

## Vector biophysics

**Crystal-clear view**

JE Rabinowitz and RJ Samulski

Gene Therapy (2002) 9, 1413–1414. doi:10.1038/sj.gt.3301884

The best way to understand the specificity and infectivity of gene therapy vectors is through their atomic structure. However for many vectors this basic blueprint is unavailable. Michael Chapman, with the help of his team at Florida State University has now determined the structure of a representative of one of the major groups of gene therapy vectors<sup>1</sup> (Figure 1).

Adeno-associated virus serotype 2 (AAV2), the subject of the new study, is a small nonpathogenic parvovirus requiring a helper virus to complete its lytic life cycle. AAV2 and other AAV serotypes are promising candidates as vehicles for therapeutic gene transfer. Recombinant (r) AAV vectors have been used to transduce numerous tissue types in many animal species resulting in long-term gene expression without severe immune consequences.<sup>2</sup>

Various types of mutagenic approaches have previously been used to genetically modify AAV2 in order to alter specificity and infectivity: random mutagenesis,<sup>3</sup> alanine scanning mutagenesis<sup>4</sup> and site-directed insertions based on the previously determined<sup>5</sup> crystal structure of the distantly related Canine parvovirus (CPV).<sup>6,7</sup> Yet it is a good bet that data collected using these approaches and others<sup>8,9</sup> have only just scratched the surface compared with what is now possible using rational modifications based on the crystal struc-

ture of AAV2. This crystal structure is the Rosetta Stone that will allow the existing mutagenesis data to be deciphered and that should direct all future genetic modifications of AAV2 and the other serotypes.

One of the challenges that Chapman's group faced in determining the structure was obtaining sufficient virus free of adenovirus helper. This was no small task since 2–5 mg of virus (equivalent to  $1-2 \times 10^{15}$  particles) was required in each preparation for X-ray crystallography. Additionally, at the high concentration of material required for crystallography the virus precipitated from solution, requiring a specific co-solvent to increase solubility.

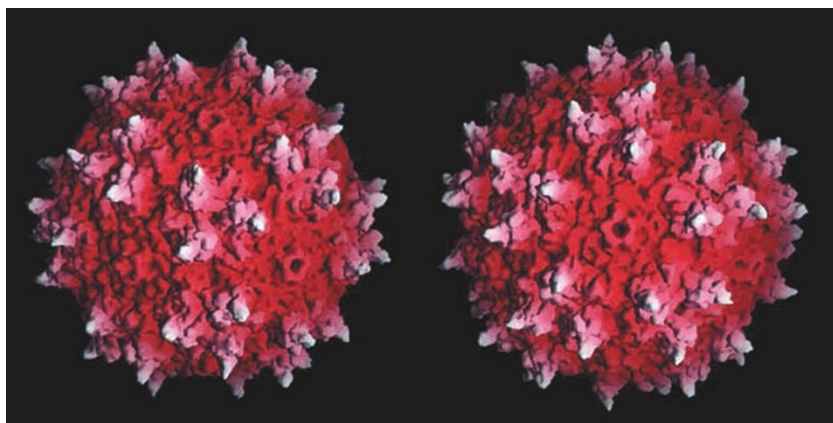
Once these problems of production and concentration were overcome, there was a seemingly intractable problem of virus biology. The virion of AAV2, like all parvoviruses, is composed of more than one capsid protein, the sequences of which share identical reading frames. The smallest of these proteins (Vp3) is the most abundant. As a result, some of the non-overlapping domains of the larger capsid proteins can not be resolved using X-ray crystallography. However genetic and infectivity data show that a Vp3-only virus will bind heparin, and compete with the normal virus in binding assays, but are non-infectious themselves.<sup>10</sup> The lack of infectivity of the Vp3 only virus may result from a lack of phospholipase

activity, previously mapped to the Vp1 non-overlapping domain (unresolved in this crystal structure).<sup>11</sup> Thus genetic and crystallographic studies are both needed to obtain a full understanding of the structure and biology of this virus.

Although several parvovirus crystal structures have been determined, this structure is the first from the dependovirus class. There are striking similarities in the core  $\beta$ -barrel motif between these parvovirus structures, which consist of anti-parallel  $\beta$ -sheets. Between  $\beta$ -sheets are looped-out domains, the longest of which is between  $\beta$ -sheets G and H, the GH-loop. This loop is approximately 230 amino acids in length for AAV serotypes and many autonomous parvoviruses.<sup>12</sup> This loop, although similar in length among AAV serotypes and autonomous parvoviruses, is very diverse in its amino acid sequence and, consequently, its surface topology and cell-surface receptor and antibody binding properties.

One of the new study's most interesting findings is the difference in topology between AAV2 and the other parvovirus crystal structures. In AAV2, centered about the three-fold axis of symmetry are three clusters of three peaks. The peaks are made up from the interaction of two adjacent subunits. However, the sequences that compose these structures are all from the GH-loop. In contrast, CPV has a large spike at the three-fold axis, which is also made up of two capsid subunits.<sup>5,12,13</sup> Interestingly, in the insect densovirus, which has 134 fewer amino acids in its GH-loop, there are few surface features near or at the three-fold axis of symmetry.<sup>14</sup> These examples illustrate that functional viruses can be assembled from very different GH-loops. Given the apparent tolerance of parvoviruses to GH-loop variability, we predict a flurry of structure/function studies based on genetic modifications of the GH-loop. In similar fashion, modifications of the HI loop in adenoviruses resulted in improved adenoviral vectors.

Extensive interactions between capsid subunits at the three-fold axis may play a crucial role during subunit assembly (Figure 1, left). In the homologous structures of CPV glycine residues at the base of these loops allow for flexibility such that loops from adjacent capsid proteins can fold over themselves.<sup>13</sup> Yet in AAV2 the loop structure between amino acids 485 and 517 (Vp1 numbering) is sandwiched between the adjacent subunit's loop structures  $\beta$ GH2-3 and  $\beta$ GH12-13. Additionally, the authors suggested that interactions at the five-fold axis are potentially driving assembly (Figure 1, right).



**Figure 1** Surface topology of AAV-2 colored according to distance from center (red closest, white furthest). Viewed down the three-fold axis (left) and five-fold axis (right) of symmetry. The authors would like to thank Michael Chapman for kindly providing these images.

However AAV2 is assembled, the important question of what influences receptor binding is now answered. Many AAV mutations, widely spaced between amino acids 509 and 591 (Vp1 numbering), have been isolated that interfere with binding and hence virus infectivity.<sup>3,4,6</sup> We now know how these mutations fit into this clear picture. The region between the peaks is lined with basic amino acids which are centered about the three-fold axis of symmetry.<sup>1</sup> This implies that three subunits are required for the interaction between AAV2 and the cell-surface receptor heparan sulfate.

The AAV2 crystal structure will have a big impact on the field of targeted gene therapy. It has already allowed informed interpretation of previously available genetic data and localization of

important functions including receptor and antibody binding to the GH-loops. With this information, a combination of site-directed and insertional mutagenesis, serotype and class domain swaps, and shuffling within the GH-loop domains could be used to produce much-improved gene transfer vectors. Additionally, regions of opposing peaks lining the pocket of each three-fold cluster can be exchanged with antibody binding sites or immune cell recognition sequences. Determination of the atomic structure of AAV2 represents a major step in rationally developing this vector, as well as the other AAV serotypes, into the safest possible targeted long-term expressing gene therapy tool. The new crystal-clear view of AAV2 will launch several years of exciting research. ■

Joseph E. Rabinowitz and R. Jude Samulski are in the University of North Carolina Gene Therapy Center Chapel Hill North Carolina, USA.

e-mail: rjs@med.unc.edu

1. Xie Q *et al.* *Proc Natl Acad Sci USA* 2002; **99**: 10405–10410.
2. Samulski R, Sally M, Muzyczka N. Cold Spring Harbor: Cold Spring Harbor Laboratory Press 1998.
3. Rabinowitz J, Xiao X, Samulski R. *Virology* 1999; **265**: 274–285.
4. Wu P *et al.* *J Virol* 2000; **74**: 8635–8647.
5. Tsao J *et al.* *Science* 1991; **251**: 1456–1464.
6. Girod A *et al.* *Nat Med* 1999; **5**: 1052–1056.
7. Grifman M *et al.* *Mol Ther* 2001; **3**: 964–975.
8. Moskalenko M *et al.* *Virol* 2000; **74**: 1761–1766.
9. Wobus C. *J Virol* 2000; **74**: 9281–9293.
10. Rabinowitz J *et al.* *J Virol* 2002; **76**: 791–801.
11. Zadori Z *et al.* *Dev Cell* 2001; **1**: 291–302.
12. Chapman M, Rossmann M. *Virology* 1993; **194**: 491–508.
13. Xie Q, Chapman M. *J Mol Biol* 1996; **264**: 497–520.
14. Simpson A *et al.* *Structure* 1998; **6**: 1355–1367.