. Sci. USA* **64**: 863–869.
- Carter, B. J., Khoury, G., and Denhardt, D. T. (1975). Physical map and strand polarity of specific fragments of adeno-associated virus DNA produced by endonuclease R.EcoRI. *J. Virol.* **16**: 559–568.
- Carter, B. J., Khoury, G., and Rose, J. A. (1972). Adeno-associated virus multiplication: extent of transcription of the viral genome *in vivo*. *J. Virol.* **10**: 1118–1125.
- Kocot, F. J., Carter, B. J., Garon, C. F., and Rose, J. A. (1973). Self-complementarity of terminal sequences within plus or minus strands of adeno-associated virus DNA. *Proc. Natl. Acad. Sci. USA* **70**: 215–219.
- Gerry, H. W., Kelly, T. J., and Berns, K. I. (1973). Arrangement of nucleotide sequences in adeno-associated virus DNA. *J. Mol. Biol.* **79**: 207–225.
- Berns, K. I., and Kelly, T. J. (1974). Visualization of the inverted terminal repetition in adeno-associated virus DNA. *J. Mol. Biol.* **82**: 267–271.
- Carter, B. J., Fife, K. H., de la Maza, L. M., and Berns, K. I. (1976). Genome localization of adeno-associated virus RNA. *J. Virol.* **19**: 1044–1053.
- Denhardt, D. T., Eisenberg, S., Bartok, K., and Carter, B. J. (1976). Multiple structures of adeno-associated virus DNA: analysis of terminally labeled molecules with endonuclease R.HaeIII. *J. Virol.* **18**: 672–684.
- Straus, S. E., Sebring, E. D., and Rose, J. A. (1976). Concatemers of alternating plus and minus strands are intermediates in adeno-associated DNA synthesis. *Proc. Natl. Acad. Sci. USA* **73**: 742–746.
- Hauswirth, W. W., and Berns, K. I. (1977). Origin and termination of adeno-associated virus DNA replication. *Virology* **93**: 57–68.
- de la Maza, L. M., and Carter, B. J. (1980). Molecular structure of adeno-associated virus variant DNA. *J. Biol. Chem.* **255**: 3194–3203.
- Lusby, E., Fife, K. H., and Berns, K. I. (1980). Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. *J. Virol.* **34**: 402–409.
- Cavalier-Smith, T. (1974). Palindromic base sequences and replication of eukaryotic chromosome ends. *Nature (London)* **250**: 467–470.
- Mendelson, E., Trempe, J. P., and Carter, B. J. (1986). Identification of *trans*-acting rep proteins of adeno-associated virus by antibodies to a synthetic oligopeptide. *J. Virol.* **60**: 823–832.
- Im, D. S., and Muzyczka, N. (1992). Partial purification of adeno-associated virus Rep78, Rep68, Rep52, and Rep40 and their biochemical characterization. *J. Virol.* **66**: 1119–1128.
- Hoggan, M. D., Thomas, G. F., and Johnson, F. B. (1972). Continuous carriage of adenovirus-associated virus genome in cell culture in the absence of helper adenovirus. In *Proceedings of the Fourth Lepetit Colloquium*, pp. 41–47. North-Holland, Amsterdam.
- Cheung, A. K. M., Hoggan, M. D., Hauswirth, W. W., and Berns, K. I. (1980). Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. *J. Virol.* **33**: 739–743.
- Myers, M. W., and Carter, B. J. (1980). Adeno-associated virus assembly. *Virology* **102**: 71–82.
- Laughlin, C. A., Westphal, H., and Carter, B. J. (1979). Spliced adeno-associated virus RNA. *Proc. Natl. Acad. Sci. USA* **76**: 5567–5571.
- Marcus, C. J., Laughlin, C. A., and Carter, B. J. (1981). Adeno-associated virus RNA transcription *in vivo*. *Eur. J. Biochem.* **121**: 147–154.
- Jay, F. T., Laughlin, C. A., and Carter, B. J. (1981). Eukaryotic translational control: adeno-associated virus protein synthesis is affected by a mutation in the adenovirus DNA-binding protein. *Proc. Natl. Acad. Sci. USA* **78**: 2927–2931.
- Laughlin, C. A., Myers, M. W., Risin, D. L., and Carter, B. J. (1979). Defective-interfering particles of the human parvovirus, adeno-associated virus. *Virology* **94**: 162–174.
- Carter, B. J. (1990). Adeno-associated virus helper functions. In *Handbook of Parvoviruses* (P. Tijssen, Ed.), Vol. 1, pp. 225–282. CRC Press, Boca Raton, FL.
- Samulski, R. J., Berns, K. I., Tan, N., and Muzyczka, N. (1982). Cloning of infectious adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. *Proc. Natl. Acad. Sci. USA* **79**: 2077–2081.
- Laughlin, C. A., Tratschin, J. D., Coon, H., and Carter, B. J. (1983). Cloning of infectious adeno-associated virus genomes in bacterial plasmids. *Gene* **23**: 65–73.
- Hoggan, M. D., Shatkin, A. J., Blacklow, N. R., Kocot, F. J., and Rose, J. A. (1968). Helper-dependent infectious deoxyribonucleic acid from adeno-associated virus. *J. Virol.* **2**: 850–852.
- Samulski, R. J., Srivastava, A., Berns, K. I., and Muzyczka, N. (1983). Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats. *Cell* **33**: 135–143.
- Senapathy, P., Tratschin, J. D., and Carter, B. J. (1984). Replication of adeno-associated virus DNA: complementation of naturally-occurring *rep* mutants by a wild-type genome or an *ori* mutant and correction of terminal palindrome deletions. *J. Mol. Biol.* **179**: 1–20.
- Srivastava, A., Lusby, E. W., and Berns, K. I. (1983). Nucleotide sequence and organization of adeno-associated virus 2 genome. *J. Virol.* **45**: 555–564.
- Hermonat, P. L., Labow, M., Wright, R., Berns, K. I., and Muzyczka, N. (1984). Genetics of adeno-associated virus: isolation and preliminary characterization of adeno-associated virus type 2 mutants. *J. Virol.* **51**: 329–339.
- Tratschin, J. D., Miller, I. L., and Carter, B. J. (1984). Genetic analysis of adeno-associated virus: properties of deletion mutants constructed *in vitro* and evidence for an adeno-associated virus replication function. *J. Virol.* **51**: 611–619.
- Tratschin, J. D., Tal, J., and Carter, B. J. (1986). Negative and positive regulation *in trans* of gene expression from adeno-associated virus vectors in mammalian cells by a viral *rep* gene product. *Mol. Cell. Biol.* **6**: 2884–2894.
- Labow, M. A., Graf, L. H., and Berns, K. I. (1987). Adeno-associated virus gene expression inhibits cellular transformation by heterologous genes. *Mol. Cell. Biol.* **7**: 1320–1325.
- Senapathy, P., and Carter, B. J. (1984). Molecular cloning of adeno-associated virus variant genomes and generation of infectious virus by recombination in mammalian cells. *J. Biol. Chem.* **259**: 4661–4666.
- Samulski, R. J., Chang, L. S., and Shenk, T. E. (1989). Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *J. Virol.* **63**: 3822–3828.
- Flotte, T. R., Barraza-Ortiz, X., Solow, R., Afione, S. A., Carter, B. J., and Guggino, W. B. (1995). An improved system for packaging recombinant adeno-associated virus vectors capable of *in vivo* transduction. *Gene Ther.* **2**: 39–47.
- Yang, Q., Chen, F., and Trempe, J. P. (1994). Characterization of cell lines that inducibly express the adeno-associated virus Rep proteins. *J. Virol.* **68**: 4847–4856.
- Clark, K. R., Voulgaropoulou, F., Fraley, D. M., and Johnson, P. R. (1995). Cell lines for the production of recombinant adeno-associated virus. *Hum. Gene Ther.* **6**: 1329–1341.
- Mendelson, E., Smith, M. G., Miller, I. L., and Carter, B. J. (1988). Effect of a viral rep gene on transformation of cells by an adeno-associated virus vector. *Virology* **166**: 612–615.
- Laughlin, C. A., Cardellicchio, C. B., and Coon, H. G. (1986). Latent infection of KB cells with adeno-associated virus type 2. *J. Virol.* **60**: 515–524.
- McLaughlin, S. K., Collis, P., Hermonat, P. L., and Muzyczka, N. (1988). Adeno-associated virus general transduction vectors: analysis of proviral structure. *J. Virol.* **62**: 1963–1973.
- Kotin, R. M., and Berns, K. I. (1989). Organization of adeno-associated virus DNA in latently infected Detroit 6 cells. *Virology* **170**: 460–467.
- Kotin, R. M., et al. (1990). Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. USA* **87**: 2211–2215.

53. Kearns, W. G., *et al.* (1996). Recombinant adeno-associated virus (AAV-CFTR) vectors do not integrate in a site-specific fashion in an immortalized epithelial cell line. *Gene Ther.* **3**: 748–755.
54. Weitzman, M. D., Kyostio, S. R. M., Kotin, R. M., and Owens, R. A. (1994). Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. *Proc. Natl. Acad. Sci. USA* **91**: 5808–5812.
55. Schnepf, B. C., Clark, K. R., Klemanski, D. L., Pacak, C. A., and Johnson, P. R. (2003). Genetic fate of recombinant adeno-associated virus vector genomes in muscle. *J. Virol.* **77**: 3495–3504.
56. Flotte, T. R., *et al.* (1993). *In vivo* delivery of adeno-associated vectors expressing the cystic fibrosis transmembrane conductance regulator to the airway epithelium. *Proc. Natl. Acad. Sci. USA* **93**: 10163–10167.