

Shared and Group-Specific Features of the Rotavirus RNA Polymerase Reveal Potential Determinants of Gene Reassortment Restriction^{▽†}

Sarah M. McDonald,¹ Daniel Aguayo,² Fernando D. Gonzalez-Nilo,² and John T. Patton^{1*}

Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20826-8026,¹ and Center for Bioinformatics and Molecular Simulations, Universidad de Talca, 2 Norte 685, Casilla 721, Talca, Chile²

Received 24 February 2009/Accepted 31 March 2009

Rotaviruses (RVs) are nonenveloped, 11-segmented, double-stranded RNA viruses that are major pathogens associated with acute gastroenteritis. Group A, B, and C RVs have been isolated from humans; however, intergroup gene reassortment does not occur for reasons that remain unclear. This restriction might reflect the failure of the viral RNA-dependent RNA polymerase (RdRp; VP1) to recognize and replicate the RNA of a different group. To address this possibility, we contrasted the sequences, structures, and functions of RdRps belonging to RV groups A, B, and C (A-VP1, B-VP1, and C-VP1, respectively). We found that conserved amino acid residues are located within the hollow center of VP1 near the active site, whereas variable, group-specific residues are mostly surface exposed. By creating a three-dimensional homology model of C-VP1 with the A-VP1 crystallographic data, we provide evidence that these RV RdRps are nearly identical in their tertiary folds and that they have the same RNA template recognition mechanism that differs from that of B-VP1. Consistent with the structural data, recombinant A-VP1 and C-VP1 are capable of replicating one another's RNA templates in vitro. Nonetheless, the activity of both RdRps is strictly dependent upon the presence of cognate RV core shell protein A-VP2 or C-VP2, respectively. Together, the results of this study provide unprecedented insight into the structure and function of RV RdRps and support the notion that VP1 interactions may influence the emergence of reassortant viral strains.

Rotaviruses (RVs), members of the *Reoviridae* family, are nonenveloped, segmented double-stranded RNA (dsRNA) viruses generally associated with enteric disease (8, 16). Most human RV isolates can be classified into one of three genetically divergent groups (A, B, or C) based on serological and phylogenetic analyses (16). Group A RVs are ubiquitous in nature, infecting the young of numerous mammalian and avian species (8, 16). In humans, group A RV infections cause acute infantile gastroenteritis, leading to approximately 500,000 deaths each year worldwide (28). Group B or C RV disease in humans is rare, but these strains can cause mild to severe diarrhea in children and adults (6, 10, 16). Despite the observation that RVs from different groups can infect the human host, there have been no reports of intergroup gene reassortment. In contrast, reassortment within group A has been well documented (9, 44). It is not clear why RV groups do not exchange gene segments, but it might reflect a failure of proteins and RNA from divergent strains to function together during replication.

The infectious RV virion is an icosahedron made up of three concentric protein layers, which encapsidate 11 segments of genomic dsRNA (8). The outermost virion layer is removed during the entry of RV into a host cell, resulting in a transcriptionally active double-layered particle (DLP). The outer layer of the DLP (i.e., the intermediate layer of the virion) is formed of

VP6 and surrounds the innermost T=1 VP2 core shell (19, 20). Tethered beneath the VP2 shell, proximal to each fivefold axis, are enzyme complexes that consist of the viral RNA-dependent RNA polymerase (RdRp; VP1) and the RNA capping enzyme (VP3) (19, 20, 36). VP1/VP3 complexes of transcriptionally active DLPs generate 11 species of capped, nonpolyadenylated plus-strand RNAs (+RNAs) using genomic dsRNAs as templates (29). The +RNA transcripts are extruded from the DLP through channels at the fivefold axes and subsequently serve as templates for protein and dsRNA synthesis (genome replication) (13). Genome replication occurs concurrently with the initial stages of virion assembly in cytoplasmic inclusions (viroplasms). Studies of group A RVs are consistent with the idea that VP1 (possibly with VP3) first recognizes viral +RNAs, yielding a set of stable, catalytically inactive protein-RNA complexes (21, 41). During or immediately after the packaging of these complexes into an assembling core, VP2 binds to VP1 and induces the enzyme to initiate minus-strand synthesis, creating 11 dsRNA genome segments (21, 41). Although in vitro genome packaging has not yet been achieved for RVs, group A VP1 (A-VP1) shows robust, template-specific in vitro RdRp activity (4, 30, 31, 41). Moreover, in vitro dsRNA synthesis by recombinant A-VP1 requires group A VP2 (A-VP2), suggesting that this assay recapitulates the process of core-associated genome replication seen during viral infection (31, 41). In comparison to group A, few studies have focused on group B or C RV replication.

High-resolution crystal structures of A-VP1, alone and in complex with RNA, reveal a compact, globular protein ~70 Å in diameter that is composed of three domains: (i) an N-terminal domain (ii) a polymerase domain, and (iii) a C-terminal domain (Fig. 1A and B) (21). The architecture of the polymerase domain is a characteristic “right hand” with finger,

* Corresponding author. Mailing address: Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20826-8026. Phone: (301) 594-1615. Fax: (301) 496-8312. E-mail: jpatton@niaid.nih.gov.

† Supplemental material for this article may be found at <http://jvi.asm.org/>.

▽ Published ahead of print on 8 April 2009.

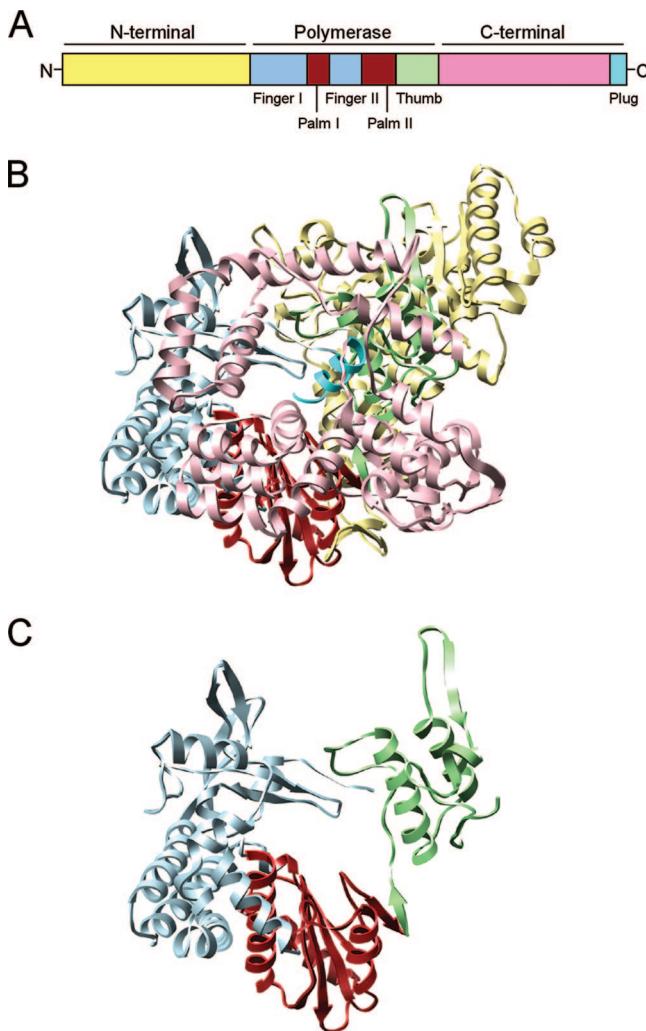


FIG. 1. Organization and structure of A-VP1. (A) Cartoon schematic illustrating the three-domain organization of A-VP1. The N-terminal domain is yellow, the C-terminal domain is pink, and the finger, palm, and thumb subdomains of the polymerase domain are pale blue, brick red, and green, respectively. The extreme C-terminal plug is cyan. (B) Ribbon representation of the A-VP1 apoenzyme crystal structure (PDB no. 2R7Q). (C) Ribbon representation of the “right-hand” A-VP1 polymerase domain. The colors in panels B and C correspond to those in panel A.

palm, and thumb subdomains (Fig. 1C). The polymerase domain also contains the canonical motifs (A to F) involved in various aspects of phosphodiester bond formation (27). The N- and C-terminal domains envelop the polymerase domain, creating a cage-like enzyme with a hollow catalytic center. Four tunnels lead into this center, serving as conduits for the entry and exit of free nucleotides (nucleoside triphosphates [NTPs]), template RNAs, and RNA products. Overall, the structural organization of A-VP1 is extraordinarily similar to that of the mammalian orthoreovirus RdRp $\lambda 3$ (38). Unlike $\lambda 3$, however, the extreme C terminus of A-VP1 forms a “plug” that protrudes into the putative dsRNA exit tunnel, possibly regulating the release of replication products. Additionally, the priming loop, a structural element important for stabilizing the priming NTP during de novo initiation, is retracted in A-VP1 compared

to $\lambda 3$. Thus, the solved A-VP1 structure represents a catalytically inactive form of the RdRp captured prior to the initiation of dsRNA synthesis.

A remarkable feature of A-VP1 revealed by soaks of crystals with RNA oligonucleotides was its capacity to specifically recognize bases in the 3'-terminal UGUGACC sequence common to nearly all group A RV +RNAs (21). The UGUGACC sequence, referred to as the 3' consensus sequence (3'CS), represents a critical *cis*-acting element promoting A-VP1 template recognition (41). Specific binding of the UGUG bases by residues in the A-VP1 template entry tunnel places the terminal nucleotide of the RNA molecule “out of register” with the active site, contributing to the formation of a stable but catalytically inactive complex (21). The terminal ACC nucleotides of the group A 3'CS exhibit stacking interactions, which are thought to promote initiation complex formation (41). Group C +RNAs end with a similar 3'CS (UGUGGGCU), sharing the UGUG recognition element but differing in the terminal residues (33). This observation suggests that group C VP1 (C-VP1) might be capable of binding to group A +RNAs and vice versa. In contrast, the 3' ends of +RNA templates of group B strains do not include UGUG. Instead, they end with a very different 3'CS (AAAACCC), suggesting that A-VP1 and C-VP1 would replicate these templates very poorly or not at all (33). A-VP1 can recognize RV +RNA in the absence of A-VP2, but its activation requires interactions with VP2 that presumably induce structural alterations within the enzyme (21, 41). For effective intergroup reassortment to occur, not only does the RdRp of one group need to recognize and replicate the +RNA of a different group, but also the proteins encoded by the exchanged genes (i.e., VP1 and VP2) must function properly together. In the present study, we performed bioinformatic and biochemical analyses of group A, B, and C RV RdRPs (A-VP1, B-VP1, and C-VP1, respectively) with the goal of gaining structural and functional insight into (i) the mechanism of RV RNA synthesis and (ii) factors influencing RV gene reassortment. Our results suggest that even divergent RV RdRPs have similar template recognition and VP2-dependent catalytic mechanisms but that group-specific differences in RdRp interactions may influence the emergence of reassortant viral strains.

MATERIALS AND METHODS

Phylogenetic analyses and amino acid alignments. Phylogenetic analyses and amino acid alignments were constructed with MacVector 10.0.2. Amino acid phylogenograms were generated by the neighbor-joining method, systematic best tree, with gaps distributed proportionately and the Poisson correction parameter. The RdRp amino acid sequences (GenBank accession numbers) used in this study were reovirus $\lambda 3$ (EF494435), J19 (DQ113897), B219 (EF453355), IDIR (M97203), WH-1 (EU490413), Bang373 (EU490415), Cal-1 (EU490414), ST3 (EF583045), IAL28 (EF583029), P (EF583037), Wi61 (EF583049), Wa (DQ490539), D (EF583021), YM (X76486), YO (DQ870497), AU-1 (DQ490533), T152 (DQ146699), SA11-5N (DQ838631), Se584 (EF583041), UK (X55444), RF (J04346), DS-1, TB-Chen (AY787653), S2 (DQ870485), L26 (EF583033), 69 M EF576937), A64 (EF583017), P0-13 (AB009629), Cowden (M74216), and Bristol (NC007547).

Modeling of the three-dimensional structure of C-VP1. A homology model of the C-VP1 structure (strain Bristol) was built with the crystallographic data of A-VP1 (strain SA11) as a reference (Protein Data Bank [PDB] no. 2R7Q) (21). An amino acid alignment of A-VP1 and C-VP1 was generated with CLUSTALW, and C-VP1 secondary structures were predicted with PsiPred (25, 40). MODELLER v8.2 was used to convert the alignment into several three-dimensional models of C-VP1, which were then ranked and refined according to DOPE and GA341 potentials (37). PROCHECK v.3.5.4 was used to identify C-VP1 regions showing

structural perturbations (18). Such regions were relaxed with simulated annealing with MODELLER v8.2. Tautomeric states of histidine residues were assigned according to the local environment.

The C-VP1 three-dimensional model was further relaxed with Molecular Dynamics software (NAMD 2.6) (35). Specifically, C-VP1 was solvated with a water box of 80 by 80 by 80 Å (TIP3P and 150 mM NaCl). To remove initial amino acid side chain clashes, undesired contacts, and filling vacuum pockets, the full system was submitted to an energy minimization, followed by molecular dynamics for 2 ns. During the relaxation, the backbone of C-VP1 was restrained with a harmonic potential of 10, 5, 2, 1, and 0 kcal/mol Å² by 200 ps at each stage. The molecular dynamic simulation was done with a time step of 1 ps with a 12-Å spherical cutoff for nonbonding interactions, a switching function from 10 Å for the Van der Waals term, and force-shifted electrostatics. For long-range interactions, the particle-mesh Ewald method was used with a grid spacing on the order of 1 Å or less. The simulation was carried out at a temperature of 310 K in the NVT ensemble. Constant temperature was forced through Langevin dynamics (damping coefficient of 1 ps⁻¹). Periodic boundary conditions were employed for all simulations. All energy calculations were done with the force field CHARMM27 (22). The final model of C-VP1 is available at <http://cbsm.utalca.cl/jpatton/>.

Electrostatic potentials for A-VP1 (PDB no. 2R7R) and C-VP1 were calculated with the Poisson Boltzmann equation through the algorithm implemented in the APBS program (1). All calculations were done using the atom radius, and partial charges were assigned with the force field CHARMM 27 parameters at pH 7.0. The protein and solvent dielectric constants were assigned values of 2 and 78.5, respectively. The dielectric boundary between the protein and the solvent was based on the molecular surface definition calculated with a probe sphere radius of 1.4 Å. The ionic strength of monovalent ions was 150 mM, with an ion exclusion radius of 2 Å. The Poisson Boltzmann equation was solved on a 225 by 225 by 225 grid with a 0.4-Å point spacing centered at the position of the +RNA.

Images were generated with the UCSF Chimera Molecular Modeling System or the Visual Molecular Dynamics program (12, 34).

Generation of VP1- and VP2-expressing baculoviruses. The BaculoDirect Expression System (Invitrogen) was used according to the manufacturer's protocol to create recombinant baculoviruses expressing group A (strain SA11) or group C (strain Bristol) VP1 and VP2 proteins. Briefly, the VP1- and VP2-encoding genes were cloned into the entry vector pENTR-1A (see below) and then inserted into the BaculoDirect C-Term Linear DNA by recombination with LR Clonase II. The baculovirus DNA was transfected into *Spodoptera frugiperda* (Sf9) cells, and recombinant virus was harvested from the medium.

To create pENTR-A-VP1, the nucleotides encoding SA11 VP1 with a carboxy-terminal hexahistidine (His) tag were PCR amplified from pCR-BacI-4 (42) and subcloned into pENTR-1A with primer-generated restriction sites (5' KpnI and 3' NotI). To create pENTR-A-VP2, the nucleotides encoding SA11 VP2 were subcloned from pCI-SA11-g2 (24) into pENTR-1A at the 5' EcoRI and 3' NotI restriction sites. The deduced amino acid sequences of A-VP1 and A-VP2 exactly match what is reported in GenBank for SA11-5N (DQ838631 and DQ838631, respectively).

The group C (strain Bristol) VP1- and VP2-encoding sequences were synthesized de novo. To create pENTR-C-VP1, the nucleotides encoding Bristol VP1 with a carboxy-terminal His tag were generated by Blue Heron Biotechnology (Bothell, WA) and then subcloned into pENTR-1A. To create pENTR-C-VP2, the nucleotides encoding Bristol VP2 were first codon optimized for protein expression and RNA stability by GENEART (Regensburg, Germany) prior to subcloning into pENTR-1A at engineered restriction sites (5' BamHI and 3' NotI). The deduced amino acid sequence of Bristol VP1 exactly matches what is reported in GenBank (AJ304859). The amino acid sequence of Bristol VP2 is nearly identical to what is reported in GenBank (AJ303139), with the exception of a single amino acid substitution (M36V) that was created to disrupt an alternative start codon.

Purification of recombinant VP1 and VP2. To prepare His-tagged A-VP1 and C-VP1, 4 × 10⁷ Sf9 cells were infected at a multiplicity of infection of approximately 5 with the appropriate baculovirus and maintained in TNH-FH medium (Invitrogen) containing 10% fetal bovine serum for 72 h at 20°C. Infected cells were pelleted, washed twice with cold phosphate-buffered saline, resuspended in 30 ml of cold VP1 lysis buffer (25 mM NaHPO₄, 200 mM NaCl [pH 7.8]), and sonicated. The insoluble fraction was removed by centrifugation at 15,000 × g for 10 min at 4°C, and His-tagged VP1 was recovered from the soluble fraction by incubation with cobalt resin (Talon) for 2 h at 4°C. The purified VP1 proteins bound to resin were washed with lysis buffer and eluted from the resin with 500 µl of lysis buffer containing 300 mM imidazole. Purified VP1 preparations were dialyzed against low-salt buffer (2 mM Tris-HCl [pH 7.5], 0.5 mM EDTA, 0.5 mM dithiothreitol) and stored at 4°C.

To prepare A-VP2 and C-VP2, 2 × 10⁸ Sf9 cells were infected at a multiplicity of infection of approximately 5 with the appropriate baculovirus and maintained in TNH-FH medium (Invitrogen) containing 10% fetal bovine serum for 72 h at 28°C. The protease inhibitors leupeptin and aprotinin at 1 µg/ml were added to the medium daily. Infected cells were pelleted, washed twice with cold phosphate-buffered saline, and resuspended in 25 ml of cold low-salt buffer. Cells were lysed by adding deoxycholic acid to a 1% final concentration and by sonication. VP2 was pelleted through a 5-ml cushion of 35% (wt/vol) sucrose in low-salt buffer by centrifugation at 80,000 × g for 90 min at 10°C. VP2 was resuspended in 500 µl of low-salt buffer containing 10 µg/ml leupeptin and aprotinin and stored at 4°C.

The concentrations of the VP1 and VP2 proteins were determined by comparison with known amounts of standards electrophoresed in sodium dodecyl sulfate (SDS)-polyacrylamide gels and stained with PageBlue (Fermentas). The average purifications yielded approximately 5 to 10 µg of VP1 (per 4 × 10⁷ cells) and 800 to 1,000 µg of VP2 (per 2 × 10⁸ cells).

Preparation of RNA templates for in vitro dsRNA synthesis. To generate cDNA templates containing authentic 5' and 3' ends for T7 promoter-driven in vitro transcription, PCRs were performed with Accuprime Pfx Supermix (Invitrogen). For group A gene 8 cDNA (encoding SA11 NSP2), the vector pSP65g8R was used as the PCR template (41). For group C gene 9 cDNA (encoding Bristol NSP2), the vector pQE60g8C (39) was first modified to include the appropriate group C 5' and 3' untranslated regions and then used as a PCR template.

The sense primer (5'-GCCCTTAAATACGACTCACTATAGGCTT-3') for all reactions included a T7 promoter sequence (underlined). The antisense primers were 5'-GGTCACATAAGCGCTTCTATTCT-3' (group A g8 cDNA) and 5'-AGCCACATATGCACAGCTGCAATTCT-3' (group C g9 cDNA). To create group A gene 8 cDNA with substitutions in the 3'-terminal nucleotides (UGUGAAAA), the antisense primer 5'-GGTTTTTAAGCGCTTCTATTCT-3' was used.

The PCR-amplified cDNAs were extracted twice with phenol-chloroform-isoamyl alcohol and once with chloroform prior to serving as templates in the T7 MEGAscript transcription system (Ambion) according to the manufacturer's instructions. The +RNA products of the transcription reactions were passed through mini Quick Spin RNA columns (Roche) to remove unincorporated nucleotides. RNA quantity was determined with a UV spectrophotometer (optical density at 260 nm), and RNA quality was assessed by electrophoresis in 7 M urea-5% polyacrylamide gels stained with ethidium bromide.

RNA replication assays. Reaction mixtures were optimized for C-VP1 activity in the presence of C-VP2. These conditions are similar to those used for A-VP1 with A-VP2 (41) and included 50 mM Tris-HCl (pH 7.1); 1.5% polyethylene glycol; 2 mM dithiothreitol; 1.5 U of RNasin; 20 mM magnesium acetate; 4 mM MnCl₂; 1.25 mM each ATP, CTP, UTP, and GTP; 10 Ci of [α -³²P]UTP (800 Ci/mmol); 8 pmol of template +RNA; 2 pmol of VP1; and 20 pmol of VP2. Reaction mixtures were incubated at 37°C for 3 h. The radiolabeled dsRNA products were separated by SDS-polyacrylamide gel electrophoresis and visualized by autoradiography.

Nucleotide sequence accession number. The nucleotide sequence of the codon-optimized Bristol VP2 gene used for this study was deposited in GenBank (FJ785474).

RESULTS

Genetic variation among group A, B, and C RV RdRPs. To define the similarities and differences among RV RdRPs, we first analyzed the phylogenetic relationship of VP1 proteins from several human and animal RV strains. We found that the VP1 proteins segregate phylogenetically according to the strain's group designation, as has been previously reported (Fig. 2) (2, 5, 7, 14, 23, 26, 45). There is an abundance of A-VP1 sequence information, and several genotypes have been defined (11, 23). In contrast, only four B-VP1 (strains WH-I, Bang373, Cal-1, and IDIR) and two C-VP1 (strains Bristol and Cowden) sequences are known (2, 5, 7, 14, 26, 45). The VP1 sequences from strains J19 and B219 are closely related to those of B-VP1 but form a separate, distinct branch, consistent with their classification into a novel RV group (Fig. 2) (15, 26). Yet, because of their genetic similarities, we will consider the RdRPs of J19 and B219 to be B-VP1 proteins in this report.

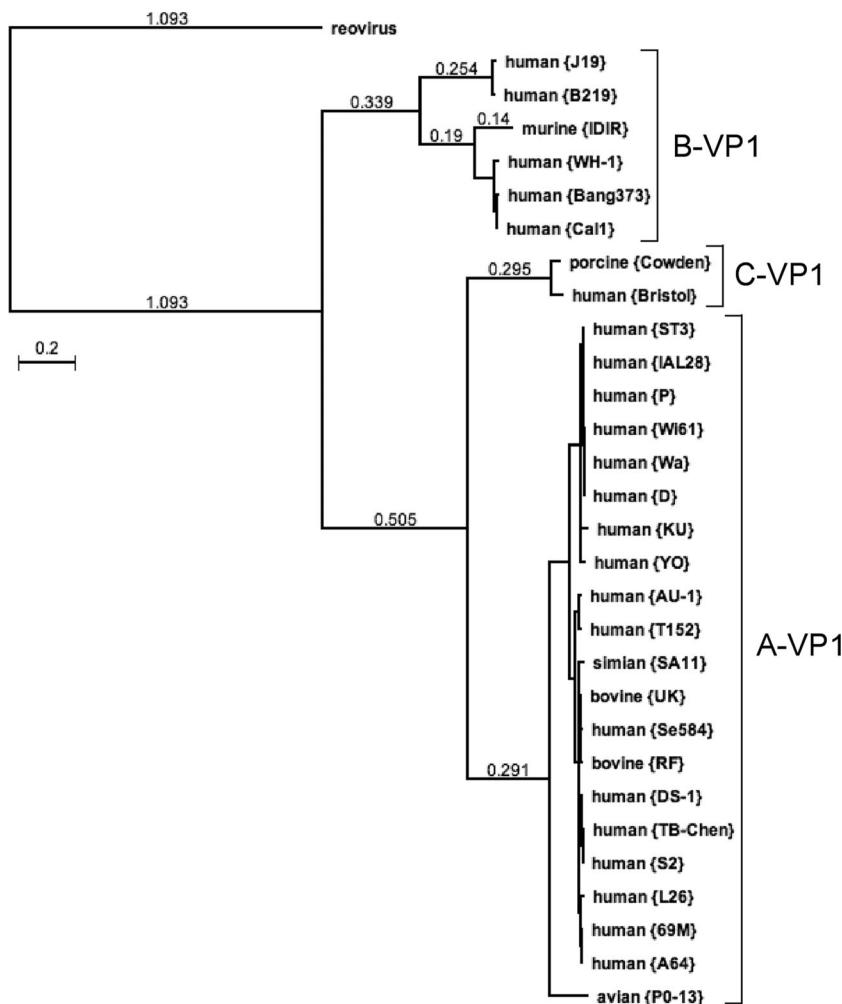


FIG. 2. Phylogenetic relationship of A-VP1, B-VP1, and C-VP1. The dendrogram was constructed by the neighbor-joining method, systematic best tree, with gaps distributed proportionately and the Poisson correction parameter. The scale bar represents 0.2 amino acid substitution per position. The RV strain is in parentheses following the name of the organism from which it was isolated. The reovirus RdRp $\lambda 3$ sequence was used as the root. Brackets indicate the VP1 group designations (A-VP1, B-VP1, and C-VP1) discussed in this report. Accession numbers are given in Materials and Methods.

As reflected by their phylogenetic relationship, VP1 proteins of RV strains belonging to the same group are more similar at the amino acid level. In fact, the intragroup percent similarities for A-VP1, B-VP1, and C-VP1 are quite high, ranging from 71.9 to 99.2% (Table 1). The intergroup percent similarities, however, are much lower, with values of approximately 35% for A-VP1 versus B-VP1 and for B-VP1 versus C-VP1 (Table 1). The closest relationship is seen between the A-VP1 and C-VP1 proteins, with an intergroup amino acid sequence similarity of approximately 64% (Table 1). This relationship is even more obvious in a comparison of the sequences of the polymerase domain (see also below), in which A-VP1 and C-VP1 have similarities as high as 74.4% (Table 1). Additionally, A-VP1 and C-VP1 are alike in size and both have a calculated molecular mass of approximately 125 kDa (Table 2). In contrast, B-VP1 proteins are slightly larger, averaging 70 amino acids more than the A-VP1 and C-VP1 proteins and having calculated molecular masses of approximately 132 kDa (Table 2). The functional significance of these extra B-VP1

residues is unknown but might suggest that RdRps from this group entertain additional or different interactions.

To identify regions of conservation and variation among the group A, B, and C RV RdRPs, we performed ClustalW alignments (40). For simplicity, we show the alignment of RdRPs from strains SA11, Cal-1, and Bristol as prototypes for A-VP1,

TABLE 1. Intra- and intergroup VP1 amino acid similarities and identities

Comparison	% of VP1		% of polymerase domain	
	Similar	Identical	Similar	Identical
A vs A	86.3–98.2	76.0–96.1	84.3–97.5	91.7–91.6
B vs B	71.9–99.2	55.5–97.2	80.6–99.8	69.5–99.1
C vs C	95.9	92.7	97.7	95.5
A vs B	34.6–36.5	18.5–20.0	46.4–48.3	29.2–32.4
B vs C	33.8–35.7	18.5–18.9	43.9–47.7	29.2–30.5
A vs C	63.3–64.7	46.5–47.0	72.9–74.4	55.2–57.2

TABLE 2. Amino acid characteristics of RV VP1 proteins

VP1 protein group	Molecular mass (kDa)	Isoelectric point	Full length (aa)	N-terminal domain (aa)	Pol domain (aa)	C-terminal domain (aa)
A	124.8–125.2	8.28–8.67	1,088	332	446	310
B	131.6–132.9	8.70–8.91	1,059–1,167	376–382	464–465	320
C	125.8	8.59–9.02	1,090	338	441	311

B-VP1, and C-VP1, respectively (Fig. 3). For a comprehensive VP1 alignment showing several RV representatives, see Fig. S1 in the supplemental material. Based on the known motifs and domains of A-VP1, the corresponding regions for B-VP1 and C-VP1 were predicted (Fig. 3). These RdRPs are each composed of an N-terminal domain, a central “right-hand” polymerase domain, and a C-terminal domain (21). The N- and C-terminal domains show less conservation across all three groups than do the polymerase domains (Fig. 3). Nonetheless, even these terminal domains are more conserved between A-VP1 and C-VP1, while B-VP1 is quite variable (Fig. 3).

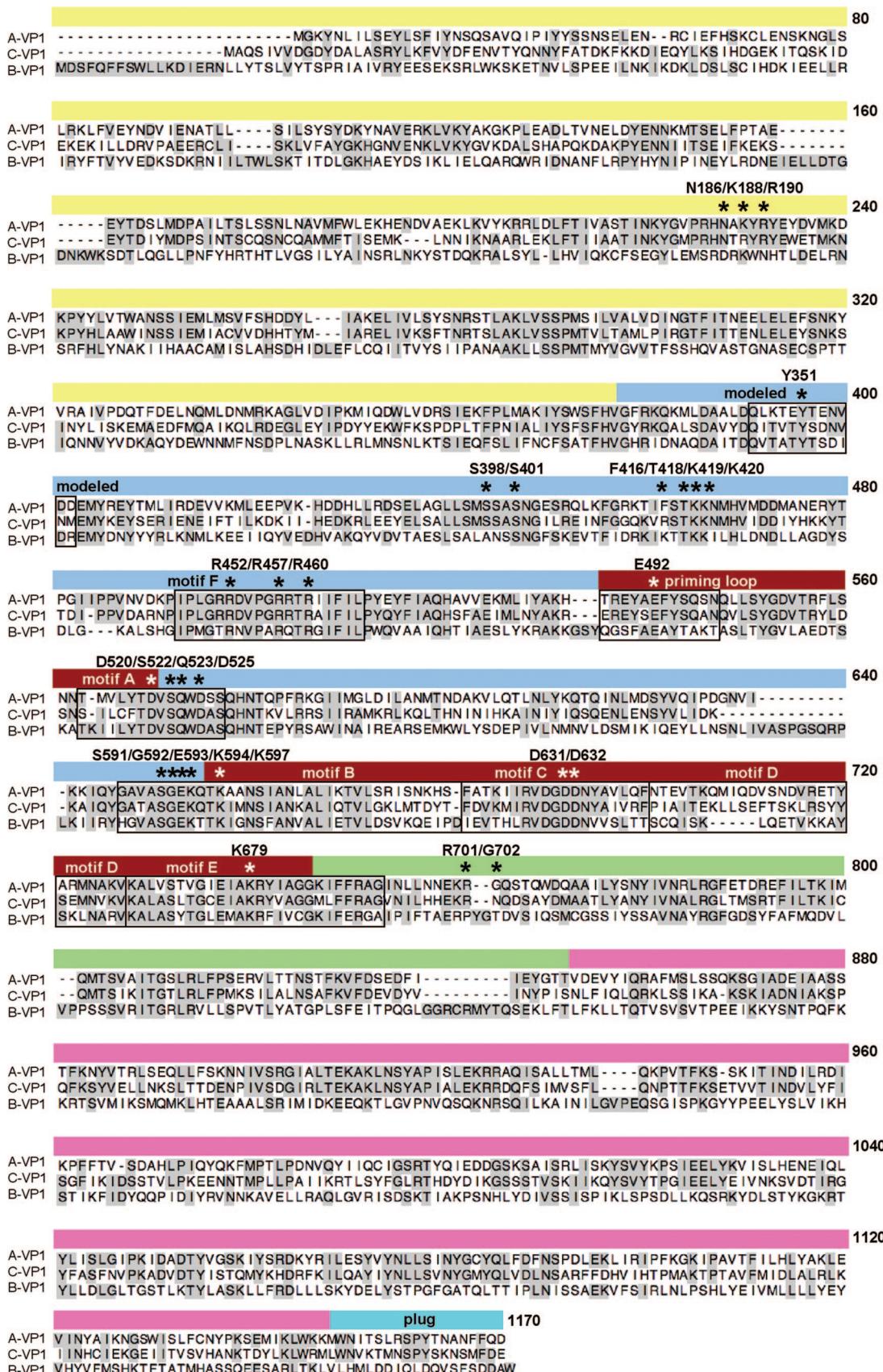
Conserved residues lie in the hollow center of the RV RdRp. The three-dimensional locations of conserved and variable residues of group A, B, and C RdRPs were extrapolated from the amino acid sequence alignment of the three enzymes and the high-resolution structure determined for A-VP1 (Fig. 4) (21). We found that very few conserved amino acids map to the surface of the protein. Of those that do, the majority are located in a region near the NTP entry tunnel, juxtaposed to a disordered loop (residues Q346 to D357) that was not seen in the solved A-VP1 crystal structure (Fig. 4A to D). A limited number of conserved surface residues also buttress the putative cap-binding site of the enzyme (Fig. 4B) (21). Still, the greatest concentration of conserved amino acids maps to the hollow center of A-VP1, with many lining the interior walls surrounding the active site and NTP entry tunnel (Fig. 4E and F). Perhaps not surprisingly, residues that make up the catalytic center, including those of the palm subdomain that form the canonical motifs (A to F), are shared among the group A, B, and C RdRPs (Fig. 4E and F). In addition, numerous residues lining the template entry tunnel, but not the RNA exit tunnels, are conserved among the RdRPs (Fig. 4E and F). Collectively, the positions of conserved residues indicate that the RV RdRPs have similar strategies of template binding and RNA catalysis.

To gain a better understanding of which VP1 residues play key roles in catalysis, we identified amino acids that (i) project toward the catalytic center based on the A-VP1 crystal structure, (ii) belong to known structural or functional motifs, and (iii) are conserved among all group A, B, and C RV RdRPs (Fig. 5; see Fig. S2 in the supplemental material). Residues satisfying these criteria include the catalytic aspartates of motifs A (D520 and D525) and C (D631 and D632), which are involved in coordinating two Mg²⁺ ions and interacting with the phosphate groups of incoming NTPs in support of nucleotidyl transfer (Fig. 5; see Fig. S2 in the supplemental material). Two additional conserved residues (S522 and Q523) are present in motif A and are positioned suitably for interaction with the γ-phosphate of an incoming NTP (Fig. 5; see Fig. S2 in the supplemental material). Motif E is thought to contribute to flexibility between the palm and thumb subdomains, allow-

ing freedom of movement during catalysis. A single conserved lysine residue (K679) within motif E projects toward the active site, where it may interact with the priming NTP or with the RNA product (Fig. 5; see Fig. S2 in the supplemental material). Motif B contains a number of conserved residues, with several (G592, E593, K594, and K597) anchoring the 3' terminus of the +RNA template inside the enzyme (Fig. 5; see Fig. 8; see also Fig. S2 in the supplemental material). Another conserved residue in motif B (S591) projects toward the base and ribose of the NTP in the nucleotide (N) site, perhaps aiding in NTP selection and stabilization (Fig. 5; see Fig. S2 in the supplemental material). The position and composition of motif B suggest a particularly important role in aligning the template strand with NTPs in the catalytic center. The residues of motif F include three conserved arginine residues positioned to interact with the γ-phosphate of the incoming NTP (R457) or with nucleosides in the P or N site (R452 and R460) (Fig. 5; see Fig. S2 in the supplemental material). The priming loop residues (T487 to N498) show a surprising degree of variation among the RV RdRPs, with the group B sequence differing the most. Still, two A-VP1 priming loop residues (E492 and Y494) are present in both the B-VP1 and C-VP1 proteins, and the glutamic acid is in a position to possibly support the priming nucleotide during initiation (Fig. 3 and 5; see Fig. S2 in the supplemental material). The conserved array of amino acids surrounding the active site of group A, B, and C RdRPs point to a shared mechanism of RNA synthesis.

Structural comparison of group A and C RdRPs. It is clear that the catalytic sites of RV RdRPs are highly conserved at the amino acid level, while the N- and C-terminal domains are more variable. To determine whether changes in the terminal domains correlate with differences in protein structures, we sought to create three-dimensional molecular models of B-VP1 and C-VP1 based on the A-VP1 crystallographic data. The sequence similarity between A-VP1 and B-VP1 was not sufficient to create an accurate model (Table 1). However, C-VP1 (strain Bristol) has 46% sequence identity and 64% sequence similarity with A-VP1 (strain SA11). Thus, we were able to generate a homology model of the C-VP1 polypeptide (Fig. 6). The final C-VP1 model showed no structural perturbations, side chain clashes, or undesired contacts, and it was relaxed at its lowest energy state (see Fig. S3 and Table S1 in the supplemental material). Ramachandran plots indicate that most of the main chain dihedral angles in the modeled structure are in allowed regions (see Fig. S3 and Table S1 in the supplemental material). Specifically, 96% of the residues are in allowed regions and bond angles and lengths are within 1 standard deviation of the mean (see Fig. S3 and Table S1 in the supplemental material).

The modeled structure of C-VP1 shows that the predicted architecture and three-domain organization of the protein are



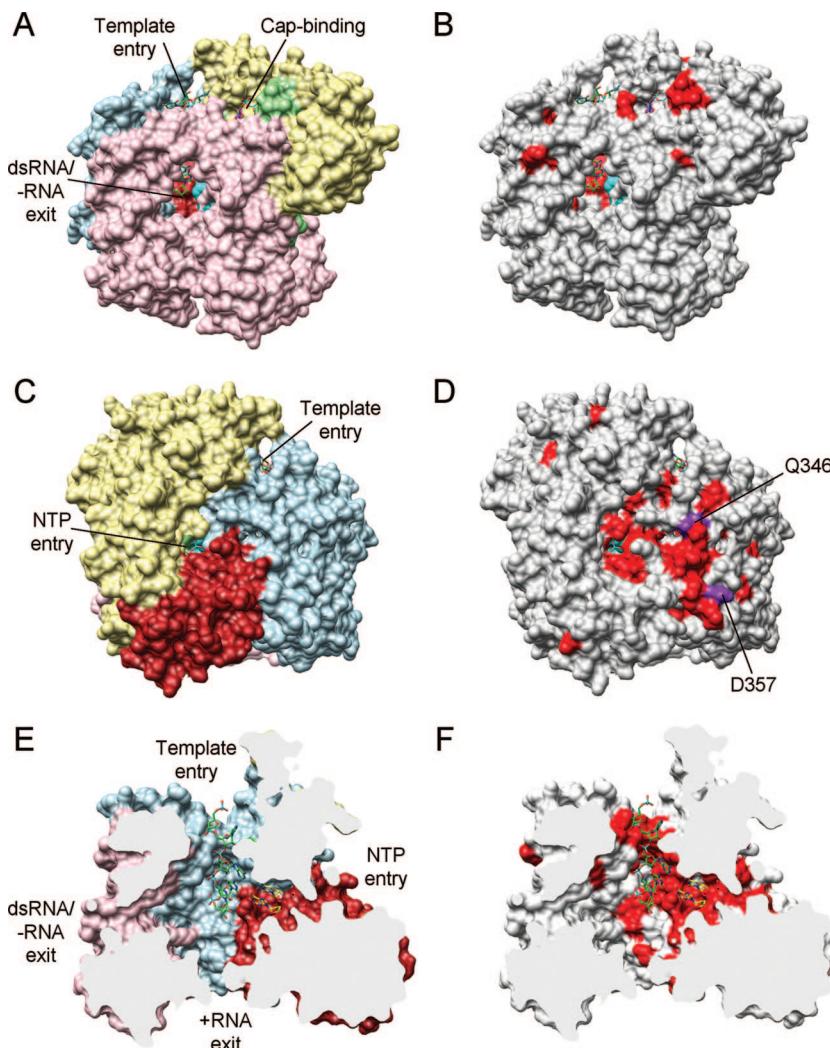


FIG. 4. Three-dimensional locations of residues conserved among A-VP1, B-VP1, and C-VP1. (A) Surface representation of the A-VP1 crystal structure (PDB no. 2R7R). Domains and subdomains are colored as in Fig. 1. Template entry and dsRNA/RNA exit tunnels are labeled for reference. The 3'CS of group A +RNA (UGUGACC) is shown (green element sticks) in the template entry tunnel, and a GTP (red element sticks) is shown in the putative cap-binding site. (B) Same image as panel A with conserved (red) and nonconserved (gray) amino acids shown. (C and D) Same image as in panels A and B but rotated 180° to the right. The NTP and template entry tunnels are shown for reference. Conserved residues (red) are shown around an unstructured loop that extends from Q346 to D357 (purple). (E and F) Same image as in panels A and B but rotated 90° to the left and shown in sagittal cutaway. All four tunnels are labeled, and free NTPs (yellow element sticks) are modeled near the active site. Conserved residues (red) are shown decorating the catalytic center and NTP and template entry tunnels.

nearly identical to A-VP1 (Fig. 6A to D). A flexible loop extending from the surface of the finger subdomain was not resolved in the A-VP1 structure (residues Q346 to D357); we have modeled the corresponding residues of C-VP1 (residues Q352 to M363) (Fig. 6E and F). A single tyrosine within this loop (Y351 of A-VP1) is conserved in all group A, B, and C RV RdRPs, but the function of this residue remains to be elucidated (Fig. 6F). In comparison to A-VP1, C-VP1 has

several missing or inserted residues, slightly altering the length and subtle positioning of some secondary structural elements (Fig. 3 and 6). Like what is seen for A-VP1, the C-VP1 priming loop is not in an appropriate position for initiation (Fig. 6B), suggesting that the activity of this RdRp might be VP2 dependent as well.

All of the A-VP1, B-VP1, and C-VP1 proteins analyzed have isoelectric points that are slightly basic (pI 8.28 to 9.02) (Table

FIG. 3. Amino acid sequence alignment of A-VP1, B-VP1, and C-VP1. The primary amino acid sequence alignment of A-VP1 (strain SA11), B-VP1 (strain Cal-1), and C-VP1 (strain Bristol) is shown. Dashes indicate gaps in the protein sequence, and shading indicates conservation of amino acid identity. The domains and subdomains of A-VP1 are represented by a line above the alignment and are colored as in Fig. 1. Motifs and C-VP1 modeled residues are identified in boxed regions. Asterisks specify amino acids of interest in the present study. For a more comprehensive VP1 alignment, see Fig. S1 in the supplemental material.

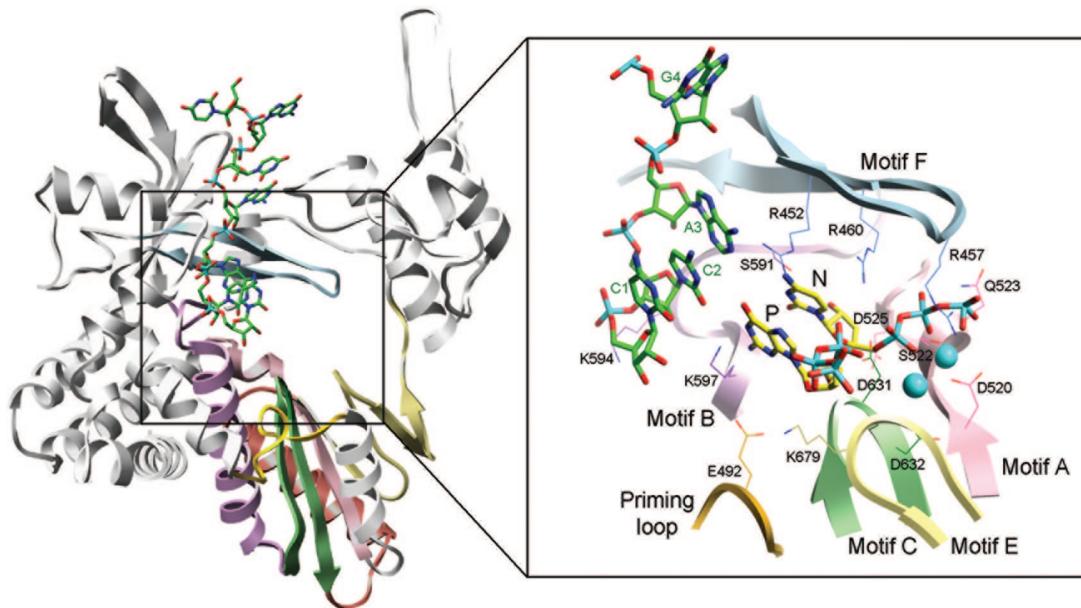


FIG. 5. Critical residues of the VP1 catalytic site. The image on the left is a ribbon representation of the A-VP1 polymerase domain (PDB no. 2R7R). Structural and functional motifs are colored pink (motif A), purple (motif B), green (motif C), red-orange (motif D), light yellow, (motif E), light blue (motif F), and gold (priming loop). The image on the right is a magnification of the catalytic center showing residues (colored element wires; labeled) that are conserved in A-VP1, B-VP1, and C-VP1. Nucleotides (GACC; green element sticks) of the group A +RNA are shown. Two Mg^{2+} ions (cyan) and free NTPs (yellow element sticks) were modeled into the priming (P) and nucleotide (N) sites based on the reovirus λ3 initiation complex (PDB no. 1MWH). For a stereo image, see Fig. S2 in the supplemental material.

2). To determine whether the surfaces of A-VP1 and C-VP1 showed differences in charge distribution, electrostatic potentials were calculated for the proteins with the APBS program (Fig. 7). The analysis revealed a similar highly electropositive zone on a face of both RdRPs that includes their template entry tunnels (Fig. 7). This feature likely attracts the backbone of template RNAs toward the entry tunnel. Interestingly, the surface-exposed electropositive zone of the RdRPs extends from the template entry tunnel to the dsRNA/minus-strand RNA (−RNA) exit tunnel, an arrangement that may help guide −RNA templates back to the entry tunnel during multiple rounds of transcription. The conservation of surface potential suggests that A-VP1 and C-VP1 maintain functionally homologous but nonidentical residues.

Group A and C RV RdRPs have the same template recognition mechanism. The crystal structure of A-VP1 in complex with an RNA oligonucleotide representing the 3'CS (UGUGACC) identified amino acids that enable the RV RdRp to recognize +RNA templates for genome replication (21). Specifically, the recognition is mediated by hydrogen bonding and stacking interactions between the N186, K188, R190, F416, R701, and G702 residues of A-VP1 and the UGUG bases of the 3'CS (Fig. 8; see Fig. S4 in the supplemental material). The phosphate backbone of the +RNA oligonucleotide is further held in place, albeit nonspecifically, by A-VP1 residues S398, S401, T418, K419, K420, G592, E593, K594, and K597 (Fig. 8; see Fig. S4 in the supplemental material). The binding of these template entry tunnel residues stabilizes the +RNA molecule in a position that is out of register by one nucleotide with the active site (Fig. 8; see Fig. S4 in the supplemental material). Because the 3'CS of group C

+RNA templates also includes the UGUG sequence, we hypothesize that residues involved in template recognition would be the same for A-VP1 and C-VP1. In a similar manner, group B RNA templates do not contain a UGUG sequence in the 3'CS, suggesting that the residues involved in template recognition would differ for B-VP1. Consistent with this idea, we found that the amino acid residues responsible for UGUG recognition were either the same for A-VP1 and C-VP1 or, although different in identity, could form equivalent hydrogen bonds and stacking interactions. Specifically, A-VP1 residues N186, K188, R190, and R702 (N192, R194, R196, and R702 of C-VP1) engage G6 and G4, respectively (Fig. 3 and 8; see Fig. S4 in the supplemental material). Recognition of U7 can be directed by hydrogen bonding with the amine and carbonyl groups of F416 for A-VP1 or those of R422 for C-VP1 (Fig. 3 and 8; see Fig. S4 in the supplemental material). Similarly, recognition of U5 can be directed by hydrogen bonding with the carbonyl group of G702 for A-VP1 or that of N703 for C-VP1 (Fig. 3 and 8; see Fig. S4 in the supplemental material). In comparison, B-VP1 shows little sequence similarity to the residues and regions of A-VP1 and C-VP1 that are involved in UGUG recognition. Nearly all of the A-VP1 residues that form hydrogen bonds with the phosphate backbone of the RNA (S401, T418, K419, K420, G592, E593, K594, and K597) are conserved in the B-VP1 and C-VP1 proteins (Fig. 3 and 8; see Fig. S4 in the supplemental material). These results suggest that RV RdRPs might diverge at residues that mediate specific template recognition but that these enzymes hold their templates in the entry tunnel with a conserved mechanism. Moreover, the results suggest that A-VP1 and C-VP1 may have the capacity to recognize each other's +RNA templates via the

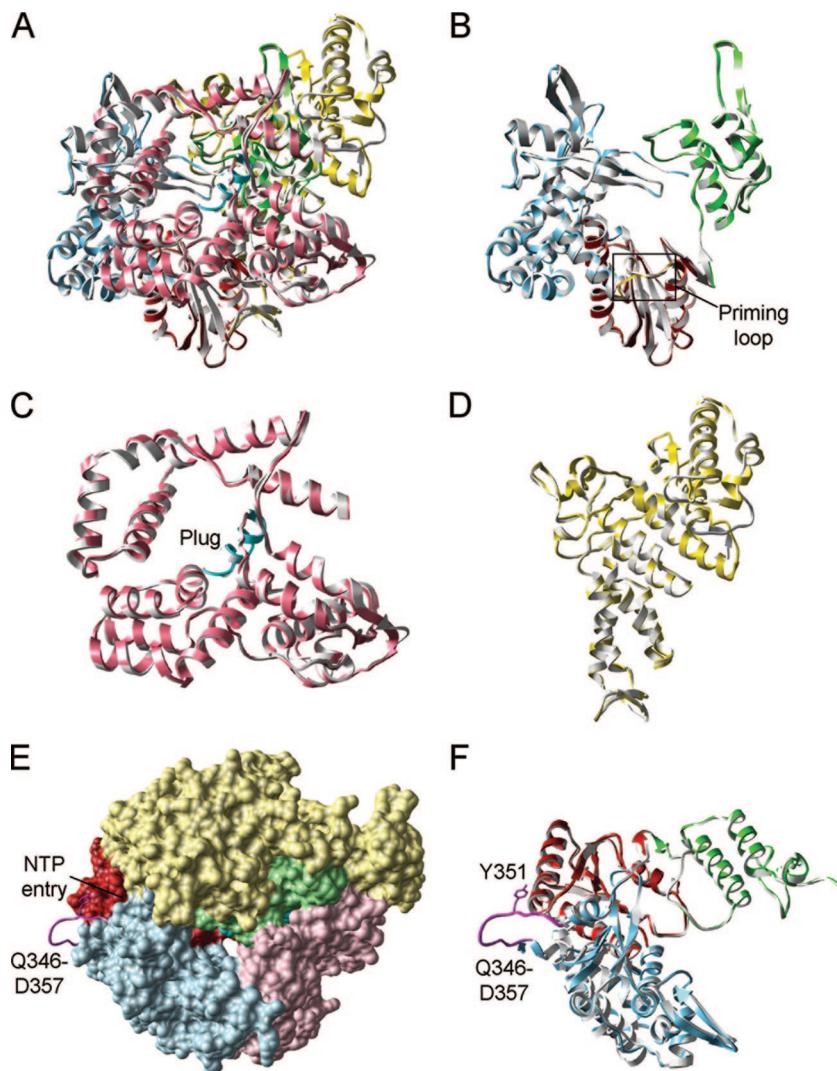


FIG. 6. Three-dimensional homology model of C-VP1. (A to D) Superimposed ribbon representations of A-VP1 (PDB no. 2R7Q; colored as in Fig. 1) and modeled C-VP1 (gray). Images are of the full-length RdRPs (A), the polymerase domains with priming loops (gold; labeled) (B), the C-terminal domains (C), and the N-terminal domains (D). (E) Surface representation of A-VP1 with a ribbon diagram showing the missing flexible loop (Q346 to D327; magenta), which was modeled for C-VP1. The NTP entry tunnel is shown for reference. (F) Ribbon representation of panel E without the N- and C-terminal domains. A single A-VP1 residue (Y357; purple wire) within the modeled loop (magenta) is also conserved for the B-VP1 and C-VP1 proteins.

UGUG sequence, an essential prerequisite of gene reassortment.

In vitro group C RV RdRp activity is VP2 dependent. The structural similarities between A-VP1 and C-VP1, in combination with the conservation of critical residues, led us to predict that these RdRps have common template recognition and VP2-dependent catalysis mechanisms. To test this prediction, we expressed recombinant C-VP1 and C-VP2 (strain Bristol) with baculovirus vectors and assayed the purified proteins for the capacity to replicate +RNA templates in vitro (Fig. 9). C-VP1 was expressed as a soluble, C-terminally hexahistidine (His)-tagged protein and purified with cobalt affinity resin (Fig. 9A). The expression and solubility profile of the protein were indistinguishable from those of the well-characterized recombinant His-tagged A-VP1 (strain SA11) (21). We found that the authentic open reading frame of C-VP2 contained several RNA instability elements

that led to rapid message turnover in insect cells and prevented detectable protein expression (data not shown). This problem was overcome by optimizing the C-VP2 coding sequence to eliminate these elements while maintaining the amino acid sequence of the protein. Following codon optimization, C-VP2 was expressed efficiently and was purified under the conditions traditionally used for A-VP2 (Fig. 9A).

Recombinant purified C-VP1 was assayed for the capacity to support in vitro dsRNA synthesis in the presence and absence of VP2 (Fig. 9B). Our results show that C-VP1 has detectable polymerase activity, but only in the presence of its cognate core shell protein C-VP2 (Fig. 9B). Likewise, A-VP1 requires A-VP2 for in vitro function and shows no detectable activity in the presence of C-VP2 (Fig. 9B). The finding that A-VP2 and C-VP2 are not functionally interchangeable suggests that incompatible protein interactions might influence the survival of

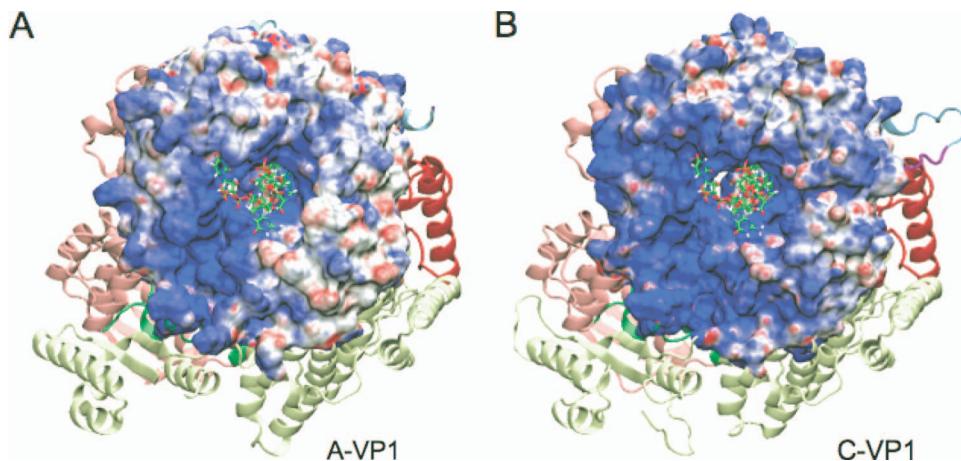


FIG. 7. Surface charge distribution for A-VP1 and C-VP1. The images show the top surfaces of A-VP1 (PDB no. 2R7R) (A) and C-VP1 (B), allowing the visualization of a UGUGACC oligonucleotide (green element sticks) in the template entry tunnel. The surface is colored according to an electrostatic potential gradient, with blue and red representing positive (10 kT) and negative (-10 kT), respectively. The visible portions of the VP1 palm, N-terminal domain, and C-terminal domain are shown in a ribbon representation. The colors correspond to those in Fig. 6.

certain reassortant strains (e.g., those that have VP1- and VP2-encoding genes belonging to different groups). As expected on the basis of their shared capacity to recognize UGUG, we found that these RV RdRPs replicate group A and C +RNAs (Fig. 9B). Interestingly, both enzymes replicate group A +RNA more efficiently than they do group C +RNA. Computer-based RNA-folding algorithms predicted that group C +RNAs might not present the UGUG element in single-

stranded form to the RdRp in vitro (data not shown). Therefore, it is possible that A-VP1 and C-VP1 can bind to the UGUG recognition element better in the context of a group A template. The level of dsRNA synthesis with C-VP1/C-VP2 was consistently much lower than with A-VP1/A-VP2 for reasons that are not yet known (Fig. 9B and data not shown).

To determine whether UGUG interaction is critical for C-VP1, like it is for A-VP1, this tetranucleotide sequence was

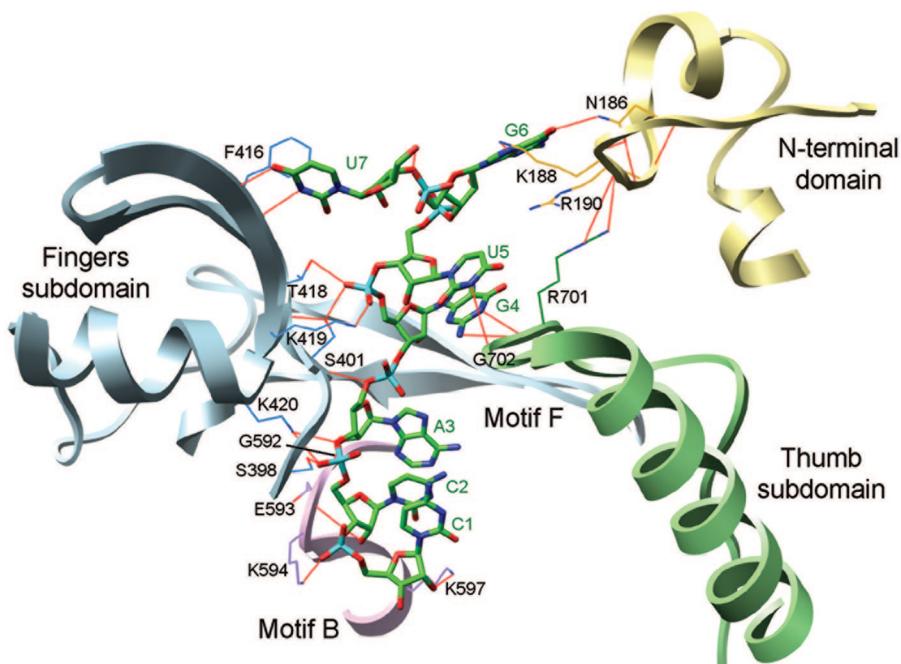


FIG. 8. RNA recognition residues of A-VP1. Partial ribbon representation of A-VP1 (PDB no. 2R7R), revealing residues (colored element sticks, labeled) that engage the 3'CS (UGUGACC) of group A +RNA. For simplicity, the nucleotides of the RNA are numbered 1 to 7 in the 3'-to-5' direction (i.e., C1, C2, A3, etc.). The image includes portions of the finger (pale blue), palm (purple), and thumb (green) subdomains and the N-terminal domain (yellow). Motifs B and F are labeled, and hydrogen bonds are red. Residues N186, K188, R190, F416, R701, and G702 are involved in specific recognition of the UGUG nucleotides. Residues lining the template entry tunnel (S398, S401, T418, K419, K420, G592, E593, K594, and K597) anchor the ribose-phosphate backbone of the RNA. For a stereo image, see Fig. S4 in the supplemental material.

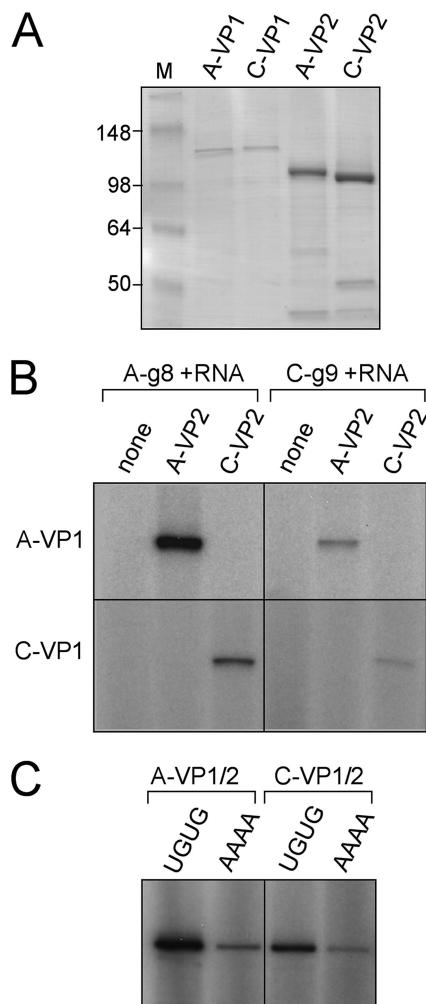


FIG. 9. C-VP1 has VP2-dependent in vitro activity and prefers a UGUG-containing template. (A) Purified VP1 and VP2. Recombinant A-VP1, C-VP1, A-VP2, and C-VP2 proteins were expressed in insect cells with baculovirus vectors and purified as described in Materials and Methods. Approximately 2 pmol of VP1 and 20 pmol of VP2 were electrophoresed in a 10% SDS-polyacrylamide gel and visualized by PageBlue staining. Lane M contains molecular size markers, and the values on the left are molecular sizes in kilodaltons. (B) In vitro dsRNA synthesis. Reaction mixtures included NTPs, Mg^{2+} , Mn^{2+} , [^{32}P]UTP, A-VP1 or C-VP1, and a group A or C template (A-g8 +RNA or C-g9 +RNA, respectively) in the absence (none) or presence of A-VP2 or C-VP2. Radiolabeled dsRNA products were resolved in a 10% SDS-polyacrylamide gel and detected by autoradiography. (C) Template mutants. Reaction mixtures included NTPs, Mg^{2+} , Mn^{2+} , [^{32}P]UTP, a group A or C protein (A-VP1/2 or C-VP1/2, respectively), and either wild-type A-g8 +RNA (UGUG) or a mutant A-g8 +RNA (AAAA) in which the VP1 recognition nucleotides UGUG were changed to AAAA. Radiolabeled dsRNA products were resolved and detected as described for panel B. The images in panels B and C are of 16-h exposures of the gels to film.

replaced with that of group B +RNA (AAAA) (Fig. 9C). We found that both A-VP1 and C-VP1 replicated the mutant templates inefficiently, supporting the conclusion that UGUG recognition is important for group A and C enzymes (Fig. 9D). Together, these results suggest that strains of group A and C (but not group B) might be capable of genetic exchange at the level of replication, but that the reassortant strains would fail

to support a productive infection if they encode incompatible proteins.

DISCUSSION

Like all segmented dsRNA viruses, RVs are capable of undergoing gene reassortment during host cell coinfection. Such exchanges result in progeny virions with gene segments belonging to more than one strain and are predicted to be a major driving force in RV evolution. Yet, there exist certain RV strain combinations seemingly incapable of reassortment under experimental conditions and in nature (3, 46). In particular, there has been no demonstration of intergroup gene reassortment among RV strains belonging to group A, B, or C. Although the reason for the observed restriction is not known, we think that multiple determinants, both direct and indirect, prevent intergroup genetic exchange. Direct determinants refer to the capacity of “foreign” RV genes to be packaged and replicated by another group’s proteins. For example, coinfection of a host cell with a group A and C RV must provide an opportunity for physical mixing of gene segments and the viral RdRp must recognize the unrelated template. On the other hand, indirect determinants of gene reassortment restriction relate to the capacity of the exchanged genes to function together in the new strain. Emerging evidence suggests that, while group A RVs readily undergo reassortment events, there are selection pressures for the maintenance of certain sets of genes (i.e., constellations) (11, 23). One of these pressures relates to how well the proteins encoded by the new RV interact during replication (11). For group A RV reassortants, the pressures are subtle and may be seen during the evolution of circulating viruses. In contrast, because RV groups encode very divergent proteins, intergroup reassortants that break preferred constellations might not be capable of carrying on a productive infection. In the present study, we combined bioinformatic, structural, and biochemical analyses to elucidate shared and group-specific features of RV RdRps. Our results provide unique insight into the functional residues of the RV RdRp and identify potential determinants influencing gene reassortment restriction. Specifically, our results suggest that template recognition by the viral RdRp VP1 might represent a direct determinant affecting gene reassortment, while interactions between VP1 and its cofactor VP2 may represent indirect determinants.

Template recognition by VP1. Numerous biochemical studies of group A RV replication have demonstrated the importance of the 3'CS (UGUGACC) in promoting +RNA recognition by A-VP1 and dsRNA synthesis (32). Electrophoretic mobility shift assays showed that the UGUG portion of the 3'CS drives high-affinity interactions with A-VP1 (41). In contrast, the terminal ACC nucleotides of the 3'CS are dispensable for specific VP1 binding but are critical for initiation of dsRNA synthesis (41). Consistent with these previous results, the high-resolution crystal structure of A-VP1 in complex with a 3'CS RNA oligonucleotide revealed direct and specific recognition of the UGUG bases by residues (N186, K188, R190, F416, R701, and G702) in the enzyme’s template entry tunnel (21). The sugar-phosphate backbone of the 3'CS is also bound, albeit nonspecifically, by several additional VP1 residues (S398, S401, T418, K419, K420, G592, E593, K594, and K597),

and the terminal nucleotides of the 3'CS (ACC) are stabilized by four motif B residues (G592, E593, K594, and K597) (21). Along with a strong electropositive charge in and surrounding the template entry tunnel, these residues allow A-VP1 to very effectively bind RNA prior to phosphodiester bond formation. Interestingly, the specific and nonspecific interactions anchor the +RNA such that its terminal nucleotide overshoots the initiation register. Thus, the VP1/+RNA structure is thought to represent a stable, catalytically inactive preinitiation complex.

Using bioinformatic analyses and homology modeling, we found that RdRps of group B and C RVs have several RNA recognition features in common with those of group A. In particular, nearly all of the residues that bind the sugar-phosphate backbone of the +RNA (S401, T418, K419, K420, S591, G592, E593, K594, and K597) are conserved for the known RV RdRps, but only C-VP1 proteins are predicted to specifically engage UGUG bases of the 3'CS in a manner very similar to what is seen for A-VP1. For instance, side chains of residues N186, K188, R190, and R701 of A-VP1 (N192, R194, R196, and R702 of C-VP1) form hydrogen bonds with the two guanines of the UGUG sequence (G4 and G6), while the amine and carbonyl groups of A-VP1 residues F416 and G702 (R422 and N703 of C-VP1) hydrogen bond with the uracils (U5 and U7). Therefore, at a direct level, it may be possible for group A and C RV gene segments, but not those of group B, to reassort during the coinfection of a cell. In support of this possibility, we found that A-VP1 and C-VP1 can replicate each other's +RNAs in vitro and that both enzymes have a strong bias toward UGUG-containing templates. In these experiments, the RdRps exhibited more robust activity on group A +RNAs than on group C +RNAs. The reason for this preference is not yet clear, but it may be due to more efficient presentation of the 3'CS in the context of group A +RNA (data not shown). Because group B +RNAs lack UGUG and B-VP1 proteins do not conserve the UGUG recognition residues, we predict that group B RVs would be incapable of reassorting genes with group A and C RVs. Thus, noncompatible interactions between the RV RdRp and RNA may represent a direct determinant influencing reassortment restriction.

It is important to note that A-VP1 is capable of using non-UGUG-containing RNA templates for multiple rounds of transcription. The 3' ends of group A -RNA show a less conserved sequence of (A/U)_nGCC that is recognized by A-VP1 but with a lower affinity than the 3'CS of +RNAs (21). Because the -RNA strand of a genomic dsRNA segment would already be paired with the RdRp inside the core, there would be no need for VP1 to be selective about which template to use during transcription. Together, the residues that bind the sugar-phosphate backbone and the electropositive charge of the template entry tunnel represent important and conserved features of nonspecific recognition that likely play a role during transcription. This nonspecific "transcription-like" polymerase activity explains the low level of activity on non-UGUG-containing templates in our in vitro assays.

Interactions between VP1 and VP2 and initiation complex formation. In vitro studies with recombinant A-VP1 have shown that +RNA recognition is necessary, but not sufficient, to form a minus-strand initiation complex (41). In fact, initiation of dsRNA synthesis in vitro is a salt-sensitive process that

requires (in addition to VP1 and +RNA) the core shell protein VP2, GTP, and Mg²⁺ (41). The need for GTP is likely a reflection of the 3'-terminal nucleotides (CC) of the group A 3'CS, and the divalent cation Mg²⁺ is a common RdRp cofactor critical for catalysis (27, 32, 41). However, the precise role VP2 plays during VP1 initiation complex formation is not well understood. The molar ratio of VP1 to VP2 required for maximum dsRNA synthesis in vitro was determined to be 1:10 (31, 41). This ratio mimics that of each decamer of the RV core, indicating that activation of the RdRp might require the formation of an assembly intermediate. We hypothesize that a VP2 decamer unit binds to a single VP1 monomer (associated with +RNA), leading to structural rearrangements in the RdRp, initiation complex formation, and dsRNA synthesis.

In this study, we show that the VP2-dependent VP1 activation mechanism also holds for group C RV proteins, which are quite divergent from group A proteins. A remarkable finding of this study is that VP1/VP2 proteins from these two groups are not functionally interchangeable in vitro. While the VP1-VP2 interaction domains remain to be elucidated, this result suggests that the nonconserved residues on the surface of VP1 may be critical for VP2 binding. Moreover, based on the location of VP1 in the core, we predict that group-specific amino acids of the VP2 inner surface are prime candidates for mediating enzyme activation. The observation that group A and C proteins undergo specific interactions supports the idea that indirect determinants influence the emergence of intergroup reassortant strains. For example, even if group A and C RVs successfully reassort their genes during coinfection, the resultant virus must encode homologous VP1 and VP2 proteins for its viability. Similar group-specific protein interactions have also been demonstrated between RV nonstructural proteins NSP2 and NSP5 (39). Group A and C VP2 and VP6 proteins have been shown to form virus-like particles in a baculovirus system, suggesting that some intergroup interactions may be allowed (17, 43). However, it remains unknown whether chimeric DLPs would have other defects downstream of assembly, such as during primary transcription.

What are the VP2-induced changes in VP1 that allow initiation complex formation? Although the answer to this question requires much more research, the structure of the reovirus RdRp (3) has provided intriguing ideas about the potential mechanism of VP1 activation. By comparison to the λ3 initiation complex, we noticed two important elements of the A-VP1 preinitiation structure that may move following VP2 binding (21, 38). First, as mentioned previously, the specific recognition of RV +RNA causes the template to overshoot its initiation register by one nucleotide. Second, the priming loop, a structural element that stabilizes the first free NTP, allowing it to base pair with the template, is disengaged (21). It is possible that VP2 interactions with VP1 cause the priming loop to shift up, the template +RNA to shift back, and dsRNA synthesis to begin. In this study, we identified several key amino acids that are conserved in all known RV RdRps and may play critical roles during catalysis. We found that residues of motifs A and C (D520, Q523, D525, D631, and D632) are likely candidates to coordinate Mg²⁺ ions and the phosphate groups of free NTPs during nucleotidyl transfer. Residues in motif B (S591, G592, E593, K594, and K597) are predicted to anchor the terminal nucleotides of the template RNA, as well as bind the

bases of N and priming (P) site NTPs. Together, residues of motif E (K679), motif F (R452, R457, and R460), and the priming loop (including E492) are also expected to support the priming NTP during de novo initiation. Ongoing studies in our laboratory aim to create recombinant A-VP1 proteins with mutations at these sites in order to define the precise function of these residues during RNA synthesis.

In conclusion, the results presented in this report provide a detailed description of the RV RdRp VP1 and identify residues that are predicted to play critical roles during RNA catalysis. Moreover, our results shed light on potential factors influencing the emergence of intergroup reassortant RV strains. In addition to the determinants of reassortment restriction discussed in this report, there are probably numerous other factors that influence this process. Indeed, a failure of divergent proteins or RNAs to interact at any stage of the RV life cycle would establish an insurmountable obstacle to productive genetic exchange. Still, given that intergroup gene reassortment has the potential to create novel and possibly more pathogenic RV strains, future studies in this area are warranted.

ACKNOWLEDGMENTS

We express our appreciation to Tamara Bar-Magen and Hongyan Yang for technical assistance and to Kristen Guglielmi, Harish Ramanathan, and Al Kapikian for insightful comments and critical reading of the manuscript.

F.D.G.-N. and D.A. are supported by grant ACT/24 of PBCT-Conicyt, Chile. J.T.P. and S.M.M. are supported by the Intramural Research Program of the NIH National Institute of Allergy and Infectious Diseases.

REFERENCES

- Baker, N. A., D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**:10037–10041.
- Bremont, M., P. Juste-Lesage, D. Chabanne-Vautherot, A. Charpilienne, and J. Cohen. 1992. Sequences of the four larger proteins of a porcine group C rotavirus and comparison with the equivalent group A rotavirus proteins. *Virology* **186**:684–692.
- Chang, K. O., P. R. Nielsen, L. A. Ward, and L. J. Saif. 1999. Dual infection of gnotobiotic calves with bovine strains of group A and porcine-like group C rotaviruses influences pathogenesis of the group C rotavirus. *J. Virol.* **73**:9284–9293.
- Chen, D., C. Q. Zeng, M. J. Wentz, M. Gorziglia, M. K. Estes, and R. F. Ramig. 1994. Template-dependent, in vitro replication of rotavirus RNA. *J. Virol.* **68**:7030–7039.
- Chen, Z., P. R. Lambden, J. Lau, E. O. Caul, and I. N. Clarke. 2002. Human group C rotavirus: completion of the genome sequence and gene coding assignments of a non-cultivable rotavirus. *Virus Res.* **83**:179–187.
- Desselberger, U., M. Iturriza-Gomara, and J. J. Gray. 2001. Rotavirus epidemiology and surveillance. *Novartis Found. Symp.* **238**:125–152.
- Eiden, J. J., and C. Hirshon. 1993. Sequence analysis of group B rotavirus gene 1 and definition of a rotavirus-specific sequence motif within the RNA polymerase gene. *Virology* **192**:154–160.
- Estes, M. K. 2001. Rotaviruses and their replication, p. 1747–1785. In D. M. Knipe and P. M. Howley (ed.), *Fields virology*, 4th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
- Gentsch, J. R., A. R. Laird, B. Bielfelt, D. D. Griffin, K. Banyai, M. Ramachandran, V. Jain, N. A. Cunliffe, O. Nakagomi, C. D. Kirkwood, T. K. Fischer, U. D. Parashar, J. S. Bressee, B. Jiang, and R. I. Glass. 2005. Serotype diversity and reassortment between human and animal rotavirus strains: implications for rotavirus vaccine programs. *J. Infect. Dis.* **192**(Suppl. 1):S146–S159.
- Haffejee, I. E. 1995. The epidemiology of rotavirus infections: a global perspective. *J. Pediatr. Gastroenterol. Nutr.* **20**:275–286.
- Heiman, E. M., S. M. McDonald, M. Barro, Z. F. Taraporewala, T. Bar-Magen, and J. T. Patton. 2008. Group A human rotavirus genomics: evidence that gene constellations are influenced by viral protein interactions. *J. Virol.* **82**:11106–11116.
- Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**:33–8, 27–8.
- Jayaram, H., M. K. Estes, and B. V. Prasad. 2004. Emerging themes in rotavirus cell entry, genome organization, transcription and replication. *Virus Res.* **101**:67–81.
- Jiang, B., L. J. Saif, J. R. Gentsch, and R. I. Glass. 2000. Completion of the four large gene sequences of porcine group C Cowden rotavirus. *Virus Genes* **20**:193–194.
- Jiang, S., S. Ji, Q. Tang, X. Cui, H. Yang, B. Kan, and S. Gao. 2008. Molecular characterization of a novel adult diarrhoea rotavirus strain J19 isolated in China and its significance for the evolution and origin of group B rotaviruses. *J. Gen. Virol.* **89**:2622–2629.
- Kapikian, A. Z., Y. Hoshino, and R. M. Chanock. 2001. Rotaviruses, p. 1787–1833. In D. M. Knipe and P. M. Howley (ed.), *Fields virology*, 4th ed. Lippincott Williams & Wilkins, Philadelphia, PA.
- Kim, Y., K. O. Chang, W. Y. Kim, and L. J. Saif. 2002. Production of hybrid double- or triple-layered virus-like particles of group A and C rotaviruses using a baculovirus expression system. *Virology* **302**:1–8.
- Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereo-chemical quality of protein structures. *J. Appl. Crystallogr.* **26**:283–291.
- Lawton, J. A., C. Q.-Y. Zeng, S. K. Mukherjee, J. Cohen, M. K. Estes, and B. V. V. Prasad. 1997. Three-dimensional structural analysis of recombinant rotavirus-like particles with intact and amino-terminal-deleted VP2: implications for the architecture of the VP2 capsid layer. *J. Virol.* **71**:7353–7360.
- Li, Z., M. L. Baker, W. Jiang, M. K. Estes, and B. V. V. Prasad. 2009. Rotavirus architecture at subnanometer resolution. *J. Virol.* **83**:1754–1766.
- Lu, X., S. M. McDonald, M. A. Tortorici, Y. J. Tao, R. Vasquez-Del Carpio, M. L. Nibert, J. T. Patton, and S. C. Harrison. 2008. Mechanism for coordinated RNA packaging and genome replication by rotavirus polymerase VP1. *Structure* **16**:1678–1688.
- Mackrell, A. D., Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fisher, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Woirkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**:3586–3616.
- Matthijssens, J., M. Ciarlet, E. Heiman, I. Arijs, T. Delbeke, S. M. McDonald, E. A. Palombo, M. Iturriza-Gomara, P. Maes, J. T. Patton, M. Rahman, and M. Van Ranst. 2008. Full genome-based classification of rotaviruses reveals a common origin between human Wa-like and porcine rotavirus strains and human DS-1-like and bovine rotavirus strains. *J. Virol.* **82**:3204–3219.
- McDonald, S. M., and J. T. Patton. 2008. Molecular characterization of a subgroup specificity associated with the rotavirus inner capsid protein VP2. *J. Virol.* **82**:2752–2764.
- McGuffin, L. J., K. Bryson, and D. T. Jones. 2000. The PSIPRED protein structure prediction server. *Bioinformatics* **16**:404–405.
- Nagashima, S., N. Kobayashi, M. Ishino, M. M. Alam, M. U. Ahmed, S. K. Paul, B. Ganesh, M. Chawla-Sarkar, T. Krishnan, T. N. Naik, and Y. H. Wang. 2008. Whole genomic characterization of a human rotavirus strain B219 belonging to a novel group of the genus rotavirus. *J. Med. Virol.* **80**:2023–2033.
- O'Reilly, E. K., and C. C. Kao. 1998. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology* **252**:287–303.
- Parashar, U. D., E. G. Hummelman, J. S. Bressee, M. A. Miller, and R. I. Glass. 2003. Global illness and deaths caused by rotavirus disease in children. *Emerg. Infect. Dis.* **9**:565–572.
- Patton, J. T. 2001. Rotavirus RNA replication and gene expression. *Novartis Found. Symp.* **238**:64–81.
- Patton, J. T. 1996. Rotavirus VP1 alone specifically binds to the 3' end of viral mRNA, but the interaction is not sufficient to initiate minus-strand synthesis. *J. Virol.* **70**:7940–7947.
- Patton, J. T., M. T. Jones, A. N. Kalbach, Y. W. He, and J. Xiaobo. 1997. Rotavirus RNA polymerase requires the core shell protein to synthesize the double-stranded RNA genome. *J. Virol.* **71**:9618–9626.
- Patton, J. T., R. Vasquez-Del Carpio, and E. Spencer. 2004. Replication and transcription of the rotavirus genome. *Curr. Pharm. Des.* **10**:3769–3777.
- Patton, J. T., R. Vasquez-Del Carpio, M. A. Tortorici, and Z. F. Taraporewala. 2007. Coupling of rotavirus genome replication and capsid assembly. *Adv. Virus Res.* **69**:167–201.
- Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**:1605–1612.
- Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**:1781–1802.
- Prasad, B. V., R. Rothnagel, C. Q. Zeng, J. Jakana, J. A. Lawton, W. Chiu, and M. K. Estes. 1996. Visualization of ordered genomic RNA and localization of transcriptional complexes in rotavirus. *Nature* **382**:471–473.
- Sali, A., and T. L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**:779–815.

38. Tao, Y., D. L. Farsetta, M. L. Nibert, and S. C. Harrison. 2002. RNA synthesis in a cage—structural studies of reovirus polymerase lambda 3. *Cell* **111**:733–745.
39. Taraporewala, Z. F., X. Jiang, R. Vasquez-Del Carpio, H. Jayaram, B. V. Prasad, and J. T. Patton. 2006. Structure-function analysis of rotavirus NSP2 octamer by using a novel complementation system. *J. Virol.* **80**:7984–7994.
40. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
41. Tortorici, M. A., T. J. Broering, M. L. Nibert, and J. T. Patton. 2003. Template recognition and formation of initiation complexes by the replicase of a segmented double-stranded RNA virus. *J. Biol. Chem.* **278**:32673–32682.
42. Tortorici, M. A., B. A. Shapiro, and J. T. Patton. 2006. A base-specific recognition signal in the 5' consensus sequence of rotavirus plus-strand RNAs promotes replication of the double-stranded RNA genome segments. *RNA* **12**:133–146.
43. Tosser, G., M. Labbe, M. Bremont, and J. Cohen. 1992. Expression of the major capsid protein VP6 of group C rotavirus and synthesis of chimeric single-shelled particles by using recombinant baculoviruses. *J. Virol.* **66**:5825–5831.
44. Unicomb, L. E., G. Podder, J. R. Gentsch, P. A. Woods, K. Z. Hasan, A. S. Faruque, M. J. Albert, and R. I. Glass. 1999. Evidence of high-frequency genomic reassortment of group A rotavirus strains in Bangladesh: emergence of type G9 in 1995. *J. Clin. Microbiol.* **37**:1885–1891.
45. Yang, J. H., N. Kobayashi, Y. H. Wang, X. Zhou, Y. Li, D. J. Zhou, Z. H. Hu, M. Ishino, M. M. Alam, T. N. Naik, and M. U. Ahmed. 2004. Phylogenetic analysis of a human group B rotavirus WH-1 detected in China in 2002. *J. Med. Virol.* **74**:662–667.
46. Yolken, R., S. Arango-Jaramillo, J. Eiden, and S. Vonderfecht. 1988. Lack of genomic reassortment following infection of infant rats with group A and group B rotaviruses. *J. Infect. Dis.* **158**:1120–1123.