# Frustration and Direct-Coupling Analyses to Predict Formation and Function of Adeno-Associated Virus

Nicole N. Thadani,[1] Qin Zhou,[2] Kiara Reyes Gamas,[1] Susan Butler,[1] Carlos Bueno,[3,9] Nicholas P. Schafer,[3,4] Faruck Morcos,[2,5,6] Peter G. Wolynes,[3,4,7,8] and Junghae Suh[1,7,9,10,*]

[1]Department of Bioengineering, Rice University, Houston, Texas; [2]Department of Biological Sciences, University of Texas at Dallas, Richardson, Texas; [3]Center for Theoretical Biological Physics and [4]Department of Chemistry, Rice University, Houston, Texas; and [5]Center for Systems Biology and [6]Department of Bioengineering, University of Texas at Dallas, Richardson, Texas; and [7]Department of Biosciences, [8]Department of Physics, [9]Department of Chemical and Biomolecular Engineering, and [10]Systems, Synthetic, and Physical Biology Program, Rice University, Houston, Texas

ABSTRACT Adeno-associated virus (AAV) is a promising gene therapy vector because of its efficient gene delivery and relatively mild immunogenicity. To improve delivery target specificity, researchers use combinatorial and rational library design strategies to generate novel AAV capsid variants. These approaches frequently propose high proportions of nonforming or noninfective capsid protein sequences that reduce the effective depth of synthesized vector DNA libraries, thereby raising the discovery cost of novel vectors. We evaluated two computational techniques for their ability to estimate the impact of residue mutations on AAV capsid protein-protein interactions and thus predict changes in vector fitness, reasoning that these approaches might inform the design of functionally enriched AAV libraries and accelerate therapeutic candidate identification. The Frustratometer computes an energy function derived from the energy landscape theory of protein folding. Direct-coupling analysis (DCA) is a statistical framework that captures residue coevolution within proteins. We applied the Frustratometer to select candidate protein residues predicted to favor assembled or disassembled capsid states, then predicted mutation effects at these sites using the Frustratometer and DCA. Capsid mutants were experimentally assessed for changes in virus formation, stability, and transduction ability. The Frustratometer-based metric showed a counterintuitive correlation with viral stability, whereas a DCA-derived metric was highly correlated with virus transduction ability in the small population of residues studied. Our results suggest that coevolutionary models may be able to elucidate complex capsid residue-residue interaction networks essential for viral function, but further study is needed to understand the relationship between protein energy simulations and viral capsid metastability.

SIGNIFICANCE Adeno-associated virus is one of the most promising gene delivery vectors today and has been approved by the U.S. Food and Drug Administration for gene replacement therapies. Despite clinical advances, improvements to vector design are sorely needed to address many other unmet medical needs. Inspired by adjacent fields of research, in which models of protein fitness derived from molecular dynamics approaches and from sequence family modeling have been used to design protein variants, we explored the potential of these models for predicting virus formation and function. A reliable and predictive computational tool would be a tremendous addition to the vector development process, enabling the generation of functional adeno-associated virus variant panels at a faster pace and at lower cost.

## INTRODUCTION

Adeno-associated virus (AAV) is widely favored as a gene therapy vector and, to date, has been tested in over 200 human clinical trials internationally. AAV was recently approved by the U.S. Food and Drug Administration as the first gene replacement therapy for an inherited genetic disorder, in part because of the vector's nonpathogenicity and generally benign safety profile ([1],[2]). AAV is an efficient delivery vector, capable of transducing both dividing and nondividing cells to produce sustained gene expression ([3],[4]). However, further improvements to the vector are highly sought after to enhance delivery efficiency in targeted cell populations, which would minimize the vector

doses required and reduce vector production costs as well as safety concerns because of dose-dependent immune responses to the vector (5–8).

As a result of successful AAV biomining efforts (9–11), gene therapy developers have had access to hundreds of isolated AAV variants. Researchers have then employed a variety of design strategies to further engineer these AAV vectors for enhanced gene delivery. The three main approaches that have emerged are 1) rational design, 2) directed evolution, and 3) computationally driven design. Rational design approaches use current knowledge of vector biology to make targeted modifications to the capsid, whereas directed evolution conducts rounds of mutation and selection to explore the capsid functional space (12,13). Bioinformatics-driven strategies have recently emerged as a method, drawing upon large data sets of AAV capsid information and applying computational models to accelerate exploration of the viral fitness landscape (14). Initial studies in this space have yielded promising results, including AAV chimera populations designed using the SCHEMA algorithm to minimize structural disruption (15,16), mapping of AAV capsid amino acids important for structure and function using high-throughput studies (17,18), and the development of ancestral AAV vectors (19). Here, we explore data-driven computational approaches with varying degrees of complexity for predicting the fitness of capsid mutants, with the future goal of incorporating these in silico approaches into the AAV vector design pipeline.

The Frustratometer approach described in Ferreiro et al. (20), the first of the two computational approaches that we applied, is based on the energy landscape theory of protein folding (21). Naturally occurring proteins generally have a smooth, funnel-like energy landscape with minimal kinetic traps, promoting robust and rapid folding into a single native structure. The shape of the landscape arose and is maintained through evolutionary optimization by strengthening residue-residue contacts in the native structure and reducing the number of energetically favorable contacts outside of the native structure. The phenomenon of the evolutionary optimization of protein energy landscapes is known as the "principle of minimal frustration" (22,23). Residual localized frustration within protein structures may be retained during evolution to facilitate protein allostery and multimeric interactions (24). Previous studies have shown that the Frustratometer, which can be used to identify these localized patches of frustration, can provide insight into protein conformational changes and association. Frustration analysis of proteins that exhibit allostery indicates that known regions of conformational cracking are highly frustrated (23). The associative memory, water-mediated, structure and energy model (AWSEM), the optimized coarse-grained model that underlies the Frustratometer, has accurately predicted homodimer and heterodimer docking sites

(25). Frustration analysis has also been applied fruitfully to the problem of mapping out structural reaction mechanisms (26). Here, we apply the Frustratometer to evaluate the change upon binding of the frustration level of residue-residue interactions that are formed during AAV capsid assembly.

Coevolutionary modeling, alternatively, seeks to identify residues with a structural or functional relationship in a protein or pair of proteins through analysis of the protein sequence family, which contains a record of the evolutionary constraints on protein function. Direct-coupling analysis (DCA) has emerged as a prominent statistical framework for identifying direct interactions of protein residues from genomic data (27,28). When applied to a family of protein sequences, this model can predict crucial pairs of interactions for protein structure and function and identify the effect of mutations by inferring the probability that a mutated sequence belongs to the family (29,30). DCA has accurately recapitulated known structural contacts in a variety of proteins and intermediate protein structures (27,28,31). DCA approaches were initially applied to the characterization of bacterial two-component system signaling networks and the design of libraries predicted to enhance signaling (32–34). More recently, DCA approaches have shown promise in identifying interacting residues in protein oligomers (35–37). Coevolutionary approaches previously applied to viruses have correctly identified interactions between viral proteins, but these models have not been evaluated for their ability to predict mutant AAV functionality (38–40). We aimed to evaluate DCA's ability to capture patterns that identify functional members of the parvovirus family by predicting the formation and function of AAV mutants.

Here, we applied the Frustratometer to AAV2 monomer and assembled capsid structures to identify residues with large shifts in frustration index between these two states. We then mutated these residues to alanine and experimentally evaluated capsid assembly, thermal stability, and transduction efficiency. We compared these experimental results to predictions from the Frustratometer and DCA models. The frustration approach shows limited promise for predicting aspects of viral capsid structure and produced counterintuitive results with regards to stability. DCA supports the possibility of predicting capsid transduction ability in this small study.

## MATERIALS AND METHODS

### Identifying residue contacts within and between AAV2 capsid subunits

The high-resolution crystal structure of AAV2 (Protein Data Bank, PDB: 1LP3) was visualized and analyzed in PyMol 2.2.0. Inter-residue contacts were identified as any pair of residues containing atoms within 4.5 Å of each other. Intrasubunit contacts were identified within an AAV2 subunit, and extrasubunit contacts were identified between AAV2 subunits.

## Sequence conservation in the parvovirus sequence family

We used the hidden Markov model (HMM) from Pfam parvovirus coat protein VP2 family (PF00740) (Data S1) to search the UniProt database and obtained a multiple sequence alignment (MSA) with 8904 protein sequences (Data S2). Sequences with gaps greater than 50 amino acids were removed, leading to an MSA of 2569 sequences, and sequences with greater than 90% identity were reweighted as described previously (28), leaving 129.77 effective sequences. Conservation was then computed for all positions modeled in the Pfam HMM using Shannon entropy and reweighted from 0 to 1 so that a value of 1 indicates full conservation (41).

## Frustratometer analysis of AAV capsid subunits

Using the crystal structure of AAV2, two types of constructs were created for analysis: a monomer form, consisting of a single AAV2 capsid subunit, and an assembly form, consisting of a central monomer and each monomer containing a residue with any atom within a radius of 4.5 Å around any of the central monomer residues. The AWSEM-MD Frustratometer Server with default parameters (sequence separation = 12 and no electrostatics) was used to calculate the single-residue frustration index (F) of the monomer and assembly structures (42). The $\Delta$-frustration index was then calculated for each residue as $F_{assembly} - F_{monomer}$ (20).

## Computational alanine mutagenesis using AWSEM

To calculate the predicted change in energy between the monomer and multimer states for alanine mutants, we applied the AWSEM, a predictive protein coarse-grained model that combines terms from energy landscape theory and information from a database of known protein structures. The physical portion of the model focuses on hydrogen bonding, hydrophobic interactions, and water-mediated interactions between hydrophilic residues. Bioinformatic data are used to bias the structure of sequence fragments (nine residues or less) toward the structure of similar sequence fragments found in other proteins. The energy function of this model is described in Eq. 1 (43).

$$V_{AWSEM} = V_{backbone} + V_{contact} + V_{burial} + V_{helical} + V_{FM} \quad (1)$$

Then, the change in energy between alanine mutants and wild type (WT) was computed for monomer and multimer structures as described in Eq. 2.

$$\Delta E = V_{AWSEM Ala} - V_{AWSEM WT} \quad (2)$$

The $\Delta\Delta E$ of alanine substitution was then computed as described in Eq. 3.

$$\Delta\Delta E = \Delta E_{Multimer} - \Delta E_{Monomer} \quad (3)$$

## Alanine frequency in the parvovirus sequence family

The parvovirus sequence family was processed and reweighted as described for the sequence conservation calculation. The frequency of alanine at each position represented in the Pfam HMM was then calculated.

## DCA of AAV2 capsid

DCA was applied to the preprocessed alignment of the parvovirus sequence family described in the sequence conservation calculation to infer a global statistical model of coevolved residue interactions. DCA infers a joint probability distribution to satisfy the statistical observations of protein family sequence, with parameters that include pairwise couplings $e_{ij}(x_i, x_j)$ and local biases (fields) $h_i(x_i)$ (28). Pairwise couplings may be restricted to occur within a certain spatial residue distance based on the structure of a given member of the sequence family (29). From this distribution, direct information values may be computed to quantify how two sites in the protein are directly coupled (28). The top 300 pairs by direct information score were plotted against measured residue-residue contacts from the AAV2 structure to assess this approach's ability to identify residue interactions (although only 300 pairs were plotted, all contacts were considered in subsequent calculations). This joint probability distribution was also used to calculate the probability $P_{DCA(seq)}$ that a given sequence is a member of the characterized sequence family, and this probability was converted into a unitless energy Hamiltonian $H_{DCA(seq)} \sim logP_{DCA(seq)}$ (29,33,44). This Hamiltonian term was computed for AAV2 sequences with each residue position mutated to alanine. The Hamiltonian score of alanine mutants was compared to WT using the $\Delta H_{DCA(Ala)}$ metric described in Eq. 4.

$$\Delta H_{DCA(Ala)} = H_{DCA Ala} - H_{DCA WT} \quad (4)$$

## Site-directed mutagenesis of AAV2 cap gene

Site-directed mutagenesis of the AAV2 *cap* gene was performed to substitute selected residues with alanines. The pXX2 plasmid containing the wtAAV2 *rep* and *cap* genes was used as the template (45). Primers containing the desired alanine mutations were purchased from Integrated DNA Technologies. 18 cycles of PCR amplification were conducted according to the QuikChange protocol using Pfu Ultra polymerase (Agilent Technologies, Santa Clara, CA). After cycling was complete, template DNA was removed by digesting with DpnI (New England Biolabs, Ipswich, MA). Resulting plasmids were sequence verified through an external vendor (Genewiz, Morrisville, NC).

## Virus production

Viruses containing the desired capsid mutations were prepared through a triple plasmid transfection of human embryonic kidney 293T (HEK293T) cells with the *rep-cap* encoding plasmid (pXX2 or pXX2-derived mutants), pSC-GFP (encodes a self-complementary GFP transgene flanked by inverted terminal repeats), and pXX6-80 (encodes adenoviral helper genes) using polyethylenimine. Cells were harvested 48 h post-transfection and lysed by three freeze-thaw cycles. 50 U/mL of benzonase (Sigma-Aldrich, St. Louis, MO) was added to the cell lysate to degrade free nucleic acids and the mixture was centrifuged to remove cell debris. The supernatant was then loaded into Quick-Seal Ultra Clear 25 × 89 mm centrifuge tubes (Beckman Coulter, Crea, CA) containing a 15–54% iodixanol step gradient. Tubes were sealed and spun in a Beckman Type 70Ti rotor at 48,000 rpm for 1 h 45 min at 18°C, and virus was extracted from the 40% iodixanol layer. For the differential scanning fluorescence assay, viruses were concentrated into gradient buffer (10 mM Tris (pH 7.6), 10 mM MgCl₂, 150 mM NaCl) using Amicon Ultra 100 kDa centrifugal filters (EMD Millipore, Burlington, MA).

## Quantification of viral particles

Viral titers were quantified using quantitative polymerase chain reaction (qPCR). Briefly, viral capsids were denatured to release their genomes using incubation in 2 M NaOH at 56°C followed by neutralization with 2 M HCl. SYBR Green Power PCR Master Mix (Thermo Fisher Scientific, Waltham, MA) along with primers against the cytomegalovirus promoter (forward, 5′-TCACGGGGATTTCCAAGTCTC-3′ and reverse, 5′-AATGGGG CGGAGTTGTTACGA-3′) were used to detect viral genomes. Samples

were analyzed on the Bio-Rad CFX96 qPCR machine (Hercules, CA) to obtain absolute titer values against a standard curve. Raw data are provided (Data S3).

## Genome protection assay

Viruses were diluted 1:10 in endo buffer (1.5 mM $MgCl_2$, 0.5 mg/mL bovine serum albumin, 50 mM Tris (pH 8.0)) and incubated at the indicated temperature for 30 min. Samples were then split into 20 $\mu L$ treatment fractions and treated with either 0.5 $\mu L$ benzonase (1:10 dilution, 250 U/$\mu L$; Sigma-Aldrich), or sham buffer (50% glycerol, 50 mM Tris-HCl, 20 mM NaCl, 2 mM $MgCl_2$ (pH 8.0)). Samples were incubated at 37°C for 30 min, and then the benzonase was inactivated through the addition of 0.5 $\mu L$ of 0.5 M EDTA. The number of viral genomes in benzonase and sham-treated fractions was then quantified using qPCR and genome protection calculated as the ratio of genomes in the benzonase-treated fraction to the sham-treated fraction. Raw data are provided (Data S3).

## Differential scanning fluorescence assay

The differential scanning fluorescence assay was adapted from previous reports (46). Viruses were diluted to a concentration of $10^{12}$ viral genomes/mL in gradient buffer (10 mM Tris (pH 7.6), 10 mM $MgCl_2$, 150 mM NaCl), and 45 $\mu L$ virus sample was mixed with 5 $\mu L$ 50× sypro orange (Thermo Fisher Scientific). Assays were conducted in a Bio-Rad CFX96 qPCR instrument using a melt curve protocol ramping from 25 to 95°C at a rate of 1° per min, with fluorescence reads every 0.2°. Lysozyme enzyme was analyzed as a control.

## Quantification of virus transduction

HEK293T cells were seeded on 48-well tissue-culture-treated poly-L-lysine coated plates ∼24 h pre-transduction. At 95% confluency, cells were transduced with virus in media with serum at 1000 multiplicity of infection. 24 h after transduction, media was changed. 48 h after transduction, cells were harvested for flow cytometry analysis on a BD FacsCanto II (BD Biosciences, San Jose, CA). Virus transduction ability was quantified using the transduction index (TI), which is the product of the percentage of GFP+ cells and geometric mean fluorescence intensity. The transduction index is a linear indicator of viral transduction efficiency (47). Counts and gating thresholds are provided (Data S4).

## RESULTS AND DISCUSSION

### Frustratometer analysis of AAV capsid

Frustration can be computationally quantified using the Frustratometer, which draws upon an optimized coarse-grained modeling that has been used to analyze protein energy landscapes (42). We applied the Frustratometer to AAV2 to predict residues that favor either the assembled capsid structure or the monomeric capsid protein. The AAV capsid consists of three capsid proteins, VP1, VP2, and VP3, that assemble in a 1:1:10 ratio to form a 60-mer capsid. Cryo-electron-microscopy-resolved structures of AAV2 do not include the unstructured VP1 and VP2 N-terminal domains, so this analysis focuses on the VP3 domain shared by all three VPs that forms the exterior capsid structure. Because it would be too computationally intensive to analyze the entire 60-mer capsid structure, we generated a

multimer assembly substructure from the 60-mer capsid structure by selecting the seven VP3 subunits within 4.5 Å of a central monomer (Fig. S1). Then, mutational frustration was calculated for every residue-residue contact in the monomer and multimer structures (Fig. 1). Because the monomer structure is derived from cryo-electron microscopy conducted on the intact capsid, any differences in frustration index between the monomer and multimer are due to multimeric residue interactions and not structural changes in the monomer. The multimer structure contains more minimally frustrated residue pairs than the monomer, particularly at capsid subunit-subunit binding interfaces. This reduction in frustration is consistent with previously observed shifts in frustration at protein binding interfaces related to the burial of hydrophobic surfaces (20).

Single-residue mutational frustration indices (looking at the frustration of each residue as opposed to residue pairs) were also computed for the monomer and the multimer assembly. The AAV2 monomer has 30.6% minimally frustrated residues and 8.1% highly frustrated residues, whereas
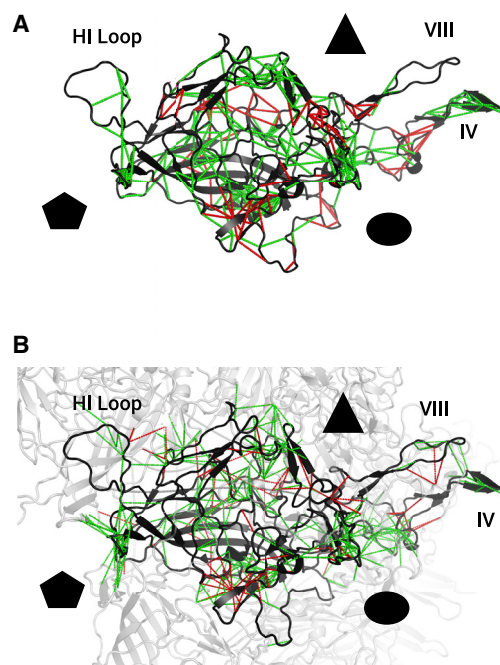


FIGURE 1 Analysis of AAV capsid using mutational frustration. Mutational frustration of all possible residue interactions was computed using the Frustratometer for the AAV2 monomer and multimer. Minimally and highly frustrated interactions were plotted on crystal structures of AAV2. (A) Capsid residue interactions that are minimally (green) or highly (red) frustrated within the AAV2 monomer are indicated by connecting lines with the protein backbone in dark gray. (B) Capsid residue interactions that are minimally (green) or highly (red) frustrated within the AAV2 monomer and between the AAV2 monomer and the other subunits in the AAV2 multimer are indicated by connecting lines with the protein backbone in dark gray and other subunits in light gray. The icosahedral (T = 1) capsid fivefold (pentagon), threefold (triangle), and twofold (circle) axes of symmetry are indicated. Variable regions IV and VIII and the HI loop are labeled.

the AAV2 multimer has 30.3% minimally frustrated residues and 6.6% highly frustrated residues. Both configurations have more neutral residues than a set of 314 monomeric proteins previously curated and analyzed from the PDB, which have on average ~40% minimally frustrated residues and >10% highly frustrated residues [20]. Single-residue mutational frustration indices were then combined to generate a $\Delta$-frustration index. This change in frustration upon protein-protein binding was previously computed for a benchmark of assemblies of homodimers and their subunits curated from the PDB [48]. In a previous study employing the Frustratometer, 25% of residues found in protein-protein interaction domains exhibited an increase of ~1.5 single-residue mutational frustration index units (i.e., a $\Delta$-frustration index of 1.5) in the bound state, whereas 7% of residues exhibited a decrease in the single-residue mutational frustration index (i.e., a negative $\Delta$-frustration index) [20]. In AAV2, ~6.0% of residues exhibit an increase in single-residue mutational frustration index in the multimer state greater than one unit ($\Delta$-frustration index >1), whereas 3.7% of residues exhibit a decrease in single-residue mutational frustration index by more than one unit ($\Delta$-frustration index <1) (Fig. S2). Molecular dynamics models of virus assembly have previously shown that interactions between viral subunits must be relatively weak to permit viral assembly while avoiding kinetic traps due to malformed structures; this property may result in a less dramatic shift toward minimal frustration upon binding than in smaller protein assemblies [49,50].

We hypothesized that this $\Delta$-frustration index metric may highlight AAV capsid residues that shift in frustration upon subunit-subunit interaction, indicating they play key roles in stability of either the monomer or the multimer assembly. By this logic, positive values of the $\Delta$-frustration index would identify residues that favor the multimer assembly state, whereas negative values identify residues that favor the monomer state.

We also compared the $\Delta$-frustration index to other residue-level parameters. There was no correlation of the change in frustration upon capsid assembly with sequence conservation in the parvovirus family (Fig. 2 C) or distance from the capsid center (Fig. 2 D). There is a slight positive correlation ($R^2 = 0.018$, $p = 0.0025$) between the number of inter-residue contacts, defined as two residues having a pair of atoms with a distance of <4.5 Å, and the $\Delta$-frustration index (Fig. 2 E). Residues with lower frustration in the multimer state tend to have more contacts with other residues, suggesting that essential capsid interactions take place in these domains that are conserved across the parvovirus family [51]. Conversely, residues with lower frustration in the monomer state have fewer contacts with other capsid residues. These residues may play a role in capsid interactions with other proteins on the virus's transduction pathway as opposed to intracapsid interactions. Indeed, mutation of R471 has previously been shown to reduce immune system
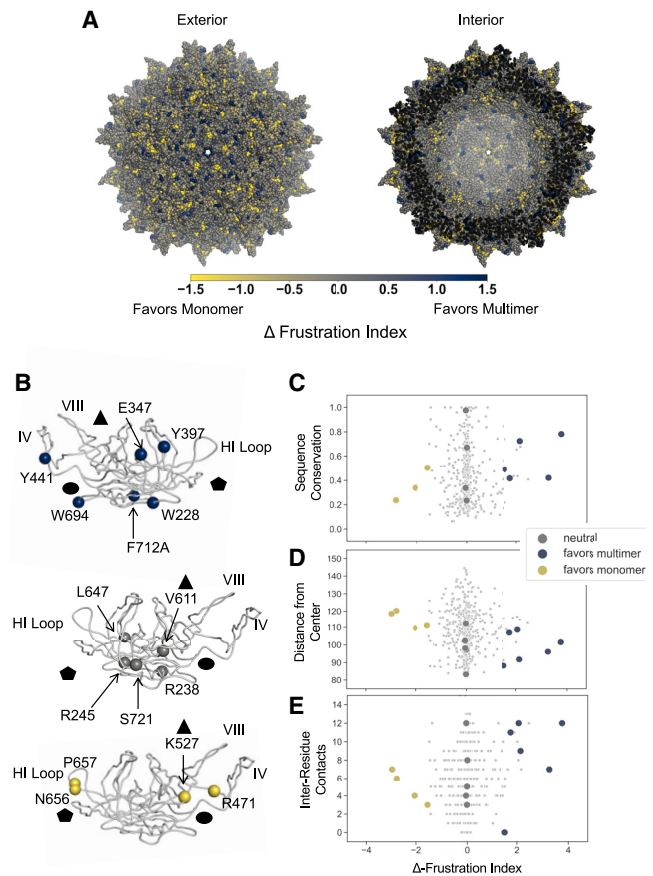


FIGURE 2 AAV2 $\Delta$-frustration index and capsid residues selected for experimental analysis. AAV2 $\Delta$-frustration index was computed as the difference between monomer and multimer single-residue frustration indices. Residues with a positive $\Delta$-frustration index are predicted to favor the multimer, whereas residues with a negative $\Delta$-frustration index are predicted to favor the monomer. (A) AAV2 60-mer capsid crystal structure colored according to $\Delta$-frustration index is given. Views of the capsid from exterior and interior are shown. (B) Residues selected for experimental analysis are shown. Residues favoring the multimer state (blue, $\Delta$-frustration index >1.5), residues favoring the monomer state (yellow, $\Delta$-frustration index <−1.5), and neutral residues (gray, $\Delta$-frustration index >−0.1 and <0.1) were selected for further study. The capsid fivefold (pentagon), threefold (triangle), and twofold (circle) axes of symmetry are indicated. Top view shows interior face of the VP3 subunit, and middle and bottom views show exterior face. $\Delta$-frustration index for AAV2 capsid residues was plotted against (C) sequence conservation in the parvovirus family computed using Shannon entropy of reweighted sequences normalized from 0 to 1 (1 indicates full conservation), (D) residue distance from center of AAV2 capsid (Å), and (E) number of residue-residue contacts according to a 4.5 Å heavy-atom distance threshold. Residues selected for mutation are large colored circles, and residues not selected are small gray dots.

interactions and K527 is known to be proximal to the AAV2 heparin binding pocket, although mutation does not impact heparin binding [52].

A set of representative residues was selected based on the $\Delta$-frustration index favoring the multimer, favoring the monomer, or neutral toward either state. Based on the threshold identified in Ferreiro et al. [20], $\Delta$-frustration

index values $>1.5$ were labeled as favoring the multimer and values $<1.5$ as favoring the monomer (20). Six residues from the 14 residues favoring the multimer were selected randomly, and all four residues that were predicted to favor the monomer were selected. Five neutral residues were selected randomly from the 229 residues with $\Delta$-frustration indices $>-0.1$ and $<0.1$. These selected residues were experimentally mutated to alanine to evaluate their role in capsid formation, thermal stability, and transduction efficiency.

In addition to applying the Frustratometer, we used the coarse-grained protein force field on which the Frustratometer is based, the AWSEM (43), directly to predict the change in energy ($\Delta E$) upon mutation of the selected residues to alanine. We developed a $\Delta\Delta E$ metric by subtracting the monomer $\Delta E$ from the multimer $\Delta E$ for each alanine mutant. Because a more negative $\Delta E$-value indicates greater energetic stability, alanine mutants with a lower $\Delta\Delta E$ score than the WT are predicted to favor the multimer state, whereas alanine mutants with a higher $\Delta\Delta E$ score than WT are predicted to favor the monomer state. Whereas the Frustratometer measures the favorability of a particular interaction or set of interactions for a particular residue compared to a distribution of decoy interactions, the $\Delta\Delta E$ metric makes predictions about the specific change in stability upon mutating to alanine, which is what was done experimentally. Because the $\Delta\Delta E$ score is highly correlated with the $\Delta$-frustration index in the studied mutants ($R^2 = 0.79$, $p << 0.001$), we included comparisons of $\Delta\Delta E$ and experimental results in our supplemental data.

## DCA of AAV capsid

We next applied DCA to make predictions about the fitness of the selected AAV2 alanine mutants. DCA is a global statistical model derived from a multiple sequence alignment of a protein family (28). In this case, we used the capsid sequences of parvoviruses drawn from Pfam (ID: PF00740) to construct the MSA that is used as input to DCA (53). We plotted the top 300 predicted residue-residue couplings from DCA against likely residue-residue interactions (including direct contacts, potential dynamic interactions, and ligand coordination interactions) defined as inter-residue distance $<12$ Å in the AAV2 crystal structure (Fig. 3; Table S1). The overlap between some DCA-predicted interactions and likely interactions identified in the AAV2 capsid crystal structure (60 monomeric pairs and eight multimeric pairs of interactions) suggests crucial pairs maintained through evolution that play important structural roles at the monomeric and oligomeric level. This model may also be able to identify interacting residue pairs that are not proximate in the AAV capsid but are functionally relevant, a strength of this approach over structure-based models.

To investigate the capsid mutants with DCA, we computed the change in DCA Hamiltonian upon alanine
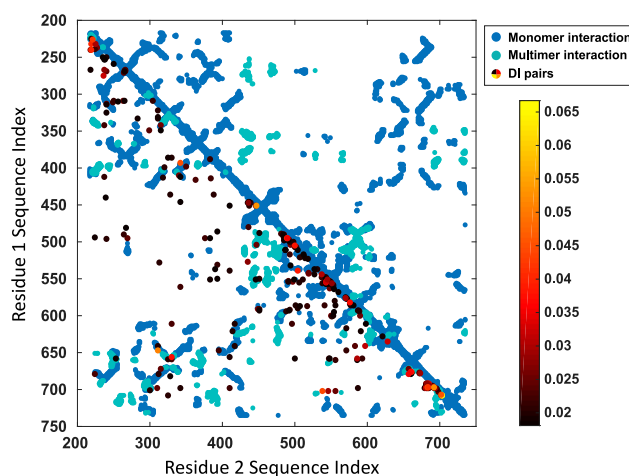


FIGURE 3 AAV2 capsid residue-residue couplings inference by direct-coupling analysis (DCA). DCA was applied to a multiple sequence alignment of sequences identified as likely parvovirus coat proteins in Pfam. Predicted residue-residue couplings from DCA and likely residue-residue interactions (inter-residue distance $<12$ Å) identified in the AAV2 capsid crystal structure (1LP3) are plotted on a map. DCA-predicted pairs (*colored by direct information score, with lighter colors suggesting stronger couplings*), structurally determined monomeric interactions (*blue*), and structurally determined multimeric interactions (*green*) are plotted at the intersection of the two contacting residues' AAV2 capsid sequence indices (VP1 numbering).

mutation for each position in the AAV2 capsid. DCA estimates a joint probability distribution that is used to generate the family-specific DCA Hamiltonian parameters that describe a relative probability that any given sequence is a member of the sequence family. This probability can be log transformed to obtain a unitless quasienergy (DCA Hamiltonian energy $H_{DCA(Ala)}$) (29,44). Lower values of the DCA Hamiltonian energy correspond to sequences that are more likely to belong to the sequence family under consideration and thus more likely to be structurally and functionally similar to the family members. The change in the DCA Hamiltonian energy for alanine mutants as compared with the WT ($\Delta H_{DCA(Ala)}$) was compared with other residue-level parameters (Fig. 4).

The energy change upon mutation of the wild-type residue to alanine, $\Delta H_{DCA(Ala)}$, is correlated with sequence conservation within the parvovirus family ($R^2 = 0.30$, $p << 0.001$), indicating that locations with higher sequence conservation are more sensitive to mutation (Fig. 4 A). DCA generates a global model of the parvovirus sequence family and makes predictions of mutant virus fitness based on the prevalence of residue identities and residue interaction coupling strength in the family. Unsurprisingly, DCA predicts that mutations of a highly conserved residue will be detrimental to fitness. $\Delta H_{DCA(Ala)}$ is negatively correlated with residue distance from the capsid center ($R^2 = 0.28$, $p << 0.001$). Residues further from the capsid center have lower (more favorable) $\Delta H_{DCA(Ala)}$ scores, suggesting
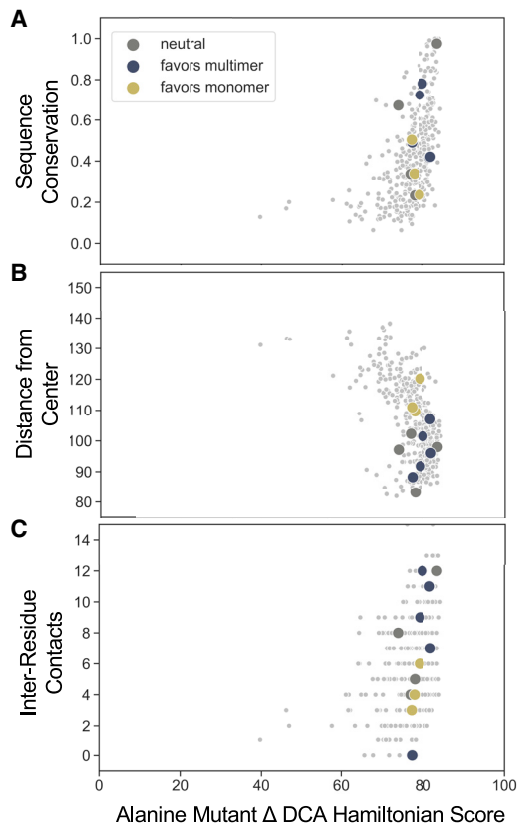
FIGURE 4 Comparison of DCA energy changes upon mutation of the wild-type residue to alanine, $H_{DCA(Ala)}$, to other residue-specific parameters. Plots of the $\Delta H_{DCA(Ala)}$ score against (*A*) sequence conservation in the parvovirus family, (*B*) three-dimensional distance of residue from center of the AAV2 capsid (Å), and (*C*) number of residue-residue contacts according to a 4.5 Å heavy-atom distance threshold are given. AAV2 capsid residues selected for mutation are plotted with large colored circles, and residues not selected for mutation are shown as small gray dots.

senting sequence fitness. To quantify the relationship between AWSEM energy and DCA quasienergy predictions for the selected alanine mutants, we computed the correlations between $\Delta E_{Ala}$ of the monomer and multimer, $\Delta\Delta E_{Ala}$, and $\Delta H_{DCA(Ala)}$ (Fig. S3). These metrics were previously found to be highly positively correlated for protein monomers (29). In the mutants analyzed, the AWSEM monomer $\Delta E$ prediction actually has a high negative correlation with $\Delta H_{DCA(Ala)}$ ($R^2 = 0.66$, $p = 0.001$), whereas there is no significant correlation between the AWSEM multimer $\Delta E$ and $\Delta H_{DCA(Ala)}$ ($R^2 = 0.24$, $p = 0.11$) or the $\Delta\Delta E$ metric and $\Delta H_{DCA(Ala)}$ ($R^2 = 0.053$, $p = 0.47$). We also quantified the relationship between $\Delta$-frustration index and $\Delta H_{DCA(Ala)}$. In the mutants analyzed, these metrics are not significantly correlated ($R^2 = 0.12$, $p = 0.26$). This surprising divergence between DCA- and AWSEM-derived predictions of the fitness of alanine mutants may be attributable to the many roles of viral proteins in transduction. This may also be a function of viral capsid metastability, as viruses must maintain stable structures to protect their genomes but also release their genomic cargo at its intended intracellular destination (49).

## Formation of AAV2 capsid alanine mutants

We mutated the selected AAV2 capsid residues to alanine and generated genome-packaging viruses (Fig. 5 *A*). Among the mutants of residues predicted to favor multimer formation, four out of six exhibit decreased genomic titers, with the remaining mutants exhibiting titers comparable to WT AAV2 capsid. Among the mutants of residues that were predicted to disfavor multimer formation, one (P657A) exhibits decreased genomic titers, whereas the other three exhibit titers comparable with the WT. Among mutants of residues predicted to be neutral with respect to multimer formation, two out of the five exhibit decreased genomic titer, whereas three exhibit titers comparable with the WT.

In an attempt to understand the factors that influence genomic titers in the alanine mutation variants, we first compared virus production titers against two simple metrics that were derived either from structural data on the AAV2 capsid (the number of contacts made by a particular residue) (Fig. 5 *B*) or from an MSA of parvovirus family sequences (the frequency at which alanine appears at the mutated site) (Fig. 5 *C*). Mutant virus production has a negative correlation with the number of residue-residue contacts made by the native residue in the capsid structure, i.e., the greater the number of native contacts a residue has, the lower the yield of virus formed when that residue is mutated to an alanine (Fig. 5 *B*). Additionally, virus production is positively correlated with the frequency of alanine residues observed at the native residue's position in an MSA of parvovirus family sequences (Fig. 5 *C*). In other words, if an alanine is frequently found in a homologous capsid position in many other parvoviruses, the alanine mutation in the

that alanine mutations of these residues are more likely to be tolerated (Fig. 4 *B*). $\Delta H_{DCA(Ala)}$ has a low but significant correlation with the number of residue-residue contacts ($R^2 = 0.13$, $p \ll 0.001$) (Fig. 4 *C*), suggesting that DCA is able to identify some residues that participate in a large number of interactions in capsid proteins. Correlations with capsid structural features were recovered despite the paucity of available parvoviral sequences for this analysis. The direct-coupling analysis joint probability distribution was inferred from an alignment of parvoviruses with 130 effective sequences ($M_{eff}$), 5.06% of the 2569 starting sequences after reweighting sequences to account for sequence homology. Although this $M_{eff}$ is lower than the number that is thought to be necessary to produce high rates of true positive contact predictions, the inferred model nonetheless shows some promise in identifying interactions in the AAV2 capsid structure (28,54).
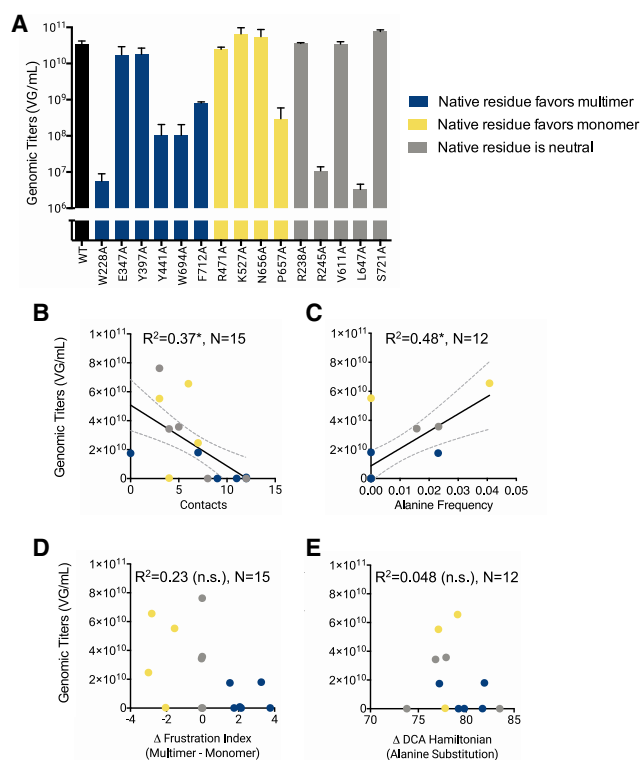
Whereas the AWSEM uses a protein force field to generate predicted energies of protein structures, DCA uses a global model of a sequence family to predict a quasienergy repre-

FIGURE 5 Ability of alanine capsid mutants to form viruses with encapsidated genomes. Selected residues were mutated to alanine, and viruses were produced. Genomic titers of one-plate virus preps were determined through qPCR. (*A*) Genomic titers of alanine mutants expressed as viral genomes per mL (VG/mL) are given. N = 3 independent virus preps. Error bars are mean ± SE. One-way ANOVA was performed with Dunnett's post hoc multiple comparison test to compare mutants to WT; no statistically significant differences were detected. Genomic titers are plotted against (*B*) number of residue-residue contacts in the native capsid structure, (*C*) alanine frequency in the parvovirus sequence family, (*D*) *Δ*-frustration index, and (*E*) *ΔH_{DCA(Ala)}*. N = 15 mutants for structure-based analyses. N = 12 mutants for sequence-based analyses because R471A, F712A, and S721A do not exist in a majority of the sequence family (majority gaps at their alignment position). Regression lines with 95% confidence interval are plotted for statistically significant correlations. $R^2$-values and significance levels are reported (*$p < 0.05$).

## AAV2 capsid is better tolerated and yields higher virus production levels.

Mutant virus formation is not significantly correlated with the native residue's *Δ*-frustration index (Fig. 5 *D*) and DCA-based *ΔH_{DCA(Ala)}* scores (Fig. 5 *E*). However, there appears to be a general trend in which mutants of residues with higher *Δ*-frustration indices have reduced formation. Proline has a unique influence on secondary structure that is not directly considered by the AWSEM Frustratometer, which can often result in challenges when attempting to predict mutational effects (55,56). If the P657A mutant is excluded from this analysis, then virus formation is correlated with the *Δ*-frustration index of native residues, but not *ΔH_{DCA(Ala)}* (Fig. S4). The results thus far demonstrate that the *Δ*-frustration index computed with AWSEM-MD

is somewhat related to capsid formation but does not fully predict assembly.

We also considered the impact of coding mutations introduced in the assembly-activating protein (AAP) and X gene *trans*-encoded elements of the AAV genome that play a role in capsid assembly and viral DNA replication, respectively (57,58). The W228A, R238A, R245A, and E347A mutations also introduce coding mutations in AAP (Table S3). These modifications may play a role in W228A and R245A's reduced viral formation because AAP is essential for AAV2 production (59). The V611A, L647A, N656A, P657A, W694A, and S721A mutations also introduce coding mutations in the X gene. These mutations may play a role in reduced viral titers for some of these mutants. However, removal of the X gene reduces AAV2 genomic titers by 33%, so X gene mutations are likely only a partial factor in L647A, P657A, and W694A's reduced genomic titers (58).

AAV formation occurs via a multistep process involving first capsid protein production and assembly into a 60-mer capsid, then insertion of the viral genome into the capsid lumen via function of other protein factors (60,61). Our findings suggest that both simple structural contact information and sequence family residue frequency information can help to guide the design of variants with the goal of modulating complete virion yield. From these observations, it is unclear whether more sophisticated computational models provide better guidance in viral variant design than do relatively simple metrics derived from the capsid structure (residue-residue contacts) and the parvovirus sequence family alignment (Fig. 5, *B* and *C*). No single model appears to capture capsid formation (i.e., combination of capsid assembly and genome packaging) fully, perhaps because none of these approaches explicitly take into account interactions with other known AAV factors and helper virus proteins that are required for capsid assembly and viral genome packaging (57,62–64). The models may potentially perform better in prediction of AAV2 VP monomer formation and empty capsid assembly as opposed to complete genome-containing virion formation because these intermediate steps require fewer interactions with protein cofactors and viral genomic DNA (65). Quantification of VP monomer yield, capsid morphology, and the ratio of empty/full viral capsids would provide some measure of these earlier stages in viral vector production.

## Thermal stability of AAV2 capsid alanine mutants

Viruses that formed with sufficient titers for further analysis ($>10^{10}$ viral genomes/mL) were screened for thermal stability through a genomic protection assay. Specifically, viruses were incubated at a range of temperatures near the WT AAV2 capsid melting temperature, which was previously reported as 72.4°C (46). The virus samples were then treated with a nuclease to degrade any uncoated viral genomes. After nuclease inactivation, the samples were
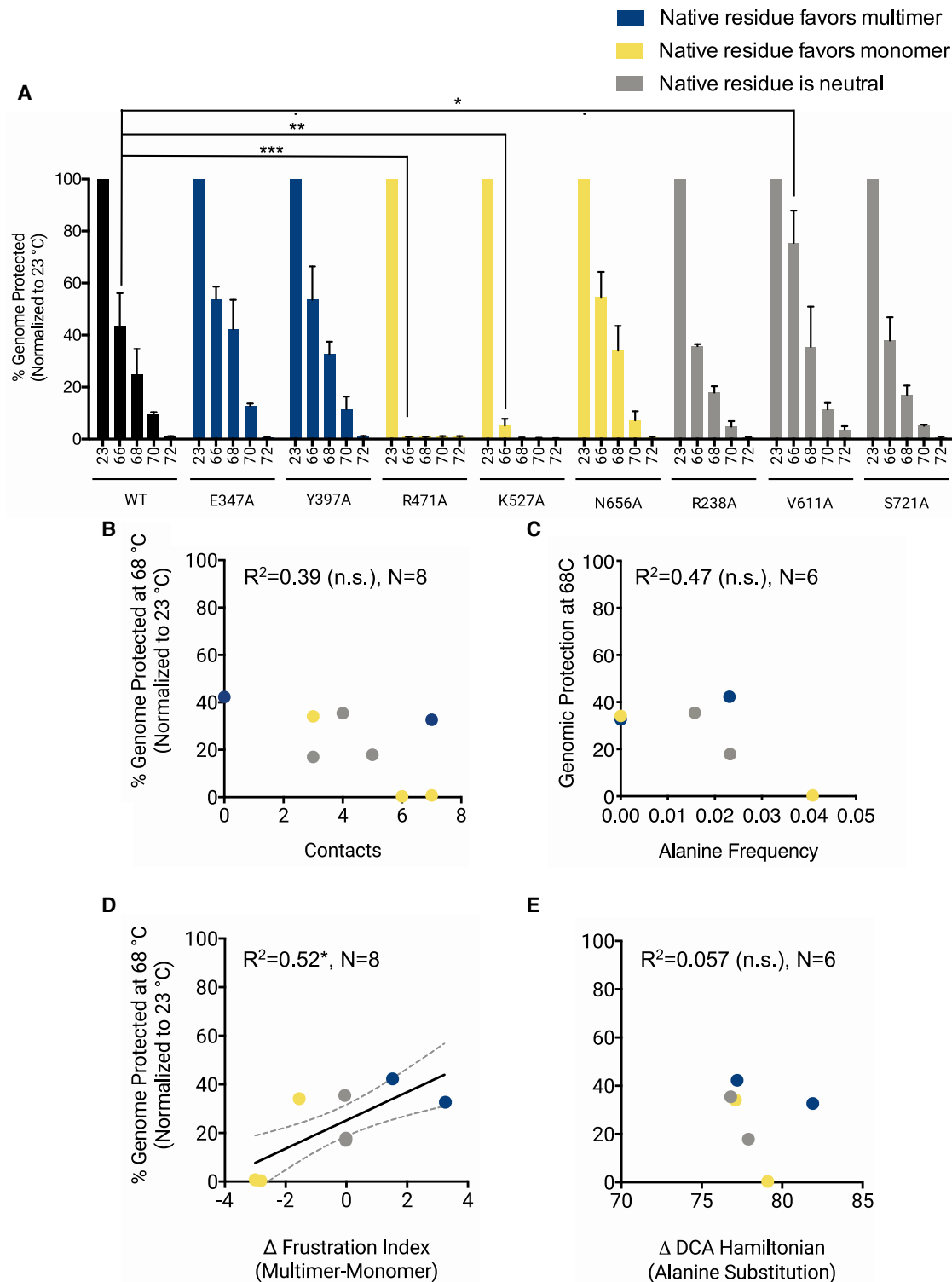
FIGURE 6 Thermal stability of alanine mutant capsids. Alanine mutants that formed at sufficient titers were exposed to various temperatures (°C) and then assayed for genomic protection. Temperature-treated samples were incubated with benzonase or sham buffer and genomes quantified using qPCR. (*A*) Genomic protection after incubation at elevated temperatures is shown. N = 3 independent experiments titered in duplicate. Error bars are mean ± SE. Two-way ANOVA was performed with Dunnett's post hoc multiple comparison test to compare mutants with WT at each tested temperature. Percentage of genome protected is the fraction of genomes in the benzonase-treated sample as compared with the sham-treated sample. Genomic protection at 68°C is plotted against (*B*) number of residue-residue contacts in the native capsid structure, (*C*) alanine frequency in the parvovirus sequence family,

*(legend continued on next page)*

then assayed for the number of remaining genomes (i.e., genomes that are protected from nuclease digestion by an intact capsid). Most capsids exhibit thermal stability comparable with the WT, with the exception of the R471A and K527A mutants (Fig. 6 A). These two mutants appear to lose genomic protection after incubation at 66°C. To further characterize these mutants' reduction in thermal stability, the melting points of these mutants were determined using a differential scanning fluorescence assay (Fig. S5 A; (46)). R471A exhibits a 6.8°C decrease in melting temperature compared with the WT, and K527A exhibits a 3.5°C decrease in melting temperature compared with the WT (Fig. S5 B).

Capsid thermal stability at 68°C was compared to various computational metrics of fitness as described above. Virus thermal stability is not correlated with the number of residue-residue contacts made by the original residue in the WT structure (Fig. 6 B) or with the frequency of alanine in an alignment of parvovirus sequences (Fig. 6 C). Thermal stability is, however, correlated with the $\Delta$-frustration index (Fig. 6, D and E). Interestingly, alanine mutants of native residues predicted to favor the monomer by $\Delta$-frustration index have lower genomic protection at 68°C (Fig. 6 D). $\Delta H_{DCA(Ala)}$ shows no correlation with capsid thermal stability (Fig. 6 E). These results indicate that metrics derived from physical energy-based models such as the $\Delta$-frustration index may capture information about thermal stability, although elucidating the reasons why capsid mutants predicted to favor the multimer have lower melting temperatures will require further studies. This relationship may be due to the stability of the capsid monomer playing an important role in overall capsid thermal stability. Indeed, parvovirus capsid denaturation or disassembly at high temperatures is thought to occur cooperatively as subunits lose tertiary structure and disassemble into trimer intermediates simultaneously; this denaturation mechanism has been demonstrated in minute virus of mice (66,67). In future work, these ideas about the relative importance of monomer and multimer stability can be evaluated by testing the melting point of mutant monomers synthesized in bacterial expression systems without other required components for capsid assembly. It is notable that this result is in contrast with capsid assembly, in which mutants of residues predicted to favor the multimer have generally higher genomic titers, in line with our initial expectations. This conflict may be resolved through a more detailed experimental examination of the stages of mutant capsid assembly, including monomer synthesis, trimer assembly, empty capsid formation, and genome packaging. It is also possible that the results

observed in this study have some dependence on the size of genomic cargo and the use of a self-complementary transgene, as transgene size and self-complementarity have previously been shown to impact virion thermostability (68). This possibility may be examined by conducting thermostability assays on mutant virions packaging genomes of different sizes and single-stranded genomes.

## Transduction efficiencies of AAV2 capsid alanine mutants

Lastly, the WT and mutant viruses were screened for their ability to transduce HEK293T cells. Variants with mutations that were predicted to favor multimer formation and variants with mutations that were predicted to disfavor multimer formation show a wide range of transduction levels, with some mutants severely deficient in transduction and others comparable to WT (Fig. 7 A). Mutants that were predicted to be neutral with respect to multimer formation all show transduction levels similar to WT. The number of residue-residue contacts (Fig. 7 B) and the $\Delta$-frustration index (Fig. 7 D) are not correlated with virus transduction ability. Because these metrics are derived from the viral capsid structure and a physical energy function, they do not take into account the interactions the AAV capsid must successfully make with its cellular environment to transduce cells. The frequency of alanine at the site of the mutated residue in an alignment of parvovirus sequences is also not correlated with transduction (Fig. 7 C). The $\Delta H_{DCA(Ala)}$ score, however, is highly correlated with virus transduction (Fig. 7 E). Capsid alanine mutants that are predicted to have higher fitness in the DCA-derived model of parvovirus coat protein sequence are more successful at genome delivery into host cells.

AAV transduction is a complex, multistage process, requiring the capsid to interact with extracellular receptors, escape the endosome, traffic to and enter the cell nucleus, and disassemble to release the transgene for transcription (69). The $\Delta H_{DCA(Ala)}$ score is the only metric found to be highly correlated with transduction efficiency, although when the number-of-contacts metric is restricted to the same set of six residues analyzed for $\Delta H_{DCA(Ala)}$, this metric also exhibits a high correlation with transduction (Table S2). Because the DCA approach is based on a global model drawing on sequences from the parvovirus family, this model may be better able to predict which mutated sequences are able to traverse the virus's infectivity pathway successfully. Although DCA is not likely to capture the wide range of receptor interactions that AAV and other parvoviruses undergo to initiate cell entry, this approach may

---

(D) $\Delta$-frustration index, and (E) $\Delta H_{DCA(Ala)}$. N = 8 mutants for structure-based analyses. N = 6 mutants for sequence-based analyses because these residues do not exist in a majority of the sequence family (these residues are considered as insertions in the hidden Markov model of the family). Regression lines with 95% confidence interval are plotted for statistically significant correlations. $R^2$-values and significance levels are reported. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
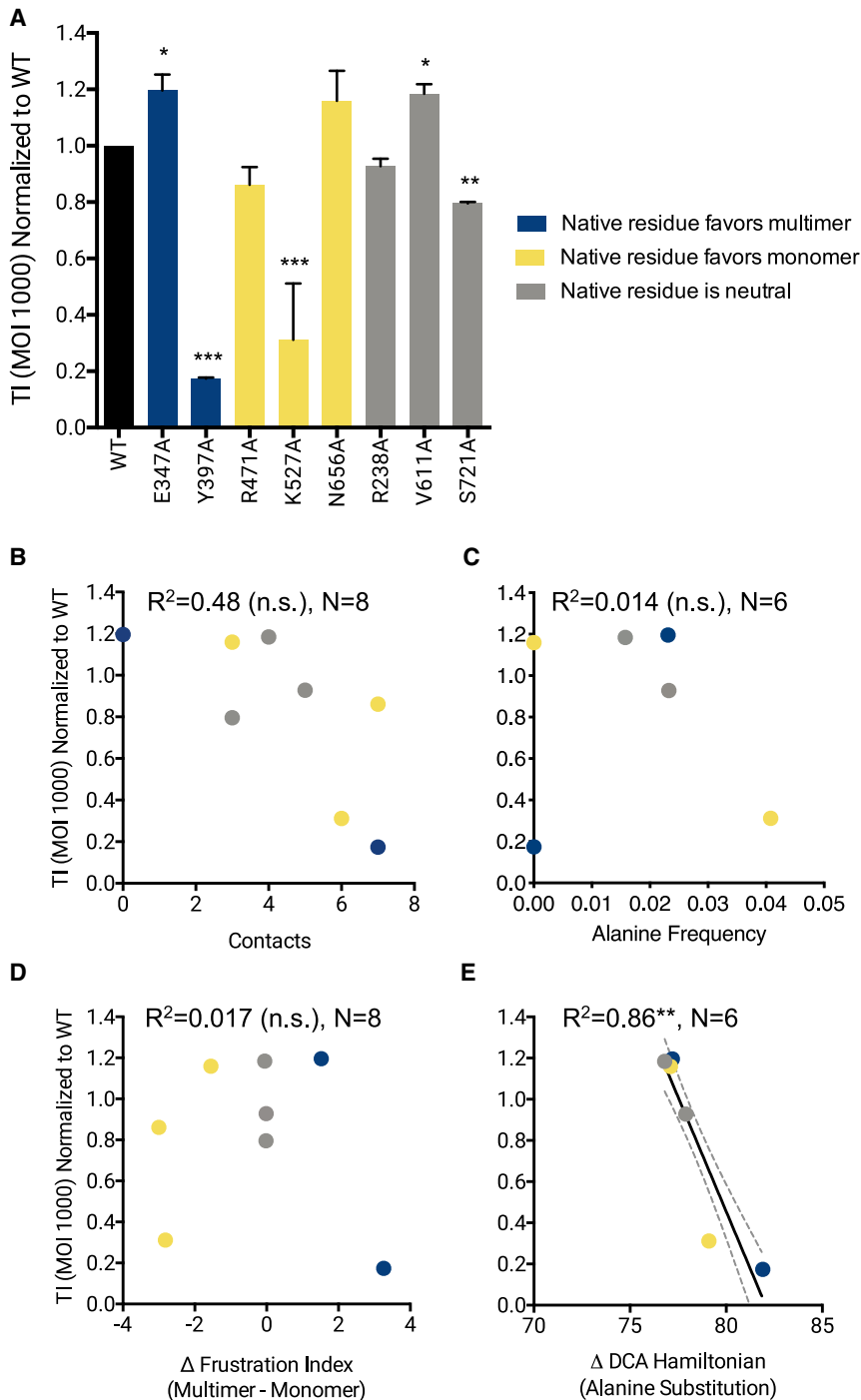
FIGURE 7 Transduction ability of alanine mutant capsids. HEK293T cells were transduced with viruses packaging scGFP transgene at a multiplicity of infection of 1000, and transgene expression was quantified at 48 h post-transduction using flow cytometry. (*A*) Transduction index (TI) of viruses. TI is the product of percentage of GFP+ cells and the geometric mean fluorescence intensity; metrics were obtained from flow cytometry. N = 3 independent experiments done in triplicate. Error bars are mean ± SE. One-way ANOVA was performed with Dunnett's post hoc multiple comparison test to compare mutants with WT. TI of alanine mutants is plotted against (*B*) number of residue-residue contacts in the native capsid structure, (*C*) frequency of alanine in the parvovirus sequence family at the site of the mutated residue, (*D*) $\Delta$-frustration index, and (*E*) $\Delta H_{DCA(Ala)}$ score. N = 8 mutants for structure-based analyses. N = 6 mutants for sequence-based analyses because R471A and S721A were excluded from these analyses because these residues do not exist in a majority of the sequence family (these residues are considered as insertions in the hidden Markov model of the family). Regression lines with 95% confidence interval are plotted for statistically significant correlations. $R^2$-values and significance levels are reported. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

capture features of alternate capsid conformations required for exposure of the viral endosomal escape and nuclear localization domains. This hypothesis may be evaluated by examining vector transduction of cell lines with different receptors, as well as specific stages of transduction such as cellular entry, endosomal escape, nuclear trafficking, and genome release. Furthermore, pairwise residue-residue interactions appear to be essential to this result, as the parvovirus family alanine frequency metric is not correlated with transduction efficiency. However, this result is derived from data collected on only six viral mutants. As a next step, the DCA approach should be applied to a larger capsid single-mutant library to validate this preliminary finding. Additionally, AAV in vitro transduction often differs significantly from in vivo gene delivery, so an assessment of the correlation of DCA predictions with in vivo infectivity is crucial for determining whether DCA has utility in informing capsid designs for gene therapy.

## CONCLUSIONS

We evaluated two different computational approaches for their abilities to predict the formation and function of AAV capsid mutants (Fig. S7). To characterize the relationship between metrics derived from these approaches and viral formation and function, we generated a series of AAV2 capsid mutants and experimentally evaluated their assembly, thermal stability, and transduction (Table 1).

Viral intracapsid interactions are governed by the same thermodynamic principles at play in all protein-protein interfaces which are typically governed by weak noncovalent interactions (70). The predominant interaction in capsid assembly is the burial of hydrophobic regions between capsid monomer binding interfaces (66,71). We sought to capture these fundamental capsid interactions through the Frustratometer energy function derived from the energy landscape theory of protein folding. This energy function is computed using the AWSEM, a coarse-grained molecular dynamics force field that has been used for simulating protein folding and protein-protein interactions. Studies drawing upon such modeling have previously captured viral assembly pathways, suggesting that this approach may prove fruitful in identifying residues crucial for virus formation (72–76).

We also explored computational approaches for translating evolutionary information about fitness constraints from the AAV2 capsid sequence family into predictions about viral mutant function. DCA is a global statistical model of sequence families that infers the parameters of an energy function with terms representing single-residue identity and residue-residue pairwise couplings. By generating a global model of sequence families rather than a local model of each residue-residue interaction, DCA can isolate directly interacting pairs from indirect correlations that appear because of the intertwined networks of primary interactions (28,29).

Both the Frustratometer and DCA capture features of the protein energy landscape; whereas the Frustratometer simulates physical interactions within proteins, DCA incorporates information from the evolutionary pressures that have shaped all members of the sequence family to fold robustly and form functional proteins. DCA also incorporates other constraints on evolution for protein function in addition to folding and stability. These two computational approaches have been shown to be correlated in their predictions in prior studies on protein monomers, but when applied to a virus capsid, they appear to capture somewhat different information about capsid stability and function. This may be due to the complex energy landscape of the metastable viral capsid or the large number of protein-protein interactions AAV encounters in its infectivity pathway.

The results of this exploratory study hint at the potential of DCA for predicting viral transduction and raise interesting questions about the relationship between the Frustratometer-based metric and capsid stability. Correlations between these predictions and experimental results are comparable to those observed in studies evaluating modern state-of-the-art models for mutation prediction in a variety of proteins and experimental outcomes (77–80). However, any conclusions are limited by the small number of mutants studied and the high propensity of AAV capsid mutations to reduce viral production yields, limiting further study. Analysis of the DCA approach is further limited by the need for coverage of each mutant position in a multiple sequence alignment. Residues that are insertions with respect to the profile hidden Markov model used to construct the multiple sequence alignment were not characterized because the DCA framework cannot make predictions about the fitness of regions that are not included in the input alignment.

This study was also limited by the selection of mutants using only one computational approach. To better understand the capacity of these models to make predictions about the AAV capsid, larger-scale mutational studies exploring a variety of mutation types are needed. Mutations should be selected that span the range of DCA mutation fitness predictions to ensure coverage of the 19% of possible alanine mutants with $\Delta H_{DCA(Ala)}$ scores more favorable than the mutants tested and 1.4% of possible alanine mutants with $\Delta H_{DCA(Ala)}$ scores less favorable than the mutants tested. Models should also be evaluated on a large range of random mutations of different designs, such as single mutants to nonalanine residues, protein chimeras, and multiresidue mutations. Prior attempts to build unbiased large libraries of AAV variants using DNA shuffling and error-prone PCR have been limited by the virus's intolerance to mutants that impact capsid architecture (17,81,82). A detailed analysis of model performance on different mutation types will clarify the value of these computational approaches for aiding viral library design. Such data may also be used to develop supervised approaches incorporating features from DCA and the Frustratometer to predict specific experimental outcomes.

Overall, the results of this study suggest that computational frameworks relying on evolutionary sequence information and force-field-based predictions may provide guidance for specific elements of AAV structure and function, but do not on their own provide a complete picture

**TABLE 1  $R^2$-Values of Linear Regression between Computational Metrics and Experimental Measures**

| | Titers | Genomic Protection | Transduction |
|---|---|---|---|
| Structural contacts | 0.37*, N = 15 | 0.39 (n.s.), N = 8 | 0.48 (n.s.), N = 8 |
| Alanine frequency | 0.48*, N = 12 | 0.47 (n.s.), N = 6 | 0.014 (n.s.), N = 6 |
| $\Delta F$ | 0.23 (n.s.), N = 15 | 0.52*, N = 8 | 0.017 (n.s.), N = 8 |
| $\Delta H_{DCA(Ala)}$ | 0.048 (n.s.), N = 12 | 0.057 (n.s.), N = 6 | 0.86**, N = 6 |

*$p < 0.05$, **$p < 0.01$. n.s, not significant.

of vector fitness. Hybrid approaches using DCA to infer likely residue-residue contacts that are given as input to constrain energy-based models may be powerful in developing a better understanding of capsid metastability (83,84). The performance of DCA may also be improved through integration of the growing body of AAV variant sequence data obtained through barcoded, directed-evolution experiments and biomining. Development and validation of these larger data-driven computational models will ideally lead to the development of an accurate in silico model of AAV fitness. The resulting refined model could then be used to prescreen engineered viruses, thereby accelerating the development of optimized vectors for more effective gene therapy.

## SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2020.12.018.

## AUTHOR CONTRIBUTIONS

## ACKNOWLEDGMENTS

## SUPPORTING CITATIONS

Reference (85) appears in the Supporting Material.

## REFERENCES

1. Weitzman, M. D., and R. M. Linden. 2011. Adeno-associated virus biology. *In* Adeno-Associated Virus: Methods and Protocols. R. O. Snyder and P. Moullier, eds. Humana Press, pp. 1–15.

2. Russell, S., J. Bennett, …, A. M. Maguire. 2017. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *Lancet.* 390:849–860.

3. Nakai, H., S. R. Yant, …, M. A. Kay. 2001. Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction in vivo. *J. Virol.* 75:6969–6976.

4. Duan, D., P. Sharma, …, J. F. Engelhardt. 1998. Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. *J. Virol.* 72:8568–8577.

5. Manning, W. C., S. Zhou, …, V. Dwarki. 1998. Transient immunosuppression allows transgene expression following readministration of adeno-associated viral vectors. *Hum. Gene Ther.* 9:477–485.

6. Halbert, C. L., T. A. Standaert, …, A. D. Miller. 1998. Successful readministration of adeno-associated virus vectors to the mouse lung requires transient immunosuppression during the initial exposure. *J. Virol.* 72:9795–9805.

7. Halbert, C. L., E. A. Rutledge, …, A. D. Miller. 2000. Repeat transduction in the mouse lung by using adeno-associated virus vectors with different serotypes. *J. Virol.* 74:1524–1532.

8. Hinderer, C., N. Katz, …, J. M. Wilson. 2018. Severe toxicity in nonhuman primates and piglets following high-dose intravenous administration of an adeno-associated virus vector expressing human SMN. *Hum. Gene Ther.* 29:285–298.

9. Gao, G. P., M. R. Alvira, …, J. M. Wilson. 2002. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl. Acad. Sci. USA.* 99:11854–11859.

10. Gao, G., L. H. Vandenberghe, …, J. M. Wilson. 2004. Clades of adeno-associated viruses are widely disseminated in human tissues. *J. Virol.* 78:6381–6388.

11. Gao, G., M. R. Alvira, …, J. M. Wilson. 2003. Adeno-associated viruses undergo substantial evolution in primates during natural infections. *Proc. Natl. Acad. Sci. USA.* 100:6081–6086.

12. Guenther, C. M., B. E. Kuypers, …, J. Suh. 2014. Synthetic virology: engineering viruses for gene delivery. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* 6:548–558.

13. Büning, H., A. Huber, …, U. Hacker. 2015. Engineering the AAV capsid to optimize vector-host-interactions. *Curr. Opin. Pharmacol.* 24:94–104.

14. Chen, M. Y., S. S. Butler, …, J. Suh. 2019. Physical, chemical, and synthetic virology: reprogramming viruses as controllable nanodevices. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* 11:e1545.

15. Ho, M. L., B. A. Adler, …, J. Suh. 2013. SCHEMA computational design of virus capsid chimeras: calibrating how genome packaging, protection, and transduction correlate with calculated structural disruption. *ACS Synth. Biol.* 2:724–733.

16. Ojala, D. S., S. Sun, …, D. V. Schaffer. 2018. In vivo selection of a computationally designed SCHEMA AAV library yields a novel variant for infection of adult neural stem cells in the SVZ. *Mol. Ther.* 26:304–319.

17. Adachi, K., T. Enoki, …, H. Nakai. 2014. Drawing a high-resolution functional map of adeno-associated virus capsid by massively parallel sequencing. *Nat. Commun.* 5:3075.

18. Ogden, P. J., E. D. Kelsic, …, G. M. Church. 2019. Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science.* 366:1139–1143.

19. Zinn, E., S. Pacouret, …, L. H. Vandenberghe. 2015. In silico reconstruction of the viral evolutionary lineage yields a potent gene therapy vector. *Cell Rep.* 12:1056–1068.

20. Ferreiro, D. U., J. A. Hegler, …, P. G. Wolynes. 2007. Localizing frustration in native proteins and protein assemblies. *Proc. Natl. Acad. Sci. USA.* 104:19819–19824.

21. Onuchic, J. N., and P. G. Wolynes. 2004. Theory of protein folding. *Curr. Opin. Struct. Biol.* 14:70–75.

22. Bryngelson, J. D., and P. G. Wolynes. 1987. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. USA.* 84:7524–7528.

23. Ferreiro, D. U., J. A. Hegler, …, P. G. Wolynes. 2011. On the role of frustration in the energy landscapes of allosteric proteins. *Proc. Natl. Acad. Sci. USA.* 108:3499–3503.

24. Zhuravlev, P. I., and G. A. Papoian. 2010. Protein functional landscapes, dynamics, allostery: a tortuous path towards a universal theoretical framework. *Q. Rev. Biophys.* 43:295–332.

25. Zheng, W., N. P. Schafer, …, P. G. Wolynes. 2012. Predictive energy landscapes for protein-protein association. *Proc. Natl. Acad. Sci. USA.* 109:19244–19249.

26. Potoyan, D. A., C. Bueno, …, P. G. Wolynes. 2017. Resolving the NFκB heterodimer binding paradox: strain and frustration guide the binding of dimeric transcription factors. *J. Am. Chem. Soc.* 139: 18558–18566.

27. Weigt, M., R. A. White, …, T. Hwa. 2009. Identification of direct residue contacts in protein-protein interaction by message passing. *Proc. Natl. Acad. Sci. USA.* 106:67–72.

28. Morcos, F., A. Pagnani, …, M. Weigt. 2011. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc. Natl. Acad. Sci. USA.* 108:E1293–E1301.

29. Morcos, F., N. P. Schafer, …, P. G. Wolynes. 2014. Coevolutionary information, protein folding landscapes, and the thermodynamics of natural selection. *Proc. Natl. Acad. Sci. USA.* 111:12408–12413.

30. Sułkowska, J. I., F. Morcos, …, J. N. Onuchic. 2012. Genomics-aided structure prediction. *Proc. Natl. Acad. Sci. USA.* 109:10340–10345.

31. Morcos, F., B. Jana, …, J. N. Onuchic. 2013. Coevolutionary signals across protein lineages help capture multiple protein conformations. *Proc. Natl. Acad. Sci. USA.* 110:20533–20538.

32. Boyd, J. S., R. R. Cheng, …, S. S. Golden. 2016. A combined computational and genetic approach uncovers network interactions of the cyanobacterial circadian clock. *J. Bacteriol.* 198:2439–2447.

33. Cheng, R. R., O. Nordesjö, …, F. Morcos. 2016. Connecting the sequence-space of bacterial signaling proteins to phenotypes using coevolutionary landscapes. *Mol. Biol. Evol.* 33:3054–3064.

34. Cheng, R. R., E. Haglund, …, J. N. Onuchic. 2018. Designing bacterial signaling interactions with coevolutionary landscapes. *PLoS One.* 13:e0201734.

35. dos Santos, R. N., F. Morcos, …, J. N. Onuchic. 2015. Dimeric interactions and complex formation using direct coevolutionary couplings. *Sci. Rep.* 5:13652.

36. Uguzzoni, G., S. John Lovis, …, M. Weigt. 2017. Large-scale identification of coevolution signals across homo-oligomeric protein interfaces by direct coupling analysis. *Proc. Natl. Acad. Sci. USA.* 114:E2662–E2671.

37. Dos Santos, R. N., S. Khan, and F. Morcos. 2018. Characterization of C-ring component assembly in flagellar motors from amino acid coevolution. *R. Soc. Open Sci.* 5:171854.

38. Altschuh, D., A. M. Lesk, …, A. Klug. 1987. Correlation of co-ordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. *J. Mol. Biol.* 193:693–707.

39. Li, G., K. Theys, …, A. M. Vandamme. 2015. A new ensemble coevolution system for detecting HIV-1 protein coevolution. *Biol. Direct.* 10:1–20.

40. Champeimont, R., E. Laine, …, A. Carbone. 2016. Coevolution analysis of Hepatitis C virus genome to identify the structural and functional dependency network of viral proteins. *Sci. Rep.* 6:26401.

41. Shenkin, P. S., B. Erman, and L. D. Mastrandrea. 1991. Information-theoretical entropy as a measure of sequence variability. *Proteins.* 11:297–313.

42. Parra, R. G., N. P. Schafer, …, D. U. Ferreiro. 2016. Protein Frustratometer 2: a tool to localize energetic frustration in protein molecules, now with electrostatics. *Nucleic Acids Res.* 44:W356–W360.

43. Davtyan, A., N. P. Schafer, …, G. A. Papoian. 2012. AWSEM-MD: protein structure prediction using coarse-grained physical potentials and bioinformatically based local structure biasing. *J. Phys. Chem. B.* 116:8494–8503.

44. Noel, J. K., F. Morcos, and J. N. Onuchic. 2016. Sequence co-evolutionary information is a natural partner to minimally-frustrated models of biomolecular dynamics. *F1000 Res.* 5:1–7.

45. Xiao, X., J. Li, and R. J. Samulski. 1998. Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J. Virol.* 72:2224–2232.

46. Rayaprolu, V., S. Kruse, …, B. Bothner. 2013. Comparative analysis of adeno-associated virus capsid stability and dynamics. *J. Virol.* 87:13150–13160.

47. Judd, J., M. L. Ho, …, J. Suh. 2014. Tunable protease-activatable virus nanonodes. *ACS Nano.* 8:4740–4746.

48. Mintseris, J., K. Wiehe, …, Z. Weng. 2005. Protein-protein docking benchmark 2.0: an update. *Proteins.* 60:214–216.

49. Hagan, M. F., and D. Chandler. 2006. Dynamic pathways for viral capsid assembly. *Biophys. J.* 91:42–54.

50. Hagan, M. F. 2008. Controlling viral capsid assembly with templating. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 77:051904.

51. Xie, Q., W. Bu, …, M. S. Chapman. 2002. The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc. Natl. Acad. Sci. USA.* 99:10405–10410.

52. Lochrie, M. A., G. P. Tatsuno, …, P. Colosi. 2006. Mutations on the external surfaces of adeno-associated virus type 2 capsids that affect transduction and neutralization. *J. Virol.* 80:821–834.

53. El-Gebali, S., J. Mistry, …, R. D. Finn. 2019. The Pfam protein families database in 2019. *Nucleic Acids Res.* 47:D427–D432.

54. Cotmore, S. F., M. Agbandje-McKenna, …, A. J. Davison. 2014. The family Parvoviridae. *Arch. Virol.* 159:1239–1247.

55. Yang, Y., J. Gao, …, Y. Zhou. 2018. Sixty-five years of the long march in protein secondary structure prediction: the final stretch? *Brief. Bioinform.* 19:482–494.

56. Papoian, G. A., and P. G. Wolynes. 2017. AWSEM-MD: from neural networks to protein structure prediction and functional dynamics of complex biomolecular assemblies. *In* Coarse-Grained Modeling of Biomolecules. G. A. Papoian, ed. CRC Press, pp. 121–190.

57. Sonntag, F., K. Schmidt, and J. A. Kleinschmidt. 2010. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc. Natl. Acad. Sci. USA.* 107:10220–10225.

58. Cao, M., H. You, and P. L. Hermonat. 2014. The X gene of adeno-associated virus 2 (AAV2) is involved in viral DNA replication. *PLoS One.* 9:e104596.

59. Maurer, A. C., S. Pacouret, …, L. H. Vandenberghe. 2018. The assembly-activating protein promotes stability and interactions between AAV's viral proteins to nucleate capsid assembly. *Cell Rep.* 23: 1817–1830.

60. Bleker, S., F. Sonntag, and J. A. Kleinschmidt. 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.

61. Wistuba, A., S. Weger, …, J. A. Kleinschmidt. 1995. Intermediates of adeno-associated virus type 2 assembly: identification of soluble complexes containing Rep and Cap proteins. *J. Virol.* 69:5311–5319.

62. Im, D. S., and N. Muzyczka. 1990. The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. *Cell.* 61:447–457.

63. Stracker, T. H., G. D. Cassell, …, M. D. Weitzman. 2004. The Rep protein of adeno-associated virus type 2 interacts with single-stranded DNA-binding proteins that enhance viral replication. *J. Virol.* 78:441–453.

64. Prasad, K. M., and J. P. Trempe. 1995. The adeno-associated virus Rep78 protein is covalently linked to viral DNA in a preformed virion. *Virology.* 214:360–370.

65. Steinbach, S., A. Wistuba, …, J. A. Kleinschmidt. 1997. Assembly of adeno-associated virus type 2 capsids in vitro. *J. Gen. Virol.* 78:1453–1462.

66. Katen, S., and A. Zlotnick. 2009. The thermodynamics of virus capsid assembly. *Methods Enzymol.* 455:395–417.

67. Carreira, A., M. Menéndez, …, M. G. Mateu. 2004. In vitro disassembly of a parvovirus capsid and effect on capsid stability of heterologous peptide insertions in surface loops. *J. Biol. Chem.* 279:6517–6525.

68. Horowitz, E. D., K. S. Rahman, …, A. Asokan. 2013. Biophysical and ultrastructural characterization of adeno-associated virus capsid uncoating and genome release. *J. Virol.* 87:2994–3002.

69. Schultz, B. R., and J. S. Chamberlain. 2008. Recombinant adeno-associated virus transduction and integration. *Mol. Ther.* 16:1189–1199.

70. Mateu, M. G. 2013. Structure and Physics of Viruses: An Integrated Textbook. Springer, Dordrecht, the Netherlands.

71. Zlotnick, A. 1994. To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. *J. Mol. Biol.* 241:59–67.

72. Arkhipov, A., P. L. Freddolino, and K. Schulten. 2006. Stability and dynamics of virus capsids described by coarse-grained modeling. *Structure.* 14:1767–1777.

73. Perilla, J. R., B. C. Goh, …, K. Schulten. 2015. Molecular dynamics simulations of large macromolecular complexes. *Curr. Opin. Struct. Biol.* 31:64–74.

74. Sitharam, M., and M. Agbandje-McKenna. 2006. Modeling virus self-assembly pathways: avoiding dynamics using geometric constraint decomposition. *J. Comput. Biol.* 13:1232–1265.

75. Reddy, T., and M. S. P. Sansom. 2016. Computational virology: from the inside out. *Biochim. Biophys. Acta.* 1858:1610–1618.

76. Hagan, M. F., and R. Zandi. 2016. Recent advances in coarse-grained modeling of virus assembly. *Curr. Opin. Virol.* 18:36–43.

77. Gapsys, V., S. Michielssens, …, B. L. de Groot. 2016. Accurate and rigorous prediction of the changes in protein free energies in a large-scale mutation scan. *Angew. Chem. Int.Engl.* 55:7364–7368.

78. Gray, V. E., R. J. Hause, …, D. M. Fowler. 2018. Quantitative missense variant effect prediction using large-scale mutagenesis data. *Cell Syst.* 6:116–124.e3.

79. Rodrigues, C. H. M., D. E. V. Pires, and D. B. Ascher. 2018. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res.* 46:W350–W355.

80. Riesselman, A. J., J. B. Ingraham, and D. S. Marks. 2018. Deep generative models of genetic variation capture the effects of mutations. *Nat. Methods.* 15:816–822.

81. Vandenberghe, L. H., E. Breous, …, J. M. Wilson. 2009. Naturally occurring singleton residues in AAV capsid impact vector performance and illustrate structural constraints. *Gene Ther.* 16:1416–1428.

82. Koerber, J. T., J.-H. Jang, and D. V. Schaffer. 2008. DNA shuffling of adeno-associated virus yields functionally diverse viral progeny. *Mol. Ther.* 16:1703–1709.

83. Tamir, S., S. Rotem-Bamberger, …, R. Nechushtai. 2014. Integrated strategy reveals the protein interface between cancer targets Bcl-2 and NAF-1. *Proc. Natl. Acad. Sci. USA.* 111:5177–5182.

84. Cheng, R. R., F. Morcos, …, J. N. Onuchic. 2014. Toward rationally redesigning bacterial two-component signaling systems using coevolutionary information. *Proc. Natl. Acad. Sci. USA.* 111:E563–E571.

85. Wright, T. A., J. M. Stewart, …, D. Konkolewicz. 2017. Extraction of thermodynamic parameters of protein unfolding using parallelized differential scanning fluorimetry. *J. Phys. Chem. Lett.* 8:553–558.
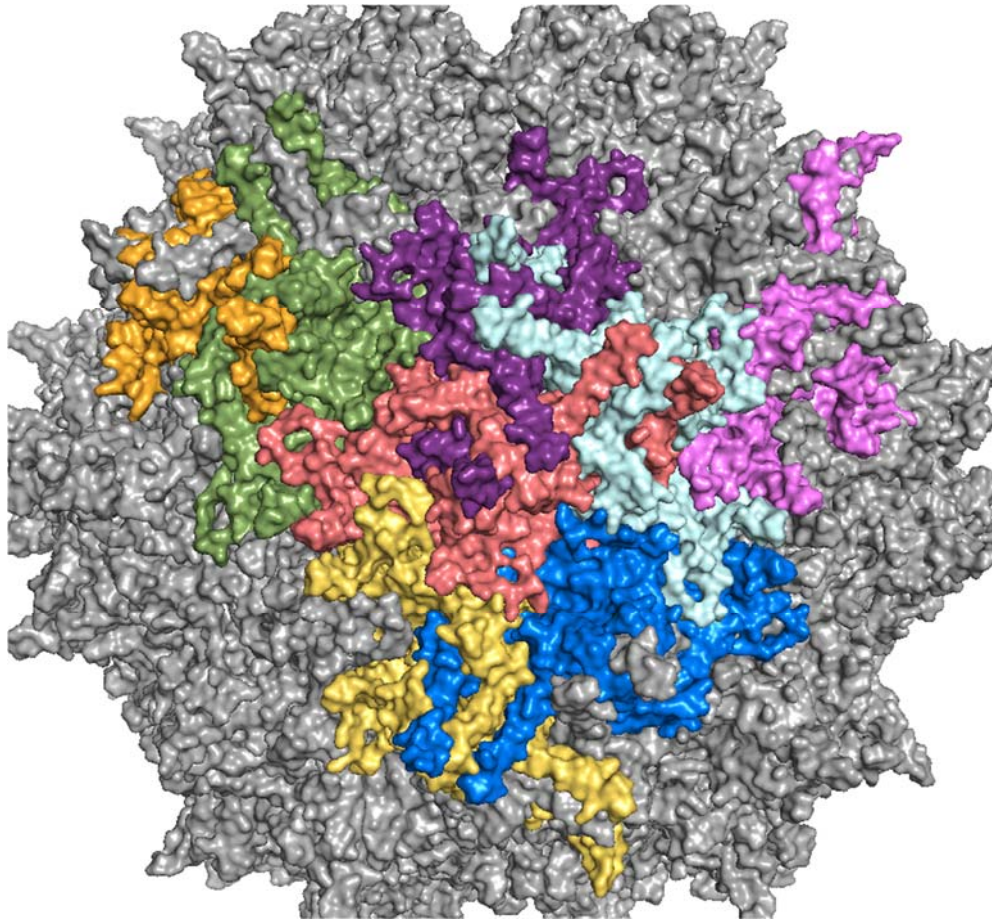
# Supplemental Information

# Frustration and Direct-Coupling Analyses to Predict Formation and Function of Adeno-Associated Virus

**Nicole N. Thadani, Qin Zhou, Kiara Reyes Gamas, Susan Butler, Carlos Bueno, Nicholas P. Schafer, Faruck Morcos, Peter G. Wolynes, and Junghae Suh**

**Figure S1: AAV2 monomer and multimer substructures (related to Fig. 1).** The AAV2 monomer (coral) is defined as one VP3 subunit. The AAV2 multimer is defined as the central monomer (coral) and the seven surrounding monomers containing a residue with any atom within a radius of 4.5Å around any of the central monomer residues. The rest of the AAV capsid (grey) was not included in this analysis.
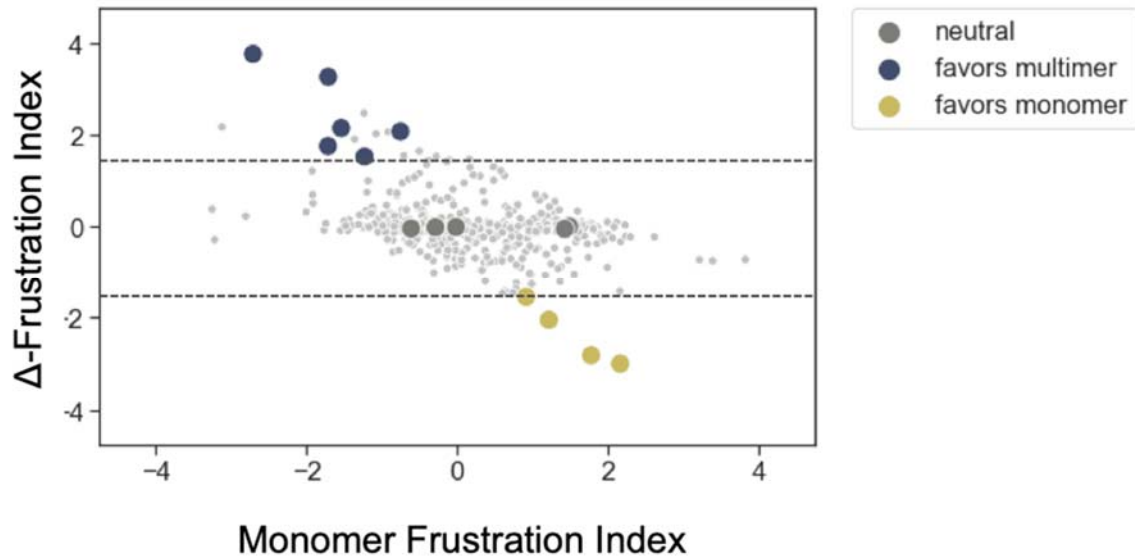
**Figure S2: AAV2 Δ-frustration index plotted against the monomer frustration index (related to Fig. 2).** The AAV2 Δ-frustration index is computed as the difference between the single-residue mutational frustration index of the multimer and the single-residue mutational frustration index of the monomer. Residues with a positive Δ-frustration index are predicted to energetically favor the assembly state, while residues with a negative Δ-frustration index are predicted to favor the monomer. Residues selected for mutation are large colored circles. Residues not selected for mutation are small grey dots. Dashed lines indicate the +/- 1.5 Δ-frustration index thresholds used to select residues for mutation.
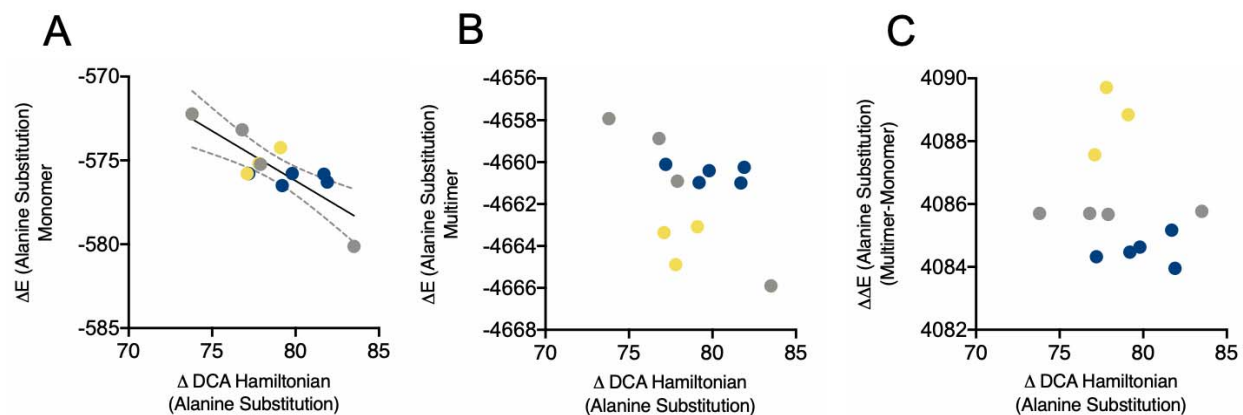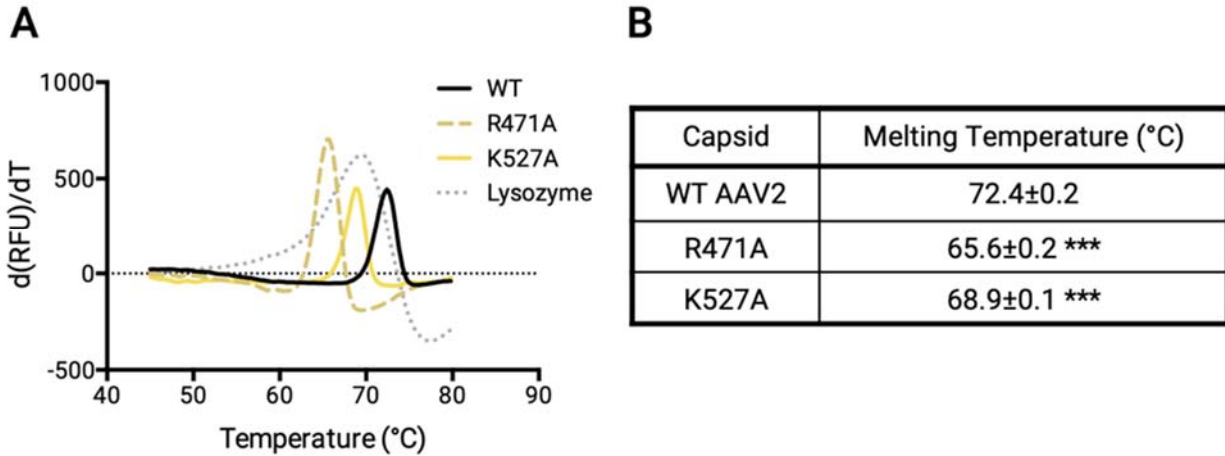
**Figure S3: AAV2 $\Delta H_{DCA(Ala)}$ plotted against ΔE-monomer, ΔE-multimer, and ΔΔE of alanine substitutions (related to Fig. 3).** AAV2 $\Delta H_{DCA(Ala)}$ is plotted against ΔE-monomer **(A)**, ΔE-multimer **(B)**, and ΔΔE **(C).** N=12. R471A, F712A, and S721A were excluded from these analyses due to a high gap frequency at their alignment position. Regression lines with 95% confidence interval are plotted for statistically significant correlations.
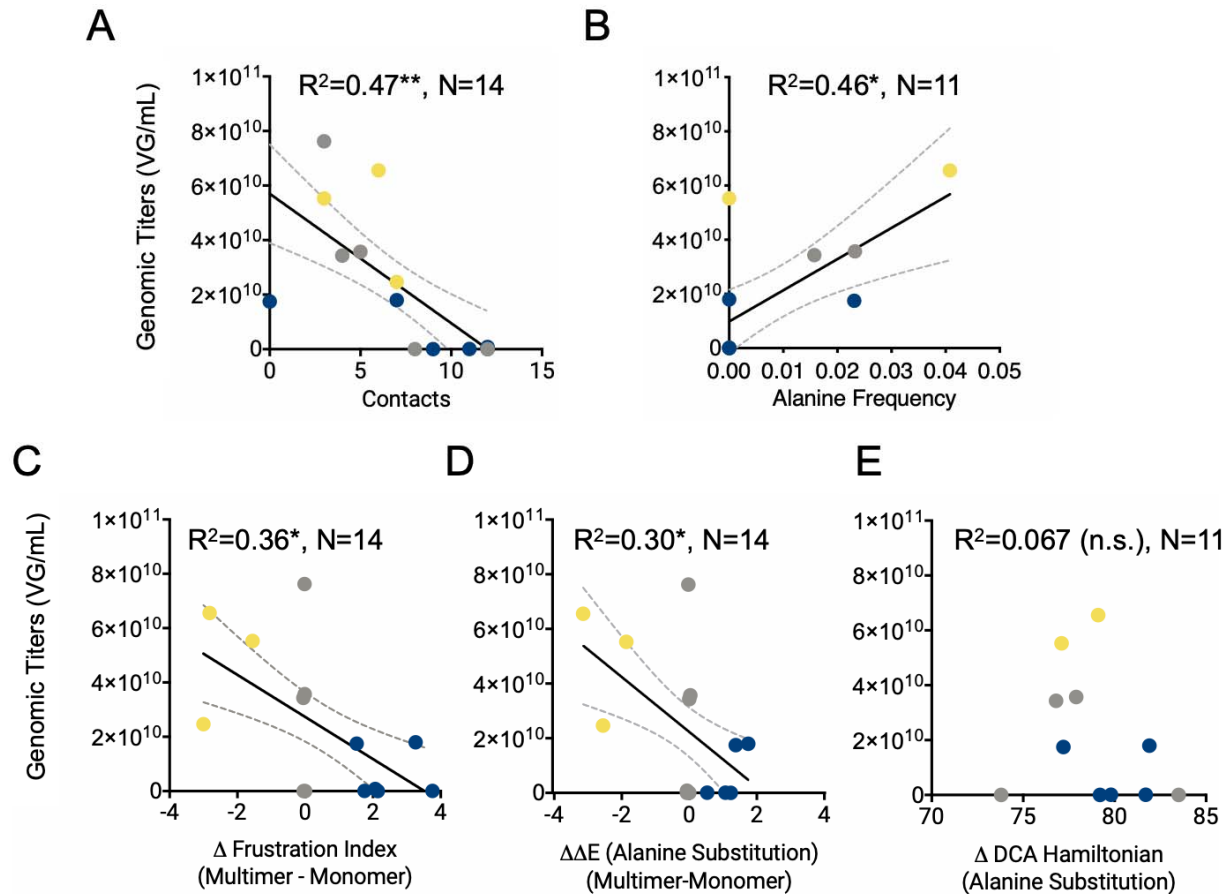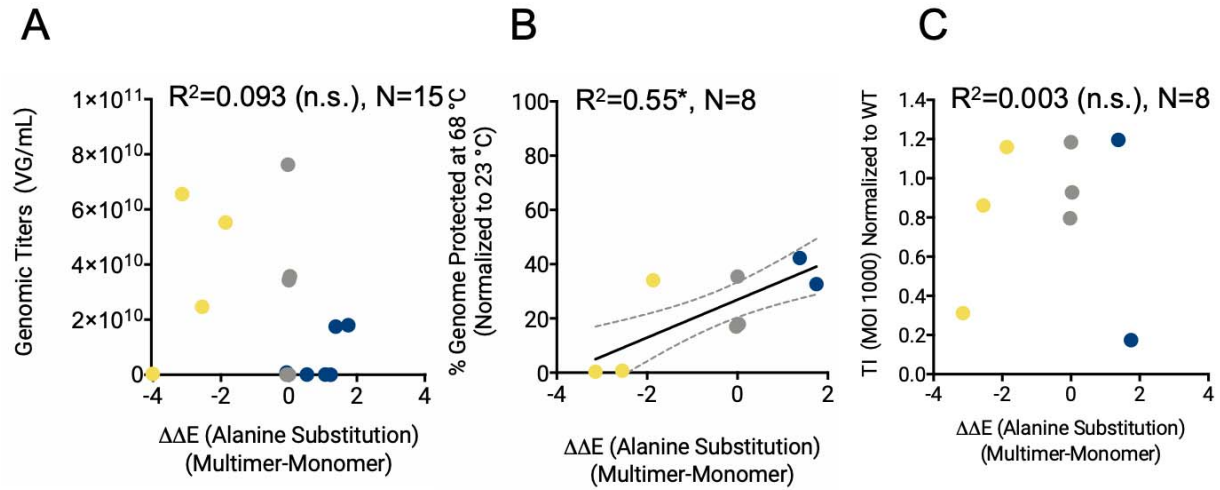
**Figure S4: Correlation of AAV2 genomic titers with computational metrics excluding the P657A mutant (related to Fig. 5)**. Genomic titers are plotted against (**A**) number of residue-residue contacts in the native capsid structure, (**B**) alanine frequency in the parvovirus sequence family, (**C**) Δ-frustration index, (**D**) ΔΔE of alanine substitutions, and (**E**) $\Delta H_{DCA(Ala)}$. N=14 mutants excluding P657A for structure-based analyses. N=11 mutants excluding P657A for sequence-based analyses as R471A, F712A, and S721A were excluded from these analyses due to a high gap frequency at their alignment position. Regression lines with 95% confidence interval are plotted for statistically significant correlations. $R^2$ values and significance levels are reported (*p<0.05, **p<0.01).

**A**

**B**

| Capsid | Melting Temperature (°C) |
|---------|--------------------------|
| WT AAV2 | 72.4±0.2 |
| R471A | 65.6±0.2 *** |
| K527A | 68.9±0.1 *** |

**Figure S5: Melting temperature of alanine mutant capsids R471A and K527A (related to Fig. 6)**. The melting temperatures of selected capsids with reductions in thermal stability were determined using differential scanning fluorescence. (**A**) Derivative in fluorescence intensity with heating of capsids. Fluorescence increases upon the binding of Sypro Orange to exposed hydrophobic pockets as capsids denature. Traces are the average of N=3 independent experiments. Lysozyme is included as a control, which denatures at 69.4°C$\pm$0.2, comparable to the previously reported melting temperature of 68.6°C (1). (**B**) Melting temperature of capsids were determined as the peak of plots shown in panel A. N=3 independent experiments in duplicate, SEM reported. One-way ANOVA was performed with Dunnett's post-hoc multiple comparison test to compare melting temperatures of mutants to WT. ***p<0.001

**Figure S6 (related to Fig. 5-7): Correlation of ΔΔE of alanine substitutions with experimental measures:** ΔΔE of alanine substitutions are plotted against (**A**) genomic titers, (**B**) % genomic protection at 68 °C, and (**C**) transduction index. Regression lines with 95% confidence interval are plotted for statistically significant correlations. $R^2$ values and significance levels are reported (*p<0.05).
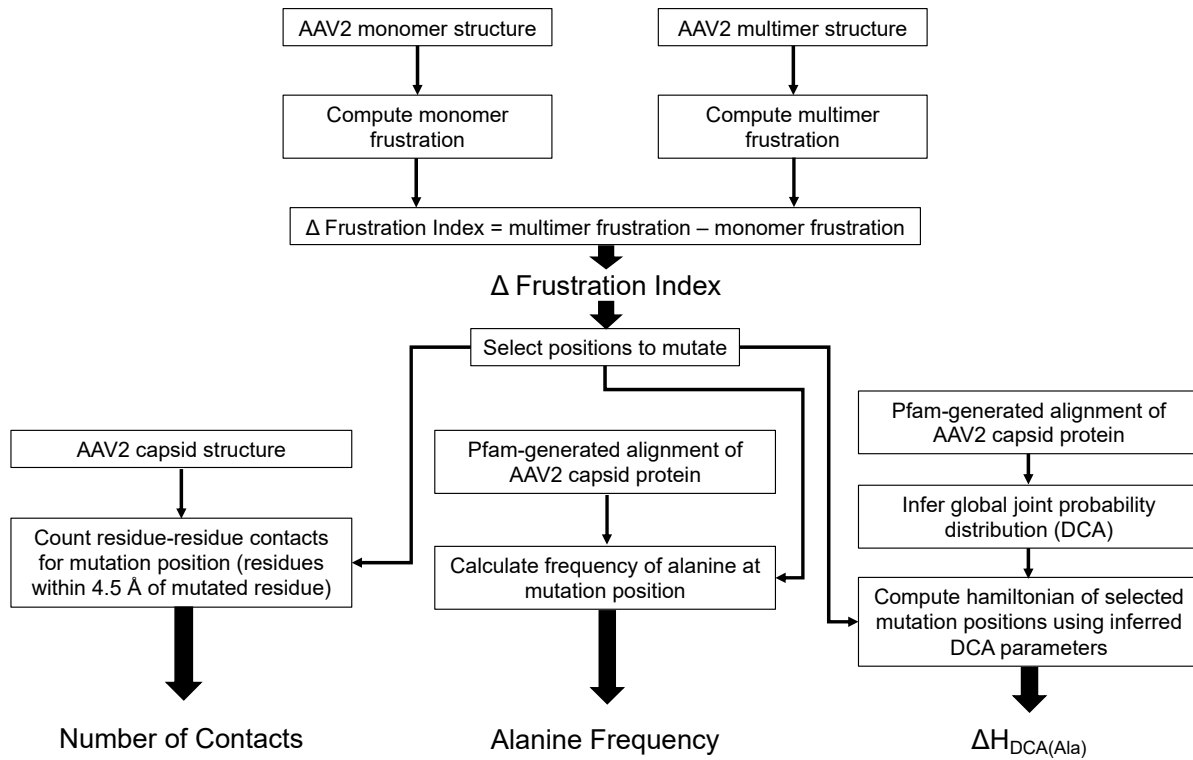
**Figure S7: Overview of computational methods and their application in this study.**

**Supporting Data File 1 (related to methods):** Hidden markov model used to find sequences in the parvovirus sequence family (Text file)

**Supporting Data File 2 (related to methods):** Multiple sequence alignment generated by using HMM to query for sequences in the parvovirus sequence family (Text file)

**Supporting Data File 3 (related to methods):** Raw qPCR fluorescence curves and plate layouts (Excel files)

**Supporting Data File 4 (related to methods):** Flow cytometry counts and thresholds (CSV files)

**Table S1 (related to Fig. 3):** List of top 300 couplings found using DCA and their direct information scores. (Excel file)

**Table S2 (related to Fig. 5-7):** R-squared values of linear regression between structure-based variables and experimental measures excluding residues that were excluded for sequence-based measures (R471A, F712A, and S721A).

| | Titers | Genomic Protection | Transduction |
|---|---|---|---|
| Structural Contacts (distance < 4.5 Å) | 0.28 *(n.s.)* | 0.36 *(n.s.)* | 0.66* |
| ΔΔE | 0.17 *(n.s.)* | 0.49 *(n.s.)* | 0.007 *(n.s.)* |
| ΔF | 0.31 *(n.s.)* | 0.40 *(n.s.)* | 0.017 *(n.s.)* |

**Table S3 (related to Fig. 5):** Mutations introduced in the AAV assembly-activating protein and x-genes in alternate reading frames on the *cap* gene.

| *Cap* gene mutation | AAP gene mutations | X gene mutations |
|---|---|---|
| W228A | I52M, G53R | N/A |
| R238A | E63Q | N/A |
| R245A | E70Q | N/A |
| E347A | S172R | N/A |
| Y397A | N/A | N/A |
| Y441A | N/A | N/A |
| R471A | N/A | N/A |
| K527A | N/A | N/A |
| V611A | N/A | C36R |
| L647A | N/A | F71L, S72P |
| N656A | N/A | I81L |
| P657A | N/A | I81M |
| W694A | N/A | G119R |
| F712A | N/A | N/A |
| S721A | N/A | I145M |

# Supporting References

1. Wright, T.A., J.M. Stewart, R.C. Page, and D. Konkolewicz. 2017. Extraction of Thermodynamic Parameters of Protein Unfolding Using Parallelized Differential Scanning Fluorimetry. *J. Phys. Chem. Lett.* 8:553–558.