A publication of
The Williamsburg BioProcessing Foundation

September/October 2003

BioProcessing™ JOURNAL

Advances & Trends In Biological Product Development

Defining a Detailed Approach to Using the Adenovirus Reference Material (ARM)

By Nancy Sajjadi
and Janice Callahan

Please see important revisions on page 86.

Vol. 2 No. 5

www.bioprocessingjournal.com

Defining a Detailed Approach to Using the Adenovirus Reference Material (ARM)

BY NANCY SAJJADI and JANICE CALLAHAN

Through the tremendous efforts of the Adenovirus Reference Material Working Group (ARMWG), an adenovirus reference material (ARM) is now available from the American Type Culture Collection (ATCC). The history and progress of the ARM production and characterization has been presented at many meetings and published in numerous journal articles.[1–8] Although general statements have been made regarding how the ARM should be used, there is no formal directive or specific set of instructions detailing its application in the field. The goals of this paper and its related presentation at The Williamsburg BioProcessing Foundation's Viral Vectors & Vaccines conference (November 10–13, 2003, Las Vegas, NV) are (1) to briefly review the objectives for development and implementation of the ARM, (2) to describe a critical assumption necessary to meet those objectives, (3) to outline specific approaches for using the ARM, and (4) to highlight the need for a working group to address the issues raised in the process.

Adenoviral vectors hold the promise of exciting therapeutic applications in many different diseases. However, the tragic death of Jesse Gelsinger in 1999 reminded us that their safety profile is not completely understood. Adenoviral vectors as a class of gene delivery vehicles may have a dose-related toxicity profile, although the mechanism of action, the route of administration, and other factors may influence the toxicity profile of an individual vector. Because adenoviral vector doses are measured and defined by particle number, there must be consistent measurement of particle number to allow for comparison of data across preclinical and clinical studies. And while toxicity may be related to the dose of intact particles, the quality of an adenoviral vector preparation in terms of contamination with replication competent adenovirus (RCA) and the total load of adenovirus related material (e.g. incomplete particles, free hexon protein) must also be appropriately characterized as their contribution to toxicity is yet to be fully understood. Thus, the primary objective in developing the ARM was to facilitate relating the results for various vectors analyzed using different methods to evaluate their quality.

FDA has outlined its rationale and role in developing the ARM and their current expectations for its use in the field.[2–4] The overall goal is to allow for comparability between studies using adenoviral vectors and ultimately to develop regulatory policy based on "standardized" measurements. To achieve this goal, FDA "recommends that sponsors use the ARM to validate their assays while using their own internal reference material" and they expect that "an internal standard, which has been calibrated against the ARM, will be used in each titer and particle assay."[4] Although they have not precisely specified how this is to be accomplished, they have stated that "it is not expected that others will duplicate the ARM titer or particle values" but "particle values are nevertheless, expected to fall easily within the range of the data collected during the ARM characterization."[4] They have also indicated that they will not specify a required method to be used in production or testing at this time. That said, let's first examine a critical assumption that is necessary when using the ARM as a calibrator.

A basic assumption in relating test data from various vectors using different analytical methods is that the contribution of the vector component being assayed, relative to the true concentration of the vector component that is being evaluated, is equivalent (or highly similar) for the vectors being compared. The meaning of this statement can perhaps be best understood by example.

When determining dose for adenoviral vectors, a particle number is measured. One definition of the term "particle number" is the number of intact particles capable of delivering a therapeutic activity. However, depending on the method used, the "particle number" unit could actually mean something different. Consider the pos-

Nancy Sajjadi, MSc, is Founder and Principal Consultant, Sajjadi Consulting (sajjadiconsulting@gmail).
Janice Callahan, PhD, is President, Callahan Associates Inc., La Jolla, CA.

Table 1: Possible Components of Adenoviral Vector Preparations

1. **Adenovirus-related material**
    a. Intact particles (individual or aggregated)
        i. Infectious vector particle
        ii. Non-infectious vector particle
            1. Inactivated vector
            2. Intact non-infectious molecular variant
        iii. Infectious RCA
        iv. Non-infectious non-vector particle
            1. Inactivated RCA
            2. Intact non-infectious molecular variant of RCA
    b. Incomplete particles
    c. Non-particle-associated material
        i. Free viral proteins
        ii. Viral nucleic acid
2. **Material not related to adenovirus**
    a. Formulation components
    b. Process-related contaminants (e.g. enzymes, FBS)
    c. Residual host cell nucleic acid (RNA, DNA)
    d. Residual host cell proteins, lipids, and metabolites

sible components in a vector preparation (Table 1) and the relative contribution that each makes to the overall composition. Then consider that a given assay will measure a subset of those components as a measure of a particular vector attribute.

For example, the particle number determined by an A_{260nm}-based method estimates a particle concentration based on the total DNA content of the adenoviral vector material. The particle number calculated will be affected by any contribution made to the absorbance of the material at 260nm. The ability of a contaminant to contribute to an A_{260nm} reading was in fact demonstrated for the ARM itself.[9] If two preparations differ with respect to their formulation, level of contaminating DNA, or other factors influencing the assay outcome, then the "particle number units" are not truly comparable. So, while it may be possible to show that two different A260-based methods have the same accuracy and precision as demonstrated by the equivalence of results they yield in analysis of the ARM, the ability to relate two different vector preparations using one method depends on how similar the vectors are. No two preparations of the same vector are identical and certainly no two preparations of different vectors are likely to be identical.

One way to overcome the inherent challenge in relating dose measurements of different vectors would be to specify a method or at least a method class. Interestingly, this was the conclusion reached by Mittereder *et al* in their assessment of several physical and biological methods to quantify adenoviral vectors.[10] In choosing a method or method class, selectivity of the method for the component of interest is paramount because the real desire is to relate the values for a component of interest rather than a surrogate measure of it. For determination of particle number, anion exchange HPLC (AE-HPLC) may be a good option.

Several different AE-HPLC methods for evaluating adenoviruses have been described and all have the distinct advantage of being able to selectively identify and quantify intact virus particles, simultaneously providing an overall purity profile.[10–12] Also, in a recent study conducted to characterize the ARM by different physical-chemical methods, AE-HPLC methods were shown to be very precise measures of particle concentration.[8] Given the inherent challenge associated with relating products of different compositions, the proposal to identify a method or method class for determining particle concentration warrants further discussion. However, evolution toward standardizing a method may be slow and thus the current focus is on using the ARM as a calibrator. In the remainder of this paper, specific approaches to using the ARM are outlined and discussed in terms of the vector attribute to be measured. Bear in mind that, in the end, it is the similarity between the vectors themselves and their relationship to the reference materials that will determine the validity with which assay results can be related.

Measurement of Particle Concentration

Methods of determining particle concentration fall into two groups — absolute methods and relative methods. Absolute methods calculate particle concentration directly from the assay readout for the sample whereas relative methods calculate particle concentration by interpolation of the test sample readout from a standard curve. The way in which the ARM is used as a "calibrator" will depend on whether the method being evaluated is an absolute or a relative method.

For an absolute method, the ARM can be used to compare the accuracy and precision of a chosen method to the ARMWG A_{260nm}/SDS method used to assign the label particle concentration for ARM. If an absolute method does not meet the requirements for accuracy and precision, the method should be subjected to further development and optimized to meet the requirements. This is a calibration approach where the ARM is the "standard" and the method is the "instrument." In other words, if the results for the ARM are not within tolerance, the method ("instrument") must be calibrated. The question then becomes, "How is the tolerance range defined for a method?"

FDA has stated its expectation that the particle (P) values "fall easily within the range of the data collected during the ARM characterization."[4] It is important to clearly define what this statement means from a statistical perspective. For example, Lehmberg *et al.* compared the mean results obtained for the ARM (n=6) in a number of methods, to the range of 4.8–6.9 x 10^{11} P/ml established for the ARM mean value of 5.8 x 10^{11} P/ml (n=14) using three times the pooled standard deviation (3.42 x 10^{10} P/ml).[6,8] However, the range of 4.8–6.9 x 10^{11} P/ml represents a prediction interval, which means that it should contain about 99% of single future observations (where one observation is the average of three replicates at each of four dilutions) for the ARM measured using the ARMWG A_{260nm}/SDS method.

It is not statistically valid to use this range as a "tolerance" range for comparing another absolute method unless the following are true: (1) the results being compared represent single observations from the method being calibrated, (2) an observation represents a mean calculated from three replicates at each of four dilutions, and (3) the standard deviation of the method being calibrated is close to the pooled standard deviation (3.42 x 10^{10} P/ml) for the ARMWG A_{260nm}/SDS method. If a measurement met all of these criteria and fell within the range, the method being evaluated could be considered within calibration and acceptable for use. But how does one evaluate whether these criteria are being met? How many single observations need to be evaluated? How should the standard deviation of the method be measured? How does one determine whether or not the standard deviation for the method being evaluated is considered similar enough to the pooled standard deviation for the ARMWG A_{260nm}/SDS method?

The answers to these questions (and some others that will be raised in the process) is somewhat complicated and cannot be addressed within the scope of this paper. But clearly, these questions must be answered if a calibration approach is to be used and, thus, they warrant immediate further discussion. Regardless of the eventual outcome, it is important to note that absolute methods other than A_{260nm}-based methods (e.g. electron microscopy) are unlikely to yield results comparable to the ARMWG A_{260nm}/SDS method either because they define particle number based on some other component and/or because the methods have different accuracy and precision. This will probably mean that particle determinations by an absolute method will, by default, be limited to A_{260nm}-based methods. However, particle number may also be determined using a relative method.

A recent publication by Lehmberg *et al* summarizes the analysis of the ARM by several laboratories using four different types of relative methods.[8] The study showed that the mean particle concentration values for one vector preparation (the ARM) using relative methods with different internal standards ranges from 3.4 to 8.1 x 10^{11} P/ml (compared to the 5.8 x 10^{11} P/ml assigned to the ARM). The variability in the accuracy of the methods is reflective of the differences between the ARM and the internal reference materials and the methods used to assign their respective particle concentrations. The accuracy of a relative method for measuring the ARM can be drastically improved (i.e. approach 100%) by assigning the particle concentration for the internal reference material using the ARM.

Specifically, a dilution series of both the ARM and the internal reference material should be tested in the chosen relative method. The assay readout for ARM is to be plotted as a function of the amount (particle concentration) to create a standard curve. For the internal reference material, plotting the assay readout as a function of the input amount will generate a response curve. It is important to verify that the internal reference material response curve is parallel to the ARM standard curve.

If the curves are parallel, a particle concentration for the internal reference material can be calculated by interpolation from the ARM standard curve. If the curves are not parallel, then the ARM and the vector must be considered significantly different in composition or in some manner that affects the kinetics of the assay readout response and as such should not be related. Similarly, test samples should be shown to be parallel to internal reference materials before they are used. Tying the particle concentration of an internal reference material to the ARM in this manner allows for the retrospective analysis of test samples measured using an internal reference standard previously assigned a particle concentration based on another method.

This approach can be applied even when the internal reference material is not an adenovirus. For example, lambda genomic DNA used as a standard for a DNA dye binding method can be assigned ARM particle/ml "equivalent" units by testing a dilution series of the

ARM in the same way that a sample would be assayed in the method and generating a standard curve. The lambda DNA would be tested as a sample and the concentration, in particle/ml equivalent units, determined.

Measurement of Infectious Titer

Using the ARM to evaluate and compare methods for determining infectious titer may prove to be the most difficult application and, as such, warrants some preliminary discussion examining why it should be done. Measurements of infectious titer are one way of demonstrating that physical particles have biological activity. Historically, virologists have related infectious virus concentration to physical virus concentration. Infectious virus concentrations are almost always significantly lower than physical virus concentrations. While in some instances this may indicate that some of the physical virus particles are "defective" and unable to produce an infection in culture, more often than not, the discrepancy in results is due to the low efficiency of infectious titer assays in detecting infectious particles.[10,14] Because dosing of adenoviral vectors is currently based on particle number determinations and there is a legitimate concern that vector lots may vary in the relative amount of defective particles, relating a particle concentration to an infectious titer (PN:IU ratio) can be used to evaluate lot consistency. However, FDA currently has established a PN:IU ratio of 30:1 for adenoviral vectors. While clearly many vector products may meet this specification, a single specification for different vector and infectious titer assay combinations may not be the best approach.

It is clear that an infectious titer value is not a true measure of the actual infectious particle concentration. The infectious titer value for a given vector preparation is highly dependent on the assay conditions used to test it and on the method used to calculate the concentration. While specific culture conditions and the kinetics of vector infection can influence the titer value obtained, the major limitation in detecting infectious particles is that most particles do not reach a target cell.[14] It should therefore be expected that titer values for different vectors and different assays vary and that comparing PN:IU values for different vectors is not informative because the infectious titer value is highly relative. A better approach may be to define a PN:IU specification for each vector. Initially the sponsor would establish this specification based on the thorough characterization of a few vector lots that meet other quality requirements. It may be necessary to adopt this approach because of the complexities involved in demonstrating the similarity of the ARM to a particular adenoviral vector such that the infectious titer in a given method can be normalized to the assigned ARM infectious titer.

The infectious titer of the ARM was measured in a defined 293 cell-based assay and calculated using a maximum likelihood estimator.[6] This calculation also included an adjustment based on a diffusion model (NAS IU)—for every infection event that was *measured*, a total of 273 infectious particles were *counted*. This adjustment value was based on 1 mL of virus assayed. Because the infectious titer assay used 0.2 mL of ARM, the correct adjustment factor is 54.6 (273×0.2). Thus, the calculated PN:NAS IU ratio of 8.3 for the ARM is equivalent to a PN:IU ratio of 453 (8.3×54.6) if the same data are not modeled for diffusion. If methods used in the field do not incorporate diffusion modeling into the calculation of infectious titer, the titer of the ARM should be expected to be 2–3 orders of magnitude lower, respectively, than the ATCC label values for the ARM material (7.3×10^{10}) NAS IU/mL assigned value. By reducing the infectious titer (*i.e.*, not adjusting to account for particles that were present but did not yield an infection event), the IU value is reduced by a factor of 54.6 and the PN:IU value increased by that same factor. How is one expected to deal with such results?

One possibility would be to consider the method unacceptable. However, it should be clear that an infectious titer method simply yields a relative number, not a true number, so judging a given method on its accuracy does not make sense. In fact, in incorporating an adjustment to account for the limitations in particle diffusion into the infectious titer calculation for the ARM, there is an implicit recognition that the "raw" infectious titer assay results do not reflect the true concentration of infectious particles. This appears to have been the first step in assessing the validity of applying a universal PN:IU limit to all adenoviral vectors. While this evaluation should continue as the ARM is implemented in the field, the possibility of normalizing test results for infectious titer must also be explored.

To normalize infectious titer results for a vector, the measured results of the test sample are multiplied by the ratio of measured to assigned titer values for the ARM in each assay. This approach is very similar to determining a relative potency for a product — each lot of product is assayed with, and measured in relation to, a reference material with a defined 100% potency. To be valid, the same types of restrictions placed on relative potency assays should apply. Namely, the vector and the ARM should trend together if there is large inter-assay variability and their dose response curves must be parallel (and have equivalent upper and lower asymptotes if an end point method is used) in the method. This assumption may not hold true for certain types of vectors and thus should be demonstrated experimentally to be a valid assumption before results are normalized in this manner. Similarly, the parallel relationship between the ARM and any internal reference material to be used routinely in the assay must clearly be established prior to its implementation.

Detection of Replication Competent Adenovirus

Currently most methods for detection of replication competent adenovirus yield only positive or negative result outcomes although some types of assays may also allow for estimating the amount of any RCA detected. To ensure that adenoviral vectors being tested in the clinic consistently meet the current requirement of less than 1 RCA / 3 x 10^{10} particles, the following need to occur: (1) a minimum sampling requirement (based on particle number) for adenoviral vectors must be defined, (2) the methods used to detect an RCA must have an acceptable limit of detection, and (3) an RCA "unit" must be defined.

To achieve equivalent statistical confidence in detecting an RCA, the probability of sampling an RCA must be equivalent across vector testing. If the minimum sampling requirement is defined as 3 x 10^9 particles, there is a 63% probability of sampling at least one RCA if it were present at a concentration of 1 RCA / 3 x 10^9 particles. To increase the probability of sampling at least one RCA to 95%, a total of 3 x 10^{10} particles would need to be tested. Once the probability of sampling is equivalent, the probability of detection can be considered equivalent if the RCA screening methods have comparable or equivalent limits of detection. Evaluating a method's limit of detection can be addressed using the ARM.

To determine the limit of detection of an RCA screening assay, diluted ARM must be spiked into a representative "test article" and tested. Specifically, the assay input amount must be defined for the test article as the maximum amount of adenoviral vector that will be applied to an individual test vessel in the assay. An individual test vessel is defined as the smallest individual container to which the test article sample will be applied. For example, 1 x 10^9 particles may be applied to a T-225 flask containing a cell line or 1 x 10^6 particles may be applied to the well of a 96-well plate for PCR analysis. Diluted ARM material must be spiked into the assay input amount, making multiple replicates to be tested at each dilution. The proportion of individual test vessels that yields a positive score is then determined for each ARM dilution spike level. The limit of detection will be determined based on the highest dilution of ARM that consistently yields positive results. To define an *acceptable* limit of detection, an RCA unit must be defined.

Currently the ARM has two units assigned — a particle concentration and an infectious titer value. The assigned particle concentration is 5.8 x 10^{11} particles/ml and the infectious titer is 7 x 10^{10} NAS IU/ml.[6] It should be clear that an RCA screening method could not be expected to detect one particle of ARM. Similarly, an RCA screening method may not detect one infectious unit of ARM because most RCA screening methods do not model the negative result outcomes to account for limitations in particle diffusion. So what are the units of an RCA?

One approach to answering this question and defining an acceptable limit of detection is to choose a single RCA screening method as a reference method. By defining a single method that is considered acceptable and demonstrating a limit of detection by assaying dilutions of ARM (not spiked into vector), a reference point will be established and, in essence, a "unit" of RCA will be defined. All other methods can then be compared for equivalence using statistically valid experimental designs. Note that designing and assigning a label concentration for an internal reference material to be used in an RCA screening assay will also require that an RCA unit is first defined in some manner. Again, it is important to remember that the limit of detection of a given assay will depend on the replication kinetics of the virus being detected. Viruses or vectors other than Ad5wt may be detected with higher or lower efficiencies.

Conclusion

When representatives from academia, industry, standard-setting organizations, and others met at a special Williamsburg BioProcessing Foundation meeting in October 2000, there was consensus that having a well-characterized Ad5Wt reference material was better than no reference material at all. Presumably, there would be consensus among that same group that using the ARM in a manner that is not scientifically or statistically valid is really no different (or possibly even worse!) than not having a reference material available. Now that the characterization of the first available lot of ARM is nearly complete, the time for developing consensus on its use is at hand. This paper has outlined some specific approaches for the use of the ARM and identified several issues that remain to be addressed. Given the tremendous success of the collaborative efforts that brought about the existence of the ARM, a new working group should be established to develop and publish specific approaches for appropriate use of the ARM.

REFERENCES

1. Hutchins B, Sajjadi N, Seaver S, Shepard A, Bauer S, Simek S, Carson K, Aguilar-Cordoba E. Working toward an adenoviral vector standard. *Mol Ther* 2000:2(6);532–534.
2. Bauer, S. FDA recommendations for adenoviral vector characterization. *Presented at the Adenoviral Vector Safety Public Meeting and Workshop*. 2001 February 1; Washington, DC.
3. Simek, S. FDA perspective on the development of an adenoviral standard. *Presented at the Adenoviral Vector Safety Public Meeting and Workshop*. 2001 February 1; Washington, DC.
4. Simek S, Byrnes A, Bauer S. FDA perspectives on the use of the adenovirus reference material. *BioProcessing J* 2002:1(3);40–42.
5. Hutchins B. Development of a reference material for characterizing adenovirus vectors. *BioProcessing J* 2002:1(1);25–28.
6. Callahan J. A statistical analysis of adenovirus reference material assay results. *BioProcessing Journal* 2002:1(3);43–47.
7. Carson K. Reference materials for complex biological products. *BioProcessing Journal* 2003:2(1); 17–21.
8. Lehmberg E, Shabram P, Schluep T, Wen S, Sugarman B, Croyle M, Koehl M, Bonfils E, Marlame D, Sharpe G, Nesbit H, Borellini F, Kamen A, Hutchins B, Vellenkamp G. Adenovirus Reference Material: Determination of particle concentrations obtained by orthogonal, physical-chemical methods. *BioProcessing J* 2003:2(3);50–55.
9. Vellenkamp G, Ravindran S, Lin M, Sluzky V, Lehmberg E. A contaminant in the adenovirus material. *BioProcessing J* 2002:1(3);57–61.
10. Mittereder N, March K, Trapnell B. Evaluation of the concentration and bioactivity of adenovirus vectors for gene therapy. *J Virol* 1996:70(11); 7498–7509.
11. Shabram P, Giroux D, Gordeau A, Gregory R, Horn M, Huyghe B, Liu X, Nunally M, Sugarman B, Sutjipto S. Analytical anion-exchange HPLC of recombinant type-5 adenoviral particles. *Hum Gene Ther* 1997:8; 453–465.
12. Blanche F, Cameron B, Barbot A, Ferrero L, Guillemin T, Guyot S, Somarriba S, Bisch D. An improved anion-exchange HPLC method for the detection and purification of adenoviral particles. *Gene Therapy* 2000:7;1055–1062.
13. Klyushnichenko V, Bernier A, Kamen A, Harmsen E. Improved high-performance liquid chromatographic method in the analysis of adenovirus particles. *J Chromotography B* 2001:755;27–36.
14. Nyberg-Hoffman C, Shabram P, Li W, Giroux D, Aguilar-Cordoba E. Sensitivity and reproducibility in adenoviral titer determination. *Nature Medicine* 1997:3(7);808–811.