

Disenfranchised DNA: biochemical analysis of mutant øX174 DNA-binding proteins may further elucidate the evolutionary significance of the unessential packaging protein A*

Samuel D. Love,¹ Sierra Posey,² April D. Burch,² Bentley A. Fane¹

AUTHOR AFFILIATIONS See affiliation list on p. 15.

ABSTRACT Most icosahedral DNA viruses package and condense their genomes into pre-formed, volumetrically constrained capsids. However, concurrent genome biosynthesis and packaging are specific to single-stranded (ss) DNA micro- and parvoviruses. Before packaging, ~120 copies of the øX174 DNA-binding protein J interact with double-stranded DNA. 60 J proteins enter the procapsid with the ssDNA genome, guiding it between 60 icosahedrally ordered DNA-binding pockets formed by the capsid proteins. Although J proteins are small, 28–37 residues in length, they have two domains. The basic, positively charged N-terminus guides the genome between binding pockets, whereas the C-terminus acts as an anchor to the capsid's inner surface. Three C-terminal aromatic residues, W30, Y31, and F37, interact most extensively with the coat protein. Their corresponding codons were mutated, and the resulting strains were biochemically and genetically characterized. Depending on the mutation, the substitutions produced unstable packaging complexes, unstable virions, infectious progeny, or particles packaged with smaller genomes, the latter being a novel phenomenon. The smaller genomes contained internal deletions. The juncture sequences suggest that the unessential A* (A star) protein mediates deletion formation.

IMPORTANCE Unessential but strongly conserved gene products are understudied, especially when mutations do not confer discernable phenotypes or the protein's contribution to fitness is too small to reliably determine in laboratory-based assays. Consequently, their functions and evolutionary impact remain obscure. The data presented herein suggest that microvirus A* proteins, discovered over 40 years ago, may hasten the termination of non-productive packaging events. Thus, performing a salvage function by liberating the reusable components of the failed packaging complexes, such as DNA templates and replication enzymes.

KEYWORDS single-stranded DNA packaging, single-stranded DNA replication, phiX174, microviridae

During packaging, icosahedral viruses condense their genomes into volumetrically constrained capsids. Capsid assembly and packaging can be simultaneous or sequential processes. Many single-stranded (ss) RNA genomes contain capsid protein binding sequences that nucleate assembly around the folded genome (1, 2). By contrast, double-stranded (ds) DNA genomes are packaged into pre-formed procapsids. In the virion, the genome assumes a cylindrical spool configuration to accommodate the high DNA density (3). SsDNA packaging exhibits characteristics found in both dsDNA and ssRNA systems. Like dsDNA, the øX174 genome is packaged into a preformed structure (4, 5). Reminiscent of the extensive capsid-nucleic acid associations seen in

Editor Kristin N. Parent, Michigan State University, East Lansing, Michigan, USA

Address correspondence to Bentley A. Fane, bfane@arizona.edu.

The authors declare no conflict of interest.

See the funding table on p. 15.

Received 26 November 2023

Accepted 10 January 2024

Published 2 February 2024

Copyright © 2024 American Society for Microbiology. All Rights Reserved.

ssRNA viruses and encapsidated microvirus genomes are partially ordered within the icosahedral symmetry *via* 60 DNA-binding pockets (6–8).

øX174 DNA synthesis and packaging

DNA packaging and ssDNA genome biosynthesis occur in the third stage of DNA replication. During the first stage (Stage I), 13 host cell proteins convert the penetrating single-stranded genome to a double-stranded molecule (9, 10). After viral protein synthesis, the øX174 A protein covalently attaches to the origin of replication to initiate Stage II rolling circle replication (11–15). The amplification of double-stranded, replicative form DNA continues until inhibited by the øX174 C protein (16). In Stage III DNA replication, ssDNA genomes are synthesized and packaged concurrently. The parental strand is displaced into the procapsid as the daughter strand is synthesized (Fig. 1A).

Packaging also requires the øX174 DNA-binding protein J, which is a small, positively charged (+12) peptide, only 37 amino acids in length. As seen in Fig. 1B, the N-terminus is extremely basic to interact with negatively charged phosphates in the DNA backbone. In the atomic structure, these domains appear to guide the DNA between the binding pockets (Fig. 1C), whereas the strongly conserved C-terminus binds to the coat protein (6–8, 17, 18). Two previous studies addressed the guiding function of the protein. J genes from G4 or α3 (Fig. 1B) were substituted for the indigenous gene to produce chimeric viruses: øXG4J and øXα3J (19, 20). Packaging the øX174 genome with the α3 J protein results in filled capsids that lack infectivity, whereas the G4 J protein causes packaging complexes to become unstable or partition into an insoluble fraction. Duplicating the strongest A* protein target site or elevated levels of the A* protein (described below) rescues øXG4J. However, it does not restore viability to the filled øXα3J virus-like particles, suggesting that protein A* is active during packaging not afterward.

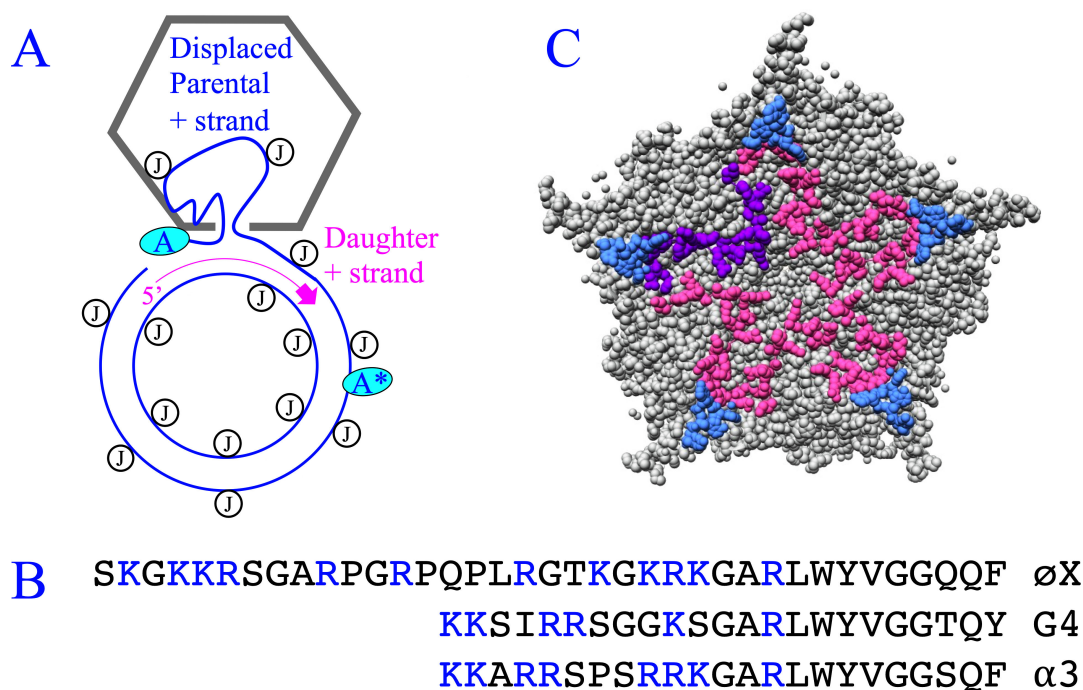


FIG 1 øX174 genome packaging and the DNA-binding protein J. (A) DNA packaging: the parental + strand is displaced into the procapsid as the daughter + strand is synthesized. (B) The amino acid sequence of the øX174, G4, and α3 J proteins. (C) Interior structure of a virion pentamer (PDB 2BPA). The coat proteins are depicted in gray. One J protein is highlighted in purple; the rest are depicted in magenta. Light blue represents the ordered DNA backbone within the DNA-binding pocket.

The essential A and unessential A* proteins

The essential A protein mediates rolling circle DNA replication. øX174-like viruses also produce an N-terminally truncated version, known as protein A*. Although strongly conserved (21), the A* protein is not essential. Comparable A/A*-like arrangements are seen in the ssDNA parvo- and circoviruses rolling circle (rep) proteins (22, 23). Eliminating A* gene expression requires mutating its ATG start codon (24), which is in the frame and nested within the large A gene. Therefore, the mutant phenotype, a modest reduction in burst size, could be caused by the lack of protein A* and/or the missense substitution in protein A.

The biochemical properties of the øX174 A and A* proteins have been investigated *in vitro*. Enzymatically, both proteins are nucleases that covalently bind DNA but re-ligate it when they dissociate. However, they differ in substrate recognition. Protein A only recognizes the origin of replication in a single-stranded DNA (25, 26). By contrast, protein A* recognizes at least 30 origin-like sequences in both single-stranded and double-stranded molecules (27). Once attached, protein A mediates rolling circle DNA replication, whereas protein A* bound to dsDNA inhibits the process by hindering host cell helicase function (28). Impeding the unwinding of dsDNA may ensure that genome biosynthesis does not outpace the packaging (Fig. 1A), which would likely be deleterious. Unpackaged ssDNA would accumulate outside the capsid where it could form secondary structures that may not pass through the 30 Å wide packaging pore.

Three J protein C-terminal residues, W30, Y31, and F37, mediate most of the J-coat protein interactions within the virion (7, 8), anchoring it to the capsid's inner surface. Herein, the packaging defects conferred by mutations at these sites were characterized. In addition, the results may expand the known functions of protein A* in both mutant and wild-type infections.

RESULTS

Phenotypes associated with substitutions for J protein residues W30, Y31, and F37

The W30, Y31, and F37 codons were randomized by site-directed mutagenesis. In addition, specific mutations were synthesized to obtain a standardized set of substitutions. These substitutions include F, Y, H, and W to assay the effects of aromatic substitutions; R, E, and H to assay the effects of charge; N and S for polar substitutions; and V, A, and M for nonpolar ones. In addition to being nonpolar, methionine residues are known to create stable interactions with aromatic rings (29) and have been shown to mimic some of the π - π stacking interactions between the coat and internal scaffolding proteins (30). Mutant nomenclature reflects the phenotype and the substitution made within the J protein. Thus, *cs(J)W30Y* conveys a cold-sensitive mutation in gene J: *cs(J)*; that arose from a W→Y substitution at amino acid 30.

The W30 site appeared to be the most sensitive to alterations (Table 1; Fig. 2A). Even the most conservative substitutions, tyrosine and phenylalanine, conferred cold-sensitive (*cs*) phenotypes, whereas the histidine substitution was lethal. However, methionine was well tolerated at this site. By contrast, the F37 site appeared to be the most tolerant. All nonpolar and aromatic substitutions were tolerated, with most conferring wild-type or near wild-type phenotypes.

Histidine, which is both charged and aromatic, conferred temperature-sensitive (*ts*) phenotypes at the Y31 and F37 sites and a lethal phenotype at W30. Temperature sensitivity is in contrast with the cold-sensitive nature of microvirus packaging (31). *Ts* phenotypes are most commonly associated with improper protein folding. However, it is unlikely that substitutions would dramatically change J protein structure or folding as protein J contains no units of secondary structure (Fig. 2B). Moreover, the *ts* phenotype is suppressed *via* an extragenic mechanism (see below).

TABLE 1 Allele specificity of the S1F and D209H suppressors in the coat protein

Parental mutant	Parental phenotype	EOP ^a with and without suppressors			Other coat suppressors ^b
		None ^c	S1F	D209H	
W30F	cs	10 ⁻⁴	10 ⁻⁶	1.6	P93F
W30Y ^d	weak cs	10 ⁻¹ –10 ⁻²	0.2	0.8	L324F
W30H	lethal	10 ⁻⁴	0.1	10 ⁻⁵	
W30V	lethal	10 ⁻⁴	10 ⁻³	10 ⁻⁵	L324F
W30N	lethal	10 ⁻⁵	10 ⁻⁶	10 ⁻⁵	
W30E	lethal	<10 ⁻³	ND ^a	ND	
W30R	Lethal	10 ⁻⁶	10 ⁻⁶	10 ⁻⁶	
Y31F	WT	1.0	NA ^a	NA	
Y31W	cs	10 ⁻³	10 ⁻⁴	0.6	Q23H
Y31H	ts	10 ⁻⁴	1.1	1.1	
Y31V	lethal	10 ⁻⁴	0.5	1.0	L324F
Y31N	lethal	10 ⁻³	0.5	10 ⁻⁶	
Y31E	lethal	10 ⁻⁴	10 ⁻⁵	10 ⁻⁵	
Y31R	lethal	10 ⁻⁵	10 ⁻⁶	10 ⁻⁵	
Y31M ^d	lethal	10 ⁻⁴	0.4	1.0	N90D, A230G, L235I
F37Y	WT	1.0	NA	NA	
F37W	WT	0.5	NA	NA	
F37H	ts	10 ⁻⁴	0.3	0.3	
F37V	WT	1.0	NA	NA	
F37N	lethal	10 ⁻³	1.2	0.5	L236F
F37E	lethal	10 ⁻⁴	0.8	0.8	D65G
F37R	lethal	10 ⁻⁵	10 ⁻⁵	10 ⁻⁵	

^aEOP (efficiency of plating restrictive titer/permissive titer). ND: not determined. NA: not applicable.

^bOther coat protein mutations that suppressed the defective phenotype.

^cThe EOP values were determined under restrictive conditions of the parental mutant.

^dThe W30Y mutant was rescued by an intragenic suppressor Q35R, an extragenic suppressor in the external scaffolding protein T111I, and a mutation in the gene E (T74M), which encodes the lysis protein. Y31M mutation was also suppressed by a substitution in the external scaffolding protein V3F.

Mutants with conservative substitutions conferring lethal and cold-sensitive phenotypes can produce infectious particles in lysis-deficient cells

Except *I(J)W30E*, the mutant assembly pathway was biochemically characterized for all of the mutants displaying defective phenotypes. Despite numerous attempts, it was not possible to generate the requisite high titer *I(J)W30E* stock with a suitably low reversion frequency. To characterize the infection products generated under restrictive conditions, lysis-resistant cells were infected at a multiplicity of infection (MOI) of 5.0 and incubated for 6 h at 24°C and 3 h at 37°C or 42°C. Soluble extracts were prepared and analyzed by rate zonal sedimentation. Procapsids (108S) and virions (114S) have similar S values. To separate these two species, the external scaffolding protein was removed from procapsids by chloroform treatment. In a wild-type infection, this typically results in unfilled degraded procapsids, which have little or no DNA, that sediment at ~70S. After gradient fractionation, assembled particles were detected by UV spectroscopy.

Despite lower yields (Fig. 3A), the *cs(J)Y31W* mutant produced particles that sediment like wild-type virions at both 24°C (open circles) and 37°C (closed circles). No significant differences in specific infectivity (pfu/A₂₆₀) between mutant and wild-type particles were observed. A similar phenomenon was observed for *cs(J)W30F* (Fig. 3B, cyan), but mutant yields were significantly lower. For graphing purposes, the A₂₆₀ values of the wild-type fractions were divided by 10. Thus, the highest absorbance values were over 0.7. Unlike the wild-type and *cs(J)Y31W* sedimentation profiles, a distinct ~70S peak was detected in the *cs(J)W30F* profile. However, particle recovery was too low to adequately analyze the protein and DNA content at this peak.

Histidine is a basic amino acid with aromatic properties. While more conservative aromatic substitutions yielded viable or cold-sensitive phenotypes (Table 1; Fig. 2A), the Y31H and F37H substitutions produced ts phenotypes, whereas the W30H substitution

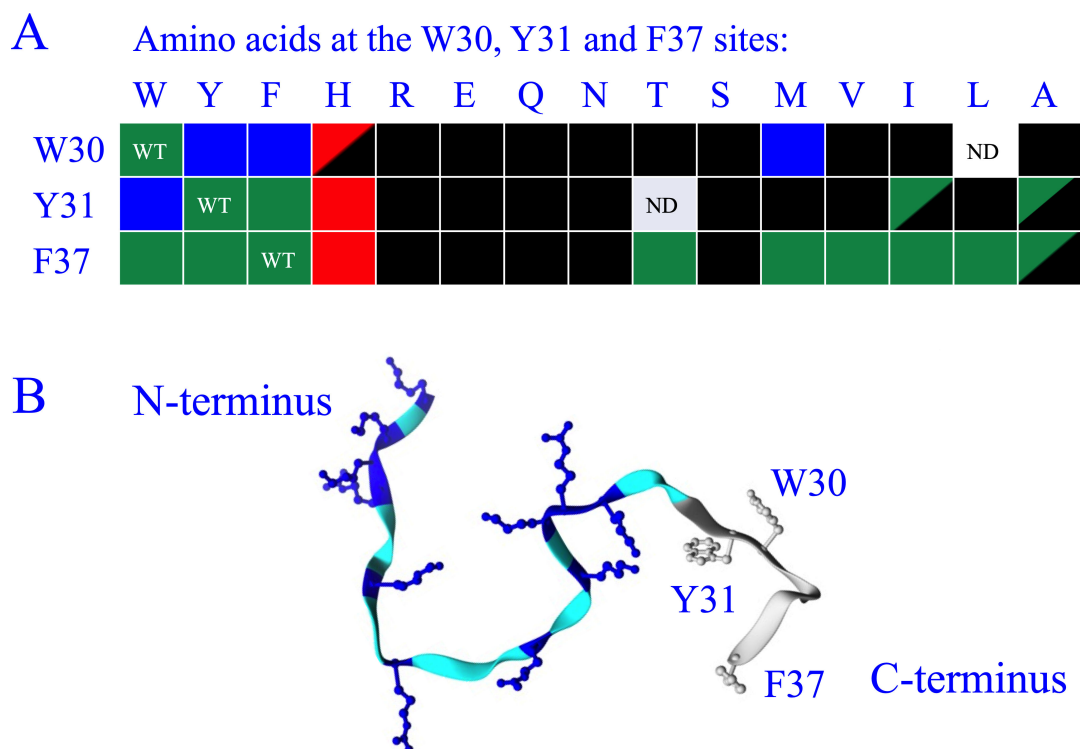


FIG 2 Phenotypes associated with amino acid substitutions for W30, Y31, and F37. (A) The amino acids found at the W30, Y31, and F37 sites are listed at the top of the grid. Color code: black, lethal; green, viable between 24–42°C; blue, cold sensitive; red, temperature sensitive. The split green/black squares indicate small plaque formation. The red/black color for H at the W30 site indicates a slightly less lethal phenotype at 24°C. See text for details. Symbols: WT: wild type, ND: not determined. (B) The atomic structure of protein J (PDB 2BPA). Basic amino acid side chains (R and K) are highlighted in dark blue. The conserved C-terminus, which contains the aromatic side chains, is depicted in gray.

conferred a lethal one. The more severe *I(J)W30H* phenotype is consistent with W30 being the least tolerant of substitutions (Table 1; Fig. 2A). The *ts(J)F37H*, *ts(J)Y31H*, and *I(J)W30H* assembly products produced at restrictive temperatures were characterized (Fig. 3 panels C and D, respectively). At 42°C, no assembled particles were detected in the extract generated from *ts(J)F37H*-infected cells (Fig. 3, Panel C). However, SDS-PAGE of infected cells indicated that the viral coat protein was present at levels comparable to the wild-type control (data not shown). This suggests that assembled particles either dissociated or partitioned into the insoluble fraction. Similar results were obtained for *I(J)W30H* at 37°C.

The lethal *I(J)W30H* phenotype was defined by plaque formation. However, *I(J)W30H* virions were detected in extracts generated at 24°C, (Fig. 3, Panel B). Thus, there was a *ts* component to the lethality. However, yields were low compared to wild type and *ts(J)F37H* at this temperature (Fig. 3, Panel B). For graphing purposes, the wild-type and *ts(J)F37H* absorbance readings were divided by 10 and 5, respectively.

Three mutants produce assembled particles more indicative of packaging defects

As exhibited in Fig. 3A through C, most mutants produced very few or no detectable assembled particles. The protein content of infected cells for all of these mutants was directly examined by SDS-PAGE, which would detect viral proteins regardless of particle solubility or stability. The viral coat protein was present at levels comparable to the wild-type control (data not shown), again suggesting that procapsid-containing intermediates either dissociate or become insoluble. However, the extracts made from *ts(J)Y31H*, *I(J)F37E*, and *I(J)Y31V* contained substantial material with S values $\geq 70S$ (Fig.

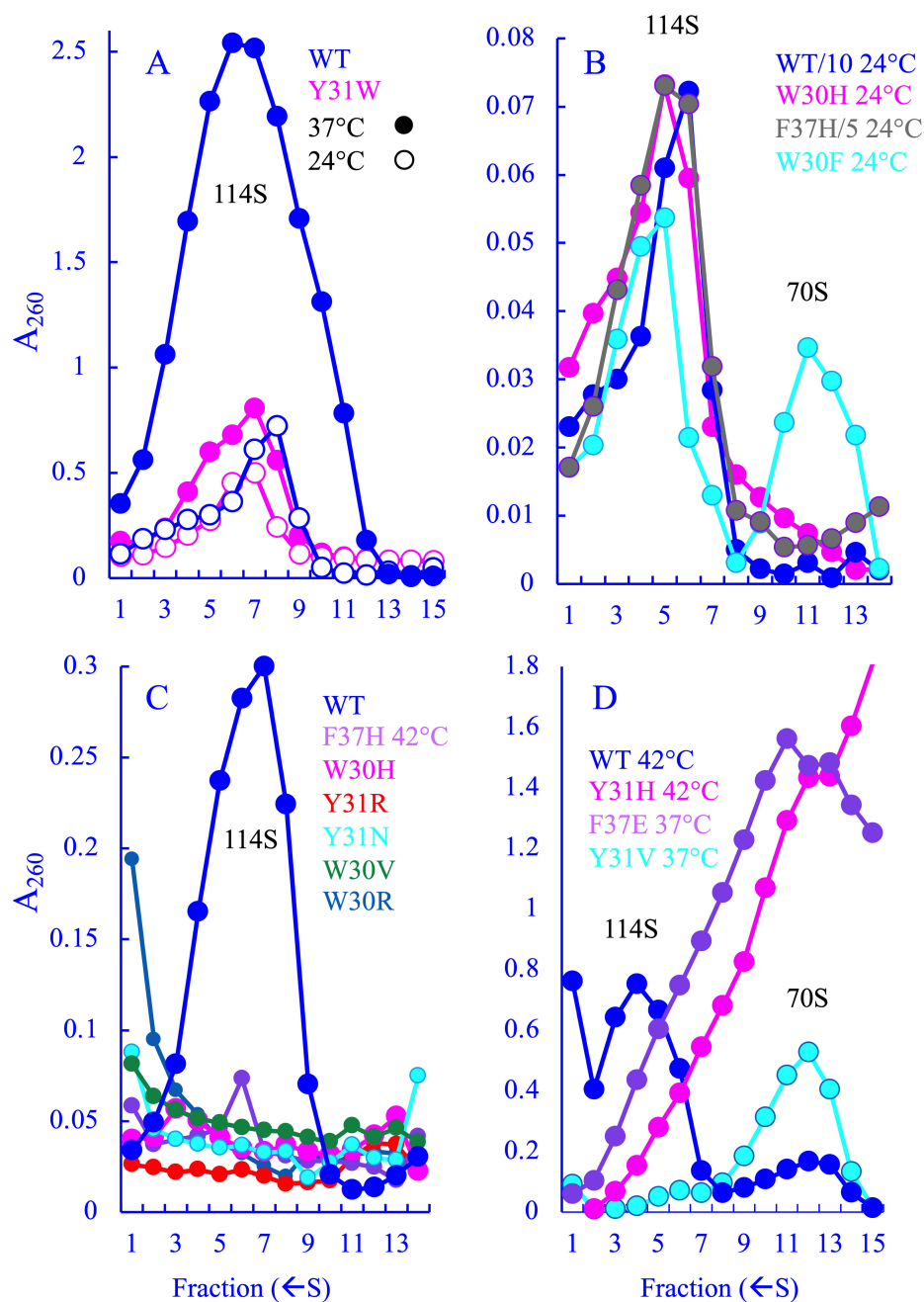


FIG 3 Assembled particle formation in wild-type and øX174 mutant infected cells. Extracts made from infected cells were analyzed by the rate of zonal sedimentation. Gradients were fractionated from the bottom of the tube. Thus, lower fractions contain particles with higher S values. Mutant names and infection temperatures are given in the figure. (A and B). Mutants that produce virions but at reduced yields compared to wild type. (C) Mutants that do not produce soluble assembled particles. The sedimentation profiles of mutants *I(J)Y31E*, *I(J)W30R*, *I(J)Y31M*, and *L(J)Y31L* are not depicted but yielded similar results. (D). Mutants that produced particles that sediment at ~70S.

3D). For *ts(J)Y31H* and *I(J)F37E*, A_{260} values steadily rose sharply from 114S ($A_{260} \approx 0.3$) to 70S and then peaked or plateaued ($A_{260} \approx 1.5$). The sedimentation profile generated from *I(J)Y31V* infected cells only contained a distinct ~70S peak.

The protein and DNA content present in the above-mentioned *ts(J)Y31H*, *I(J)Y31V*, and *I(J)F37E* peaks were analyzed along with a wild-type control. To determine protein content, gradient fractions were examined by SDS-PAGE. As can be seen in Fig. 4A, the

intensity pattern of the coat protein band mirrors the absorbance values in Fig. 3D. In wild-type fractions, the coat protein band intensity peaked in fractions 4–6, whereas in mutant fractions, intensity peaked in fractions 10–12. This, along with the lack of clear peaks in the *ts(J)Y31H* and *I(J)F37E* sedimentation profiles, suggests that particles with S values between 70 and 114S were heterogeneous.

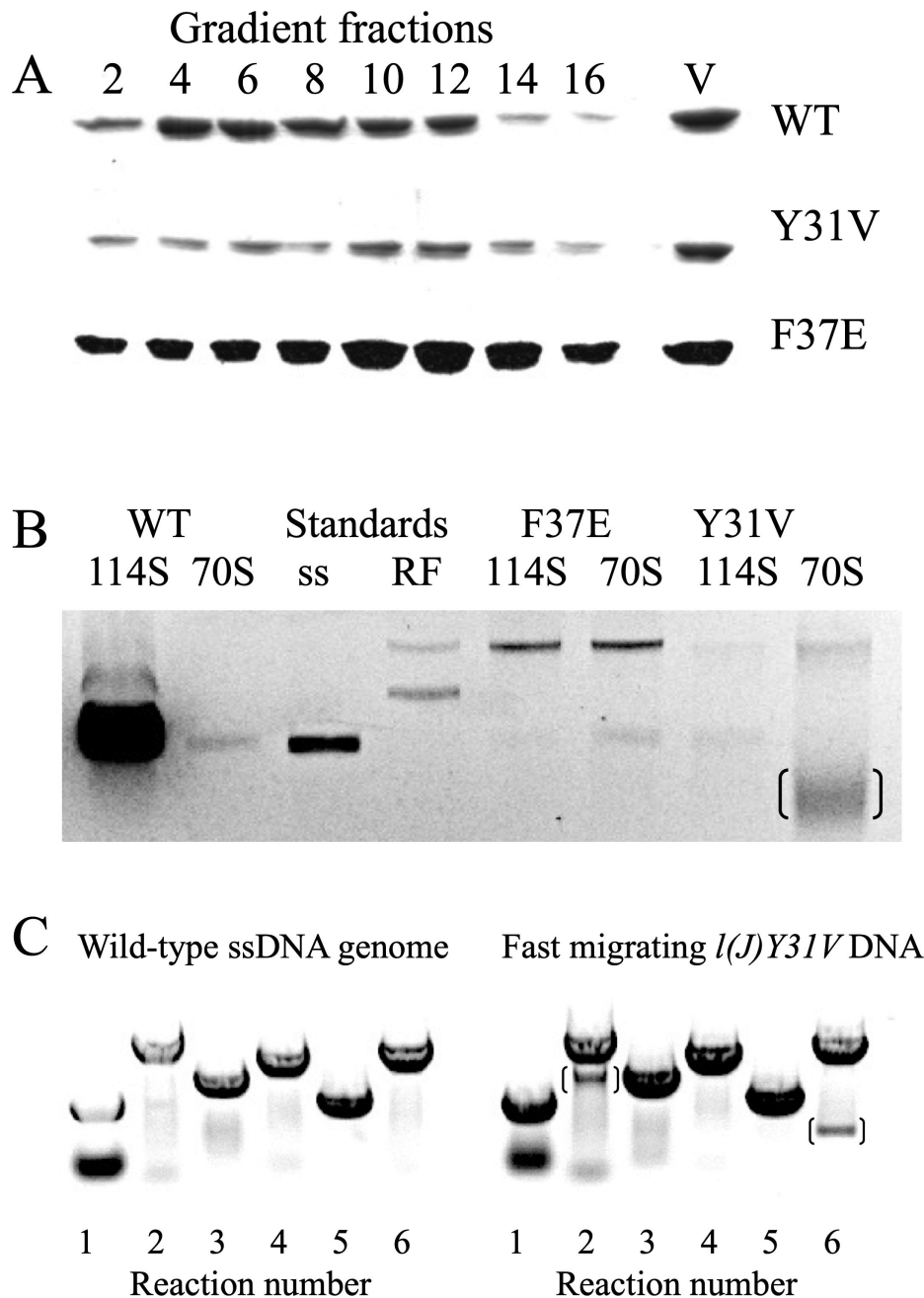


FIG 4 Composition of assembled particles generated by the *I(J)F37E* and *I(J)Y31V* packaging mutants. (A) SDS-PAGE of gradient fractions from Fig. 3D. Only the region of the gel with the viral coat protein is depicted. Band intensity corresponds to the A260 readings reported in Fig. 3D. V indicates purified viral coat protein. (B) DNA isolated from 114S and 70S gradient fractions. Symbols: ss, purified single-stranded DNA genomes; RF, replicative form dsDNA. The bracketed band identifies the small aberrant ssDNA associated with the *I(J)Y31V* 70S material. (C) PCRs were conducted with wild-type and *I(J)Y31V* 70S DNA. The six PCRs span the entire genome. Reaction numbers utilized the same primers for each template. Bracketed bands identify PCR fragments unique to the *I(J)Y31V* 70S DNA template.

The DNA content associated with wild-type and mutant particles was also examined (Fig. 4B). Gradient fractions surrounding the virion and ~70S peaks from the *I(J)F37E*, *I(J)Y31V*, and wild-type gradients were individually pooled, and the associated DNA was extracted. The DNA isolated from wild-type 114S particles was primarily single stranded. The sample was intentionally overloaded to demonstrate that a small amount of replicative form (RF) dsDNA was also detected. By contrast, the material associated with both the *I(J)F37E* 114S and 70S fractions was mostly RF DNA. However, single-stranded genomes were detected albeit at much lower levels. The DNA extracted from the *I(J)Y31V* 70S fractions was unique. It contained a faster migrating heterogeneous pool of undefined DNAs (Fig. 4B, bracketed band).

Aberrant fast-migrating DNAs contain internal deletions likely generated by a protein A*-mediated process

To investigate the genome sequences represented in the faster-migrating DNA pool associated with the *I(J)Y31V* particles, PCRs, which spanned the entire genome, were conducted and compared to those generated using purified, full-length, circular ssDNA genomes. As can be seen in Fig. 4C, every PCR product generated from the wild-type template was also present in the *I(J)Y31V* reactions. This indicates that the entire genome sequence is represented within the aberrant DNA pool associated with slow-migrating particles. However, secondary PCR products, not seen in the wild-type reactions, were often present. Several series of PCR reactions were conducted using different primers. Depending on the primers used, different secondary bands were detected (data not shown). These observations suggest that the *I(J)Y31V* small DNA band contains heterogeneous molecules with varying internal deletions, which was tested by analyzing the sequence of the unique *I(J)Y31V* PCR products.

Some of the PCR products generated unambiguous sequences, clearly defining the juncture formed by an internal deletion. For example, the sequence presented in Fig. 5A clearly indicated that the DNA found between nucleotides 990 and 2,262 has been deleted. The 5' end of the juncture involved the strongest protein A* target site. By contrast, the sequences of some fragments were clear up to a specific nucleotide (Fig. 5B) but afterward, the chromatograms suggested that the output was generated from multiple sequences. This could occur if the excised PCR band contained DNA fragments of similar lengths but with different internal deletions. Alternatively, it could reflect the degenerate nature of the protein A* cleavage sites (27) and the enzyme activity (see Discussion). As illustrated in Fig. 5C, the DNA polymerase would generate a homogeneous sequence until it reached $\Delta 1$. Beyond this point, the templates differ, which would account for the multiple peaks in the chromatogram (27). This was particularly apparent with sequences involving the A* target site at 5239 (Fig. 5B), which was implicated in three deletions. Figure 5D only presents the juncture sequences derived from chromatograms for which both sides of the juncture were clearly defined in the chromatogram. Many involved motifs are known to interact with protein A* (27). However, the participation of protein A cannot be eliminated. Some junctures were derived from the origin of replication, a substrate shared by both A and A* proteins (25–27).

The small DNA-containing particles were clearly an off-pathway product that may be unique to the Y31V mutant. Alternatively, the mutation may exacerbate a naturally occurring off-pathway reaction, which may not be readily detectable in wild-type sedimentation profiles due to a low signal blending in with the background. To determine whether the smaller DNAs were made in a wild-type infection, DNA was directly isolated out of wild-type and *I(J)Y31L* mutant infected cells. The *(J)Y31L* mutant was selected as it also produced ~70S peak, but it was extremely small. A very faint smaller DNA band was detected in wild-type infected cells, which migrated at approximately the same rate as the more abundant smaller DNAs from the mutant infection (data not shown). The smaller DNAs isolated from the wild-type and *I(J)Y31L* infections also contained deletions (Fig. 5D).

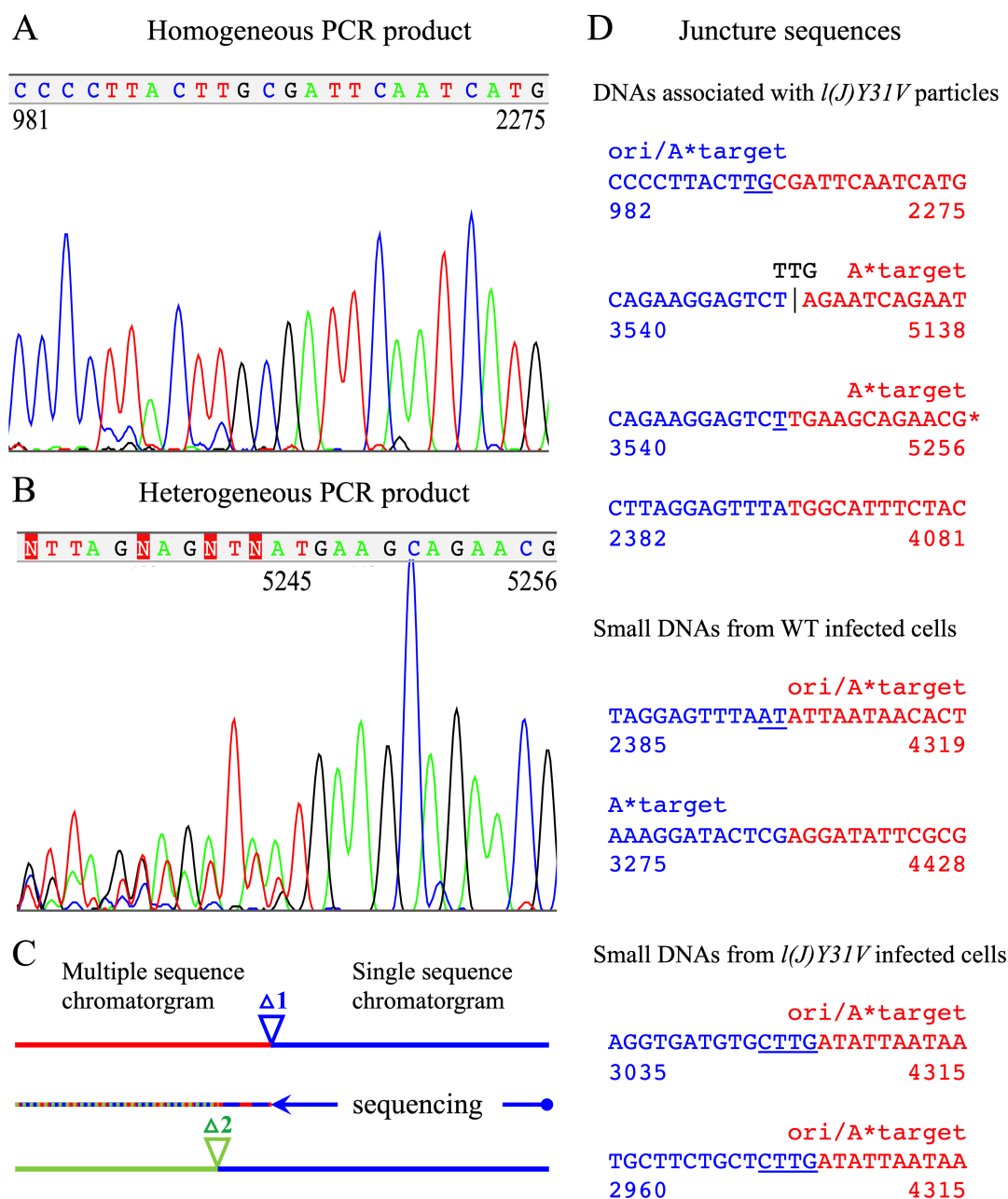


FIG 5 Internal deletions found in small viral DNAs. (A) Sequencing chromatogram of a PCR product containing a single homogeneous deletion. Numbers indicate the location of the nucleotides within the ϕ X174 sequence, which is 5386 nucleotides in length and circular. The chromatogram corresponds to the top sequence given in part D. (B) Sequencing chromatogram of a PCR product containing heterogeneous deletions. (C) Schematic for the generation of a mottled chromatogram as illustrated in part B. (D) Sequences of internal homogeneous deletions. The 5' nucleotide sequence before the deletion is depicted in blue. The 3' sequence is depicted in red. Numbers indicate the location of the nucleotides within the ϕ X174 sequence. Underlined nucleotides could have been donated by either the 5' or 3' junction sequences. Nucleotides in black text represent those found in the junction that are not present in either the 5' or 3' sequences surrounding it. "A* target" indicates a junction derived from a known target site for the A* protein. "Ori/A*target" indicates a junction derived from the origin of replication, which is also an A* target site.

Common second-site suppressors rescue mutations at different sites

A second-site suppressor analysis was conducted with 21 mutants displaying lethal, temperature-sensitive (*ts*), and cold-sensitive (*cs*) phenotypes. Due to the large number of mutants, the analysis was broad in scope, designed to obtain a general view of possible rescue mechanisms. Thus, only one source of each mutant was used, limiting

possible conclusions (32). The mutant with the least severe defective phenotype *cs(J)W30Y* was the only mutant for which an intragenic suppressor was isolated (Q35R). This parental mutant was also suppressed by a mutation in the lysis protein, as well as one within the external scaffolding protein. By contrast, 10 different mutants yielded second-site suppressors, identifying nine substitutions in the viral coat protein. All of the isolated suppressors are listed in Table 1, as well as the results of allele specificity assays (see below). As would be expected, more conservative substitutions appeared to be more easily suppressed. Beyond this observation, a compelling pattern regarding suppressor activity was not readily apparent.

Several of the mutants did not yield second-site suppressors in the initial broad screen. This may reflect a particularly severe defect, one for which suppressors cannot be isolated *via* a single mutational event. Alternatively, it may be a consequence of using only one source (slot machine) for each parental mutant in the analysis (32). To partially distinguish between these two possibilities, the two of the most commonly isolated suppressors, coat protein substitutions S1F and D209H, were crossed into the strains that did not yield suppressors. The resulting double mutants were assayed for the loss of the complementation-dependent phenotype.

Finally, the mutants W30N, W30R, Y31R, Y31E, and F37R could not be rescued by any of the employed strategies (Table 1). These substitutions are polar and/or charged. In the atomic structure of the virion, W30, Y31, and F37 are located adjacent to the positively charged coat protein residues that constitute the nucleic acid binding pocket (Fig. 6). The pocket is composed of four arginine residues, which interact with the negatively charged, phosphate backbone. Thus, charges at these sites may be too detrimental to suppress *via* single amino acid changes in the coat protein.

DISCUSSION

Microvirus DNA binding proteins: anchoring and guiding functions

The N-termini of the ϕ X174, G4, and α 3 J proteins contain positively charged, DNA-interacting amino acid residues but differ in length and charge distribution (Fig. 1B). These residues appear to guide the genome between the DNA-binding pockets within the coat proteins (Fig. 1C). Two previous studies have addressed this guiding function (19, 20). The mutations described herein affect the strongly conserved C-terminus, altering the three aromatic amino residues that most extensively interact with the viral coat protein (6–8). These residues anchor the J protein to the capsid's inner surface. The molecular phenotypes of these mutants did not resemble the guiding mutants used in previous studies.

Biochemical characterization of the mutant assembly pathways and the location of second-site suppressors suggest defects in particle stability

While mutants with conservative substitutions formed infectious virions, assembled particles and/or virions were not detected in most infections, likely due to particle instability during packaging. First, protein J is not required for procapsid formation. Second, particles consistent with packaging complexes were observed in some infections. This sediment is between 50 and 70S, producing a broad peak. The peak declines with increasing S values. Accordingly, very few infectious particles were present in the 114S virion fraction. Common suppressors restore viability to mutants for which soluble assembled particles could or could not be detected, suggesting similar defects. If instability occurs soon after packaging commences, complexes may quickly dissociate or become insoluble. Consequently, no assembled particles would be detected. If instability occurs later, on-pathway packaging may initially be supported, but as it continues the intermediates become progressively more prone to dissociation or insolubility.

Many of the second-site suppressors affect coat protein surface residues residing on the twofold axis of symmetry. T = 1 ϕ X174 assembly proceeds through pentameric intermediates. Thus, these suppressors may confer particle stability both during

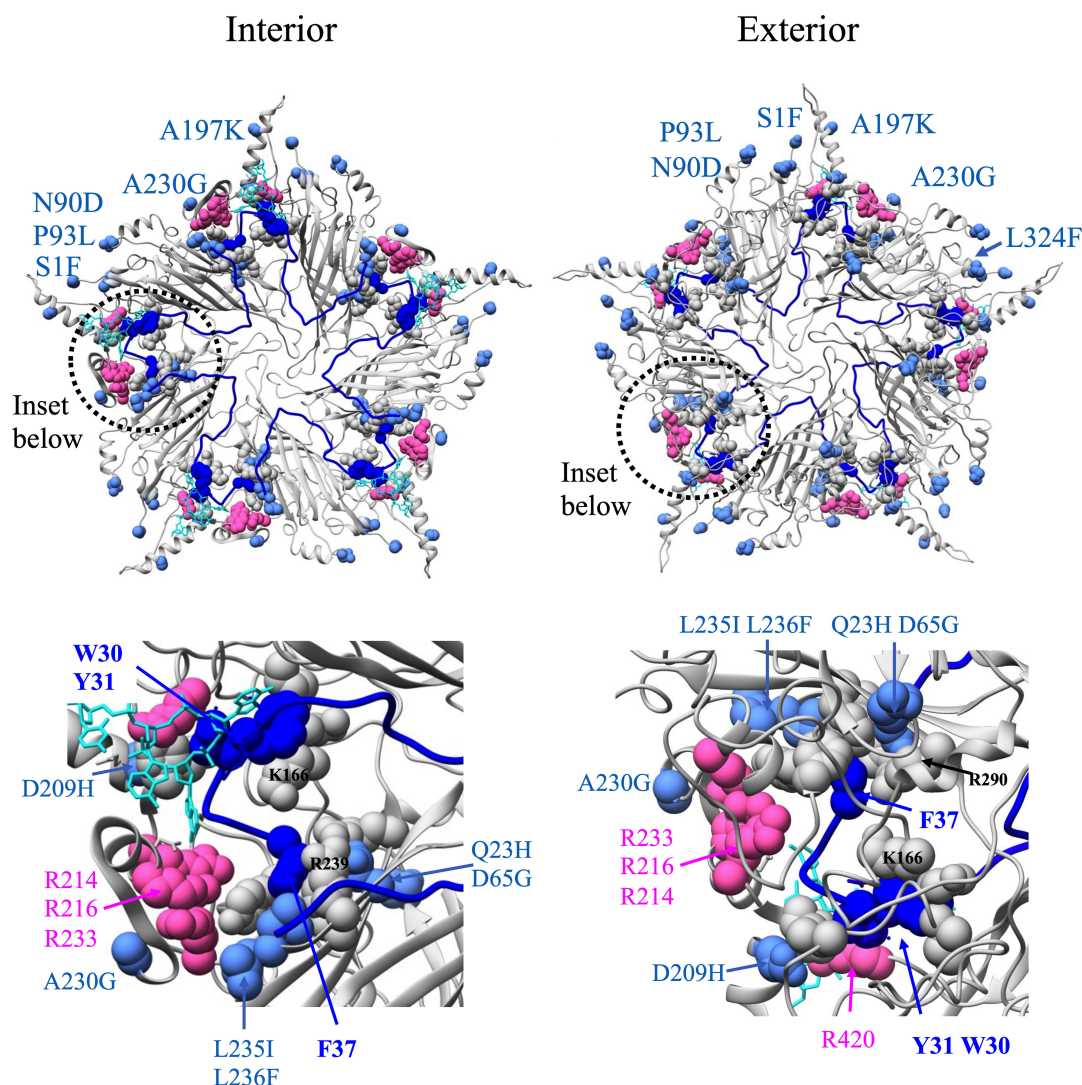


FIG 6 The location of the aromatic J residues and the suppressors that restore viability to missense mutations at those sites (PDB 2BPA). The J protein is depicted in dark blue with the W30, Y31, and F37 residues highlighted with space-filled rendering. The body of the coat protein is depicted in gray. The coat protein residues that contact J protein residues W30, Y31, and F37 are depicted with space-filled rendering. The location of the second-site suppressors is highlighted in light blue, space-filled rendering. They are labeled to reflect the wild-type residue, before the number, and the change, after the number. The residues constituting the DNA binding pocket are depicted in magenta, space-filled rendering. Icosahedrally ordered DNA is depicted in cyan ball-and-stick rendering.

packaging and in the final product. However, four suppressors, L236F (leucine→phenylalanine at amino acid 236), D65G, L235I, and D209H, are adjacent to other J-coat protein contact sites (Fig. 6). These suppressors cluster near the binding site of the very C-terminus of the J protein and may allow mutant J proteins to be positioned in a manner more similar to the wild-type positioning.

The function(s) of protein A*

When bound to dsDNA, protein A* impedes DNA synthesis by inhibiting host cell helicase function (28). Slowing genome biosynthesis likely prevents synthesis from outpacing packaging and, consequently, the detrimental accumulation of genomic ssDNA outside the capsid (see Introduction). Increasing the number of times ssDNA biosynthesis is slowed or paused, *via* protein A* binding to dsDNA, may be sufficient to rescue some packaging mutants, such as ϕ XG4J (20). This retardation may allow the

already encapsidated ssDNA to obtain a more optimal confirmation, one which is more amenable to continuing the reaction.

Although the mutants in this study were not rescued by protein A* over-expression, the sequences of particle-associated small DNAs suggest that protein A* was active during these non-productive infections. One mutant, *I(J)Y31V* produced aberrant particles at high enough levels to be characterized. These particles appear to contain small DNAs with a variety of internal deletions (Fig. 5). Some of the junctures are surrounded by DNA sequences known to be protein A* interaction sites (Fig. 5), including both the origin of replication and the strongest site located in an intergenic region. However, some of the juncture sequences were generated from sequences not previously identified as protein A* target sites. This may reflect differences between *in vivo* and *in vitro* approaches as well as protein A*'s very relaxed substrate specificity (see below)

Only deletions that gave unambiguous sequence results are given in Fig. 5. However, many sequenced fragments could define only one of the sites that contributed to a juncture but afterward, the chromatograms suggested that fragments were heterogeneous. Two factors may explain this phenomenon: (i) fragments of very similar length but containing different deletions could be present in the band isolated by agarose gel electrophoresis and (ii) enzymatically, protein A* function lacks high specificity. For example, protein A requires two stringent motifs: CAACTTG and GATATT for DNA recognition, whereas protein A* only requires one of these sequences and they exhibit considerable variation (27). Moreover, protein A requires a fairly stringent hepta-nucleotide TTCTGGT motif for cleavage and ligation. By contrast, the cleavage requirement for protein A* may be as simple as PyPyTGPuT (27).

Broad substrate specificity and differences between *in vitro* and *in vivo* approaches may explain why some junctures were derived from previously unidentified sequences of protein A* target sites. In the *in vitro* approach (27), naked ssDNA was used. However, naked ssDNA can form both extensive short- and long-range secondary structures (33), which could explain why some sites were not previously identified as targets. For example, the juncture at nucleotide 4405 (Fig. 5) contains a perfect GATATT motif but was not previously identified. Similarly, a nearly perfect TTCTGGT motif (TTCTGCT) occurs in the juncture starting at nucleotide 2961. Moreover *in vitro*, the binding of protein A* to a strong substrate could sterically hinder the protein's recognition of weaker sites, especially if it involves an uncharacterized motif. The juncture sequences starting at nucleotides 3356 and 2379/2382 were each independently identified twice, contributing to the 5' ends of junctures. Each contains the sequence AGGAGT. Relaxed specificity may also explain the addition of 1–3 nucleotides found at some junctures. With ~30 known substrates; several A*-mediated events may occur within one packaging complex. The excess nucleotides could have arisen from an A* substrate used in a previous reaction.

Any model explaining the evolutionary conservation of gene A* must address the advantage of maintaining a protein that appears to target the integrity of the viral genome. When bound to ssDNA, protein A* may conduct a salvage operation (Fig. 7), by hastening the termination of a failing packaging reactions to liberate reusable components such as DNA templates and replication enzymes. Numerous unpackaged single-stranded A* target sites may signal that a packaging reaction will not produce a viable virion. Thus, it would be advantageous to hasten termination, which could be accomplished by generating internal deletions in the non-encapsidated ssDNA. An enzyme with very relaxed substrate specificity could more aggressively perform this function.

A simple cleavage event on the excess unpackaged ssDNA would terminate the reaction. However, it may not salvage the replication machinery, which would include the essential protein needed for rolling circle replication, protein A. First, protein A would remain covalently attached to the origin, its only substrate, and be associated with the liberated partially filled procapsid. For release, protein A requires the generation of a second origin after the entire genome is replicated (34)]. Second, once ssDNA

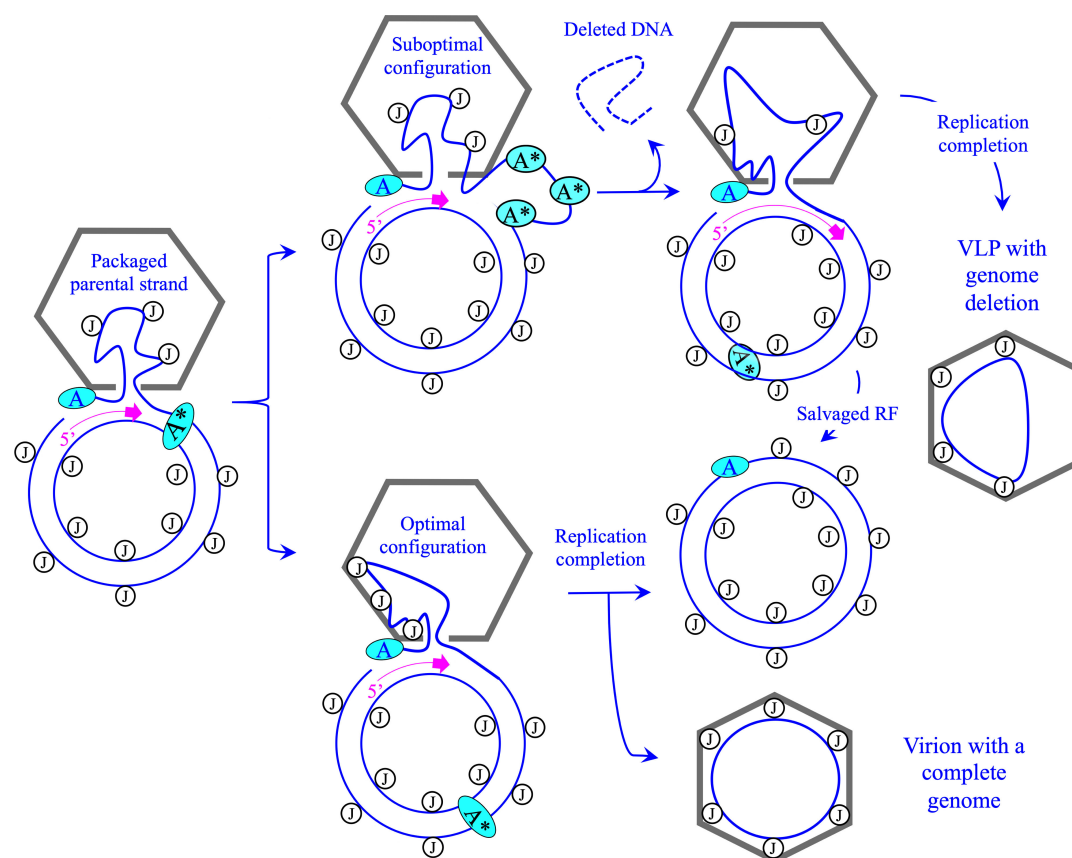


FIG 7 A speculative model of protein A* functions during concurrent ssDNA synthesis and packaging. Protein A* binds to target sites on dsDNA, which inhibits helicase function. Slowing genome biosynthesis ensures that it does not outpace packaging. (Bottom line) Packaged ssDNA assumes a configuration that supports full genome packaging. After one procapsid is filled, the RF DNA template and associated enzymes can be used to fill another procapsid. (Top line) Protein A* inhibition is not sufficient to ensure synthesis-packaging synchrony. ssDNA accumulates outside the capsid. Protein A* binds ssDNA and generates an internal deletion. Rolling circle replication continues until the origin is regenerated. Although this results in a particle without a full genome, the RF DNA template and associated enzymes are salvaged.

commences, the replicative form dsDNA becomes a poor transcription template (35, 36). Thus, there is no efficient means to replenish protein A, but the show must go on.

MATERIALS AND METHODS

Phage plating, media, buffers, stock preparation, and bacterial and phage strains

The plating protocol, media, buffers, stock preparation, and the wild-type *Escherichia coli* C122 strain have been previously described (37). The RY7211 cell line contains a mutation in the *mraY* gene, conferring resistance to viral E protein-mediated lysis (38). BAF30, which was used to express cloned genes, is a *recA* derivative of C122 (39).

Cloned genes used in these studies

The cloning of the ϕ X174 J and A* genes has been previously described (20). The J gene was cloned into pSE420 (Invitrogen). Gene expression was induced by the addition of IPTG (isopropyl- β -D-thiogalactopyranoside) to a concentration of 0.1 mM. The A* gene was cloned into pBAD33 (Thermo Fisher). Gene expression was induced by supplementing media with 0.2% arabinose and then repressed by supplementing media with 0.2% glucose.

Site-directed mutagenesis of gene J

Wild-type DNA was mutagenized with randomizing primers in PCRs with Q5 DNA polymerase (New England Biolabs), as previously described (40). To recover mutants, BAF30 (recA) cells containing the cloned J gene were transfected and incubated at 37°C. The resulting plaques were then stabbed into C122 indicator lawns (no plasmid) incubated at 24°C, 37°C, and 42°C; and BAF30 pøX174J lawns at 37°C. The results of this assay were used to assign preliminary phenotypes: lethal, *cs*, *ts*, or resembling wild-type, that is, no temperature-dependent or complementation-dependent phenotypes. Preliminary phenotypes were verified in direct plating assays. In addition, each defective mutant was assayed for protein A* rescue by plating on cells expressing the cloned A* gene.

Second-site genetic analyses and allele specificity

Second-site suppressors were isolated by plating 10^5 – 10^6 pfu (plaque-forming units) under restrictive conditions. Single plaques were isolated after a 3- to 4-h incubation at 37 or 42°C. For isolating suppressors at 24°C, plates were incubated for 16 h. Plaques were resuspended in 50–100 μ L of H₂O and PCRs were conducted with this liquid. To examine allele specificity, the S1F and D209H suppressors were placed into various gene J' backgrounds by PCR reactions with Q5 DNA polymerase (New England Biolabs) as previously described (40). Reactions were transfected in BAF30 pøX174J. Single plaques were sequenced to verify genotypes. Afterward, phage was plated under the parental mutants' restrictive conditions to assay for suppressor activity.

ssDNA, replicative-form (RF) dsDNA, assembled particle-associated DNA purification, sucrose gradient sedimentation, and SDS-PAGE

Wild-type ssDNA for site-directed mutagenesis was generated and purified with the OLT (Our Little Trick) protocol (30). Briefly, 10 mL of infected lysis-resistant cells, incubated at 37°C for 3 h, was concentrated and resuspended in 1.0 mL of Wizard-plus SV miniprep kit resuspension buffer (Promega). Our little trick: the room temperature incubation in the Wizard-kit lysis buffer with the kit-provided alkaline protease was extended from 5 minutes to 1 h to allow for capsid digestion. To isolate DNA associated with assembled particles, peak-containing gradient fractions were diluted fivefold in 10 mM Tris-HCl (pH 7.5)/10 mM NaCl/1.0 mM EDTA. Sodium dodecyl sulfate (SDS), pronase, and additional EDTA were added to respective final concentrations of 0.5% (wt/vol), 0.5 mg/mL, and 10 mM. Samples were incubated at 37°C for 2.5 h before extraction as previously described (20).

Sucrose gradient preparation, particle isolation, and infected cell extractions have been previously described (41, 42). Briefly, 100 mL of lysis-resistant cells were infected at a multiplicity of infection (MOI) of 3.0 and incubated for 3 h at 37°C for lethal mutants, 3 h at 42°C for temperature-sensitive mutants, or 6 h at 24°C for cold-sensitive mutants. Infected cells were concentrated and resuspended in 3.0 mL sucrose gradient buffer [SGB: 100 mM NaCl, 5.0 mM EDTA, 6.4 mM Na₂HPO₄, 3.3 mM KH₂PO₄ (pH 7.0)], then lysed by an overnight incubation with lysozyme (2.0 mg/mL). After removing cellular debris, the supernatant was concentrated to 200 μ L in a Corning Spin-X spin column (100 kDa cutoff) and loaded atop a 5.0 mL, 5–30% sucrose gradient (wt/vol). Gradients were spun at $192,000 \times g$ for 1 h and then fractionated into 125 μ L fractions. Assembled particles were detected by UV spectroscopy (A_{260} and A_{280}).

ACKNOWLEDGMENTS

The authors acknowledge Elizabeth Ogunbunmi, aka "EO," for discussion.

This research was supported by the National Science Foundation grants MCB 2013653 (B.A.F.)

AUTHOR AFFILIATIONS

¹The BIO5 Institute, University of Arizona, Tucson, Arizona, USA²Berkshire School, Advanced Math/Science Research Program, Sheffield, Massachusetts, USAAUTHOR ORCID*s*Bentley A. Fane  <http://orcid.org/0000-0002-3984-1726>

FUNDING

Funder	Grant(s)	Author(s)
National Science Foundation (NSF)	MCB 2013653	Bentley A. Fane

REFERENCES

- Gopal A, Egecioglu DE, Yoffe AM, Ben-Shaul A, Rao ALN, Knobler CM, Gelbart WM. 2014. Viral RNAs are unusually compact. *PLoS One* 9:e105875. <https://doi.org/10.1371/journal.pone.0105875>
- Speir JA, Johnson JE. 2012. Nucleic acid packaging in viruses. *Curr Opin Struct Biol* 22:65–71. <https://doi.org/10.1016/j.sbi.2011.11.002>
- Cerritelli ME, Studier FW. 1996. Assembly of T7 capsids from independently expressed and purified head protein and scaffolding protein. *J Mol Biol* 258:286–298. <https://doi.org/10.1006/jmbi.1996.0250>
- Aoyama A, Hamatake RK, Hayashi M. 1981. Morphogenesis of phi X174: *in vitro* synthesis of infectious phage from purified viral components. *Proc Natl Acad Sci U S A* 78:7285–7289. <https://doi.org/10.1073/pnas.78.12.7285>
- Mukai R, Hamatake RK, Hayashi M. 1979. Isolation and identification of bacteriophage phi X174 prohead. *Proc Natl Acad Sci U S A* 76:4877–4881. <https://doi.org/10.1073/pnas.76.10.4877>
- Bernal RA, Hafenstein S, Esmeralda R, Fane BA, Rossmann MG. 2004. The phiX174 protein J mediates DNA packaging and viral attachment to host cells. *J Mol Biol* 337:1109–1122. <https://doi.org/10.1016/j.jmb.2004.02.033>
- McKenna R, Ilag LL, Rossmann MG. 1994. Analysis of the single-stranded DNA bacteriophage phi X174, refined at a resolution of 3.0 Å. *J Mol Biol* 237:517–543. <https://doi.org/10.1006/jmbi.1994.1253>
- McKenna R, Xia D, Willingmann P, Ilag LL, Krishnaswamy S, Rossmann MG, Olson NH, Baker TS, Incardona NL. 1992. Atomic structure of single-stranded DNA bacteriophage phi X174 and its functional implications. *Nature* 355:137–143. <https://doi.org/10.1038/355137a0>
- Shlomai J, Kornberg A. 1980. An *Escherichia coli* replication protein that recognizes a unique sequence within a hairpin region in phi X174 DNA. *Proc Natl Acad Sci U S A* 77:799–803. <https://doi.org/10.1073/pnas.77.2.799>
- Shlomai J, Polder L, Arai K, Kornberg A. 1981. Replication of phi X174 dna with purified enzymes. I. Conversion of viral DNA to a supercoiled, biologically active duplex. *J Biol Chem* 256:5233–5238. [https://doi.org/10.1016/S0021-9258\(19\)69392-0](https://doi.org/10.1016/S0021-9258(19)69392-0)
- Baas PD, Teertstra WR, van Mansfeld AD, Jansz HS, van der Marel GA, Veeneman GH, van Boom JH. 1981. Construction of viable and lethal mutations in the origin of bacteriophage 'phi' X174 using synthetic oligodeoxyribonucleotides. *J Mol Biol* 152:615–639. [https://doi.org/10.1016/0022-2836\(81\)90120-0](https://doi.org/10.1016/0022-2836(81)90120-0)
- Eisenberg S, Griffith J, Kornberg A. 1977. phiX174 cistron A protein is a multifunctional enzyme in DNA replication. *Proc Natl Acad Sci U S A* 74:3198–3202. <https://doi.org/10.1073/pnas.74.8.3198>
- Eisenberg S, Kornberg A. 1979. Purification and characterization of phiX174 gene A protein. A multifunctional enzyme of duplex DNA replication. *J Biol Chem* 254:5328–5332. [https://doi.org/10.1016/S0021-9258\(18\)50599-8](https://doi.org/10.1016/S0021-9258(18)50599-8)
- Eisenberg S, Scott JF, Kornberg A. 1976. An enzyme system for replication of duplex circular DNA: the replicative form of phage phi X174. *Proc Natl Acad Sci U S A* 73:1594–1597. <https://doi.org/10.1073/pnas.73.5.1594>
- van Mansfeld AD, Langeveld SA, Weisbeek PJ, Baas PD, van Arkel GA, Jansz HS. 1979. Cleavage site of phiX174 gene-A protein in phiX and G4 RFI DNA. *Cold Spring Harb Symp Quant Biol* 43:331–334. <https://doi.org/10.1101/sqb.1979.043.01.039>
- Aoyama A, Hayashi M. 1986. Synthesis of bacteriophage phi X174 *in vitro*: mechanism of switch from DNA replication to DNA packaging. *Cell* 47:99–106. [https://doi.org/10.1016/0092-8674\(86\)90370-3](https://doi.org/10.1016/0092-8674(86)90370-3)
- Bernal RA, Hafenstein S, Olson NH, Bowman VD, Chipman PR, Baker TS, Fane BA, Rossmann MG. 2003. Structural studies of bacteriophage alpha3 assembly. *J Mol Biol* 325:11–24. [https://doi.org/10.1016/S0022-2836\(02\)01201-9](https://doi.org/10.1016/S0022-2836(02)01201-9)
- Jennings B, Fane BA. 1997. Genetic analysis of the phi X174 DNA binding protein. *Virology* 227:370–377. <https://doi.org/10.1006/viro.1996.8351>
- Ogunbunmi ET, Roznowski AP, Fane BA. 2021. The effects of packaged, but misguided, single-stranded DNA genomes are transmitted to the outer surface of the phiX174 capsid. *J Virol* 95:e0088321. <https://doi.org/10.1128/JVI.00883-21>
- Roznowski AP, Doore SM, Kemp SZ, Fane BA. 2020. Finally, a role befitting A^{star}: the strongly conserved, unessential microvirus A* proteins ensure the product fidelity of packaging reactions. *J Virol* 94:e01593-19. <https://doi.org/10.1128/JVI.01593-19>
- Rokyta DR, Burch CL, Caudle SB, Wichman HA. 2006. Horizontal gene transfer and the evolution of microvirid coliphage genomes. *J Bacteriol* 188:1134–1142. <https://doi.org/10.1128/JB.188.3.1134-1142.2006>
- Chejanovsky N, Carter BJ. 1989. Mutagenesis of an AUG codon in the adeno-associated virus rep gene: effects on viral DNA replication. *Virology* 173:120–128. [https://doi.org/10.1016/0042-6822\(89\)90227-4](https://doi.org/10.1016/0042-6822(89)90227-4)
- Mankertz A, Hillenbrand B. 2001. Replication of porcine circovirus type 1 requires two proteins encoded by the viral rep gene. *Virology* 279:429–438. <https://doi.org/10.1006/viro.2000.0730>
- Colasanti J, Denhardt DT. 1985. Expression of the cloned bacteriophage phi X174 A* gene in *Escherichia coli* inhibits DNA replication and cell division. *J Virol* 53:807–813. <https://doi.org/10.1128/JVI.53.3.807-813.1985>
- Van Mansfeld AD, Baas PD, Jansz HS. 1984. Gene A protein of bacteriophage phi X174 is a highly specific single-strand nuclease and binds via a tyrosyl residue to DNA after cleavage. *Adv Exp Med Biol* 179:221–230. https://doi.org/10.1007/978-1-4684-8730-5_23
- van Mansfeld AD, van Teeffelen HA, Baas PD, Veeneman GH, van Boom JH, Jansz HS. 1984. The bond in the bacteriophage phi X174 gene A protein–DNA complex is a tyrosyl-5'-phosphate ester. *FEBS Lett* 173:351–356. [https://doi.org/10.1016/0014-5793\(84\)80804-2](https://doi.org/10.1016/0014-5793(84)80804-2)
- Langeveld SA, van Mansfeld AD, van der Ende A, van de Pol JH, van Arkel GA, Weisbeek PJ. 1981. The nuclease specificity of the bacteriophage phi X174 A* protein. *Nucleic Acids Res* 9:545–562. <https://doi.org/10.1093/nar/9.3.545>
- Eisenberg S, Ascarelli R. 1981. The A* protein of phi X174 is an inhibitor of DNA replication. *Nucleic Acids Res* 9:1991–2002. <https://doi.org/10.1093/nar/9.8.1991>
- Valley CC, Cembran A, Perlmutter JD, Lewis AK, Labello NP, Gao J, Sachs JN. 2012. The methionine-aromatic motif plays a unique role in

- stabilizing protein structure. *J Biol Chem* 287:34979–34991. <https://doi.org/10.1074/jbc.M112.374504>
30. Blackburn BJ, Li S, Roznowski AP, Perez AR, Villarreal RH, Johnson CJ, Hardy M, Tuckerman EC, Burch AD, Fane BA. 2017. Coat protein mutations that alter the flux of morphogenetic intermediates through the phiX174 early assembly pathway. *J Virol* 91:e01384-17. <https://doi.org/10.1128/JVI.01384-17>
 31. Ogunbunmi ET, Love SD, Rhodes KA, Morales A, Wilch MH, Jonas J, Fane BA. 2022. Low-temperature adaptation targets genome packing reactions in an Icosahedral single-stranded DNA virus. *J Virol* 96:e0197021. <https://doi.org/10.1128/jvi.01970-21>
 32. Luria SE, Delbrück M. 1943. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28:491–511. <https://doi.org/10.1093/genetics/28.6.491>
 33. Benevides JM, Stow PL, Ilag LL, Incardona NL, Thomas Jr GJ. 1991. Differences in secondary structure between packaged and unpackaged single-stranded DNA of bacteriophage phi X174 determined by Raman spectroscopy: a model for phi X174 DNA packaging. *Biochemistry* 30:4855–4863. <https://doi.org/10.1021/bi00234a004>
 34. Hayashi M, Aoyama A, Richardson D. L, Hayashi N. M. 1988. Biology of the bacteriophage Øx174, p 1–71. In Calendar R (ed), *The bacteriophages*. Vol. 2. Plenum Press, New York.
 35. Hayashi Y, Hayashi M. 1971. Template activities of the phi X-174 replicative allomorphic deoxyribonucleic acids. *Biochemistry* 10:4212–4218. <https://doi.org/10.1021/bi00799a009>
 36. Puga A, Tessman I. 1973. Mechanism of transcription of bacteriophage S13. I. Dependence of messengerRNA synthesis on amount and configuration of DNA. *J Mol Biol* 75:83–97. [https://doi.org/10.1016/0022-2836\(73\)90530-5](https://doi.org/10.1016/0022-2836(73)90530-5)
 37. Fane BA, Hayashi M. 1991. Second-site suppressors of a cold-sensitive prohead accessory protein of bacteriophage phi X174. *Genetics* 128:663–671. <https://doi.org/10.1093/genetics/128.4.663>
 38. Bernhardt TG, Struck DK, Young R. 2001. The lysis protein E of phi X174 is a specific inhibitor of the MraY-catalyzed step in peptidoglycan synthesis. *J Biol Chem* 276:6093–6097. <https://doi.org/10.1074/jbc.M007638200>
 39. Fane BA, Head S, Hayashi M. 1992. Functional relationship between the J proteins of bacteriophages phi X174 and G4 during phage morphogenesis. *J Bacteriol* 174:2717–2719. <https://doi.org/10.1128/jb.174.8.2717-2719.1992>
 40. Roznowski AP, Fane BA. 2016. Structure-function analysis of the varphiX174 DNA-piloting protein using length-altering mutations. *J Virol* 90:7956–7966. <https://doi.org/10.1128/JVI.00914-16>
 41. Uchiyama A, Fane BA. 2005. Identification of an interacting coat-external scaffolding protein domain required for both the initiation of phiX174 procapsid morphogenesis and the completion of DNA packaging. *J Virol* 79:6751–6756. <https://doi.org/10.1128/JVI.79.11.6751-6756.2005>
 42. Hafenstein S, Fane BA. 2002. phi X174 genome-capsid interactions influence the biophysical properties of the virion: evidence for a scaffolding-like function for the genome during the final stages of morphogenesis. *J Virol* 76:5350–5356. <https://doi.org/10.1128/jvi.76.11.5350-5356.2002>