The EMBO Journal (2006) 25, 880–888 | © 2006 European Molecular Biology Organization | All Rights Reserved 0261-4189/06
www.embojournal.org

THE EMBO JOURNAL

The structure of a protein primer–polymerase complex in the initiation of genome replication

Cristina Ferrer-Orta[1,3], Armando Arias[2,3], Rubén Agudo[2], Rosa Pérez-Luque[1], Cristina Escarmís[2], Esteban Domingo[2] and Nuria Verdaguer[1,*]

[1]Institut de Biologia Molecular de Barcelona (CSIC), Parc Científic de Barcelona, Barcelona, Spain and [2]Centro de Biologia Molecular 'Severo Ochoa' (CSIC-UAM), Cantoblanco, Madrid, Spain

Picornavirus RNA replication is initiated by the covalent attachment of a UMP molecule to the hydroxyl group of a tyrosine in the terminal protein VPg. This reaction is carried out by the viral RNA-dependent RNA polymerase (3D). Here, we report the X-ray structure of two complexes between foot-and-mouth disease virus 3D, VPg1, the substrate UTP and divalent cations, in the absence and in the presence of an oligoadenylate of 10 residues. In both complexes, VPg fits the RNA binding cleft of the polymerase and projects the key residue Tyr3 into the active site of 3D. This is achieved by multiple interactions with residues of motif F and helix α8 of the fingers domain and helix α13 of the thumb domain of the polymerase. The complex obtained in the presence of the oligoadenylate showed the product of the VPg uridylylation (VPg-UMP). Two metal ions and the catalytic aspartic acids of the polymerase active site, together with the basic residues of motif F, have been identified as participating in the priming reaction.

The EMBO Journal (2006) **25,** 880–888. doi:10.1038/sj.emboj.7600971; Published online 2 February 2006
Subject Categories: structural biology
Keywords: foot-and-mouth disease virus; primer–protein; replication; uridylylation; VPg

Introduction

A number of DNA viruses, RNA viruses and linear plasmids use small proteins as primers for DNA or RNA synthesis (Salas, 1991). These proteins are covalently attached to the 5′ end of the DNA or RNA genomes. During initiation of replication, they provide the hydroxyl group required by the polymerase to start polynucleotide synthesis. In the process of replication, the first step is the linkage of a nucleotide to the hydroxyl group of an amino acid in the primer protein. This process is catalyzed by the DNA or RNA polymerase and the same enzymes are also responsible for the elongation of the DNA or RNA strands. Thus, these polymerases are structurally adapted for using either a protein or a nucleic acid as a primer.

Picornaviruses use a protein of 20–24 amino acids, termed VPg, to initiate viral RNA synthesis (Paul, 2002). The different picornaviruses show partial amino-acid sequence identities in their VPgs that range from 17 to 26% (http://www.iah.bbsrc.ac.uk/virus/picornaviridae/SequenceDatabase/alignments/pico2c3d_pep.txt). Foot-and-mouth disease virus (FMDV), a member of the Picornaviridae family, possesses a linear plus strand RNA genome about 8500 nucleotides long. It contains a long 5′-untranslated region (UTR), a single open reading frame and a short 3′-UTR with a poly(A) tail. The 5′-terminal uridine of the RNA is covalently linked to the hydroxyl group of a tyrosine in the terminal protein VPg. The genome of FMDV encodes three different VPgs (VPg1, VPg2 and VPg3) in tandem (Forss and Shaller, 1982), and all three are active as primers for RNA replication. Deletion of any individual copy of VPg has a deleterious effect on RNA replication (Falk *et al*, 1992). The replication of this RNA requires all of the nonstructural proteins of the virus and also several *cis*-replicating elements (Nayak *et al*, 2005). The viral proteins most directly involved in RNA synthesis are the template- and primer-dependent RNA polymerase (3D), the terminal protein VPg and 3CD (unprocessed 3C–3D), which acts as a proteinase and an RNA binding protein (Paul, 2002; Nayak *et al*, 2005).

We have previously reported the structure and interactions of the FMDV 3D polymerase with a template–primer RNA substrate, providing insights of the RNA elongation process and the interactions of the FMDV 3D with the exiting duplex product (Ferrer-Orta *et al*, 2004). In addition, modeling studies based on alignments of the FMDV 3D/RNA complex with the complexes HIV-RT/DNA and phi6-RdRP/DNA (Huang *et al*, 1998; Butcher *et al*, 2001) allowed us to identify several key residues in the FMDV-3D polymerase, which may be involved in nucleotide (NTP) binding and catalysis. In our model, the conserved residues Asp338 and Asp339 of motif C, together with Asp245 in motif A, are in the adequate position to bind to the divalent cations, helping to orient the 5′ phosphate group of the incoming nucleotide for nucleophilic attack by the 3′ end of the primer. The basic residues Arg168, Lys172 and Arg179 of motif F are also in an ideal position to interact with the negatively charged phosphates of the nucleotide substrate (Ferrer-Orta *et al*, 2004). The mechanism of the critical event of initiation of RNA synthesis during picornavirus replication is largely unknown owing to the absence of structural information. Here, we report the 3D structure of two complexes between the FMDV polymerase (3D) and its protein–primer VPg1 in its nonuridylylated and uridylylated forms. The structures reveal critical VPg and 3D amino acids and the interactions involved in the positioning and addition of the first nucleotide (UMP) to the VPg molecule. In addition, the complex with the uridylylated

*Corresponding author. Institut de Biologia Molecular de Barcelona (CSIC), Parc Científic de Barcelona, Josep Samitier 1-5, Barcelona 08028, Spain. Tel.: +34 93 403 49 52; Fax: +34 93 403 49 79;
E-mail: nvmcri@ibmb.csic.es
[3]These authors contributed equally to this work

Received: 9 August 2005; accepted: 2 January 2006; published online: 2 February 2006

VPg1 form clarifies the role of the metal ions in the priming reaction and provides new insights into the mechanism of initiation of RNA genome replication. Amino-acid replacements at several positions of 3D and VPg resulted in the modification of uridylylation activity, in agreement with the structure of the complexes.

Results

Structure of the VPg primer protein bound to the 3D polymerase

Two different crystal forms of the complex between the FMDV 3D polymerase, the primer protein VPg, UTP and metal ions, in the absence and in the presence of an RNA oligonucleotide of 10 adenine bases (A10) were obtained and analyzed by X-ray crystallography. The structures of the complexes were obtained at 3.0 and 2.9 Å, respectively (Table I). In the two crystal forms, the quality of the electron density maps, after several rounds of manual modeling and refinement, allowed the accurate positioning of 15 of the 23 amino acids of the VPg protein (Figure 1). Weak and discontinuous density was also visible to accommodate one or two additional VPg residues poorly ordered, exiting from the polymerase. The remaining six carboxy-terminal residues of VPg appeared completely disordered.

The 15 VPg amino acids determined in both structures showed almost the same conformation with little secondary structure. The N-terminal portion of the protein is located close to the NTP entry cavity and projects the side chain of residue Tyr3 into the active site. Then, the peptide chain snakes through the large RNA binding cleft towards the thumb domain of the 3D protein, following a similar trajectory to that taken by the RNA primer and duplex product in the 3D–RNA complex (Ferrer-Orta *et al*, 2004; Figure 2).

In the crystals obtained when the oligonucleotide A10 was absent, no density was observed to accommodate the UTP substrate. However, in the complex obtained in the presence of A10 and $MgCl_2$, the difference in electron density maps ($2|F_o|-|F_c|$ and $|F_o|-|F_c|$) showed the product of the VPg uridylylation (VPg-UMP) and revealed the presence of two metal ions in the polymerase catalytic site (Figures 1A and B). To distinguish whether Mg^{2+} or Mn^{2+} was at the active site was not obvious because both cations were present during the complex formation and the crystallization processes. We have recently reported the ability of FMDV 3D polymerase to catalyze the uridylylation of VPg using poly(A) as a template in the presence of Mn^{2+} ions (Arias *et al*, 2005). In contrast, the VPg uridylylation in the presence of Mg^{2+} was undetectable (A Arias and R Agudo, unpublished results). In order to provide an experimental evidence of the presence of Mn^{2+} in the active site of our complex, we collected a new data set of the 3D–VPg–UTP–metal complex cocrystals using the wavelength of the radiation tuned to the Mn^{2+} K edge, 1.89 Å (see Materials and methods section and Table I). The anomalous scattering difference Fourier map showed one peak (4σ level), indicating the presence of a single Mn^{2+} ion bound to the active site of the polymerase. The second strong peak seen in conventional difference Fourier maps was then interpreted by the presence of one additional Mg^{2+} cation (Figure 1). No ordered density was visible in the template entry channel to position the oligonucleotide that was expected to act as a template in the uridylylation. Therefore,

Table I Data collection and refinement statistics

	3D + VPg	3D + VPg + UMP	3D + VPg + UMP[a]
Data collection details			
Space group	$P3_221$	$P3_221$	$P3_221$
Unit cell (Å)	$a=b=95.09$, $c=100.63$	$a=b=94.38$, $c=99.71$	$a=b=94.530$, $c=100.403$
Wavelength (Å)	0.93	0.98	1.89
Resolution range[b] (Å)	30–3.0 (3.1–3.0)	30–2.9 (3.0–2.9)	30–3.1 (3.2–3.1)
$\langle I/\sigma \rangle$	6.8 (2.3)	7.5 (2.8)	12.2 (7.5)
R_{merge}[c]	9.5 (40.5)	9.5 (45.5)	6.1 (37.2)
Completeness (%)	100 (100)	99.9 (99.9)	100 (99.9)
Refinement statistics			
Unique data	10 830	11 795	
R_{work}[d] (%)	27.7	26.9	
R_{free}[d] (%)	29.4	28.5	
No residues			
3D pol	474	474	
VPg	15	15	
Metal ions	—	Mn^{2+} and Mg^{2+}	
Average temperature factors (Å^2)			
All atoms	76.9	65.8	
3D pol	76.8	65.5	
VPg protein[e]	81.9	66.4	
Metal ions[e]	—	66.7	

[a]Data collected at the Mn K edge and used only to calculate the anomalous scattering difference Fourier map.
[b]Data within parenthesis are for the highest resolution shell.
[c]$R_{merge}=\sum_j\sum_h (|I_j,h-\langle I_h\rangle|)/\sum_j\sum_h (\langle I_h\rangle)$, where h are unique reflections indices, I_j,h are intensities of symmetry-related reflections and $\langle I_h\rangle$ is the mean intensity.
[d]R_{work} and R_{free} are defined by $R=\sum_{hkl}||F_{obs}|-|F_{calc}||/\sum_{hkl}|F_{obs}|$, where h, k and l are the indices of the reflections (used in refinement for R_{work}; 5%, not used in refinement, for R_{free}), and F_{obs} and F_{calc} are the structure factors, deduced from measured intensities and calculated from the model, respectively.
[e]The VPg protein and metal ions were refined with occupancy of 70%.

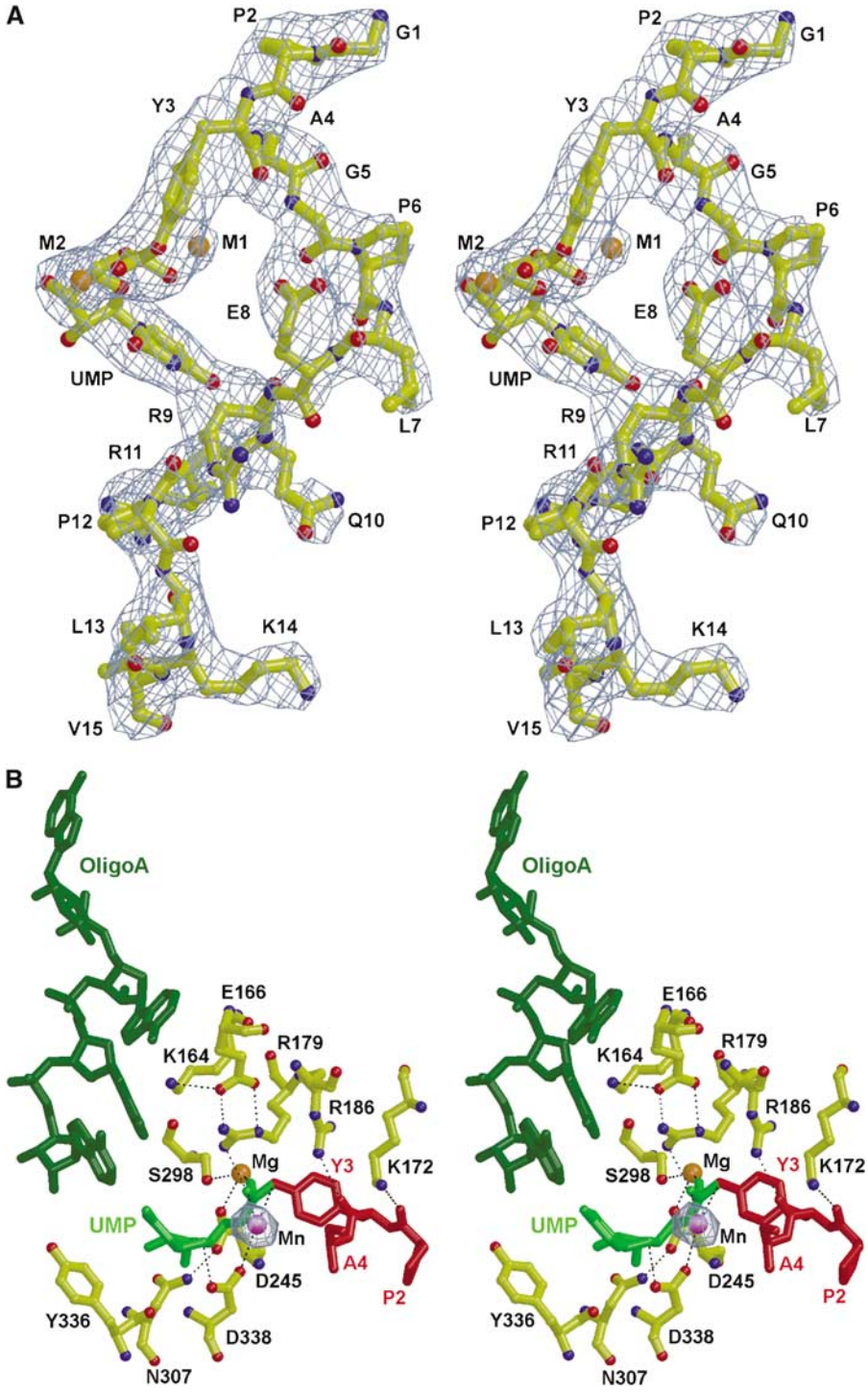

Figure 1 Structure of the primer protein VPg in a complex with 3D. (**A**) Stereo view of a sigma A weighted $|F_o|-|F_c|$ electron density map at 2.9 Å resolution and contoured at 3.0σ around the VPg–UMP molecule (The VPg–UMP and ions were omitted from the phasing model). The 15 amino acids of VPg, the UMP covalently linked to the protein and the metal ions are placed inside the density in ball and stick representation colored in atom type code. Names for all residues are explicitly labeled in one letter code. (**B**) Details of the interactions seen in the active site of the 3D polymerase during the uridylylation reaction. The residues Pro2, Tyr3 and Ala4 of VPg are shown in sticks in red and the UMP, covalently linked to the hydroxyl group of Tyr3, in light green. The divalent cations Mn^{2+} and Mg^{2+} are shown as magenta and orange spheres, respectively, and the anomalous difference Fourier map is shown as a chicken wire in blue. The 3D amino acids involved in direct hydrogen bonds with ions and the uridylylated tyrosine are shown in ball and sticks in atom type code, and the hydrogen bonds appear as dashed lines. All residues are explicitly labeled. The predicted position of the oligo(A) template strand (dark green) was determined using the 3D–RNA template–primer complex (PDB entry 1WNE) as a guide.

it seems that uridylylation reaction took place before crystal formation. The pyrophosphate product was also absent. The position of the oligo(A) template was further modeled using the FMDV 3D–RNA template/primer complex (Ferrer-Orta *et al*, 2004) as a guide (Figure 1B). The resulting model shows how the first five nucleotides of the oligo(A) template

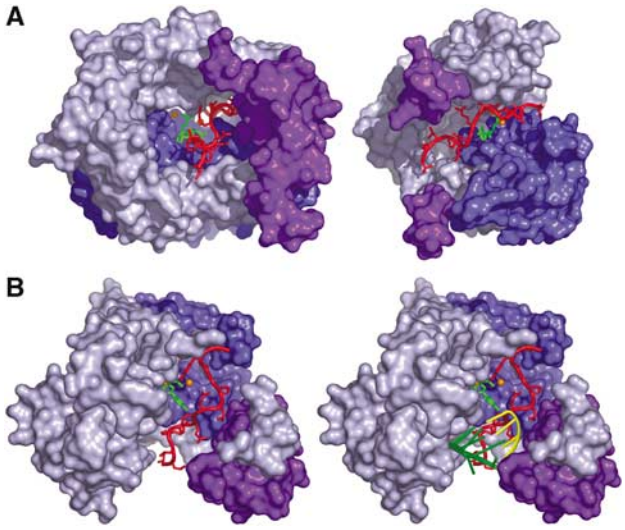

Figure 2 Structure of VPg bound to the FMDV 3D polymerase. (**A**) Molecular surface of the 3D polymerase shown in two different views: the conventional orientation, as if looking into a right hand (left) and a side view (right). The polymerase domains: fingers, palm and thumb are colored in light-blue, dark-blue and purple, respectively. The VPg protein that binds across the RNA binding cleft is represented as a red ribbon with side chains shown in stick representation in red, the UMP molecule covalently linked to the hydroxyl group of the VPg residue Tyr3 is shown in green, and the metal ions as orange spheres. In the right panel, the fingers residues at the top of the NTP tunnel and most of the thumb domain are removed to allow the visualization of VPg. (**B**) Top-down views of the polymerase molecule showing the trajectory of the VPg protein (left) compared to the trajectory of RNA template–primer (right) in the structure of the complex FMDV 3D–RNA template–primer (PDB entry 1WNE; Ferrer-Orta *et al*, 2004). Comparisons were carried out by the structural superimposition of both polymerase complexes that gave a root mean square deviation of 0.33 Å for the superimposition of all Cα atoms. The RNA is shown as a ribbon representation in green (template chain) and yellow (primer chain). Roughly, the N-terminal moiety of the bound VPg protein occupies part of the NTP entry channel and the position of the primer, whereas the carboxy-terminus of the protein superimposes with the trajectory of the RNA duplex product. In both panels, the N-terminal residues (from 34 to 48) and residues at the top of the NTP tunnel (from 163 to 180) of 3D are omitted to better show the VPg protein.

traverse the template entry channel pointing the base, A5 towards the catalytic site in a correct position, to be paired with the UMP molecule covalently attached to VPg (Figure 1B).

VPg–3D polymerase interactions

The structures determined show how VPg accesses the polymerase catalytic site from the front part of the molecule through the RNA binding cleft (Figure 2). No important conformation changes of the 3D polymerase conformation were induced by the VPg binding; only local rearrangements in the side chain of residues Arg179 and Asp338, which participate directly in the uridylylation reaction, were seen (see below). The largest network of VPg–3D interactions was observed between residues Glu166, Ile167, Arg168, Lys172 and Arg179, in motif F of the fingers domain, which, together with residues Thr407, Ala410 and Ile411 of helix α13 of the thumb domain, contacted the N-terminal moiety of VPg, stabilizing the conformation of Tyr3 in the active site cavity (Figure 3). Two main chain–side chain hydrogen bonds were

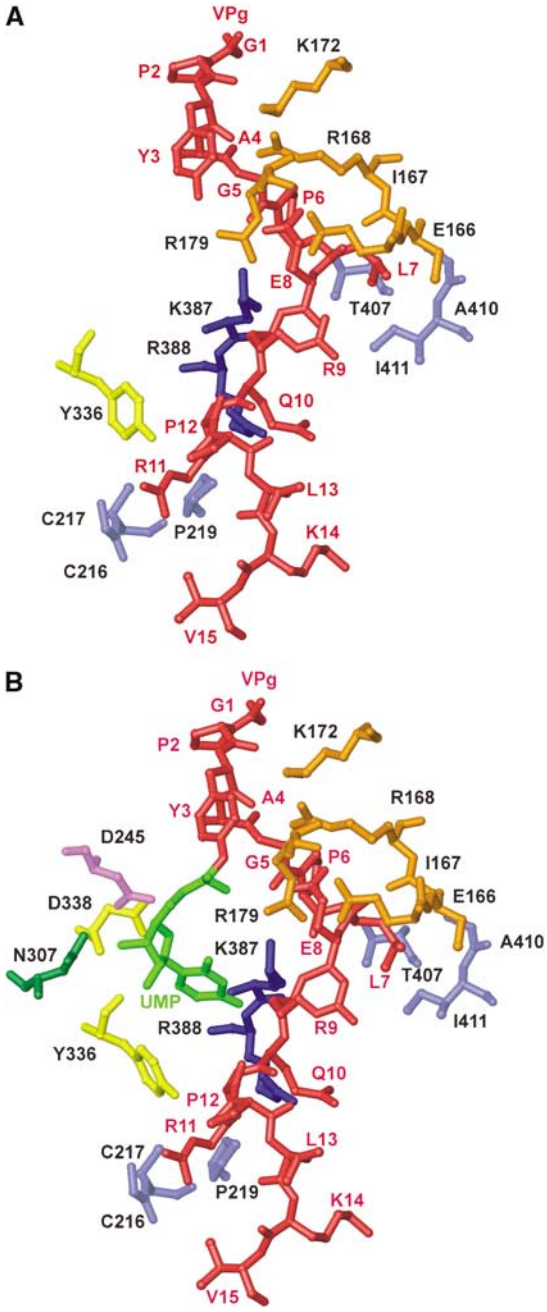

Figure 3 VPg–3D polymerase interactions. (**A**) Structure of the VPg primer protein (red) with the contacting residues of the 3D polymerase shown in different colors. Four different regions of the polymerase molecule contact VPg residues E166, I167, R168, K172 and R179, belonging to motif F of fingers (orange), together with residues T407, A410 and I411 of the thumb domain (light blue), interact with the N-terminal moiety of VPg, stabilizing the conformation of Y3 in the active site cavity. In addition, residues E166, I167 of motif F (orange), K387 and R388 of motif E (dark blue) and T407, A410 and I411 of helix α13 (light blue) interact with the central part of the VPg protein. Finally, the 3D residues G216, C217 and P219, located at the beginning of helix α8 (light blue) in the fingers domain, together with the side chain of Y336 within the C motif (yellow) of the palm domain, establish hydrophobic contacts with R11 at the exit of the polymerase cavity. (**B**) Structure of the uridylylated VPg protein (shown in red and the linked UMP in green) with the contacting residues of the 3D polymerase shown in blue. In addition to the interactions described in (A), amino acids D245 of motif A (pink) and D338 of motif C (yellow) are placed in the correct orientation for the catalysis of the phosphodiester linkage in the active site of the 3D protein.

observed between the backbone oxygen atoms of Pro2 and Tyr3 of VPg, and the amino groups of the side chains of Lys172 and Arg168 of the polymerase. The guanidinium group of Arg179, which forms a double salt bridge with Glu166 within the same motif F of 3D is also hydrogen bonded with the hydroxyl group of Tyr3 in the nonuridylylated VPg. In the complex with the uridylylated form of VPg, the side chain of Arg179 changes slightly its conformation, maintaining the interaction with Glu166 and forming an additional salt bridge with the phosphate oxygen O2 of UMP (Figures 1B and 3). The main chain oxygen atoms of Pro6 and Leu7 of VPg were hydrogen bonded to the side chain of the 3D residue Lys387, located in motif E of the palm domain. The side chain of Leu7 is also located in a hydrophobic cavity formed by the side chains of residues Glu166 and Ile167 of motif F and the side chains of Thr407, Ala410 and Ile411 of helix α13 (Figure 3). Additional contacts were observed between Arg388 of motif E and residues Glu8, Arg9 and Gln10 of VPg. The side chain of Tyr336, within the catalytic motif C, also participates in interactions with VPg forming a hydrogen bond with the guanidinium group of Arg11 of the peptide primer. Finally, the 3D residues Gly216, Cys217 and Pro219, located at the beginning of the helix α8 in the fingers domain, also establish hydrophobic contacts with VPg residue Arg11 at the exit of the polymerase cavity (Figure 3).

Insights into the mechanism of VPg uridylylation

The comparison of the 3D–VPg and the 3D–VPg–UMP structures reveal how the side chain of Tyr3 is oriented towards the active site close to the catalytic aspartic acids 245 of motif A and 338 of motif C. In the 3D–VPg–UMP complex, the hydroxyl group of Tyr3 side chain was found covalently attached to a UMP molecule by a phosphodiester linkage (Figure 1B). Two metal ions, Mn^{2+} and Mg^{2+}, participate in the uridylylation reaction that appears to follow a similar mechanism to that described for the nucleotidyl transfer reaction in other polymerases (Steitz, 1998). Mn^{2+} bridges the carboxylate group of Asp338 and the O^{-} of tyrosine side chain, now covalently bonded to phosphate α of UMP. The side chain of Asp338 changes its rotamer conformation to optimize the described interactions. Mg^{2+} coordinates the carboxylic group of Asp245, the O1 oxygen of phosphate α and the hydroxyl group of Ser298 side chain. Tyr336 of motif C, which is hydrogen bonded to the VPg residue Arg11, also participates in hydrophobic contacts with the UMP molecule covalently attached to VPg (Figure 3). Finally, the positively charged residues of motif F (Arg168, Lys172 and Arg179) also participate in the uridylylation process (Figure 1B). These basic amino acids of motif F, together with Tyr336 of motif C, seem to play a role stabilizing Tyr3 and UMP in a proper conformation for the catalytic reaction.

Biochemical analysis of structure-based mutants of polymerase and VPg molecules

To investigate the effect of amino-acid replacements in 3D and VPg in the uridylylation reaction, five single replacements (Glu166-Ala, Glu166-Arg, Arg168-Ala, Arg179-Ala, Asp338-Ala) and two double replacements (Lys387-Ala/Arg388-Ala, Thr407-Ala/Ile411-Ala) were introduced in 3D by site-directed mutagenesis of plasmid pET-28a3Dpol (Materials and methods). The uridylylation activity on VPg1 was tested and compared with that of the wild-type 3D. The results (Figure 4A) show at least a 10-fold reduction for all mutants examined, except for 3D mutant Arg168-Ala that showed about 60% of the activity relative to wild-type 3D, and the double mutation Thr407-Ala/Ile411-Ala, which did not produce any detectable effect on the uridylylation activity.

VPg replacements Tyr3-Phe, Glu8-Ala, Arg9-Ala, Arg9-Glu and, to a lesser extent, Pro6-Ala resulted in significant decreases in uridylylation activity, as compared with wild-type VPg1, VPg2 or VPg3 (Figure 4B). Thus, modification of 3D and VPg residues, predicted to be involved in critical interactions according to the 3D structure of the uridylylation complex, have remarkable effects in the uridylylation reaction.

Discussion

Implications for the mechanism of initiation of RNA synthesis

The 3D structures of RdRPs, unbound to ligands, are available for four other members of the Picornaviridae family (PV, HRV1B, HRV14 and HRV16) (Love *et al*, 2004; Thompson and Peersen, 2004; Appleby *et al*, 2005). The structure-based alignment of this five picornaviral polymerases shows that

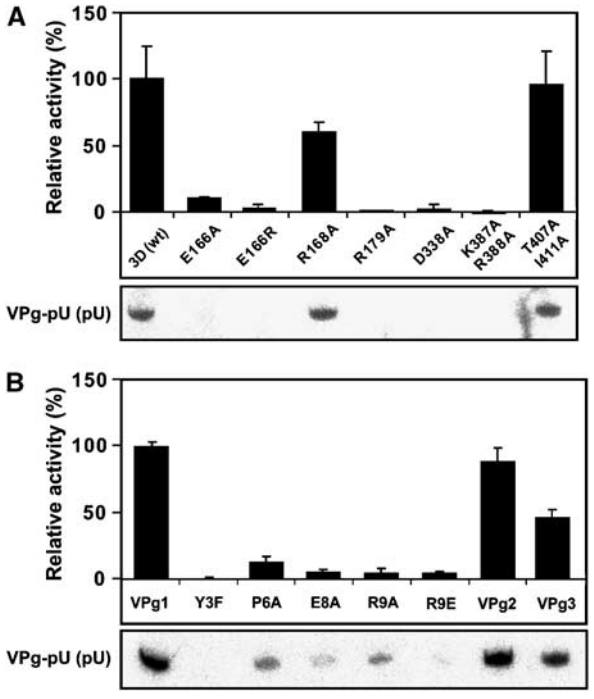

Figure 4 Effect of amino acids replacements in 3D and VPg on the VPg-uridylylation activity. (**A**) VPg1 uridylylation activity by substituted FMDV 3Ds, including one or two amino-acid replacements, relative to the wild-type enzyme. Values are the average of at least three independent experiments. The mono- and di-uridylylated band of a representative electrophoretic separation is shown. No significant differences were seen when the reaction products were treated with RNAse A before electrophoresis. Preparation of mutant 3Ds and assay procedures are described in Materials and methods. (**B**) Uridylylation activity of wild-type FMDV 3D on VPg1, VPg2, VPg3 and VPg1 with single amino-acid substitution. Results and procedures are as in (A).

12 out of the 16 residues of 3D polymerase involved in VPg binding are strictly conserved (Figure 5A). Sequence alignments of the VPg proteins from these picornaviruses also revealed a partial conservation of the 3D-interacting residues (Figure 5B). Mapping of the 12 VPg-interacting residues onto the surface of the FMDV polymerase is shown in Figure 6A. Structural comparisons between FMDV–3D–VPg complex with the other four picornaviral RdRPs showed that the size and the shape of the central cavity, where VPg binds, is almost identical in all five polymerases (Figure 6B), and FMDV VPg protein can be modeled in the central cavity of all picornaviral polymerases, without important steric hindrances. The modeled VPg in the central cavity of PV and rhinovirus RdRPs retained almost the same interactions described in the FMDV–3D–VPg complex. In particular, those relevant to the uridylylation reaction would be strictly conserved (Figures 5 and 6B).

The FMDV mutants at the conserved polymerase residues that strongly interact with VPg—Glu166 and Arg179 of motif F, Asp338 of motif C and Lys387/Arg388 of motif E—show a drastic defect in VPg uridylylation (Figure 4). Substitution of Arg168 resulted in ~40% reduction in uridylylation. This residue contacts VPg through a hydrogen bond with the backbone oxygen atom of Tyr3. The lack of the Arg168 side chain can be, at least in part, compensated by the presence in the vicinity of other positively charged side chains, like that of Lys172 (Figure 3). The double mutation Thr407/Ile411 to alanine shows no defect in VPg uridylylation (Figure 4). These residues form part of a hydrophobic cavity that accommodates Leu7 of VPg (Figure 3). It seems that alanine residues do not disturb this hydrophobic cavity. Both residues are not conserved among picornaviruses.

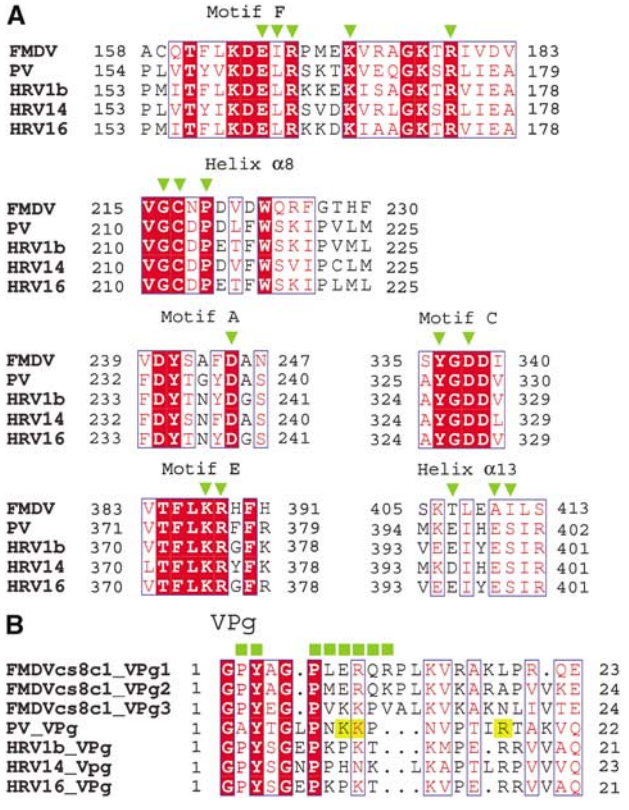

Figure 5 (**A**) Structure-based alignment of the picornavirus polymerase residues that make contact with VPg. Aligned are domains for the different picornaviral polymerases whose structure is known. The strictly conserved residues are in red blocks and similar residues in red characters and blue boxes. The FMDV 3D residues interacting with the VPg primer protein are marked by green inverted triangles (**B**) Sequence alignment of the VPg protein of the different picornaviruses. The FMDV VPg1 residues interacting with 3D are marked by green squares. Residues of PV VPg previously shown to interact with PV 3D polymerase by yeast two-hybrid analysis (Xiang *et al*, 1998) are highlighted in yellow blocks.

Mutational analysis in PV has identified two groups of residues at the surface of the polymerase whose substitutions affected either the rate of VPg uridylylation or the 3AB binding affinity and the ability of membrane-bound 3AB recruitment. Uridylylation was also partially affected by this second group of mutants (Lyle *et al*, 2002; Boerner *et al*, 2005; Figure 6C). Substitution of the first group of residues (Tyr326, Asp358 and Lys359) corresponding to FMDV (Tyr336, Asp368 and Lys369) results in a drastic loss of uridylylation activity (Lyle *et al*, 2002). These residues are located at or near the VPg binding site (Figure 6C). The second group of residues (Phe377, Arg379, Glu382 and Val391) that would correspond to His389, His391, Tyr394 and Val402 in FMDV are located on the back side of the polymerase and lie far from VPg to allow direct contacts with this primer protein (Figure 6C). Substitution of most amino acids of this second group resulted only in a moderate decrease of VPg uridylylation activity, perhaps because this major effect is on binding of the 3AB precursor rather than of VPg (3B). Additional studies are needed to clarify this point. It can also be considered that these second group of residues, which lie within and around motif E at the hinge region of the palm and thumb domains, could play a role in maintaining the structural integrity and proper positioning of key polymerase residues during VPg uridylylation.

Replacement of the conserved VPg residues interacting with 3D (Tyr3, Pro6, Glu8 and Arg9) resulted in a dramatic reduction in the rate of VPg uridylylation (Figure 4B). Previous mutational studies in PV, using an inducible yeast two-hybrid system, identified three positively charged residues of VPg (Lys9, Lys10 and Arg17; Figure 5B) as interacting with the 3D polymerase (Xiang *et al*, 1998). The mutational and structural results obtained in the present work (Figures 3 and 4) correlate well with this previous observation with PV.

In light of the structural and mutational results, we suggest a general mode of VPg binding and uridylylation through the front side of the picornaviral polymerases. A similar VPg binding mode was previously hypothesized for HRV16, on the basis of the crystallographic structure of the RdRP of this virus and molecular modeling (Appleby *et al*, 2005).

Comparisons among different RNA-dependent RNA polymerases whose structures have been solved show that those viruses that follow a primer-dependent mechanism of initiation of replication, such as picornaviruses and caliciviruses, have a more accessible active site than viruses with a *de novo* initiation mechanism, such as flaviviruses and bacteriophage ϕ6 (van Dijk *et al*, 2004; Ferrer-Orta *et al*, 2004, 2006). An extra C-termini domain has been determined for polymerases with a *de novo* initiation of RNA synthesis (Butcher *et al*, 2001; van Dijk *et al*, 2004). The C-terminal protrusions of flaviviruses and ϕ6 polymerases partially occlude the active site, resulting in a more compact molecule where two narrow positively charged tunnels allow the access of RNA template and NTP substrate to the active site. The template tunnel is

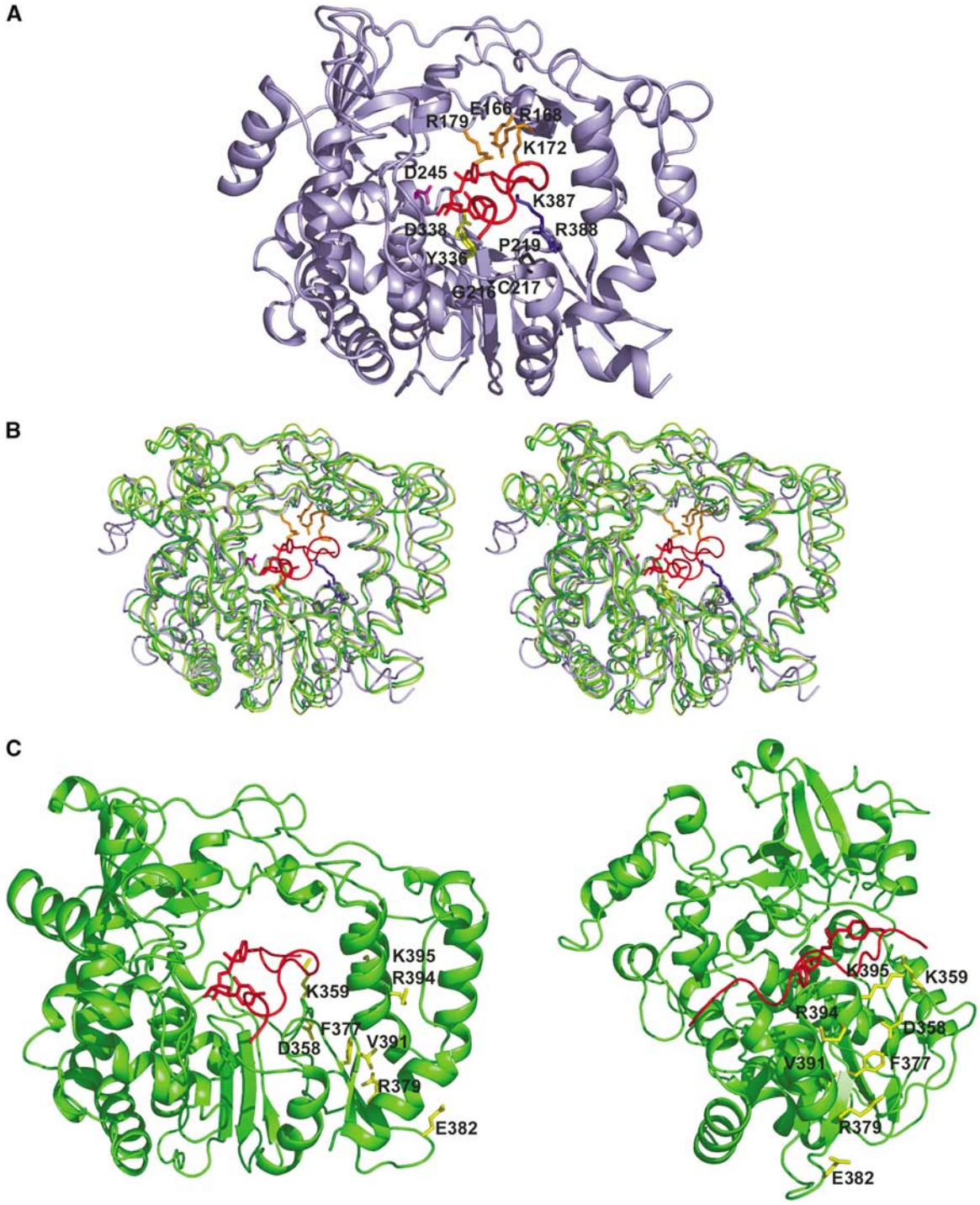

Figure 6 Comparison of FMDV polymerase bound to VPg with other picornaviral polymerases. (**A**) Mapping of the conserved FMDV 3D residues that contact the VPg primer protein onto the polymerases structure (blue). The VPg-interacting residues are colored according to the different polymerase motifs: D245 of motif A is in magenta; E166, R168, K172 and R179 belonging to motif F are in orange; residues G216, C217 and P219 located in or near helix α8 in dark gray; amino acids Y336 and D338 of motif C are in yellow; and residues K387 and R388 of motif E in dark blue. The VPg protein occupying the central cleft of the polymerase is shown in red. (**B**) Stereoview of the structural superimposition of the coordinates of PV (green) and HRV16 (yellow) RdRPs onto the FMDV 3D–VPg complex (blue), showing the positioning of the VPg primer protein (red) in the central cavity of the polymerases. The Cα root mean square deviation is 1.24 Å for the superimposition of 362 atoms between PV and FMDV RdRPs, and 1.29, 1.25 and 1.29 Å for the superimpositions of 254, 370 and 353 Cα atoms, between FMDV and HRV1B, HRV14 and HRV16 RdRPs, respectively. The conserved VPg-interacting residues are shown as in (A). (**C**) Ribbon diagram of the structure of PV 3D polymerase (green) with the VPg of FMDV modeled inside (red). Two different views of the PV polymerase are shown: the conventional orientation (left) and a side view (right; the finger residues at the top of the NTP tunnel and most of the thumb domain are removed to allow the visualization of VPg). The residues that affected VPg binding (yellow) and uridylylation (blue), as determined by mutational analysis (Lyle *et al*, 2002), are shown as stick representation and are explicitly labeled. In the model, Tyr326 has the correct orientation to interact directly with the UMP substrate, and Lys359 is properly oriented to interact directly with the N-terminal moiety of the VPg peptide.

wide enough to accommodate ssRNA but not dsRNA (van Dijk *et al*, 2004; Ferrer-Orta *et al*, 2006). Structural and biochemical studies indicated that the protruding extensions of the thumb domains of phi6 and flavivirus polymerases play two distinct functions: acting as priming platforms that stabilize the initiation complexes, and preventing undesirable back-priming reaction by physically separating the template binding site from the room reserved for the daughter RNA chain. Furthermore, these initiation platforms block the path of the elongating RNA product at the level of two or three nucleotides and large conformational rearrangements are required to accommodate longer product chains (Van Dijk *et al*, 2004; Ferrer-Orta *et al*, 2006).

In contrast, the structure of FMDV 3D in complex with RNA determined recently shows how the wide central cleft of picornavirus polymerases is able to accommodate a template–primer duplex during the phase of RNA elongation (Ferrer-Orta *et al*, 2004), and the structure presented in this work demonstrated that this cavity can also accommodate the protein primer during the initiation step of picornavirus replication.

Evolutionary implications

RNA synthesized by DNA primases is involved in the initiation of cellular DNA replication (Alberts *et al*, 2002). RNAs act as primers for replication of some DNA viruses, and for transcription of several RNA viruses; capped RNA structures, captured from cellular mRNAs, serve to initiate orthomyxovirus and bunyavirus replication by 'cap snatching' (Alberts *et al*, 2002; Flint *et al*, 2004).

The position of VPg in complex with the FMDV 3D polymerase is remarkably similar to the position of the primer and RNA duplex product found in the complex with the same enzyme (Figure 2B) (Ferrer-Orta *et al*, 2004). Most of the amino acids of 3D seen in contact with the RNA primer and duplex product are also involved in interactions with VPg (Figures 1B and 3). In fact, the structure shows how the VPg protein accesses the active site cavity from the front of the molecule through the large RNA binding cleft mimicking, at least in part, the RNA molecule (Figure 2). The N-terminal position of VPg projects into the active site where the hydroxyl moiety of the residue Tyr3 is in good proximity to the catalytic aspartates 245 of motif A and 338 of motif C. In this position, Tyr3 essentially mimics the 3′OH of the primer strand during the RNA elongation. All the observed structural features suggest a conservation of the catalytic mechanism described for all polymerases (Steitz, 1998).

For viruses, whose survival must rely in the efficient production of many progeny copies, stability and velocity of replication are essential, and these features may have favored, in prolonged evolutionary times, protein versus RNA, as an alternative to other chemical modifications to protect the genome termini.

Materials and methods

Crystallization, data collection and processing

FMDV polymerase (3D) was obtained and purified as described in Ferrer-Orta *et al* (2004) and stored in a buffer containing 40 mM Tris–HCl (pH 7.5), 0.5 M NaCl, 0.8 mM DTT, 0.8 mM EDTA and 8% glycerol at a concentration of 4.6 mg/ml. The VPg used for crystallization was prepared by solid-phase peptide synthesis of the FMDV C-S8c1 VPg1 sequence (GPYAGPLERQRPLKVRAKLP RQE), purified by G25 Sephadex chromatography and HPLC, and analyzed by mass spectrometry.

The 3D–VPg complex was obtained and crystallized as follows: 9.5 nmol of VPg were mixed with 2 mM UTP and then the 3D polymerase was slowly added in an equimolar proportion in the presence of 2 mM $MnCl_2$. The solution was left 12 h at 4°C before the crystallization trials. Crystals were grown by vapor diffusion at 20°C from a solution containing 33% PEG 4K, 0.2 M ammonium acetate, 0.1 M sodium citrate (pH 5.6) and 4% γ-butyrolactone. Trigonal $P3_221$ crystals ($a=b=95.09$ Å, $c=100.63$ Å) typically appeared after 2 days.

The 3D–VPg–UMP–oligoA10 complex was obtained and crystallized as follows: firstly, 5 nmol of 3D polymerase were mixed with an equimolar proportion of the A10 oligonucleotide in the presence of 2 mM $MgCl_2$. A second solution was prepared containing 5 nmol of VPg and UTP (2 mM) in the presence of 2 mM $MnCl_2$. Then, both solutions were slowly mixed and incubated for 12 h at 4°C. Trigonal crystals ($P3_221$, $a=b=94.38$ Å, $c=99.71$ Å) were grown from a solution containing 32% PEG 4K, 0.2 M ammonium acetate, 0.1 M MES (pH 6.0) and 4% γ-butyrolactone.

The two crystal forms were flash frozen in liquid nitrogen, using cryoprotectants that contained 20% glycerol in the crystallization buffers. Three data sets were collected at 100K using synchrotron radiation at the ESRF (Table I). The X-ray diffraction data for the 3D–VPg complex were collected at 3.0 Å resolution on beam line ID14 EH2 ($\lambda=0.93$ Å). For the 3D–VPg–UMP crystals, two different data sets were obtained. The first data set was collected at 2.9 Å on beam line ID14 EH4 ($\lambda=0.98$ Å) and the second X-ray data (3.2 Å resolution) were collected on BM16 beam line at the K edge of manganese, 6547 KeV ($\lambda=1.89$ Å). All X-ray data were processed and reduced using DENZO/SCALEPACK package (Table I) (Minor and Otwinowski, 1997).

Structure determination and refinement

The initial maps were obtained after a rigid body fitting of the coordinates of 3D protein (PDB entry 1WNE; Ferrer-Orta *et al*, 2004) to the new unit cells (Table I), using the program CNS (Brunger *et al*, 1998). In both structures, these initial $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ difference maps showed an elongated extra density covering part of the RNA binding groove that was interpreted by the presence of the VPg primer protein occupying the cavity. The density was clear to interpret with confidence the first 15 amino acids of VPg (Figure 1A), but the carboxy-terminal part of the oligopeptide was disordered and could not be interpreted.

In the structure of the complex obtained in the presence of UTP, A10, $MnCl_2$ and $MgCl_2$, the difference density maps showed the product of the VPg uridylylation (VPg–UMP), plus two strong peaks of density that would be interpreted by the presence of two metal ions in the polymerase catalytic site (Figures 1A and B). Data collected at the manganese K edge were used to detect one Mn^{2+} ion, via anomalous difference Fourier maps (Figure 1B). The second strong peak seen in conventional difference Fourier maps was then interpreted by the presence of one additional Mg^{2+} ion. No ordered density for the pyrophosphate product was seen in the vicinity of the polymerase active site. The A10 oligonucleotide added as a template in the uridylylation reaction was also disordered in the electron density maps. Several cycles of automatic refinement, performed with program CNS, were alternated with manual model rebuilding using the program TURBO (Roussel and Cambillau, 1991). The statistics of the refinement in both complexes are summarized in Table I. Coordinates have been deposited at the PDB with accession codes 2D7S and 2F8E for the 3D–VPg and 3D–VPg–UMP, respectively.

Directed mutagenesis and expression and purification of mutant polymerases

Mutant polymerases with substitutions Glu166-Ala, Glu166-Arg, Arg168-Ala, Arg179-Ala, Lys387-Ala/Arg388-Ala or Thr407-Ala/Ile411-Ala were obtained using expression and purification protocols previously described for the wild-type 3D (Arias *et al*, 2005). Different vectors encoding for all mutant polymerases have been constructed by site-directed mutagenesis by using the QuickChange Site-Directed Mutagenesis Kit (Stratagene) as indicated by the manufacturer. Briefly, for each construction, a different pair of primers (sense and antisense) coding for an amino-acid substitution (Glu166-Ala, Glu166-Arg, Arg168-Ala, Arg179-Ala) or two-amino-acid substitutions (Lys387-Ala/Arg388-Ala, Thr407-Ala/Ile411-Ala)

was used. The vector that contains the wild-type 3D-coding region (pET-28a3Dpol) was amplified with the appropriate pair of primers to introduce the desired mutation. The PCR product was digested with *Dpn*I and used to transform XL1-Blue Supercompetent Cells (Stratagene). Positive colonies were isolated, plasmid DNA was extracted and transfected into *Escherichia coli* BL21 cells. The DNA was extracted again, and the entire 3D-coding region was sequenced to confirm the presence of the expected mutation(s). The cells were grown, IPTG-induced, collected, lysed, and the mutant 3Ds purified as described previously (Ferrer-Orta *et al*, 2004; Arias *et al*, 2005). All mutant polymerases were >95% pure by SDS–PAGE analysis.

Mutant VPgs synthesis and VPg uridylylation assay

All mutant peptides were prepared by solid-phase peptide synthesis, purified by G25 Sephadex chromatography and HPLC, and analyzed by mass spectrometry as described previously (Arias *et al*, 2005). The peptides prepared are the FMDV VPg1 (GPYAGPLERQRPLKV RAKLPRQE), VPg2 (GPYAGPMERQKPLKVKARAPVVKE), VPg3 (GPYAGPVKKPVALKVKAKNLIVTE) and the VPg1 mutants Tyr3-Phe, Pro6-Ala, Glu8-Ala, Arg9-Ala and Arg9-Glu.

The VPg uridylylation assay was performed as described in Arias *et al* (2005). The uridylylation mixture contained 30 mM MOPS (pH 7.0), 33 mM NaCl, 0.6 mM $MnCl_2$, 40 ng/μl poly(A) (300 residues on average), 150 μM VPg, 0.4 mg/ml BSA (Boehringer), 8% glycerol and 50 μM [α-^{32}P]UTP (0.01 mCi/ml; 200 mCi/mmol and 0.4–0.8 μM 3D; 46 μl of the mixture were prewarmed at 37°C and the reaction was started by adding 4 μl of 3D). The standard reaction was carried out for 30 min at 37°C, and stopped by adding 10 μl of 500 mM EDTA; the product was analyzed by SDS–PAGE in a Tris–Tricine buffer (16% (w/v) acrylamide, 10% (v/v) glycerol).

Sequence alignments

Amino-acid sequences of picornaviral 3D and VPg proteins were retrieved from GeneBank, aligned using the program ClustalW (Higgins *et al*, 1996) and rendered with ESPript (Gouet *et al*, 1999).

Acknowledgements

We are indebted to I Fita for critically reading the manuscript, and work in Barcelona was supported by Grants BIO2002-00517 and BFU2005-02376/BMC. Work in Madrid by Grants BMC 2001.1823.C02-01, BFU2005-00863/BMC and Fundación R Areces. CF and AA were supported by I3P fellowships from Ministerio de Educación y Ciencia. RA was supported by an FPI fellowship from Comunidad de Madrid. Data were collected at the EMBL protein crystallography beam lines ID14.2, ID14.4 and BM16 at ESRF (Grenoble) within a Block Allocation Group (BAG Barcelona). The financial support was provided by the ESRF.

References

Alberts B, Johnson J, Lewis J, Raff M, Roberts K, Walter P (2002) *Molecular Biology of the Cell*. New York: Garland Science

Appleby TC, Luecke H, Shim JH, Wu JZ, Cheney IW, Zhong W, Vogeley L, Hong Z, Yao N (2005) Crystal structure of complete rhinovirus RNA polymerase suggests front loading of protein primer. *J Virol* **79:** 277–288

Arias A, Agudo R, Ferrer-Orta C, Perez-Luque R, Airaksinen A, Brocchi E, Domingo E, Verdaguer N, Escarmis C (2005) Mutant viral polymerase in the transition of virus to error catastrophe identifies a critical site for RNA binding. *J Mol Biol* **353:** 1021–1032

Boerner JE, Lyle JM, Daijogo S, Semler BL, Schultz SC, Kirkegaard K, Richards OC (2005) Allosteric effects of ligands and mutations on poliovirus RNA-dependent RNA polymerase. *J Virol* **79:** 7803–7811

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D* **54:** 905–921

Butcher SJ, Grimes JM, Makeyev EV, Bamford DH, Stuart DI (2001) A mechanism for initiating RNA-dependent RNA polymerization. *Nature* **410:** 235–240

Falk MM, Sobrino F, Beck E (1992) VPg gene amplification correlates with infective particle formation in foot-and-mouth disease virus. *J Virol* **66:** 2251–2260

Ferrer-Orta C, Arias A, Escarmís C, Verdaguer N (2006) A comparison of viral RNA-dependent RNA polymerases. *Curr Opin Struct Biol* **16** (in press)

Ferrer-Orta C, Arias A, Perez-Luque R, Escarmís C, Domingo E, Verdaguer N (2004) Structure of foot-and-mouth disease virus RNA-dependent RNA polymerase and its complex with a template–primer RNA. *J Biol Chem* **279:** 47212–47221

Flint SJ, Enquist LW, Racaniello VR, Skalka AM (2004) *Principles of Virology. Molecular Biology, Pathogenesis, and Control of Animal Viruses*. Washington, DC: ASM Press

Forss S, Shaller H (1982) A tandem repeat gene in a picornavirus. *Nucleic Acid Res* **10:** 6441–6450

Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: multiple sequence alignments in PostScript. *Bioinformatics* **15:** 305–308

Higgins DG, Thompson JD, Gibson TJ (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* **266:** 383–402

Huang H, Chopra R, Verdine GL, Harrison SC (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282:** 1669–1675

Love RA, Maegley KA, Yu X, Ferre RA, Lingardo LK, Diehl W, Parge HE, Dragovich PS, Fuhrman SA (2004) The crystal structure of the RNA-dependent RNA polymerase from human rhinovirus: a dual function target for common cold antiviral therapy. *Structure* **12:** 1533–1544

Lyle JM, Clewell A, Richmond K, Richards OC, Hope DA, Schultz SC, Kirkegaard K (2002) Similar structural basis for membrane localization and protein priming by an RNA-dependent RNA polymerase. *J Biol Chem* **277:** 16324–16331

Minor W, Otwinowski Z (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276:** 307–326

Nayak A, Goodfellow IG, Belsham GJ (2005) Factors required for the uridylylation of the foot-and-mouth disease virus 3B1, 3B2, and 3B3 peptides by the RNA-dependent RNA polymerase ($3D^{pol}$) *in vitro*. *J Virol* **79:** 7698–7706

Paul AV (2002) Possible unifying mechanism of picornavirus genome replication. In *Molecular Biology of Picornaviruses*, Semler BL, Wimmer E (eds), pp 227–246. Washington, DC: ASM Press

Roussel A, Cambillau C (1991) Turbo-Frodo. In *Silicon Graphics Geometry Partners Directory*, Graphic S (ed), p 86. Mountain View, CA: Silicon Graphics

Salas M (1991) Protein–priming of DNA replication. *Annu Rev Biochem* **60:** 39–71

Steitz TA (1998) A mechanism for all polymerases. *Nature* **391:** 231–232

Thompson AA, Peersen OB (2004) Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO J* **23:** 3462–3471

van Dijk AA, Makeyev EV, Bamford DH (2004) Initiation of viral RNA-dependent RNA polymerization. *J Gen Virol* **85:** 1077–1093

Xiang W, Cuconati A, Hope D, Kirkegaard K, Wimmer E (1998) Complete protein linkage map of poliovirus P3 proteins: interaction of polymerase 3Dpol with VPg and with genetic variants of 3AB. *J Virol* **72:** 6732–6741

