

# Enhancement of recombinant adeno-associated virus activity by improved stoichiometry and homogeneity of capsid protein assembly

Takayuki Onishi,<sup>1</sup> Michika Nonaka,<sup>1</sup> Takahiro Maruno,<sup>1,2</sup> Yuki Yamaguchi,<sup>1</sup> Mitsuko Fukuhara,<sup>1,2</sup> Tetsuo Torisu,<sup>1</sup> Masaharu Maeda,<sup>3</sup> Susan Abbatiello,<sup>4</sup> Anisha Haris,<sup>4</sup> Keith Richardson,<sup>5</sup> Kevin Giles,<sup>4</sup> Steve Preece,<sup>5</sup> Noriko Yamano-Adachi,<sup>1</sup> Takeshi Omasa,<sup>1</sup> and Susumu Uchiyama<sup>1</sup>

<sup>1</sup>Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan; <sup>2</sup>U-Medico Inc, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan; <sup>3</sup>Osaka Consolidated Laboratory, Manufacturing Technology Association of Biologics, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan; <sup>4</sup>Waters Corporation, Milford, MA 01757, USA; <sup>5</sup>Waters Corporation (Micromass UK Ltd), Stamford Avenue, Altrincham Road, Wilmslow SK9 4AX, UK

**Studies of recombinant adeno-associated virus (rAAV) revealed the mixture of full particles with different densities in rAAV. There are no conclusive results because of the lack of quantitative stoichiometric viral proteins, encapsidated DNA, and particle level analyses. We report the first comprehensive characterization of low- and high-density rAAV serotype 2 particles. Capillary gel electrophoresis showed high-density particles possessing a designed DNA encapsidated in the capsid composed of  $(VP1 + VP2)/VP3 = 0.27$ , whereas low-density particles have the same DNA but with a different capsid composition of  $(VP1 + VP2)/VP3 = 0.31$ , supported by sedimentation velocity-analytical ultracentrifugation and charge detection-mass spectrometry. *In vitro* analysis demonstrated that the low-density particles had 8.9% higher transduction efficacy than that of the particles before fractionation. Further, based on our recent findings of VP3 clip, we created rAAV2 single amino acid variants of the transcription start methionine of VP3 (M203V) and VP3 clip (M211V). The rAAV2-M203V variant had homogeneous particles with higher  $(VP1+VP2)/VP3$  values (0.35) and demonstrated 24.7% higher transduction efficacy compared with the wild type. This study successfully provided highly functional rAAV by the extensive fractionation from the mixture of rAAV2 full particles or by the single amino acid replacement.**

## INTRODUCTION

Recombinant adeno-associated virus (rAAV) vectors have been used for *in vivo* gene therapy. Several therapeutic products based on rAAV have already been approved for use in human gene therapy.<sup>1–5</sup> Adeno-associated virus (AAV) capsid consists of three viral proteins (VPs), namely, VP1, VP2, and VP3, which assemble to form the T = 1 icosahedral capsid composed of 60 VPs. VP1, VP2, and VP3 are present in the AAV capsid at the average molar ratio of 1:1:10,<sup>6,7</sup> as determined through gel densitometry. Meanwhile, we have also recently reported the existence of VP3 clip in some serotypes of rAAV, which is eight amino acid residues shorter than

VP3.<sup>8,9</sup> The icosahedral capsid is assembled by the VP3 common region, whereas the N-terminal extensions of VP1 and VP2 play crucial roles in endosomal trafficking<sup>10,11</sup> and escape,<sup>12,13</sup> nuclear localization, and genome release. Therefore, AAV particles with high VP1 and VP2 stoichiometry are required to develop AAVs that can be safely used in gene therapy because high doses of low-potency AAVs can cause adverse immune reactions, making their application in gene therapy problematic.

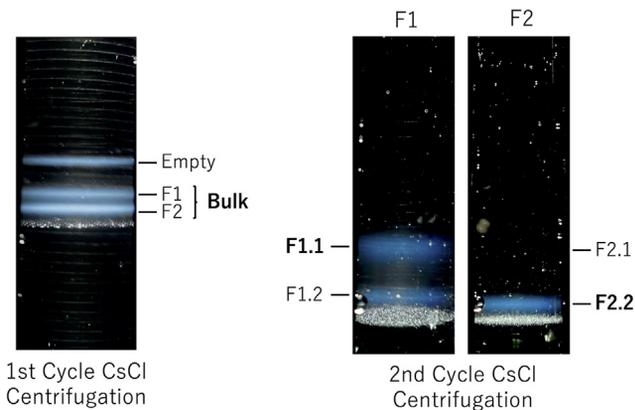
The VP subunits of the AAV capsid are randomly incorporated, and mass spectrometry and mass spectral simulation have suggested that they exhibit particle heterogeneity.<sup>14,15</sup> Cesium chloride (CsCl) isopycnic density centrifugation could be used for the fractionation of heterogeneous particles. Wang et al. reported that rAAV8 preparations showed low-density and high-density full particles in CsCl isopycnic centrifugation.<sup>16</sup> The two types of particles contained similar packaged DNA, but the high-density particles had a higher DNA:protein ratio than the low-density particles. In addition, the high-density particles have been found to exhibit lower VP1 and VP2 ratio than the low-density particles through gel densitometry. Previous studies have reported the existence of low- and high-density particles that represent heterosis of AAVs in the CsCl isopycnic gradient, by focusing not only on capsid protein<sup>7,16,17</sup> but also on DNA length as the cause of heterosis.<sup>18,19</sup> On the other hand, in the wild-type (WT) AAVs, contradictory findings have been reported, i.e., that both types of particles were reported to have the same DNA:protein ratio and density in sucrose metrizamide gradients.<sup>19</sup> As described earlier, the VP ratio affects transduction efficacy.<sup>12,20–24</sup> Therefore, accurate quantification and control of the stoichiometry of VPs is becoming increasingly important. Moreover, in the

Received 17 July 2023; accepted 21 October 2023;  
<https://doi.org/10.1016/j.omtm.2023.101142>.

**Correspondence:** Susumu Uchiyama, Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan.

**E-mail:** [suchi@bio.eng.osaka-u.ac.jp](mailto:suchi@bio.eng.osaka-u.ac.jp)





**Figure 1. Illustration of two-cycle CsCl isopycnic density gradient ultracentrifugation**

High- and low-density rAAV2 particles were separated by performing two cycles of CsCl isopycnic density gradient ultracentrifugation after affinity chromatography purification. Purified rAAV2-CMV-EGFP vectors were developed using a suspended HEK293T production system. Two full particle bands were observed, denoted as F1 and F2 (left). In both fractions, considerable mixing of the two types of particles was observed after the first cycle; thus, a second cycle of CsCl ultracentrifugation (right) was conducted for F1 and F2, where the two sets of particles with different densities were observed in both fractions, leading to further fractionation into F1.1, F1.2, F2.1, and F2.2. In another preparation, the F1 and F2 sample fractionated in a mixture denoted as “Bulk”.

manufacturing of rAAV, the focus is currently on the scaling up production.<sup>25–27</sup> In the future, it will be necessary to resolve heterogeneity, which receives little attention from a manufacturing perspective.

Here, we comprehensively characterized low- and high-density particles that were fractionated using a high-resolution fractionation method. Particle heterogeneity given by particles of two different densities could be the packaged DNA or the VP ratio or post-translational modifications of VPs. Recently, highly sensitive and quantitative protein and DNA analysis methods based on capillary gel electrophoresis (CGE) have been developed. Moreover, we reported a quantitative evaluation method for VP stoichiometry using CGE and liquid chromatography-mass spectrometry (LC-MS).<sup>9</sup> This method allows accurate and reliable estimation of the stoichiometry of VPs in rAAV particles. We have used this quantitation method to evaluate the fractions of high- and low-density particles during rAAV2 production. Typically, one cycle of CsCl isopycnic ultracentrifugation is used for the fractionation; however, the two types of particles remain mixed together even after the fractionation because of the marginal difference in their density. We thus utilized two-cycle purification through ultracentrifugation with separation conditions optimized according to CsCl density gradient simulation. Using this approach, comprehensive assessments of the physicochemical properties and *in vitro* potency of the rAAV particles in the well-separated fractions at the particle level were possible.

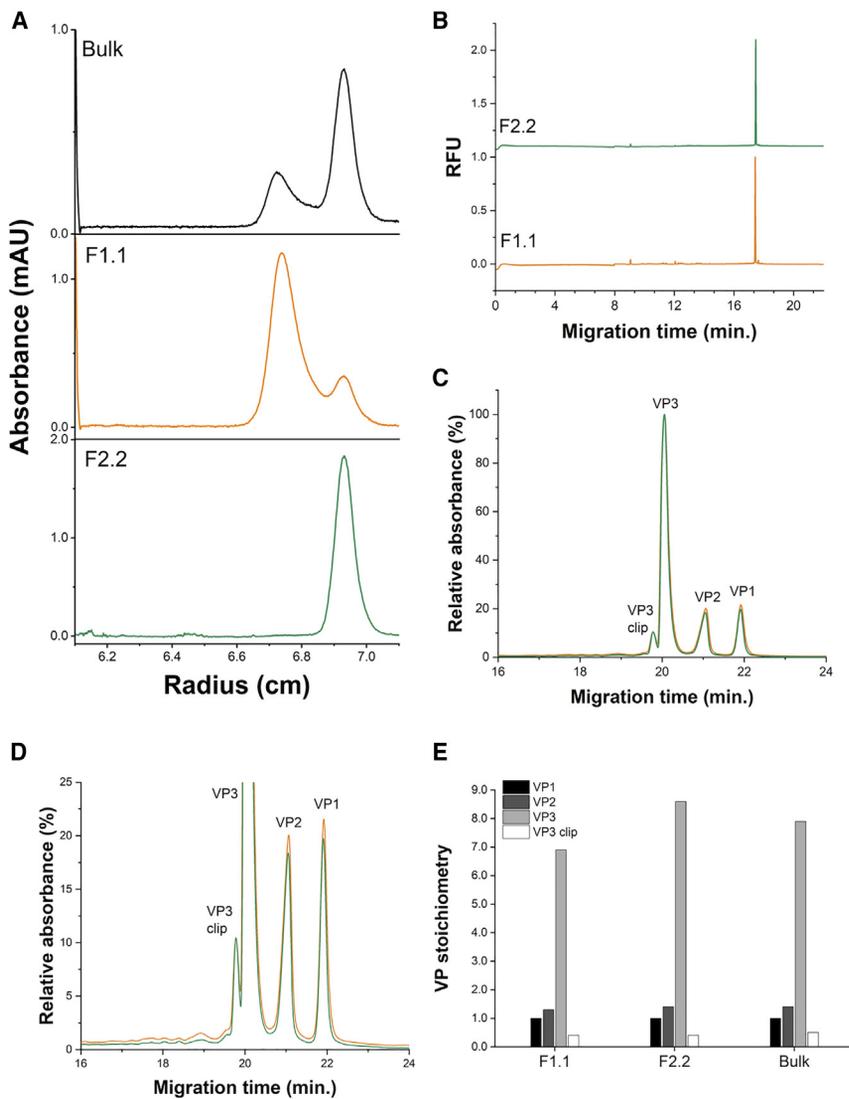
Meanwhile, we successfully obtained rAAV variants with increased VP1 and VP2 stoichiometry compared with that of WT rAAV by pro-

ducing amino acid substitutions at positions M203 and M211, the translation initiation sites of VP3 and VP3 clip, which were previously considered to be associated with a lack of particle formation. Furthermore, the distribution of rAAV particles became highly homogeneous. In this study, we successfully obtained functional rAAV particles with high potency and less protein heterogeneity through two cycles of CsCl ultracentrifugation or amino acid substitution. Based on these findings, this study proposes a potentially novel approach to capsid-based rAAV development.

## RESULTS

### Comprehensive physicochemical characterization of high- and low-density particles

During rAAV production, it would be expected that both low- and high-density full particles would be observed.<sup>16</sup> Figure 1 illustrates two bands for both particles in CsCl ultracentrifugation, forming two fractions due to the piston fractionator with an online monitoring apparatus, denoted as F1 and F2. In both fractions, significant mixing of the two types of particles was observed; thus, a second cycle of CsCl ultracentrifugation was conducted for F1 and F2, where two sets of particles with different densities were found in both fractions, leading to further fractionation into F1.1, F1.2, F2.1, and F2.2. Based on the results of SDS-PAGE and absorbance measurements, which showed an absence of protein and DNA (data not shown), we concluded that the white layer was precipitated CsCl. The purified low- and high-density full particles were subjected to comprehensive characterization. CsCl density gradient-analytical ultracentrifugation (CsCl-DG-AUC) analysis showed that the two full particles had estimated buoyant densities of 1.352 g/cm<sup>3</sup> for F1.1 and 1.361 g/cm<sup>3</sup> for F2.2, with Ab260/280 ratios of 1.277 for F1.1 and 1.280 for F2.2. After the second fractionation, F1.1 still contained a small fraction of high-density full particles; however, low-density full particles could not be detected in F2.2 through CsCl-DG-AUC (Figure 2A). CGE was then performed to analyze the packaged DNA and VP components of F1.1 and F2.2. CGE enables greater separation and sensitivity than those by conventional SDS-PAGE and agarose gel electrophoresis.<sup>28,29</sup> CGE for single-stranded DNA (ssDNA) showed that the length of the contained transgene DNA was the same between the high- and low-density particles, indicating no contribution of the packaged DNA genome to the difference in density between these particles (Figure 2B). Electropherograms from CGE for VP proteins with detection at 214 nm and normalized intensity with respect to the VP3 peak are shown in Figures 2C and 2D. All the structural capsid proteins were present in both F1.1 and F2.2 particles. In addition to VP1, VP2, and VP3 peaks, we also detected VP3 clip generated by ribosomal leaky scanning of the first initiation codon of VP3. The analysis suggested that the F1.1 capsids contained a higher ratio of VP1 and VP2 to VP3 than that of the F2.2 capsids. The VP stoichiometry for these particles was determined by calculating the peak area, using the molar absorption coefficient at 214 nm estimated from the amino acid sequence of each VP protein<sup>9,30</sup> (Figures 2C and 2D). The ratio of VP1:VP2:VP3:VP3<sub>clip</sub> was 1.0:1.3:6.9:0.4 for FP1.1 and 1.0:1.4:8.6:0.4 for FP2.2, respectively (Figure 2E), which correspond to (VP1 + VP2)/VP3<sub>total</sub> of 0.32 for FP1.1 and 0.25 for



**Figure 2. Comprehensive physicochemical characterization of high- and low-density particles**

(A) CsCl density-gradient-analytical ultracentrifugation (CsCl-DG-AUC) profile of Bulk, F1.1 and F2.2 fractions. (B) Capillary gel electrophoresis (CGE) electropherograms of encapsidated DNA in F1.1 and F2.2 particles. (C) CGE electropherogram of F1.1 and F2.2 particles for capsid viral protein (VP) components. (D) Magnified image of electropherogram shown in (C). (E) VP stoichiometry of Bulk, F1.1, and F2.2 fractions. To determine the VP molar stoichiometry, areas of the CGE peaks detected at 214 nm were divided by the molar extinction coefficient of VPs at 214 nm, reflecting the UV absorbance of peptide bonds and amino acids under denatured conditions.

tively (Figure 3A). According to Equations S1 and S2, the increase in the sedimentation coefficient of these F1.1 particles indicates that the average molecular weight of the F1.1 particles was increased compared with that of the F2.2 particles. Additionally, CDMS measurements revealed the single mass of F1.1 and F2.2 particles as 4.70 MDa and 4.58 MDa, respectively (Figure 3B). Considering the possible influence of counterions, salt adducts and trapped solvent as suggested by Jarrold<sup>32</sup> and that CGE provided average VP stoichiometry, a reasonably good correspondence was found between experimental and theoretical values calculated by converting the VP ratio from CGE to 60-mer and multiplying by the mass of each VP (Table S1). A low charge population was also presented in charge versus mass scatterplot of CDMS (Figure 3B). According to the CDMS study, the ejection of encapsidated DNA from AAV will result in the higher charged state while compact, near spherical geometries of AAV will lead to the

lower charged state.<sup>33</sup> The relative comparison of the molecular mass obtained from both BS-AUC and CDMS showed higher value in F1.1 than F2.2. Because the DNA content was the same, the increase in mass could be attributed to the higher VP1 and VP2 ratios in the 60-mers. This in turn supported the assertion that the VP1 and VP2 stoichiometry in the F1.1 particles is higher than that in the F2.2 particles.

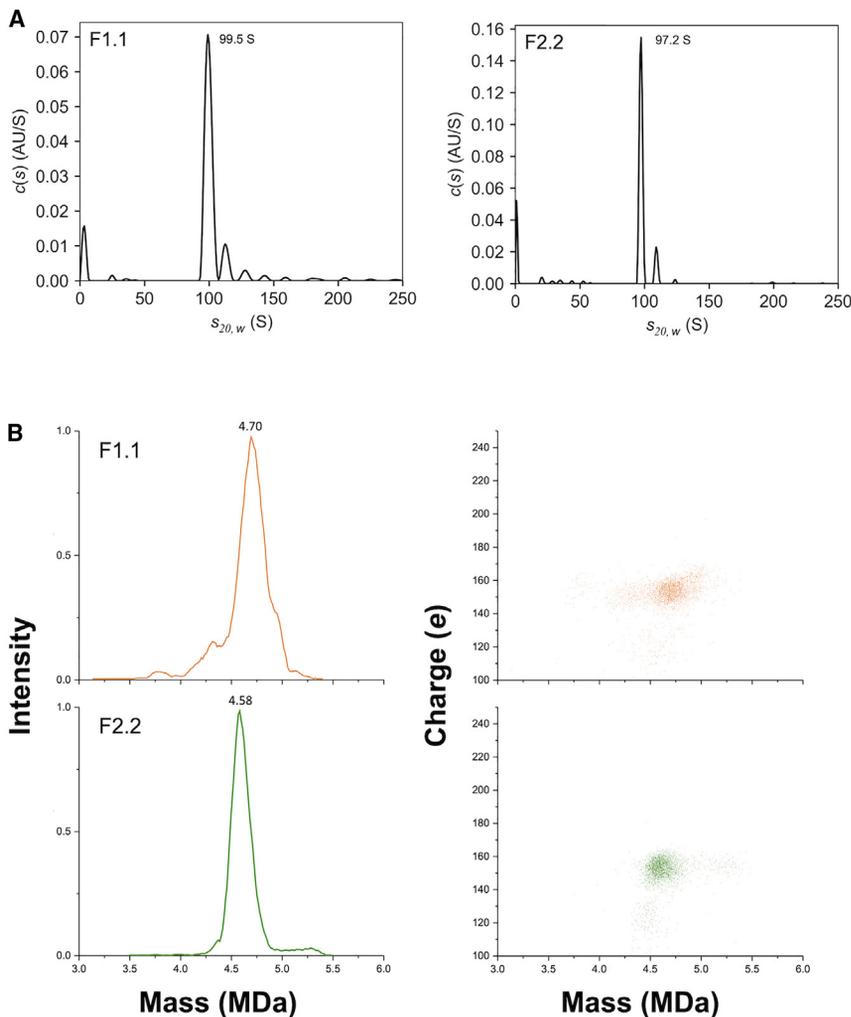
### F1.1 particles have higher mass than that of F2.2 particles

As F2.2 particles contained low VP1 and VP2 stoichiometry, it is reasonable to suspect that the molecular weight of the high-density particles could differ from that of the low-density particles. Therefore, band sedimentation-analytical ultracentrifugation (BS-AUC) and charge detection-mass spectrometry (CDMS) measurements were conducted for the particles in F1.1 and F2.2. BS-AUC enables evaluation of the size distribution in solution with only 1/25 of the sample volume compared with that required for sedimentation velocity-analytical ultracentrifugation.<sup>31</sup> The sedimentation profile showed that F1.1 and F2.2 included no empty particles, and the sedimentation coefficients of F1.1 and F2.2 particles were 99.5 S and 97.2 S, respec-

lower charged state.<sup>33</sup> The relative comparison of the molecular mass obtained from both BS-AUC and CDMS showed higher value in F1.1 than F2.2. Because the DNA content was the same, the increase in mass could be attributed to the higher VP1 and VP2 ratios in the 60-mers. This in turn supported the assertion that the VP1 and VP2 stoichiometry in the F1.1 particles is higher than that in the F2.2 particles.

### In vitro transduction efficacy of F1.1 and F2.2

We hypothesized that the virus could require fewer particles for infectivity if the stoichiometry is biased toward VP1, which is directly related to the presence of the phospholipase A2 (PLA2) domain on the VP1 unique (VP1u) sequence.<sup>11,23</sup> Furthermore, the VP1/VP2 common region has a nuclear localization sequences (NLSs).<sup>10</sup> Therefore, these N-terminal extensions of VP1 and VP2 play crucial roles in endosomal trafficking and escape, nuclear localization, and genomic release, and the total amount of VP1 and VP2 significantly affects



**Figure 3. Particle mass analysis of F1.1 and F2.2 particles**

(A) Sedimentation profiles of F1.1 and F2.2. (B) CDMS mass histogram (right) and charge versus mass scatterplot (left) for F1.1 and F2.2 particles.

VP3 and VP3 clip, located at positions M203 and M211, have been reported to prevent particle recovery during production, suggesting that these residues play a crucial role in the assembly and/or stability of the particles of rAAV2.<sup>34</sup> We then focused on AAV7 and AAV9, which have highly homologous N-terminal sequences of VP3 among the serotypes and conserved sequences, other than ATG at the start codon (Figure 5A). Through serotype sequence alignment, we noticed that, in AAV7, position M203 is CTG (valine), whereas in AAV9, position M211 is CTG (valine). We thus created constructs with CTG substitution at the ATG translation start site for each (rAAV2-M203V and rAAV2-M211V). Interestingly, these rAAV variants exhibited productivity comparable to that of WT rAAV2 (Figure 5B), and they showed less variation in productivity across several batches (data not shown). Meanwhile, DNA packaging ratio of rAAV variants were no significant difference compared with that of WT (Figure S1).

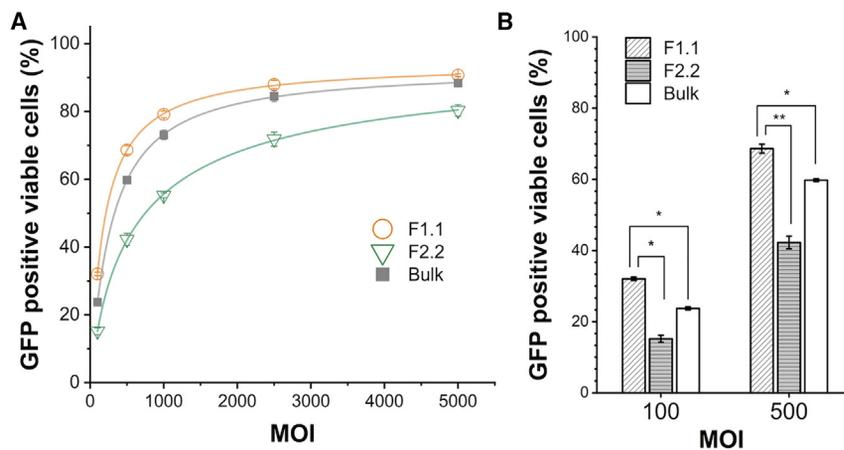
CGE and LC-MS were performed to evaluate the VP components of both rAAV2-M203V and rAAV2-M211V. rAAV2-M203V particles significantly incorporated VP3 clip, whereas rAAV2-M211V particles incorporated full-length VP3 besides VP1 and VP2 (Figure 5C). The deconvoluted mass from LC-MS analysis that provided the mass of each VP in these rAAV variants corresponded to the theoretical values by <25 ppm (Figures 5D and S2; Table S2). VP stoichiometries of rAAV2-M203V and rAAV2-M211V were calculated from the molar absorption coefficient at 214 nm of each denatured VP and the peak area of CGE (Figure 5E). VP1 and VP2 stoichiometry was indicating a predominant trend in rAAV2-M203V and rAAV2-M211V such that  $(VP1 + VP2)/VP3_{total} = 0.35$  and  $0.32$ , respectively (Figure S3).

The denaturation temperatures ( $T_m$ ) of these rAAV variants were assessed by differential scanning fluorimetry (DSF) to be 67.1°C for WT rAAV2, 68.2°C for rAAV2-M203V, and 67.7°C for rAAV2-M211V, indicating that these mutations had negligible impact on the conformational stability of the particles (Figure 5F). rAAV2-M203V and rAAV2-M211V showed 21.9% (MOI  $5 \times 10^2$ ) to 24.7% (MOI  $1 \times 10^2$ ) and 8.2% (MOI  $2.5 \times 10^3$ ) to 11.9% (MOI  $1 \times 10^3$ ) higher *in vitro* transduction effects on HeLaRC32 cells compared with WT rAAV, respectively (Figure 6). These results were consistent with

transduction efficiency. The biological activity was evaluated using the expression level of a transgene, GFP, encoded in the rAAV2 genome. The proportion of GFP-positive viable cells was evaluated by flow cytometry at five different multiplicity of infection (MOI) levels. The infectivity of F1.1 was up to 8.9% (MOI  $5 \times 10^2$ ) higher and F2.2 up to 17.8% (MOI  $1 \times 10^3$ ) lower compared to Bulk (Figures 4A and 4B). These results confirmed that VP1 and VP2 stoichiometry plays a crucial role in infectivity *in vitro*, consistent with previous studies of Bosma et al., who reported an interplay between VP1 and VP2 stoichiometry and infectivity *in vitro*.<sup>24</sup>

#### Change in the translation initiation site of VP3 and VP3 clip alters the stoichiometry of VP1:VP2:VP3<sub>total</sub>

rAAV variants by site-directed mutagenesis at the translation initiation sites of VP3 and VP3 clip were created to determine whether these rAAV variants could alter the level of VP3 incorporation into particles. This was based on the report by Bosma et al. that showed that the level of VP3 incorporation affects the incorporation of VP1 and VP2.<sup>24</sup> However, mutations at the translation initiation site of



**Figure 4. *In vitro* transduction efficacy of F1.1 and F2.2 particles in HeLaRC32 cells**

GFP reporter gene assay results are presented as the mean from triplicate wells in one representative experiment out of two independent runs. Error bars represent  $\pm$  SD of the three independent wells. (A) GFP-positive viable cells were measured using a fluor cytometer at five MOI points (MOI  $1 \times 10^2$ ,  $2.5 \times 10^3$ ,  $1 \times 10^3$ ,  $2.5 \times 10^3$ , and  $5 \times 10^3$ ). (B) *In vitro* transduction efficacy of selected MOIs ( $1 \times 10^2$  and  $5 \times 10^2$ ).

high VP1 and VP2 stoichiometry showing high transduction efficiency and replicated with different genes of interest (Figure S4).

#### rAAV2-M203V and rAAV2-M211V possess relatively homogeneous capsid stoichiometry

We noticed that the particle heterogeneity of these rAAV variants was altered during rAAV purification. CsCl-DG-AUC was performed to measure the particle density distribution for each variant in a bulk state. As a result, the distribution of the particles into two different densities in WT rAAV converged to a single density in rAAV2 variants (Figures 7A and S5). The estimated densities from CsCl-DG-AUC of both rAAV2-M203V and rAAV2-M211V were  $1.359 \text{ g/cm}^3$ .

The theoretical molecular weights of rAAV2-WT, -M203V, and -M211V calculated using the VP stoichiometry derived from CGE were 4.54, 4.56, and 4.57 MDa, respectively. Particle measurements with CDMS for the bulk samples of rAAV2-WT, -M203V, and -M211V gave molecular weights of 4.75, 4.73, and 4.71 MDa, respectively (Figure 7B), and these values were in close agreement with the theoretical values. Notably, compared with bulk WT rAAV2, the convergence of molecular mass to homogeneity was confirmed for the two rAAV2 variants by the full width at half maximum (FWHM) values of mass histogram. The FWHM values for rAAV2-WT, -M203V, and -M211V were 0.18 MDa, 0.13 MDa, and 0.16 MDa, with the rAAV variant having smaller values than WT. The sharpness of the peak in CDMS mass histogram was in the following order: rAAV2-WT > -M211V > -M203V. These results revealed that the two rAAV2 variants have reduced particle heterogeneity.

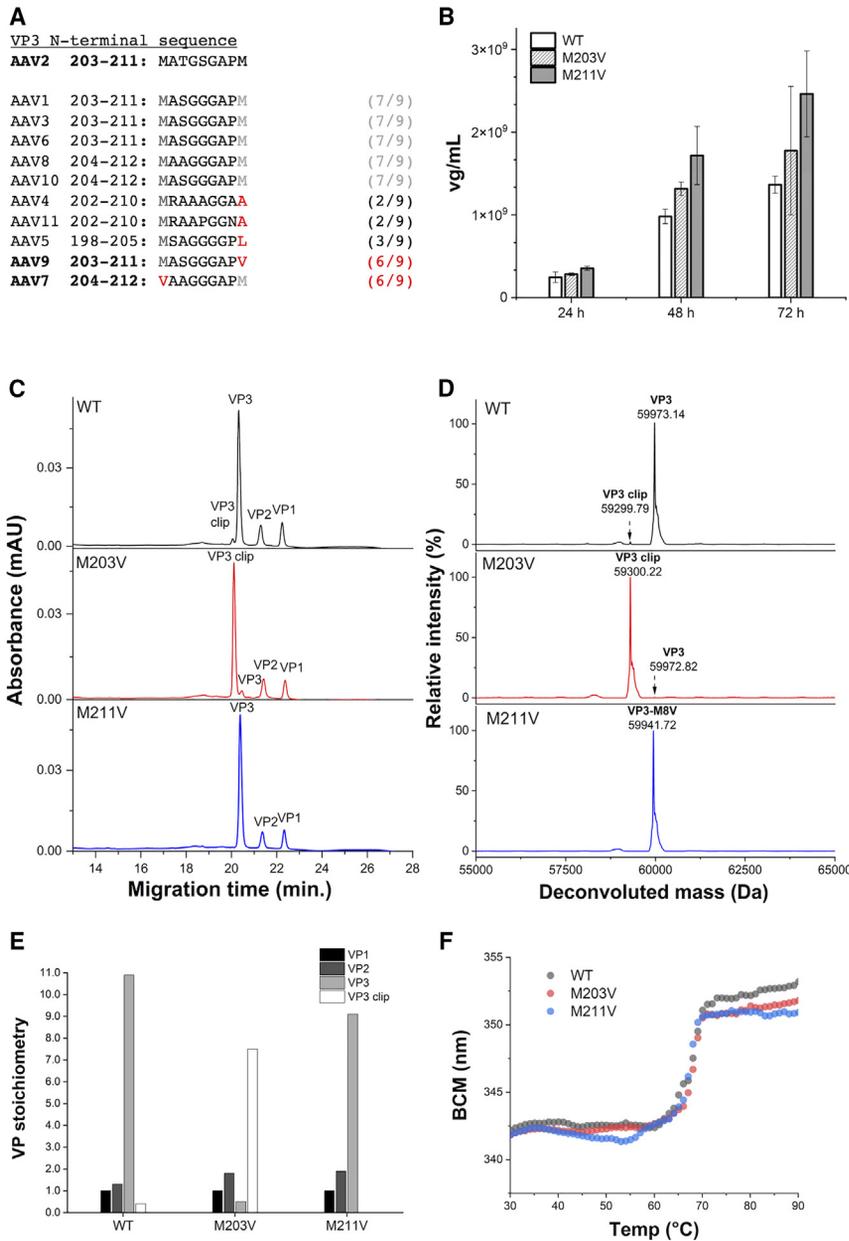
#### Total VP1 and VP2 stoichiometry affects *in vitro* transduction efficacy

Based on GFP reporter gene expression results, we normalized the gene transduction efficiencies of F1.1, F2.2, rAAV2-M203V, and rAAV2-M211V using the efficiency of gene transduction of bulk rAAV-WT as measured by flow cytometry. Further, the correlation among the relative VP1 stoichiometry ( $\text{VP1}/[\text{VP2} + \text{VP3}_{\text{total}}]$ ), VP2 stoichiometry ( $\text{VP2}/[\text{VP1} + \text{VP3}_{\text{total}}]$ ), VP1 and VP2 stoichiometry

( $[\text{VP1} + \text{VP2}]/\text{VP3}_{\text{total}}$ ) and the normalized transduction efficiency based on the VP stoichiometry calculated from the area ratio of the CGE and the molar absorption coefficient was depicted (Figure 8). The results showed that there was no good correlation between “VP1 only” and “VP2 only,” indicating good correlation between the VP1 and VP2 stoichiometry and relative transduction efficiency. This was consistent with previous reports that showed the correlation based on the quantitation by gel densitometry, while our current study revealed a clearer numerical relationship about the correlation. This is directly related to the presence of a phospholipase A2 domain on the VP1u sequence and nuclear transfer signals on the VP1 and VP2 sequences. Naturally, compared with those of F1.1, rAAV2-M203V, and -M211V, the significantly lower transduction efficacy at MOI  $1 \times 10^3$  of F2.2, which contains only high-density rAAV particles, could be explained by the lower VP1 and VP2 stoichiometry.

#### DISCUSSION

In this study, we demonstrated the two different distributions of the density of full particles observed in rAAV in the CsCl density gradient originated from the heterogeneity of VP subunits and not from the size difference of the viral genome. The  $(\text{VP1} + \text{VP2})/\text{VP3}_{\text{total}}$  values were 0.31 for F1.1 and 0.27 for F2.2, corresponding to an increase of 1.8% and a decrease of 2.8% compared with Bulk ( $[\text{VP1} + \text{VP2}]/\text{VP3}_{\text{total}}$  value, 0.28), respectively. These quantitative relationships could be clarified by the accurate CGE method that we recently developed by combining CGE and mass spectrometry.<sup>9</sup> The BS-AUC and CDMS results supported the higher VP1 + VP2 number in F1.1 than that in Bulk and F2.2. The estimation of the VP1, VP2, and VP3 combinations in 60-mer based on particle mass determined by CDMS suggested that VP1 and VP2 stoichiometry was obviously higher for the F1.1 particles than for the F2.2 particles (Figure S6). Correspondingly, F1.1 particles purified through two cycles of CsCl ultracentrifugation showed high transduction efficacy *in vitro*, as evidenced by the flow cytometry analysis of GFP reporter gene expression in HeLaRC32 cells and transduction efficiency that showed higher transgene expression for F1.1 than Bulk (up to 8.9%). However, considering the reason for the different densities of AAV in CsCl isopycnic density gradient centrifugation was unclear from the contradictory reports published to date, no definitive conclusions about the relationship between AAV density and potency could be drawn.<sup>7,16–19,35</sup> As described above, an accurate and reliable



**Figure 5. Characterization of rAAV2-M203V and -M211V**

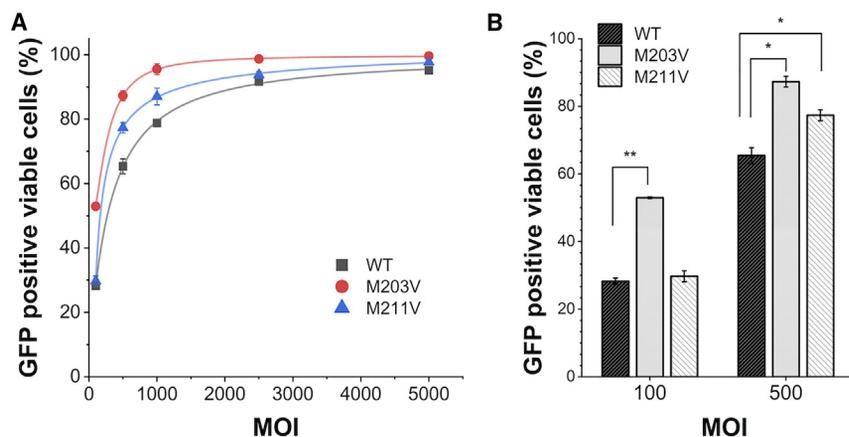
(A) Amino sequence alignment of the VP3 N-terminal region of AAV1-11 and the number of matches with the AAV2 sequence. AAV7 and AAV9 had a the highly conserved sequence with AAV2 for the VP3 N-terminal region. (B) Productivity comparison between rAAV2 wild-type (WT) and rAAV2 variants at 24, 48, and 72 h post-transfection. Productivity assay results are presented as the means of three independent production batches. Error bars represent  $\pm$  SD of titer in three independent production batches. (C) CGE electropherograms of rAAV2-WT and rAAV2 variants. (D) Deconvoluted mass values of rAAV2-WT and rAAV2 variants for VP3 and VP3 clip. (E) VP stoichiometry of rAAV2-WT and rAAV2 variants. To determine the VP molar stoichiometry, areas of the CGE peaks detected at 214 nm were divided by the molar extinction coefficient of VPs at 214 nm, reflecting the UV absorbance of peptide bonds and amino acids under denatured conditions. (F) Thermal unfolding plot of rAAV2-WT and rAAV2 variants.

using adenoviruses as helpers and other parvoviruses (H-1 parvovirus<sup>37</sup> and mouse minute virus [MVM]<sup>38</sup>) preparations. High-density particles of WT AAV2 were reported to have a higher DNA:protein ratio than low-density particles.<sup>7</sup> However, other studies reported that both types of particles have the same DNA:protein ratio and sedimentation velocity.<sup>19</sup> Furthermore, the ratio of infectious units to physical particles has been reported to be 16- to 300-fold higher for low-density particles of AAV2 than that for high-density particles.<sup>19</sup> Conversely, no significant difference was observed in the infectivity between low-density and high-density particles of AAV4, MVM, and H1 parvovirus,<sup>18,37,38</sup> although some studies attributed that the difference in density originated from the difference in the DNA contents of such particles. Meanwhile, the rAAV study concluded that the difference in density originated from the difference in the VP ratio of such particles. This discrepancy

regarding the properties of high- and low-density particles of AAV could be attributed to differences in sample preparation method, as some studies used single CsCl ultracentrifugation for fractionation. Especially in the case of the same encapsidated DNA, and density difference derived from VP stoichiometry, as in this case, a single CsCl ultracentrifugation is insufficient to separate the two types of particles into fully distinct fractions, since the difference in estimated density between the two types of particles is marginal ( $\sim 0.009$  g/cm<sup>3</sup>). In fact, F1.1, which was prepared by CsCl isopycnic gradient centrifugation after affinity chromatography purification, still possessed residual high-density particles. In this study, we applied second CsCl isopycnic gradient centrifugation to acquire F2.2, in which

method for quantifying the VP stoichiometry was recently developed. For example, SDS-PAGE with a dye staining approach provides less accurate quantitation results because of the dependence of the band intensity on the amino acid composition. In the development of rAAV vector, the VP stoichiometry has been recognized as a critical quality attribute and quality control as a specification for rAAV products as a biopharmaceutical. Meanwhile, how VP ratio of rAAV produced by HEK cells has an impact on the transduction efficacy was unknown so far.

Low- and high-density particles have also been previously observed in WT AAVs (AAV1,<sup>17</sup> AAV3,<sup>7</sup> AAV4,<sup>18</sup> and avian AAV<sup>36</sup>) produced



**Figure 6. *In vitro* transduction efficacy of rAAV2-M203V and -M211V in HeLaRC32 cells**

GFP reporter gene assay results are presented as the mean from triplicate wells in one representative experiment out of two independent runs. Error bars represent  $\pm$  SD of the three independent wells. (A) GFP-positive viable cells were measured using a fluor cytometer at five MOI points (MOI  $1 \times 10^2$ ,  $5 \times 10^3$ ,  $1 \times 10^3$ ,  $2.5 \times 10^3$ , and  $5 \times 10^3$ ). (B) *In vitro* transduction efficacy of selected MOIs ( $1 \times 10^2$  and  $5 \times 10^2$ ).

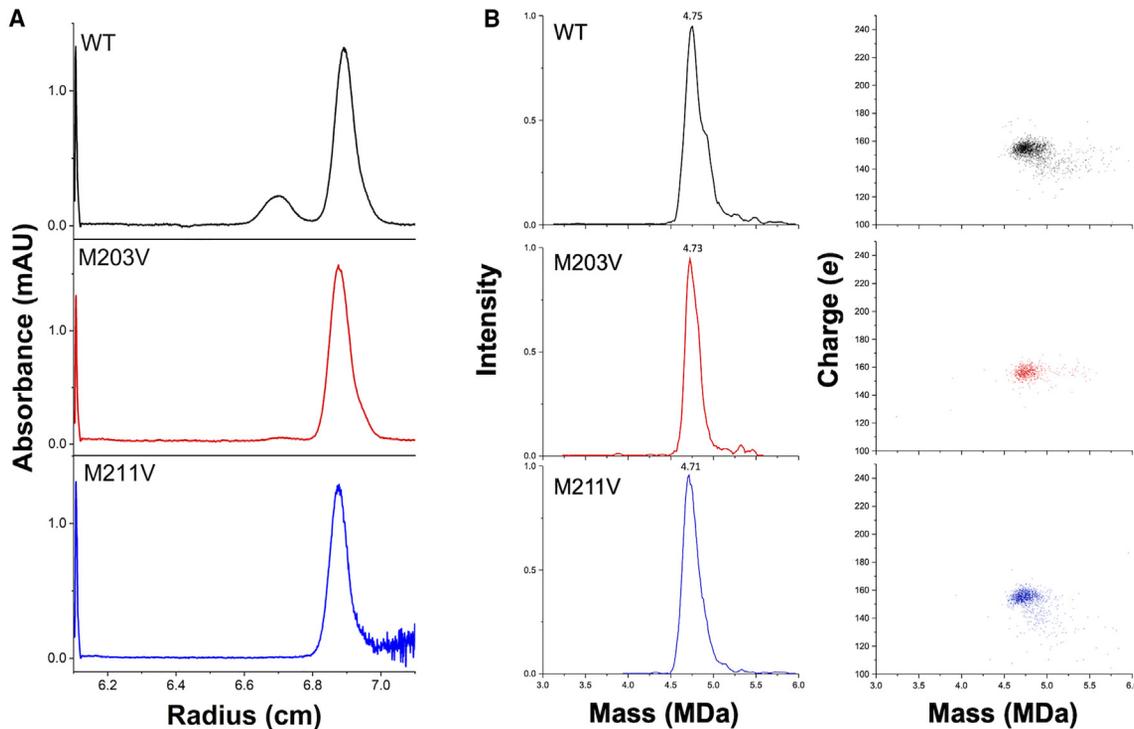
low-density rAAV is no longer involved in the preparation. Thus, highly purified rAAV particles with high potency were thus acquired, indicating that the repetition of CsCl density gradient centrifugation or its effective optimization is necessary for the preparation of AAV with higher potency from a heterogeneous AAV full-particle sample that is in fact a mixture of AAV full particles with different VP stoichiometries.

Amino acid replacement at both the VP3 and VP3 clip transcription start sites in rAAV2 resulted in homogeneous rAAV capsids with high VP1 and VP2 stoichiometry. Leucine (rAAV2-M203L and rAAV2-M211L)<sup>34</sup> and alanine (rAAV2-M211A) substitutions did not confirm particle formation, even though some serotypes were conserved at residues 203 and 211, although the N-terminal of VP3 homology was less (Figures 5A and S7). Homogeneous particle formation was seen with valine substitution, suggesting that a very sophisticated balance of 60-mer AAV capsid particles is being formed. The stoichiometry of VP1 + VP2 was higher for rAAV2-M203V, rAAV2-M211V, and WT rAAV in the bulk state, in descending order. The changes in stoichiometry of rAAV2-M203V and rAAV2-M211V are probably related to the translation process itself. Bosma et al. reported that different contexts of VP1 translation initiation sites alter the amount of VP3 incorporated by altering the level of VP1 translation.<sup>24</sup> In such cases, a decrease in VP1 expression could result in less VP1 and VP2 incorporation and more VP3 incorporation. Meanwhile, in this study, we introduced an amino acid replacement at the start site of VP3; the decrease in VP3 translation that resulted from this could have led to the increased VP1 and VP2 incorporation for rAAV2-M203V, and similarly the lack of VP3 clip translation volume could have resulted in the slight increase in VP1 and VP2 incorporation for AAV2-M211V, based on the observation of the VP1 and VP2 stoichiometry being higher in rAAV2-M203V, rAAV2-M211V, and rAAV2-WT, in that order.

The observation of VP3 in rAAV2-M203V could be explained by the recognition of the CTG sequence by the ribosome at the original VP3 translation start position as a non-canonical start codon (Figures 3C and 3D). Once the ribosome binds to an mRNA, it moves forward un-

til it finds the translation initiation site (ATG in the case of the canonical site) in the appropriate context and initiates protein synthesis. When weak initiation sites (e.g., ACG and CTG) are surrounded by a favorable nucleotide context, they can initiate protein synthesis in a non-canonical manner. This mechanism, called leaky ribosome scanning, has recently attracted substantial attention in the engineering of protein expression.<sup>39,40</sup> In fact, a study on nucleosomes confirmed that CTG is a non-canonical translation start site in approximately 7.5% of all translation start sites in HEK293T cells,<sup>41</sup> which were used for the production of rAAV in this study. As discussed in the previous study,<sup>8,9</sup> Kozak sequences regulate the expression levels of VP3 and VP3 clip. If there are two possible initiation codons to be translated, the initiation codon that has A at the  $-3$  position besides G at the  $+4$  position (A in the initiation codon AUG is counted as  $+1$ ) is preferentially used for expression.<sup>42–44</sup> In the case of VP3 of some AAV serotypes, the first ATG at M203 has A at the  $-3$  position and G at the  $+4$  position (strong Kozak sequence), whereas the second ATG at M211 has C at  $-3$  and G at  $+4$  (weak Kozak sequence). Therefore, the population of VP3 is much larger than that of VP3 clip. In summary, this study showed experimentally that a portion of the ribosomal subunit reaches the VP3 second start codon at M211 after missing the VP2 non-canonical start codon and the VP3 start codon, resulting in VP3 clip expression, while the strong Kozak sequence at M203 results in significant VP3 translation in the WT. VP3 clip was shown experimentally to be significantly translated in the WT. The ATG to CTG mutation at position 203 could cause the low translation of VP3 of WT rAAV starting from M203 because of the lower frequency of the ribosome being trapped by the strong Kozak sequence, resulting in the translation at a non-canonical translation initiation codon. This was indicated by the deletion of the valine residue at position 203 and the acetylation of the alanine residue at the N terminus (see Table S1). On the other hand, the change from ATG to CTG at position 211 resulted in a weak initiation codon and a weak Kozak sequence that did not translate the VP3 clip, which was generated through the mechanism mediated by the non-canonical initiation codon, resulting in rAAV2-M211V for the three VPs.

The use of flow cytometry to evaluate the *in vitro* transduction efficacy showed that rAAV2-M203V and rAAV2-M211V could transduce genes into HeLa cells with higher efficiency (up to 24.7% and 21.9%, respectively) than the bulk WT rAAV2. Considering the clear



**Figure 7. Particle homogeneity of rAAV2-M203V and -M211V**

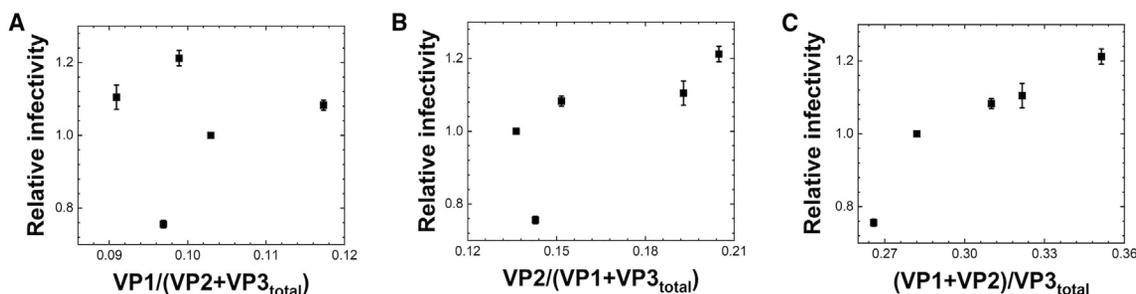
(A) CsCl-DG-AUC profile of rAAV2-WT, -M203V, and -M211V for bulk state. (B) CDMS mass histogram (right) and charge versus mass scatterplot (left) of rAAV2-WT, -M203V, and -M211V for bulk state.

relationship between the transduction efficiency and the VP1 and VP2 stoichiometry seen in F1.1 and F2.2, this result is attributed to the increased VP1 and VP2 stoichiometry in the homogeneous particles of these rAAV variants. In a previous report, mutations at the M203 and M211 positions were shown to affect particle stability and assembly.<sup>34</sup> In this study, we examined the thermostability of the particles that corresponds to the structural stability of the capsid. The results indicated no significant difference in denaturation temperatures in DSF between WT rAAV2 and these rAAV2 variants. The fact that the  $T_m$  values were similar to that of WT rAAV suggests the increase in VP1 and VP2 stoichiometry, rather than the impact on the capsid structural stability, is solely responsible for the increased transduction efficacy.

Wörner et al. proposed a VP stoichiometry model based on native mass spectrometry data and parallel spectral simulations, which predicted that the compositional stoichiometry of VP subunits of AAV capsids is broadly distributed.<sup>15</sup> Our CsCl isopycnic gradient ultracentrifugation succeeded in the fractionation of particles with different VP stoichiometries that constitute the WT bulk. CsCl-DG-AUC analysis also showed that the difference in the theoretical  $vbar$  in water between F1.1 and F2.2 based on VP stoichiometry determined by CGE with the same encapsidated ssDNA was 0.0003 cm<sup>3</sup>/g (F1.1: 0.6839 cm<sup>3</sup>/g and F2.2: 0.6837 cm<sup>3</sup>/g), while the difference in  $vbar$  in CsCl isopycnic point in the experiment was 0.0049 cm<sup>3</sup>/g (F1.1: 0.7396 cm<sup>3</sup>/g and F2.2: 0.7347 cm<sup>3</sup>/g). This suggests that the

difference in  $vbar$  due to differences in VP stoichiometry was significantly enhanced in the CsCl condition. However, particle stoichiometry distributions of rAAV2-M203V and rAAV2-M211V converged to homogeneity in the CsCl DG-AUC. Consistent with this, rAAV2-M203V and rAAV2-M211V showed only one band during purification using CsCl density gradient ultracentrifugation (DG-UC). The CDMS results also indicated particle VP homogeneity of the variants, mass histograms of rAAV2-M203V and rAAV2-M211V exhibited FWHM of 0.13 MDa and 0.16 MDa, respectively, indicating a narrower mass distribution than that of rAAV2-WT (0.18 MDa) (Figure 7B). Recently, it has been shown that VP ratios depend on the production system, with VP1 expression generally being lower in baculovirus production systems.<sup>45,46</sup> The VP1 and VP2 ratio has also been found to be higher in rAAV produced by HEK production systems than in that produced in baculovirus production systems.<sup>15</sup> It would be reasonable to infer that the distribution of VP stoichiometry in rAAV particles was affected by the levels of VP3 and VP3 clip translation, although all rAAV2 was produced in HEK293T cells. The exact mechanism of AAV particle assembly and the reduced particle heterogeneity remains unresolved.

In this study, the VP stoichiometry was well associated with *in vitro* transduction efficiency, as confirmed by the quantitative analysis of VP stoichiometry regarding the VP1 and VP2 stoichiometry and the transduction efficiency analysis. Two regions in VP1, one of which



**Figure 8. Correlation between VP stoichiometry and transduction efficacy**

(A–C) Plot of VP molar stoichiometry and transduction efficiency of Bulk, F1.1, F2.2, rAAV2-M203V, and -M211V. Relative potencies indicate GFP-positive viable cell counts at MOI  $1 \times 10^2$ , all normalized to the bulk state WT value of 1. Error bars represent  $\pm$  SD for three independent wells in the GFP reporter gene assay. (A) Correlation between VP1 stoichiometry and relative potency. (B) Correlation between VP2 stoichiometry and relative potency. (C) Correlation between the total stoichiometry of VP1 and VP2 and relative potency.

is also present in VP2, are postulated to be responsible for the transduction efficiency of AAV. The N-terminal region of VP1, normally embedded in the capsid and exposed during intracellular transport of the vector, plays an important role in the translocation of vector to the nucleus.<sup>10,13,20,22,23</sup> Phospholipase domains in this region specifically hydrolyze the 2-acyl ester (sn-2) bonds of phospholipids that form the lipid bilayer of the endosome, resulting in endosomal escape of AAV. The high level of VP1 incorporation into the capsid is expected to improve the efficiency of AAV-based DNA transport into the cell nucleus. Here, when we looked only at the stoichiometry of VP1, this did not explain the *in vitro* transduction efficacy. Therefore, we also considered the contribution of NLS, another functional region of VP1 that is also present in VP2. Upon applying this approach, we found a complete correlation between the total sum of VP1 and VP2 and the transduction efficacy. Thus, VP1 and VP2 stoichiometry is important for rAAV to achieve high transduction efficiency.

In this study, the AAV particles, originally known to be heterogeneous, differed in VP stoichiometry, and two cycles of ultracentrifugation enabled their fractionation into rAAV particles with high VP1 and VP2 stoichiometry and high potency. The important roles played by the total increase in VP1 and VP2 is associated with both possessing an NLS, which is responsible for the transportation of rAAV and a PLA2 domain, responsible for endosomal escape. This study showed the successful creation of variants with a high VP1 and VP2 stoichiometry and with particle homogeneity. These variants with high VP1 and VP2 stoichiometry are better suited to the more scalable LC-based purification. This is currently a focus of development for achieving large-scale production. In recent years, while capsid engineering for well-directed tropism has been conducted, manufacturability is critical. Therefore, variants with high potency and high manufacturability, such as those presented in this study, will become more important from a manufacturing perspective in the future.

## MATERIALS AND METHODS

### Cell culture

Suspended HEK293T cells were used for AAV vector production. In addition, HeLaRC32 Cells were used for *in vitro* characterization. Sus-

pended HEK293T cells were maintained with BalanCD HEK293 (FUJIFILM Irvine Scientific, Inc., Santa Ana, CA) with 1% penicillin-streptomycin. HeLaRC32 cells were maintained in DMEM (Sigma-Aldrich, St. Louis, MO) supplemented with 10% fetal bovine serum (FBS, Hyclone [GE Healthcare Life Sciences], Madison, WI) and 1% penicillin-streptomycin (Gibco [Thermo Fisher Scientific], Grand Island, NY). Cells were grown as adherent cultures in 5% CO<sub>2</sub> at 37°C.

### Construction of variant plasmid

The translational start of the VP3 and VP3 clip was modified through site-directed mutagenesis. For each variant plasmid, two complementary PCR primers, containing a missense mutation in the individual capsid protein in the start codons were used to introduce changes in the *cap* open reading frame. These plasmids were screened for restriction sites inserted by silent mutations, and the mutations were confirmed through DNA sequencing.

### rAAV preparation

All rAAV vectors were generated using the triple plasmid, co-transfection. Briefly, pAAV-Rep&Cap (serotype 2), pAd helper, and transgene (CMV-EGFP) plasmids (Vector Builder, Vector ID: VB010000-9394npt) were co-transfected into suspended HEK293T cells cultured in a bioreactor at a ratio of 1:1:1. Vectors from the transfected cells and the medium were harvested 72 h post-transfection and purified through affinity chromatography using AAVx columns (Thermo Fisher Scientific, Waltham, MA). Bulk rAAV samples were purified using affinity chromatographic purification followed by a single CsCl ultracentrifugation to separate full and empty particles. The rAAV samples purified through affinity chromatography (>2 mL) were transferred to 13.2 mL ultra-clear ultracentrifuge tubes containing 2.5 M CsCl/PBS with 0.001 w/v% poloxamer-188 solution, bringing the final volume to 12 mL. The tubes were centrifuged at 34,000 rpm in an Optima XE-90 (Beckman Coulter, Inc., Brea, CA) using a Beckman SW41Ti rotor at 20°C for 72 h. The virus band was collected with an online monitoring apparatus and dialyzed in Slide-A-Lyzer 10K (Thermo Fisher Scientific).

### Two-cycle CsCl DG-UC purification

We conducted two-cycle CsCl DG-UC to separate and fractionate high- and low-density full particles. The purified rAAV samples (>2 mL) were transferred to 13.2-mL ultra-clear ultracentrifuge tubes containing 2.5 M CsCl/PBS with 0.001 w/v% poloxamer-188 solution, bringing the final volume to 12 mL. The tubes were then centrifuged at 34,000 rpm in an Optima XE-90 (Beckman Coulter) using a Beckman SW41Ti rotor at 20°C for 72 h. The virus band containing the enriched complete particles was extracted and mixed with 2.5 M CsCl/PBS with 0.001 w/v% poloxamer-188 solution and subjected to another round of ultracentrifugation at 24,000 rpm for 72 h. The virus band was collected with an online monitoring apparatus and dialyzed in Slide-A-Lyzer 10K (Thermo Fisher Scientific).

### CGE for ssDNA

We treated 10  $\mu$ L AAV solution ( $1.0 \times 10^{12}$  viral genome) with DNAase and ProteaseK to extract the ssDNA from the capsid using the following protocol. To prepare the rAAV sample, 20  $\mu$ L of nuclease-free water, 3  $\mu$ L of  $10\times$  DNase buffer, 1.5  $\mu$ L of Benzonase, and 5.5  $\mu$ L of  $1\times$  PBS with 0.001% P-188 were mixed to obtain a final volume of 30  $\mu$ L. Then, the sample was incubated at 37°C for 30 min. Thereafter, 30  $\mu$ L of DNase-treated AAVs were transferred to a new tube, to which 10  $\mu$ L of 500 mM EDTA, 55  $\mu$ L of  $1\times$  PBS with 0.001% P-188, and 5  $\mu$ L of Proteinase K (20 mg/mL) were added to obtain a final volume of 100  $\mu$ L. Next, the mixture was incubated at 55°C for 60 min, after which the mixture was heated at 95°C for 20 min, followed by centrifugation to collect the lysate. Then, the ssDNA was purified according to the protocol mentioned in the QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany) and used as the final collected sample. CGE was measured using a PA800Plus system (Sciex, Framingham, MA). The prepared samples were injected via electrokinetic injection. Detection was performed using a 488-nm laser excitation fluorescence with an emission filter of 520 nm.

### CGE for VP components

AAV samples for CGE measurement were prepared mostly in accordance with a previously reported procedure.<sup>29</sup> AAV solutions with a volume of 10  $\mu$ L ( $5.0 \times 10^{10}$  viral genomes) were denatured and buffer-exchanged following the protocol and the final collected sample was diluted with 50  $\mu$ L of deionized water for injection. CGE measurement was performed using a PA800Plus system (Sciex). Prepared samples were injected with water plug sample stacking. Detection was performed at 214 nm using a photo diode array detector.

### CDMS analysis with intact capsid

Prior to analysis by CDMS, 25  $\mu$ L of each sample was directly buffer-exchanged into 200 mM aqueous ammonium acetate (Invitrogen, AM9070G) with 0.01% pluronic F-68 (ThermoFisher Scientific, 24040032) using Micro Bio-Spin P-6 gel columns (Bio-Rad, 7326221).

Mass analysis was performed using a prototype Charge Detection-Mass Spectrometer with an electrostatic linear ion trap that is based on the system built by Megadalton Solutions, which has been described previously.<sup>47–50</sup> Ions were generated by positive mode

nanoelectrospray ionization using a TriVersa NanoMate (Advion, Ithaca, NY), equipped with a 5- $\mu$ m ID nozzle and standard A chip. The mass range of CDMS system was calibrated using L-glutamate dehydrogenase (GDH) from bovine liver (Sigma-Aldrich, G7882). To generate this calibration, the  $m/z$  peaks in the GDH spectra were plotted against the theoretical  $m/z$  values of GDH to give a correction factor that was used to accurately determine the mass and charge from the CDMS system. The NanoMate aspirated 5  $\mu$ L ( $5.0 \times 10^{12}$  vg/mL) of each sample to be infused into CDMS system with 1.75 kV applied to the nozzle. The spectra were collected until approximately 3,000 ions were captured within the mass range of 3–5 MDa, with a total acquisition time of approximately 10–12 min. Signal processing and data visualization were performed using prototype software developed in-house.

### LC-MS with intact VPs

rAAV samples were denatured with 10% acetic acid (FUJIFILM Wako Chemicals, Hyogo, Japan) and incubated at room temperature for 15 min. For the on-column denaturation experiment, AAV original solutions were directly used without acetic acid treatment. AAV samples were injected to Nexra HPLC (Shimadzu, Kyoto, Japan) coupled with a maXis II ETD ESI-QTOF mass spectrometer (Bruker). The separation was performed on an ACQUITY BEH C4 column (300  $\text{\AA}$ , 1.7  $\mu$ m, 2.1 mm  $\times$  150 mm; Waters, Milford, MA) at a flow rate of 0.2 mL/min and temperature of 80°C. Mobile phases A and B were 0.1% difluoroacetic acid (Waters) in MS-grade water (Kanto Kagaku Co., Ltd., Tokyo, Japan) and in acetonitrile (Thermo Fisher Scientific), respectively. The pH values of solution A and a mixture of 95% solution A and 5% solution B were measured as 1.93 and 1.90, respectively. Intact VPs were eluted from a 32% B to 36% B gradient in 15 min. UV adsorption at 280 nm and intrinsic fluorescence of aromatic amino acids with the excitation wavelength set at 280 nm and the emission wavelength set at 350 nm were used for detection of the LC chromatogram. Deconvolution analysis was performed in the mass range of 40,000–90,000 Da to determine the mass values of VP components in the eluted peaks. Visualization and processing of mass spectra were performed using Bruker Compass Data Analysis software ver. 5.1 (Bruker) and the maximum entropy was used for deconvolution.

### CsCl-DG-AUC

Samples were dissolved in a CsCl solution of the set prepared concentration in PBS with 0.001 w/v% poloxamer-188 as the solvent. Each rAAV stock was diluted to a final absorbance at a 1-cm path length of approximately 0.1 at 230 nm. A total of 390  $\mu$ L of the sample was loaded into the sample sector equipped with sapphire windows and a 12-mm double-sector charcoal-filled epon centerpiece (Beckman Coulter, USA). Further, 400  $\mu$ L of a corresponding CsCl solvent was loaded into each reference sector.

Data were collected at 20°C using Optima AUC (Beckman Coulter) at 42,000 rpm with a UV-Visible absorption detection system every hour for up to 72 h with a radial increment of 10  $\mu$ m.

### Determination of $T_m$

Structural stabilities of rAAV capsid were assessed by DSF using Uncle (Unchained Labs, Pleasanton, CA). Nine microliters of the rAAV2 samples were added to special measurement wells called Uni. Intrinsic fluorescence of aromatic amino acids (mainly from Trp<sup>51</sup>) excited by a 266-nm laser was measured as the temperature was increased at a rate of 1 °C/min. A redshift, involving the wavelength shifting to a greater value because of an environmental change of Trp residues from hydrophobic surroundings to hydrophilic ones during protein denaturation, was detected by plotting temperature on the y axis and barycentric mean (BCM) of the fluorescence from 300 to 430 nm on the x axis. The point at which the integral value of BCM peaked was defined as  $T_m$ .

### BS-AUC

BS-AUC experiments and analysis were performed in accordance with our previously reported method.<sup>31</sup> Briefly, AAV stocks were diluted to a final absorbance at a 1-cm path length of 0.25 at 260 nm. Then, 15  $\mu$ L of the AAV vector solution and solvent were loaded into a sample or reference reservoir well with a 12-mm band-forming centerpiece (Spin Analytical, Berwick, ME) equipped with sapphire windows. A volume of 240  $\mu$ L or 250  $\mu$ L of PBS/D<sub>2</sub>O with 0.001% w/v poloxamer-188 was loaded into the sample or reference sector, respectively. Data were collected at 20°C using the Optima AUC (Beckman Coulter) at 20,000 rpm with a UV detection system. Data were collected immediately with a radial increment of 10  $\mu$ m. The BS-AUC sedimentation data were analyzed using the analytical zone centrifugation  $c(s)$  model of the program SEDFIT (version 16.2b), where the lamella width, frictional ratio, meniscus, time-invariant noise, and radial-invariant noise were fitted using a regularization level of 0.68. The  $s$ -value range of 0–175 S was evaluated with a resolution of 350, and the buffer density and viscosity of PBS/H<sub>2</sub><sup>18</sup>O were calculated using the program SEDNTERP. The figures of the  $c(s)$  distribution were generated using the program GUSI (version 1.3.2). The  $s$ -value was described as  $sw$  ( $s$ -value as an apparent value under the experimental conditions).

### *In vitro* transduction efficacy assay

HeLaRC32 expressing the *Rep* and *Cap* genes for rAAV replication. After culture, HeLaRC cells were seeded in 24-well culture plates (Corning Inc., NY) at  $5 \times 10^4$  cells/well. The next day, cells were observed using a microscope to check that they were uniformly attached to the surface in all plates. Cells were infected with the AAV vectors at MOI  $1 \times 10^2$ ,  $2.5 \times 10^3$ ,  $1 \times 10^3$ ,  $2.5 \times 10^3$ , and  $5 \times 10^3$  in triplicate. The amounts of rAAV solutions were determined according to the calculated cell counts in three randomly selected wells. After 5 h, complete medium (DMEM, 10% FBS, 1% P/S) was added for further culture. At 72 h post-infection, cells were washed with DPBS (Wako, Osaka, Japan) and detached using TrypLE Select Enzyme (1X), with no phenol red (Thermo Fisher Scientific). Trypsin digestion was terminated by adding complete medium. Analysis of GFP expression was performed in a CytoFlex II Flow Cytometer (Beckman Coulter) using a fluorescein isothiocya-

nate (excitation: 498 nm, emission: 522 nm) channel with a threshold of 10,000 being used to identify GFP-positive cells.

### Real-time qPCR

rAAV vectors were quantified by qPCR in QuantStudio 3 Real-Time PCR System (Thermo Fisher Scientific) using AAVpro Titration Kit (for Real-Time PCR) Ver. 2 (Takara, Tokyo, Japan). Almost all steps were processed according to the manual. Briefly, DNase I treatment was performed to digest DNA outside the capsid, following heat inactivation of enzyme and capsid denaturation. The extracted viral genome was diluted and mixed with solution containing TB Green intercalating dye. A dilution series of the positive control ( $2 \times 10^7$  to  $2 \times 10^2$  copies/ $\mu$ L) was measured to generate a standard curve. Two minutes of initial denaturation at 95°C was followed by 35 cycles of denaturation at 95°C for 5 s and annealing at 60°C for 30 s, with final melting curve analysis. Viral titers were calculated in technical duplicates as genomic particles per milliliter.

### Productivity assay

WT rAAV and rAAV variants were produced in 20-mL flasks (Corning, Corning, NY, USA) using suspended HEK293T cells (three flasks per rAAV type). The cells were subjected to triple transfection according to established protocols. Then, the cells were allowed to grow for 24, 48, and 72 h in an incubator. At each time point, the cells were collected through centrifugation and the cell pellet was resuspended in AAV-MAX lysis buffer to obtain a lysate containing rAAV particles. Real-time qPCR was used to quantify the rAAV titer in each lysate. Finally, the productivity of WT and rAAV variants was compared based on their respective titers.

### Image processing of graphical abstract figure

The rAAV2 structure, PDB ID: 1LP3 (60-mer), was used as template. The visualization of X-ray crystallography maps and image making was performed using 3D Protein Imaging.<sup>52</sup>

### DATA AND CODE AVAILABILITY

Data will be made available on request.

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtm.2023.101142>.

### ACKNOWLEDGMENTS

The authors thank Ms. Toshie Kuwahara for the help with sample preparations. This work was supported by a grant-in-aid from the “Research and development of core technologies for gene and cell therapy” supported by the Japan Agency for Medical Research and Development (AMED) (JP20ae0201001 and JP20ae0201002) and JST, the establishment of university fellowships toward the creation of science technology innovation, Grant Number JPMJFS2125.

### AUTHOR CONTRIBUTIONS

Conceptualization, T.O., T.T., and S.U.; Investigation, T.O., M.N., T.M., M.F., M.M., S.U., A.H., K.R., K.G., and S.P.; Data curation,

T.O.; Resources, N.Y. and T.Oamasa; Writing – original draft, T.O. and S.U.; Writing – review & editing, Y.Y., T.T., and S.U.

## DECLARATION OF INTERESTS

The authors declare the following financial interests/personal relationships that may be considered as potential competing interests: T.M. and M.F. report a relationship with U-Medico Inc. that includes employment. M.M. reports a relationship with Osaka Consolidated Laboratory that includes employment. S.A., A.H., K.R., K.G., and S.P. report a relationship with Waters corporation that includes employment.

## REFERENCES

- Blair, H.A. (2022). Valoctocogene Roxaparovec: First Approval. *Drugs* 82, 1505–1510. <https://doi.org/10.1007/s40265-022-01788-y>.
- Deng, C., Zhao, P.Y., Branham, K., Schlegel, D., Fahim, A.T., Jayasundera, T.K., Khan, N., and Besirli, C.G. (2022). Real-world outcomes of voretigene neparovec treatment in pediatric patients with RPE65-associated Leber congenital amaurosis. *Graefes Arch. Clin. Exp. Ophthalmol.* 260, 1543–1550. <https://doi.org/10.1007/s00417-021-05508-2>.
- Keam, S.J. (2022). Eladocogene Exuparovec: First Approval. *Drugs* 82, 1427–1432. <https://doi.org/10.1007/s40265-022-01775-3>.
- Keeler, A.M., and Flotte, T.R. (2019). Recombinant Adeno-Associated Virus Gene Therapy in Light of Luxturna (and Zolgensma and Glybera): Where Are We, and How Did We Get Here? *Annu. Rev. Virol.* 6, 601–621. <https://doi.org/10.1146/annurev-virology-092818-015530>.
- Schwartz, M., Likhite, S., and Meyer, K. (2021). Onasemnogene abeparovec-xioi: a gene replacement strategy for the treatment of infants diagnosed with spinal muscular atrophy. *Drugs Today* 57, 387–399. <https://doi.org/10.1358/dot.2021.57.6.3264117>.
- Buller, R.M., and Rose, J.A. (1978). Characterization of adenovirus-associated virus-induced polypeptides in KB cells. *J. Virol.* 25, 331–338. <https://doi.org/10.1128/JVI.25.1.331-338.1978>.
- Johnson, F.B., Ozer, H.L., and Hoggan, M.D. (1971). Structural proteins of adenovirus-associated virus type 3. *J. Virol.* 8, 860–863. <https://doi.org/10.1128/JVI.8.6.860-863.1971>.
- Jin, X., Liu, L., Nass, S., O'Riordan, C., Pastor, E., and Zhang, X.K. (2017). Direct Liquid Chromatography/Mass Spectrometry Analysis for Complete Characterization of Recombinant Adeno-Associated Virus Capsid Proteins. *Hum. Gene Ther. Methods* 28, 255–267. <https://doi.org/10.1089/hgtb.2016.178>.
- Oyama, H., Ishii, K., Maruno, T., Torisu, T., and Uchiyama, S. (2021). Characterization of Adeno-Associated Virus Capsid Proteins with Two Types of VP3-Related Components by Capillary Gel Electrophoresis and Mass Spectrometry. *Hum. Gene Ther.* 32, 1403–1416. <https://doi.org/10.1089/hum.2021.009>.
- Sonntag, F., Bleker, S., Leuchs, B., Fischer, R., and Kleinschmidt, J.A. (2006). Adeno-associated virus type 2 capsids with externalized VP1/VP2 trafficking domains are generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus. *J. Virol.* 80, 11040–11054. <https://doi.org/10.1128/JVI.01056-06>.
- Venkatakrishnan, B., Yarbrough, J., Domsic, J., Bennett, A., Bothner, B., Kozyreva, O.G., Samulski, R.J., Muzyczka, N., McKenna, R., and Agbandje-McKenna, M. (2013). Structure and dynamics of adeno-associated virus serotype 1 VP1-unique N-terminal domain and its role in capsid trafficking. *J. Virol.* 87, 4974–4984. <https://doi.org/10.1128/JVI.02524-12>.
- Bleker, S., Sonntag, F., and Kleinschmidt, J.A. (2005). Mutational analysis of narrow pores at the fivefold symmetry axes of adeno-associated virus type 2 capsids reveals a dual role in genome packaging and activation of phospholipase A2 activity. *J. Virol.* 79, 2528–2540. <https://doi.org/10.1128/JVI.79.4.2528-2540.2005>.
- Girod, A., Wobus, C.E., Zádori, Z., Ried, M., Leike, K., Tijssen, P., Kleinschmidt, J.A., and Hallek, M. (2002). The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. *J. Gen. Virol.* 83, 973–978. <https://doi.org/10.1099/0022-1317-83-5-973>.
- Pierson, E.E., Keifer, D.Z., Asokan, A., and Jarrold, M.F. (2016). Resolving Adeno-Associated Viral Particle Diversity With Charge Detection Mass Spectrometry. *Anal. Chem.* 88, 6718–6725. <https://doi.org/10.1021/acs.analchem.6b00883>.
- Wörner, T.P., Bennett, A., Habka, S., Snijder, J., Friese, O., Powers, T., Agbandje-McKenna, M., and Heck, A.J.R. (2021). Adeno-associated virus capsid assembly is divergent and stochastic. *Nat. Commun.* 12, 1642. <https://doi.org/10.1038/s41467-021-21935-5>.
- Wang, Q., Firrman, J., Wu, Z., Pokiniewski, K.A., Valencia, C.A., Wang, H., Wei, H., Zhuang, Z., Liu, L., Wunder, S.L., et al. (2016). High-Density Recombinant Adeno-Associated Viral Particles are Competent Vectors for In Vivo Transduction. *Hum. Gene Ther.* 27, 971–981. <https://doi.org/10.1089/hum.2016.055>.
- Lipps, B.V., and Mayor, H.D. (1982). Characterization of heavy particles of adeno-associated virus type 1. *J. Gen. Virol.* 58, 63–72. <https://doi.org/10.1099/0022-1317-58-1-63>.
- Torikai, K., Ito, M., Jordan, L.E., and Mayor, H.D. (1970). Properties of light particles produced during growth of Type 4 adeno-associated satellite virus. *J. Virol.* 6, 363–369. <https://doi.org/10.1128/JVI.6.3.363-369.1970>.
- de la Maza, L.M., and Carter, B.J. (1980). Heavy and light particles of adeno-associated virus. *J. Virol.* 33, 1129–1137. <https://doi.org/10.1128/JVI.33.3.1129-1137.1980>.
- Grieger, J.C., Johnson, J.S., Gurda-Whitaker, B., Agbandje-McKenna, M., and Samulski, R.J. (2007). Surface-exposed adeno-associated virus VP1-NLS capsid fusion protein rescues infectivity of noninfectious wild-type Vp2/Vp3 and Vp3-only capsids but not that of fivefold pore mutant virions. *J. Virol.* 81, 7833–7843. <https://doi.org/10.1128/JVI.00580-07>.
- Grieger, J.C., Snowdy, S., and Samulski, R.J. (2006). Separate basic region motifs within the adeno-associated virus capsid proteins are essential for infectivity and assembly. *J. Virol.* 80, 5199–5210. <https://doi.org/10.1128/JVI.02723-05>.
- Xiao, P.J., and Samulski, R.J. (2012). Cytoplasmic trafficking, endosomal escape, and perinuclear accumulation of adeno-associated virus type 2 particles are facilitated by microtubule network. *J. Virol.* 86, 10462–10473. <https://doi.org/10.1128/JVI.00935-12>.
- Popa-Wagner, R., Porwal, M., Kann, M., Reuss, M., Weimer, M., Florin, L., and Kleinschmidt, J.A. (2012). Impact of VP1-specific protein sequence motifs on adeno-associated virus type 2 intracellular trafficking and nuclear entry. *J. Virol.* 86, 9163–9174. <https://doi.org/10.1128/JVI.00282-12>.
- Bosma, B., du Plessis, F., Ehlert, E., Nijmeijer, B., de Haan, M., Petry, H., and Lubelski, J. (2018). Optimization of viral protein ratios for production of rAAV serotype 5 in the baculovirus system. *Gene Ther.* 25, 415–424. <https://doi.org/10.1038/s41434-018-0034-7>.
- Mietzsch, M., Grasse, S., Zurawski, C., Weger, S., Bennett, A., Agbandje-McKenna, M., Muzyczka, N., Zolotukhin, S., and Heilbronn, R. (2014). OneBac: platform for scalable and high-titer production of adeno-associated virus serotype 1–12 vectors for gene therapy. *Hum. Gene Ther.* 25, 212–222. <https://doi.org/10.1089/hum.2013.184>.
- Srivastava, A., Mallela, K.M.G., Deorkar, N., and Brophy, G. (2021). Manufacturing Challenges and Rational Formulation Development for AAV Viral Vectors. *J. Pharmaceut. Sci.* 110, 2609–2624. <https://doi.org/10.1016/j.xphs.2021.03.024>.
- Florea, M., Nicolaou, F., Pacouret, S., Zinn, E.M., Sanmiguel, J., Andres-Mateos, E., Unzu, C., Wagers, A.J., and Vandenberghe, L.H. (2023). High-efficiency purification of divergent AAV serotypes using AAVX affinity chromatography. *Mol. Ther. Methods Clin. Dev.* 28, 146–159. <https://doi.org/10.1016/j.omtm.2022.12.009>.
- Lechner, A., Giorgetti, J., Gahoual, R., Beck, A., Leize-Wagner, E., and François, Y.N. (2019). Insights from capillary electrophoresis approaches for characterization of monoclonal antibodies and antibody drug conjugates in the period 2016–2018. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* 1122–1123, 1–17. <https://doi.org/10.1016/j.jchromb.2019.05.014>.
- Zhang, C.X., and Meagher, M.M. (2019). Highly Sensitive SDS Capillary Gel Electrophoresis with Sample Stacking Requiring Only Nanograms of Adeno-Associated Virus Capsid Proteins. *Methods Mol. Biol.* 1972, 263–270. [https://doi.org/10.1007/978-1-4939-9213-3\\_19](https://doi.org/10.1007/978-1-4939-9213-3_19).

30. Kuipers, B.J.H., and Gruppen, H. (2007). Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis. *J. Agric. Food Chem.* 55, 5445–5451. <https://doi.org/10.1021/jf070337l>.
31. Maruno, T., Ishii, K., Torisu, T., and Uchiyama, S. (2023). Size Distribution Analysis of the Adeno-Associated Virus Vector by the c(s) Analysis of Band Sedimentation Analytical Ultracentrifugation with Multiwavelength Detection. *J. Pharmaceut. Sci.* 112, 937–946. <https://doi.org/10.1016/j.xphs.2022.10.023>.
32. Jarrold, M.F. (2022). Applications of Charge Detection Mass Spectrometry in Molecular Biology and Biotechnology. *Chem. Rev.* 122, 7415–7441. <https://doi.org/10.1021/acs.chemrev.1c00377>.
33. Barnes, L.F., Draper, B.E., and Jarrold, M.F. (2022). Analysis of thermally driven structural changes, genome release, disassembly, and aggregation of recombinant AAV by CDMS. *Mol. Ther. Methods Clin. Dev.* 27, 327–336. <https://doi.org/10.1016/j.omtm.2022.10.008>.
34. Warrington, K.H., Jr., Gorbatyuk, O.S., Harrison, J.K., Opie, S.R., Zolotukhin, S., and Muzyczka, N. (2004). Adeno-associated virus type 2 VP2 capsid protein is nonessential and can tolerate large peptide insertions at its N terminus. *J. Virol.* 78, 6595–6609. <https://doi.org/10.1128/JVI.78.12.6595-6609.2004>.
35. de la Maza, L.M., and Carter, B.J. (1980). Molecular Structure of Adeno-associated Virus Variant DNA. *J. Biol. Chem.* 255, 3194–3203.
36. Pronovost, A.D., Yates, V.J., and Fry, D.E. (1979). Inhibition and enhancement of avian adenovirus plaque production by heavy and light avian adenovirus associated viral particles. *Am. J. Vet. Res.* 40, 549–551.
37. Paradiso, P.R. (1981). Infectious process of the parvovirus H-1: correlation of protein content, particle density, and viral infectivity. *J. Virol.* 39, 800–807. <https://doi.org/10.1128/JVI.39.3.800-807.1981>.
38. Clinton, G.M., and Hayashi, M. (1976). The parvovirus MVM: a comparison of heavy and light particle infectivity and their density conversion in vitro. *Virology* 74, 57–63. [https://doi.org/10.1016/0042-6822\(76\)90127-6](https://doi.org/10.1016/0042-6822(76)90127-6).
39. Ferreira, J.P., Noderer, W.L., Diaz de Arce, A.J., and Wang, C.L. (2014). Engineering ribosomal leaky scanning and upstream open reading frames for precise control of protein translation. *Bioengineered* 5, 186–192. <https://doi.org/10.4161/bioe.27607>.
40. Hinnebusch, A.G. (2011). Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol. Mol. Biol. Rev.* 75, 434–467. first page of table of contents. <https://doi.org/10.1128/MMBR.00008-11>.
41. Ichihara, K., Matsumoto, A., Nishida, H., Kito, Y., Shimizu, H., Shichino, Y., Iwasaki, S., Imami, K., Ishihama, Y., and Nakayama, K.I. (2021). Combinatorial analysis of translation dynamics reveals eIF2 dependence of translation initiation at near-cognate codons. *Nucleic Acids Res.* 49, 7298–7317. <https://doi.org/10.1093/nar/gkab549>.
42. Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene* 234, 187–208. [https://doi.org/10.1016/s0378-1119\(99\)00210-3](https://doi.org/10.1016/s0378-1119(99)00210-3).
43. Becerra, S.P., Rose, J.A., Hardy, M., Baroudy, B.M., and Anderson, C.W. (1985). Direct mapping of adeno-associated virus capsid proteins B and C: a possible ACG initiation codon. *Proc. Natl. Acad. Sci. USA* 82, 7919–7923. <https://doi.org/10.1073/pnas.82.23.7919>.
44. Firth, A.E., and Brierley, I. (2012). Non-canonical translation in RNA viruses. *J. Gen. Virol.* 93, 1385–1409. <https://doi.org/10.1099/vir.0.042499-0>.
45. Kohlbrenner, E., Aslanidi, G., Nash, K., Shklyae, S., Campbell-Thompson, M., Byrne, B.J., Snyder, R.O., Muzyczka, N., Warrington, K.H., Jr., and Zolotukhin, S. (2005). Successful production of pseudotyped rAAV vectors using a modified baculovirus expression system. *Mol. Ther.* 12, 1217–1225. <https://doi.org/10.1016/j.ymthe.2005.08.018>.
46. Mietzsch, M., Casteleyn, V., Weger, S., Zolotukhin, S., and Heilbronn, R. (2015). OneBac 2.0: Sf9 Cell Lines for Production of AAV5 Vectors with Enhanced Infectivity and Minimal Encapsulation of Foreign DNA. *Hum. Gene Ther.* 26, 688–697. <https://doi.org/10.1089/hum.2015.050>.
47. Draper, B.E., and Jarrold, M.F. (2019). Real-Time Analysis and Signal Optimization for Charge Detection Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 30, 898–904. <https://doi.org/10.1007/s13361-019-02172-z>.
48. Todd, A.R., and Jarrold, M.F. (2020). Dynamic Calibration Enables High-Accuracy Charge Measurements on Individual Ions for Charge Detection Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 31, 1241–1248. <https://doi.org/10.1021/jasms.0c00081>.
49. Todd, A.R., Alexander, A.W., and Jarrold, M.F. (2020). Implementation of a Charge-Sensitive Amplifier without a Feedback Resistor for Charge Detection Mass Spectrometry Reduces Noise and Enables Detection of Individual Ions Carrying a Single Charge. *J. Am. Soc. Mass Spectrom.* 31, 146–154. <https://doi.org/10.1021/jasms.9b00010>.
50. Hogan, J.A., and Jarrold, M.F. (2018). Optimized Electrostatic Linear Ion Trap for Charge Detection Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 29, 2086–2095. <https://doi.org/10.1007/s13361-018-2007-x>.
51. Ghisaidoobe, A.B.T., and Chung, S.J. (2014). Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Forster resonance energy transfer techniques. *Int. J. Mol. Sci.* 15, 22518–22538. <https://doi.org/10.3390/ijms151222518>.
52. Tomasello, G., Armenia, I., and Molla, G. (2020). The Protein Imager: a full-featured online molecular viewer interface with server-side HQ-rendering capabilities. *Bioinformatics* 36, 2909–2911. <https://doi.org/10.1093/bioinformatics/btaa009>.

**Supplemental information**

**Enhancement of recombinant adeno-associated  
virus activity by improved stoichiometry  
and homogeneity of capsid protein assembly**

**Takayuki Onishi, Michika Nonaka, Takahiro Maruno, Yuki Yamaguchi, Mitsuko Fukuhara, Tetsuo Torisu, Masaharu Maeda, Susan Abbatiello, Anisha Haris, Keith Richardson, Kevin Giles, Steve Preece, Noriko Yamano-Adachi, Takeshi Omasa, and Susumu Uchiyama**

### **Molecular weight (Mw) estimation**

The theoretical Mw values of AAV2-Empty were determined from the amino acid composition of capsid protein determined by CGE measurement using SEDNTERP. Furthermore, the theoretical Mw values of ssDNA were calculated from the DNA compositions.

In BS-AUC experiment, the Mw values of AAV2-Empty and AAV2-Full were calculated using the Svedberg equation:

$$\frac{s}{D} = \frac{M(1-\bar{v}\rho)}{RT} \quad (\text{Equation S1})$$

where  $s$  is the sedimentation coefficient,  $D$  is the diffusion coefficient,  $M$  is the Mw,  $\bar{v}$  is the particle partial-specific volume,  $\rho$  is the solvent density,  $R$  is the gas constant, and  $T$  is the absolute temperature.

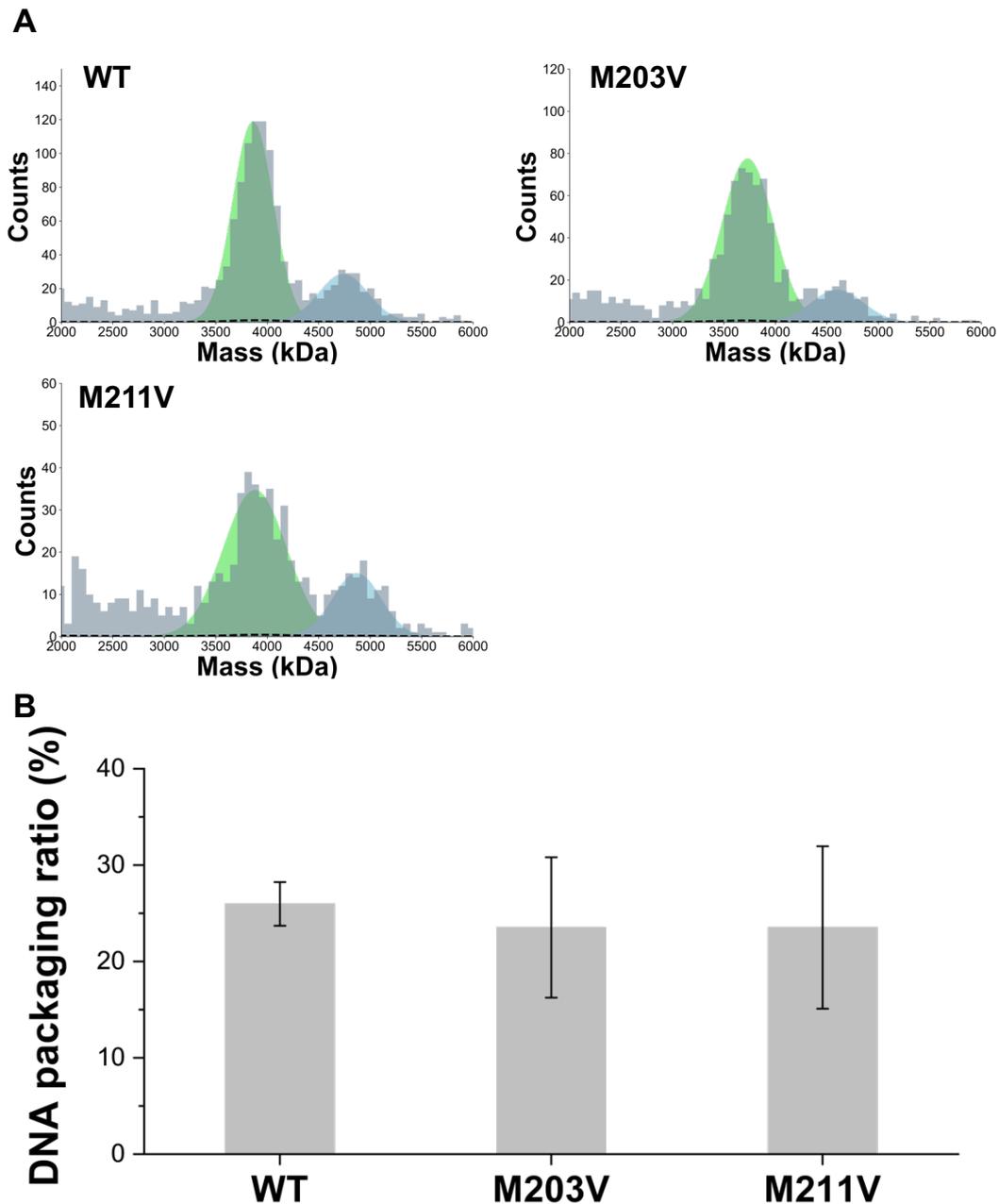
### **Partial specific volume ( $\bar{v}$ ) calculation**

The  $\bar{v}$  values of AAV-Empty were calculated from the amino acid composition of capsid protein determined by CGE measurement.

The  $\bar{v}$  values of F1.1 and F2.2 in water can be theoretically calculated using the following equation:

$$\bar{v}_{AAV-FP} = \frac{\bar{v}_{AAV-Empty} \times M_{AAV-Empty} + \bar{v}_{ssDNA} \times M_{ssDNA}}{M_{AAV-Empty} + M_{ssDNA}} \quad (\text{Equation S2})$$

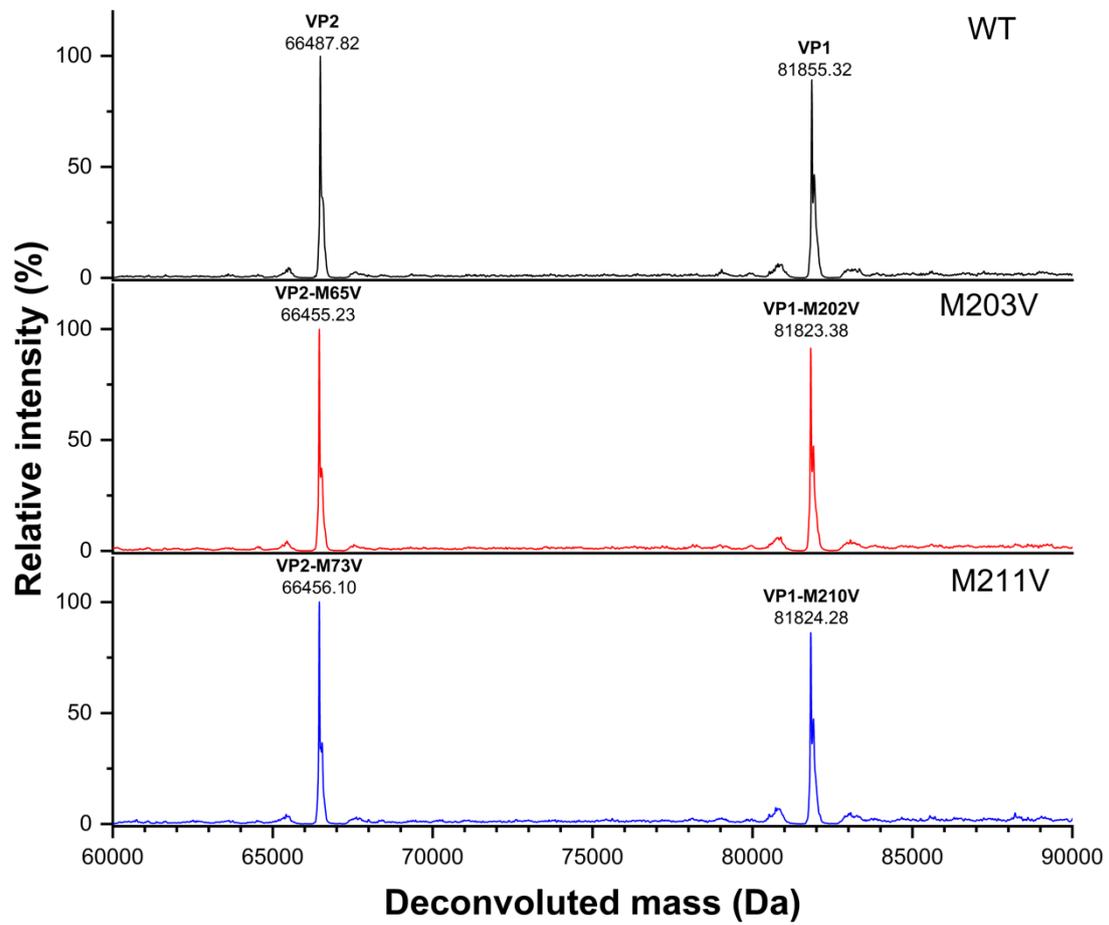
where  $\bar{v}_{AAV-F1.1 \text{ or } F2.2}$ ,  $\bar{v}_{AAV-Empty}$ , and  $\bar{v}_{ssDNA}$  are the partial-specific volume of AAV-F1.1 or F2.2, AAV-Empty, and ssDNA, respectively.  $0.52 \text{ cm}^3 \text{ g}^{-1}$  was used as  $\bar{v}_{ssDNA}$  based on the previous study.<sup>5</sup>  $M_{AAV-F1.1 \text{ or } F2.2}$ ,  $M_{AAV-Empty}$ , and  $M_{ssDNA}$  are the Mw of AAV-FP1.1 or AAV-FP2.2, AAV-EP, and ssDNA, respectively. , using partial specific volumes of  $0.6840 \text{ cm}^3/\text{g}$  and  $0.6837 \text{ cm}^3/\text{g}$  (Equation S2), respectively, and a frictional ratio ( $f/f_0$ ) of 1.46 for both particles.



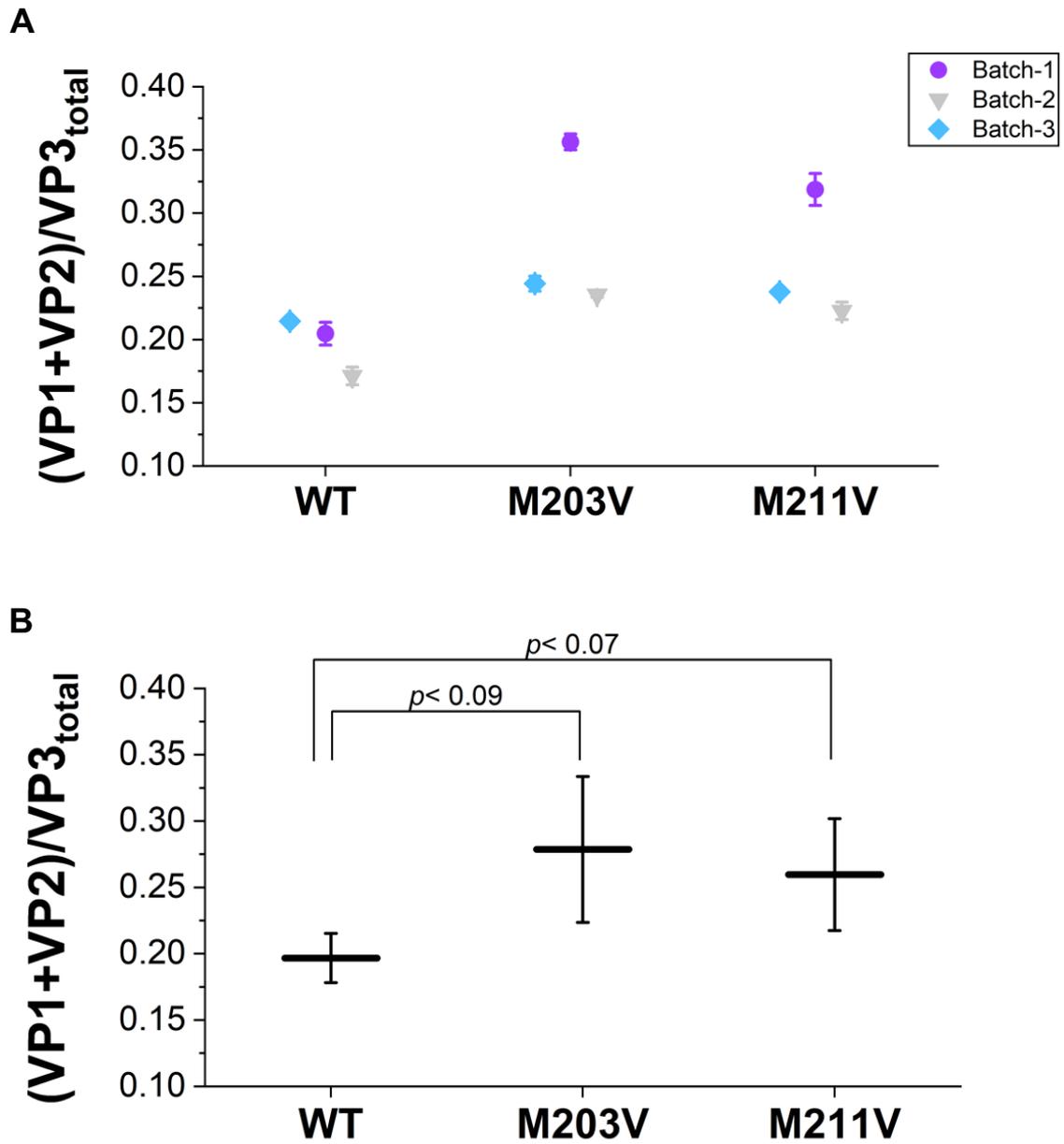
**Figure S1. DNA packaging ratio of rAAV-WT, -M203V and -M211V evaluated by mass photometry (MP).**

DNA packaging ratio before centrifugal purification were measured by MP. MP experiments were conducted using the TwoMP instrument (Refeyn, Oxford, UK).

- A)** One of the representative mass distributions of rAAV-WT, -M203V and -M211V. Show Gaussian fitting for empty particle in green for full particle in blue.
- B)** DNA packaging ratio of rAAV-WT, -M203V and -M211V.



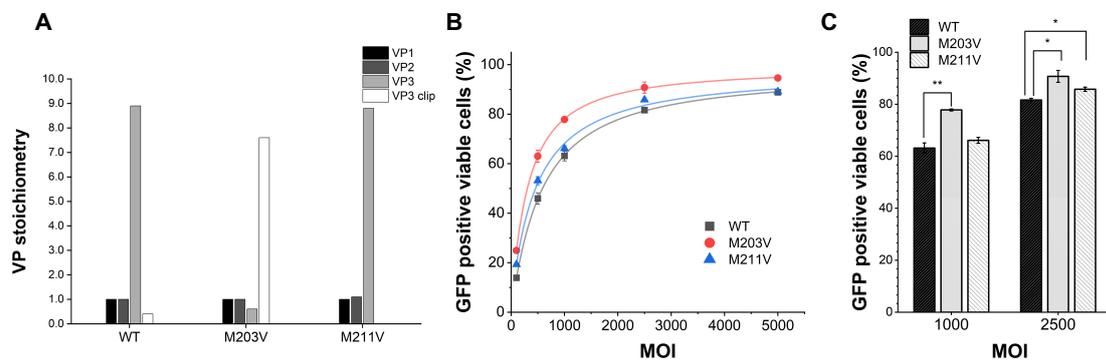
**Figure S2. Deconvoluted mass values of WT, M203V, and M211V for VP1 and VP2 by liquid chromatography-mass spectrometry (LC-MS).**



**Figure S3. Variability of VP stoichiometry of rAAV2-WT, -M203V, -M211V.**

A) Variability of VP stoichiometry for rAAV2-WT, -M203V, -M211V in the three different batches.

B) Means and standard deviations in the three different batches.



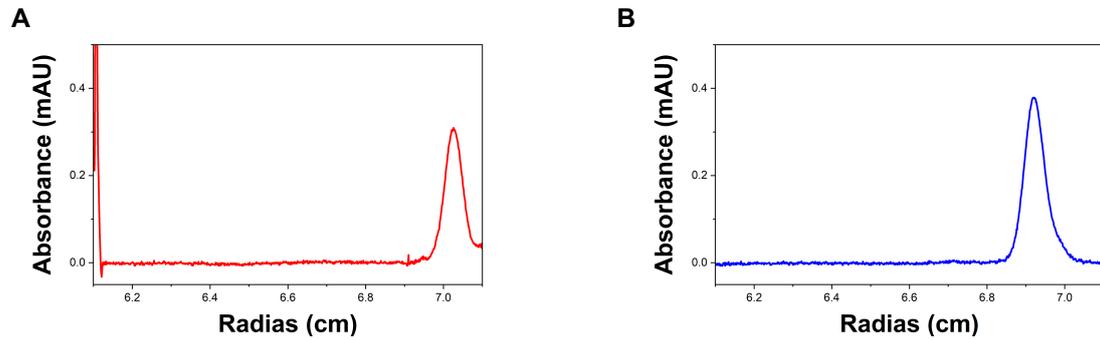
**Figure S4. Characterization of rAAV2-WT, -M203V and -M211V on differential transgene.**

rAAV2-M203V and -M211V with the transgene changed to CMV-zsGreen1 were prepared and evaluated VP stoichiometry and *In vitro* transduction efficacy.

**A)** VP stoichiometry of WT, M203V and M211V.

**B)** GFP positive viable cells were measured by fluor cytometer at five points of MOI (MOI  $1 \times 10^2$ ,  $5 \times 10^2$ ,  $10 \times 10^2$ ,  $25 \times 10^2$  and  $50 \times 10^2$ ).

**C)** *In vitro* transduction efficacy of selected MOI  $10 \times 10^2$  and  $25 \times 10^2$ .

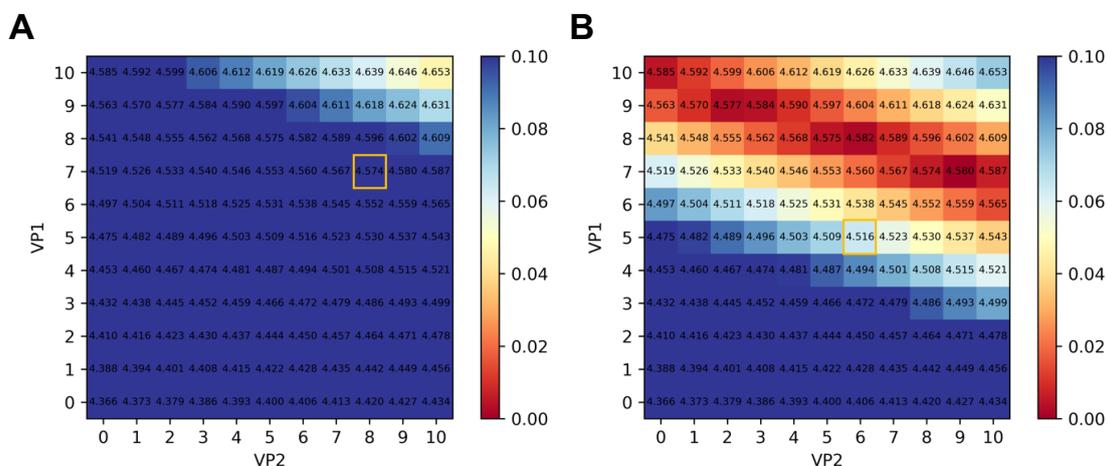


**Figure S5. CsCl-density gradient analytical ultracentrifugation (CsCl-DG-AUC) profile of rAAV2-M203V and -M211V on differential transgene.**

A-B were CsCl-DG-AUC profile of rAAV2 variants packaged zsGreen1 gene downstream of cytomegalovirus promoter (CMV-zsGreen1), and the single distribution was observed in both variants.

**A)** CsCl-DG-AUC profile of rAAV2-CMV-zsGreen1-M203V.

**B)** CsCl-DG-AUC profile of rAAV2-CMV-zsGreen1-M211V.

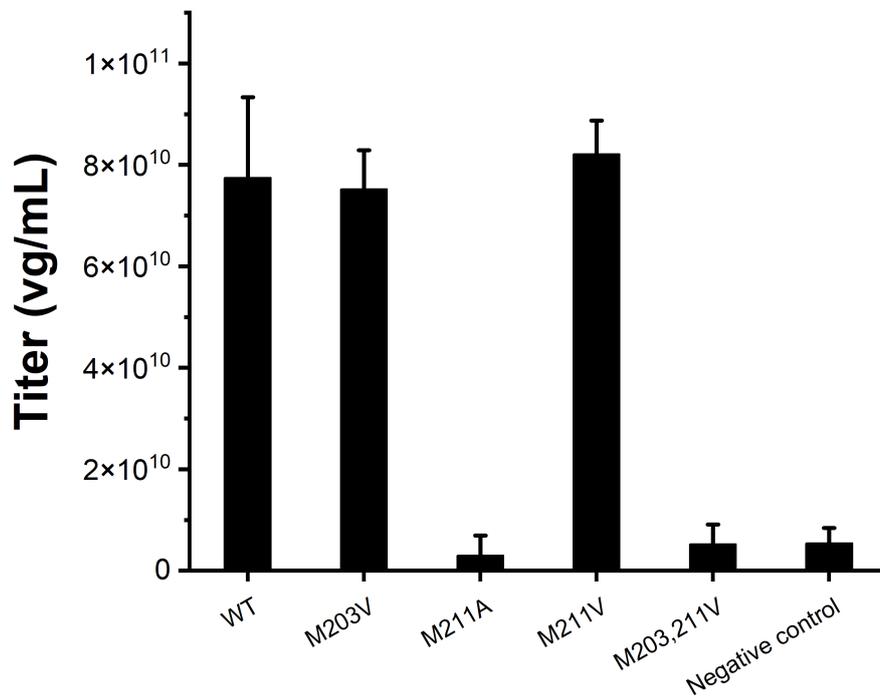


**Figure S6. Estimation of VP combination in 60-mer based on particle mass by CD-MS.**

Heat maps of VP matrix-mass value were created to assess the correspondence of VP1 and VP2 stoichiometry between theoretical mass values and measured mass values for particles from CD-MS analysis. Calculated mass value based on CGE experiments were also indicated.

The theoretical mass values of the corresponding VP combination were indicated in each cell, and the color chart indicates the difference from the CD-MS mass value. The mass values of all VP combinations were calculated as a 60-mer and all components except VP1 and VP2 were calculated as VP3. Yellow squares indicate mass value using VP1 and VP2 stoichiometry obtained from CGE measurement.

- A) Heat map of mass difference between CD-MS mass value and each VP combination for F1.1 particles.
- B) Heat map of mass difference between CD-MS mass value and each VP combination for F2.2 particles.



**Figure S7. Particle assembly assessment.**

To evaluate rAAV particle formation, cell lysates at 72 hours post-transfection were assayed by qPCR. Cell lysates transfected with plasmids except for the pHelper plasmid were used as negative controls, considering the possibility of insufficient digestion with Benzonase prior to qPCR measurement. Particle formation was considered to have failed in cases of titers comparable to those of the negative control.

**Table S1. Comparison of theoretical and measured molecular mass.**

	<b>VP ratio from CGE (VP1:VP2:VP3:VP3clip)</b>	<b>Theoretical mass from CGE result* (MDa)</b>	<b>Molecular mass from CDMS (MDa)</b>	<b>Molecular weight from BS-AUC (MDa)</b>
<b>AAV</b>	1.0 : 1.0 : 10.0 :-	4.54	-	-
<b>F1.1</b>	1.0 : 1.3 : 6.9 : 0.4	4.56	4.70	5.24
<b>F2.2</b>	1.0 : 1.4 : 8.6 : 0.4	4.53	4.58	5.11
<b>rAAV2-WT</b>	1.0 : 1.3 : 10.9 : 0.4	4.54	4.75	-
<b>rAAV2-M203V</b>	1.0 : 1.8 : 0.5 : 7.5	4.56	4.73	-
<b>rAAV2-M211V</b>	1.0 : 1.9 : 9.1 : 0	4.57	4.71	-

\*Total VP number was assumed to be 60 in a single AAV particle.

\*Molecular weights used for the calculation of theoretical mass value from amino acid sequence: VP1: 81855.24, VP2: 66488.20, VP3: 59974.02, VP3<sub>clip</sub>: 59301.27 for wild-type, VP1: 81823.1875, VP2: 66456.14, VP3: 59974.02, VP3<sub>clip</sub>: 59301.27 for rAAV2-M203V, VP1: 81823.1875, VP2: 66456.14, VP3: 59941.96 for rAAV2-M211V and the calculation of theoretical mass value from nucleic acid sequence: ITR-CMV-EGFP-ITR (2,521 bp): 777502.35.

**Table S2. Estimated amino acid sequence and mass accuracy by LC-MS measurements.**

	<b>Estimated amino acid sequence</b>	<b>Observed mass (Da)</b>	<b>Theoretical mass (Da)</b>	<b>Mass accuracy (ppm)</b>
<b>AAV2-WT</b>	VP1 (A2(Ac)-L735)	81855.32	81855.24	0.95
	VP2 (A139-L735)	66487.82	66488.20	5.76
	VP3 (A204(Ac)-L735)	59973.14	59974.02	14.73
	VP3 clip (A212(Ac)-L735)	59299.79	59301.27	24.95
<b>AAV2-M203V</b>	VP1-M202V (A2(Ac)-M203V-L735)	81823.38	81823.19	2.35
	VP2-M65V (A139-M203V-L735)	66455.23	66456.14	13.70
	VP3 (A204(Ac)-L735)	59972.82	59974.02	20.07
	VP3 clip (A212(Ac)-L735)	59300.22	59301.27	17.70
<b>AAV2-M211V</b>	VP1-M210V (A2(Ac)-M211V-L735)	81824.28	81823.19	13.35
	VP2-M73V (A139-M211V-L735)	66456.10	66456.14	0.61
	VP3-M8V (A204(Ac)-M211V-L735)	59941.72	59941.96	4.08
	VP3 clip (A212(Ac)-L735)	ND	-	-