Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Structure of a backtracked state reveals conformational changes similar to the state following nucleotide incorporation in human norovirus polymerase

Dmitry Zamyatkin,[a] Chandni Rao,[a] Elesha Hoffarth,[a] Gabriela Jurca,[a] Hayeong Rho,[a] Francisco Parra,[b] Pawel Grochulski[c] and Kenneth Kai-Sing Ng[a]*

[a]Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada, [b]Departamento de Bioquimica y Biologia Molecular, Universidad de Oviedo, Instituto Universitario de Biotecnologia de Asturias, 33006 Oviedo, Spain, and [c]Canadian Light Source, 44 Innovation Boulevard, Saskatoon, Saskatchewan S7N 2V3, Canada

Correspondence e-mail: ngk@ucalgary.ca

Received 26 June 2014
Accepted 29 September 2014

PDB reference: human norovirus polymerase replication complex, 4qpx

The RNA-dependent RNA polymerase (RdRP) from norovirus (NV) genogroup II has previously been crystallized as an apoenzyme (APO1) in multiple crystal forms, as well as as a pre-incorporation ternary complex (PRE1) bound to Mn^{2+}, various nucleoside triphosphates and an RNA primer-template duplex in an orthorhombic crystal form. When crystallized under near-identical conditions with a slightly different RNA primer/template duplex, however, the enzyme–RNA complex forms tetragonal crystals (anisotropic data, $d_{\min} \simeq 1.9$ Å) containing a complex with the primer/template bound in a backtracked state (BACK1) similar to a post-incorporation complex (POST1) in a step of the enzymatic cycle immediately following nucleotidyl transfer. The BACK1 conformation shows that the terminal nucleotide of the primer binds in a manner similar to the nucleoside triphosphate seen in the PRE1 complex, even though the terminal two phosphoryl groups in the triphosphate moiety are absent and a covalent bond is present between the α-phosphoryl group of the terminal nucleotide and the 3′-oxygen of the penultimate nucleotide residue. The two manganese ions bound at the active site coordinate to conserved Asp residues and the bridging phosphoryl group of the terminal nucleotide. Surprisingly, the conformation of the thumb domain in BACK1 resembles the open APO1 state more than the closed conformation seen in PRE1. The BACK1 complex thus reveals a hybrid state in which the active site is closed while the thumb domain is open. Comparison of the APO1, PRE1 and BACK1 structures of NV polymerase helps to reveal a more complete and complex pathway of conformational changes within a single RdRP enzyme system. These conformational changes lend insight into the mechanism of RNA translocation following nucleotidyl transfer and suggest novel approaches for the development of antiviral inhibitors.

1. Introduction

Human noroviruses (NVs) belong to the *Caliciviridae* family and are a major cause of outbreaks of gastroenteritis throughout the world (Blanton *et al.*, 2006). More generally, NVs and the *Caliciviridae* belong to the positive-strand RNA virus superfamily, the members of which all possess a small genome consisting of one or a small number of single-stranded RNA segments which have the same sense as messenger RNA. In all of the members of the positive-strand RNA virus superfamily, a virally encoded RNA-dependent RNA polymerase (RdRP) is the central enzyme required for replication (Ortín & Parra, 2006). Because of the central importance of this enzyme to the viral life cycle, this enzyme has become one of the key targets for the development of novel antiviral

© 2014 International Union of Crystallography

agents for hepatitis C (Gentile *et al.*, 2014), dengue (Caillet-Saguy *et al.*, 2014), hand, foot and mouth disease (Tan *et al.*, 2014) and norovirus gastroenteritis (Eltahla *et al.*, 2014).

RdRPs, like other nucleic acid polymerases, are believed to adopt several distinct conformations as they pass through the multiple binding and catalysis steps required to complete a single enzymatic cycle (Ng *et al.*, 2008; Doublié *et al.*, 1999; Rothwell & Waksman, 2005; Showalter *et al.*, 2006; Beard & Wilson, 2006; Ferrer-Orta *et al.*, 2006; Steitz, 2006; Gong & Peersen, 2010; Sholders & Peersen, 2014). In addition to crystallographic studies, kinetics, site-directed mutagenesis, NMR and molecular-dynamics simulations have helped to define some of the conformational changes that may be important for different steps of catalysis in RdRPs, especially the picornaviral polymerases (Castro *et al.*, 2009; Cameron *et al.*, 2009; Moustafa *et al.*, 2011; Yang *et al.*, 2012; Shen, Sun *et al.*, 2012; Verdaguer & Ferrer-Orta, 2012; Garriga *et al.*, 2013; Sholders & Peersen, 2014). Different conformations have been observed in RdRPs from various viruses, but the determination of high-resolution structures of a complete series of conformational changes within a single enzyme has been elusive. The most complete sets of structures from single-enzyme RdRP systems are from *Foot-and-mouth disease virus* (FMDV; Ferrer-Orta *et al.*, 2004, 2007) and *Poliovirus* (PV) polymerase (Gong & Peersen, 2010). However, in both of these systems most of the structures were only determined to modest resolution ($d_{\rm min}$ = 2.25–3.0 Å) and sometimes with highly anisotropic data. Also, it has not yet been possible to determine the structure of a closed complex with nucleoside triphosphate bound in either of these systems, although structures have been determined for open complexes with RNA bound in both FMDV and PV, and the structure of a closed post-incorporation complex has also been determined in PV. As a result of these limitations, conformational changes between different states and the closed ternary state with nucleoside triphosphate bound requires comparisons between the picornaviral polymerase structures and higher resolution ($d_{\rm min}$ = 1.7 Å) structures of closed ternary complexes formed by NV polymerase (Zamyatkin *et al.*, 2008, 2009). Unfortunately, the significant level of sequence difference (∼20% sequence identity overall, with very poor similarity in the thumb domain) and hence structural differences between the picornaviral and NV enzymes compromises the strength of the conclusions that can be drawn by comparing differences in conformation between the picornaviral and NV enzymes. Structural differences between the states seen in picornaviral enzymes and NV polymerase may be owing to sequence differences between evolutionarily divergent enzymes as opposed to conformational changes intrinsic to different states in the enzymatic cycle of each enzyme. To better distinguish between these two possibilities, it is desirable to compare different conformational states within a single-enzyme system.

To increase our understanding of the role of conformational changes in the structure and function of viral RdRPs, we have pursued structural studies of NV polymerase in complex with various RNA sequences and nucleoside triphosphates. Conformational changes in the norovirus polymerase have previously been observed when structures of the apoenzyme (APO1; Ng *et al.*, 2004; Högbom *et al.*, 2009) were compared with structures of ternary complexes crystallized in a pre-incorporation state with divalent metal ions, a primer-template duplex RNA and various nucleoside triphosphates (PRE1; Zamyatkin *et al.*, 2008, 2009). At least three key changes in conformation were seen: (i) closure of the active-site region to allow Asp side chains from sequence motif C, as well as an Asp side chain and a main-chain carbonyl group from motif A, to form a tight coordination complex with the two divalent metal ions that also coordinate to the 3′-hydroxyl group nucleophile of the primer and the electrophilic P atom in the α-phosphate group of the nucleoside triphosphate, (ii) rotation of the central helix (the second helix or helix II) in the thumb domain (residues 435–449) by 22° to help form a binding pocket for the primer RNA strand and (iii) displacement of the C-terminal tail region away from the central active-site groove, which also allows the rotation of helix II of the thumb domain.

During the course of studies to crystallize ternary complexes of NV polymerase with different RNA primer/template duplexes and nucleoside triphosphates, we fortuitously crystallized NV polymerase in complex with two manganese ions and a primer/template duplex but no nucleoside triphosphate in the active site. Surprisingly, the

Table 1
Crystallographic statistics.

Values in parentheses are for the outermost resolution shell.

Data collection	
Space group	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 81.2$, $c = 188.1$
Wavelength (Å)	0.97934
Resolution (Å)	40–1.86 (1.91–1.86)
Total reflections	683293 (17334)
Unique reflections	50992 (3852)
Completeness (%)	92.3 (72.0)
$\langle I/\sigma(I)\rangle$	15.6 (1.8)
$R_{\rm meas}$†	0.098 (1.98)
Refinement	
$R_{\rm work}$‡	0.205 (0.325)
$R_{\rm free}$§	0.253 (0.354)
No. of atoms	
Protein	3884
RNA	317
Solvent and ions	340
R.m.s. deviations from ideal geometry¶	
Bond lengths (Å)	0.008
Bond angles (°)	1.12
Average temperature factors (Å^2)	
Wilson plot	35.1
Protein	34.8
RNA	39.1
Solvent	41.5
Ramachandran angles†† (%)	
Favored	94.1
Outliers	0.2

† $R_{\rm meas} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith integrated intensity of a given reflection and $\langle I(hkl)\rangle$ is the weighted mean of all measurements of $I(hkl)$ (Diederichs & Karplus, 1997). ‡ $R_{\rm work} = \sum_{hkl}||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}|$ for the 95% of reflection data used in refinement. § $R_{\rm free} = \sum_{hkl}||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl}|F_{\rm obs}|$ for the 5% of reflection data excluded from refinement. ¶ Root-mean-square deviations from ideal geometry calculated by *REFMAC* (Murshudov *et al.*, 2011). †† Ramachandran plot analysis carried out using *MolProbity* (Chen *et al.*, 2010).

terminal nucleotide at the 3′-end of the primer strand occupies the location occupied by the nucleoside triphosphate in the pre-incorporation complex. As a result, the RNA duplex is bound in the manner expected for a product after nucleotide has been added to the primer. This structure thus reveals the conformation of the enzyme in a post-incorporation state and helps to provide insight into the nature of the conformational changes in NV polymerase at different steps of the enzymatic cycle. This complex could also be considered a backtracked complex in which the productive 'ER_n' complex [state (i) in Fig. 1] is returning back to a previous, post-incorporation-like state [state (vii) in Fig. 1] as a result of a 'backwards' movement of the primer/template duplex after translocation of the primer/template duplex had occurred.

2. Materials and methods

2.1. Protein expression and purification

NV polymerase was expressed and purified as previously described (Ng *et al.*, 2004; Zamyatkin *et al.*, 2008), except for the following modification. After the clarified cell extract had been loaded onto 2 ml glutathione Sepharose 4B columns (GE Healthcare) and washed, an additional wash step with 25 ml 1 *M* sodium/potassium phosphate was added to help to remove negatively charged contaminants such as nucleotides or nucleic acids.

2.2. RNA oligonucleotide synthesis and preparation

The self-complementary RNA oligonucleotide (5′-UAC-CCGGG-3′) was synthesized by the University Core DNA Services at the University of Calgary and was purified by gel filtration. After lyophilization, the RNA was dissolved in filtered, deionized water to a concentration of 10 m*M*. The RNA was heated to 85°C in a thermal cycler and gradually cooled to 5°C over 15 min using a step-gradient program that lowered the temperature at a rate of 5°C min^{-1}. The annealed RNA was kept on ice prior to crystallization with NV polymerase.

2.3. Crystallization

CHAPS {3[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonate; 0.2%(*w*/*v*)} and Superase-In ribonuclease inhibitor (1 U μl^{-1}; Ambion) were added to NV polymerase prior to crystallization. The annealed RNA duplex (18 µl) was mixed with 4.5 µl UTP (25 m*M*) and 31.5 µl water to prepare a mixture containing 1.2 m*M* RNA duplex and 2.1 m*M* UTP. A reservoir solution consisting of 19% PEG 8000, 25% glycerol, 100 m*M* Tris–HCl pH 7.0, 50 m*M* KCl, 10 m*M* $MnCl_2$, 0.1% β-mercaptoethanol was prepared. The reservoir solution (3 µl) was mixed with the RNA–UTP mixture (6 µl) and NV polymerase (4 µl). This mixture was suspended as a hanging drop and equilibrated against 1 ml reservoir solution at 22°C. Crystals appeared within a week and grew to full size (∼0.2 × 0.1 × 0.05 mm) after two to three weeks.

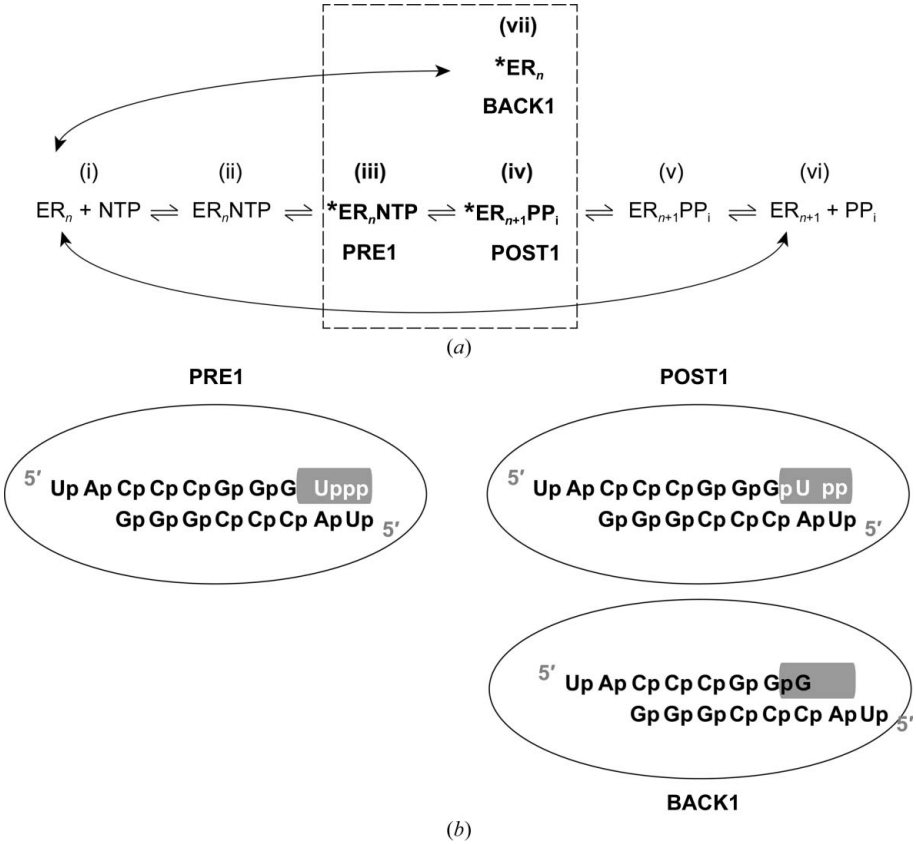

Figure 1
(*a*) General kinetic scheme for nucleotide incorporation in RdRPs and other polymerases (Kuchta *et al.*, 1987; Arnold & Cameron, 2004). E, RdRP enzyme; R_n, RNA oligonucleotide containing *n* residues; R_{n+1}, RNA oligonucleotide containing *n* + 1 residues; NTP, nucleoside triphosphate; PP_i, pyrophosphate. The dashed box indicates the three complexes in the kinetic scheme where a 'closed' conformation with full coordination of both divalent metal cations is seen in the active site. (*b*) Schematic representation of the PRE1, POST1 and BACK1 complexes with the polymerase drawn as an oval and the binding site for UTP drawn in gray.

2.4. Crystal structure determination

A single crystal (0.2 × 0.1 × 0.05 mm) was quickly mounted in a polymer fiber loop (Hampton Research) and flash-cooled in a nitrogen-gas stream at ∼100 K. The crystal was stored in liquid nitrogen and transferred to Canadian Light Source beamline 08-ID-1 for data collection using a MAR Mosaic CCD 225 detector. Data were processed and scaled using *XDS* (Kabsch, 2010). (Data-quality and refinement statistics are given in Table 1.) The diffraction pattern showed some anisotropy. As a result, data extending to 1.86 Å resolution were included for refinement, even though the data statistics are marginal for

the 1.91–1.86 Å resolution range. The structure of NV polymerase from the ternary complex (PDB entry 3bso; Zamyatkin *et al.*, 2008) was used as a search model for molecular-replacement calculations using *Phaser* (Read, 2001). Very clear solutions were obtained for the rotational (Z = 20.5) and translational (Z = 45.2) placement of a single polymerase molecule, yielding a solvent content of 55% and a V_M of 2.8 Å^3 Da^{-1}. *Coot* (Emsley & Cowtan, 2004) was used for model building and *REFMAC* (Murshudov *et al.*, 2011) was used for refinement. Initial electron-density maps clearly indicated the presence of an RNA duplex and metal ions, but NTPs were not present. The structure of the RNA was built using the structure from the ternary complex as a starting point. The final model includes residues 4–466 and 474–505. Because the electron density corresponding to the loop between residues 467 and 473 was weak and not clearly interpretable, it was inferred that this portion of the protein may be dynamically disordered, and a model for this segment of the protein was not included in the final refined model. The quality of the geometric parameters in the model was evaluated by *MolProbity* (Chen *et al.*, 2010). Coordinates were superimposed using *LSQMAN* (Kleywegt, 1996) and *LSQKAB* (Kabsch, 1976). Difference distance matrix plots were generated using *DDMP* (P. J. Fleming, unpublished work; Richards & Kundrot, 1988).

3. Results

3.1. The binding mode of the RNA primer-template duplex resembles a post-incorporation complex

The complex of NV RdRP bound to manganese ions and an RNA primer-template duplex was crystallized in a novel tetragonal crystal form that reveals a conformation with the 3′-terminal nucleotide of the primer strand occupying a position nearly identical to that seen for the base and ribose of the nucleoside triphosphates in the closed ternary pre-incorporation complexes (PRE1) that were previously crystallized in an orthorhombic crystal form (Zamyatkin *et al.*, 2008, 2009; Fig. 1). The position of this terminal nucleotide clearly indicates that the primer/template duplex is binding in a manner distinct from that seen for the normal productive state, where the 3′-end of the primer is located adjacent to the binding pocket for the nucleoside triphosphate. The binding mode captured in this structure resembles both a post-incorporation/pre-translocation 'product' complex (Fig. 1) as well as a 'backtracked' complex that appears to be important for proofreading and pausing in multisubunit polymerases (Weixlbaumer *et al.*, 2013; Cheung & Cramer, 2011; Wang *et al.*, 2009). This structure could be considered as an inhibited state that closely resembles a post-incorporation product state (POST1) except for the absence of the pyrophosphate group. In this paper, this state will be referred to as BACK1 to indicate that it is probably most similar to a backtracked state, although it also has many similarities to the POST1 state.

The coordination spheres of the two manganese ions at the active site are clearly defined by electron density, and the coordination geometry is very similar to that seen in the PRE1 complex, with the obvious exception that the terminal residue of the primer clearly lacks the β and γ phosphate groups seen in the bound nucleoside triphosphate (Fig. 2). Because the α-P atom of the 3′-terminal residue in the primer strand of the BACK1 complex is covalently bonded to the 3′-O atom of the preceding nucleotide residue in the primer strand, the distance between this P atom and the 3′-O atom in the preceding nucleotide is 1.2 Å

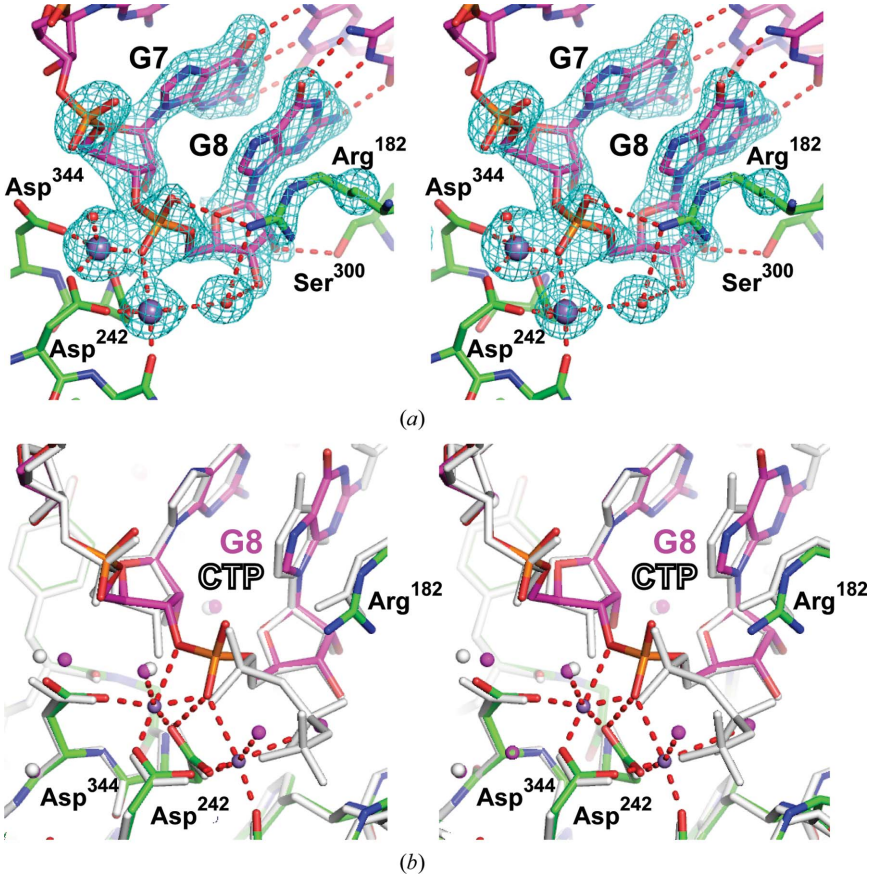

Figure 2
Stereoscopic views of the active site. (*a*) $|F_o| - |F_c|$ OMIT map (contour level 2.6σ) calculated after the terminal two nucleotides of the primer strand, the two Mn^{2+} ions and the side chain of Arg182 were removed from the model and 30 cycles of refinement were performed using *REFMAC*. Coordination bonds (red dashes) with Mn^{2+} ions *A* and *B* (bluish-gray spheres) and hydrogen bonds (red dashes) between the RNA primer (C atoms colored magenta), key water molecules (red spheres) and the protein (C atoms colored green) are shown. O atoms are colored red, N atoms blue and P atoms orange. (*b*) Superposition of the PRE1 (PDB entry 3bso, white spheres and sticks) and BACK1 [atoms and coordination bonds colored as in (*a*)] complexes. Molecular-graphics figures were prepared using *PyMOL* (DeLano, 2002).

shorter than the corresponding distance between the α-P atom of the NTP substrate and the 3′-OH group at the 3′-terminus of the primer in the PRE1 complex. However, the O atoms of the α-phosphate group of the nucleoside triphosphate of the PRE1 complex are located less than 0.5 Å away from the positions occupied by the O atoms belonging to the terminal bridging phosphate group at the 3′-end of the primer of the BACK1 complex (Fig. 2*b*). As a result, the O atoms from the corresponding phosphate groups form very similar coordination bonds with the two metal ions, being 1.9 Å from metal ion *A* and 2.6 Å from metal ion *B*. Three side-chain carboxylate O atoms (from Asp residues 242, 343 and 344), the 3′-O atom in the ribose moiety of the penultimate nucleotide of the primer RNA strand and a single water molecule complete the coordination sphere around metal ion 1, whereas two side-chain carboxylate O atoms (from Asp242 and Asp343), the main-chain carbonyl O atom from Tyr243 and two water molecules complete the coordination sphere around metal ion 2. All of these groups are clearly defined in the electron-density maps for the PRE1 (Zamyatkin *et al.*, 2008) and BACK1 (Fig. 2*b*) complexes.

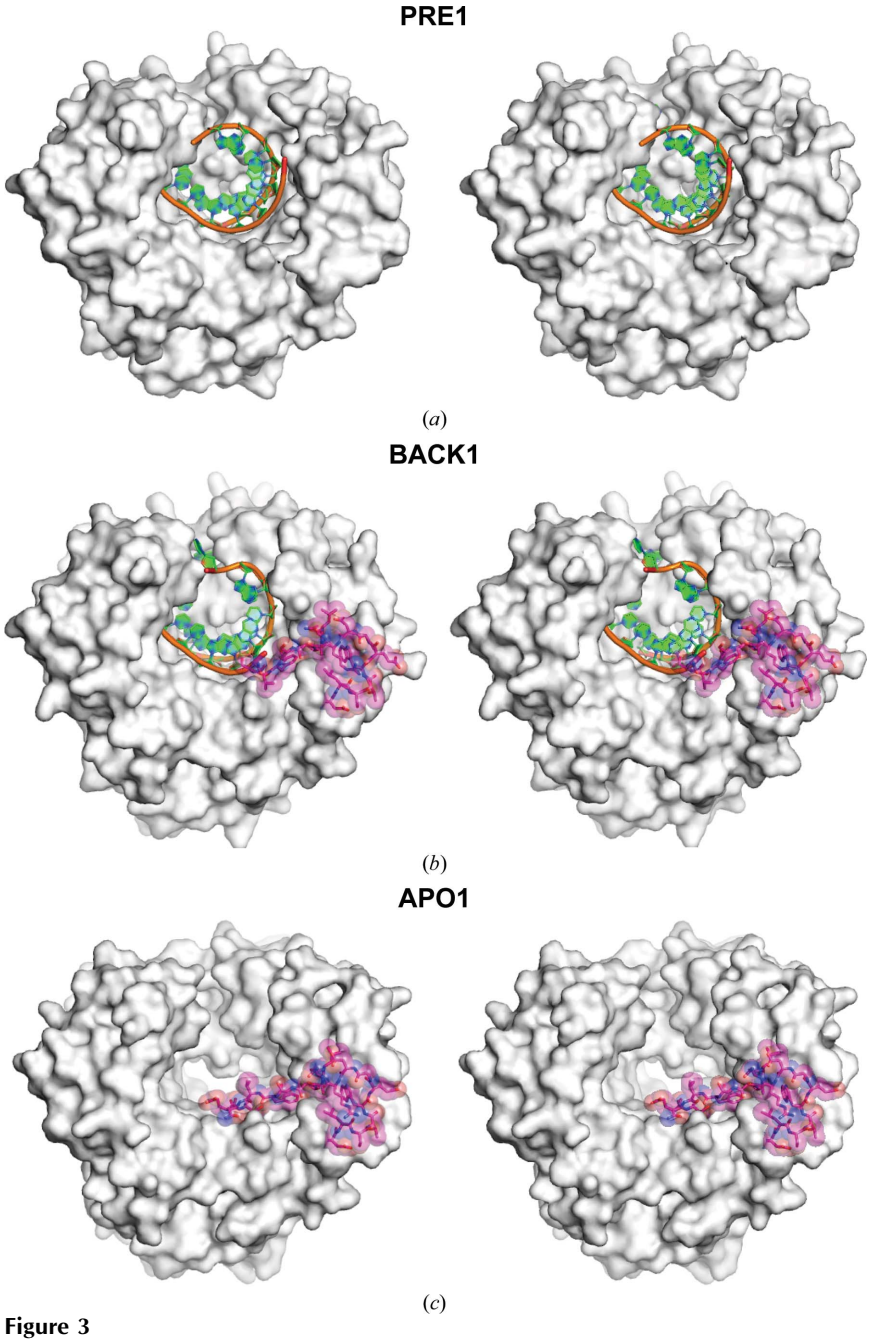

Figure 3
Stereoscopic views of the overall structure, highlighting the conformation of the C-terminal tail. Residues 5–488 are shown as a solvent-accessible surface and the primer-template RNA duplex is shown in cartoon representation. (*a*) In the PRE1 complex (PDB entry 3bso), the C-terminal tail comprising residues 489–516 is disordered and is not defined by electron density. (*b*) In the BACK1 complex, residues 489–505 (magenta) are well ordered and are drawn as sticks and as a semi-transparent space-filling representation (magenta). (*c*) In the APO1 structure (PDB entry 1sh0), residues 489–507 are well ordered and are defined by electron density: this region is drawn as in (*b*).

It is notable that the arrangement of active-site residues seen in the BACK1 complex closely resembles the structures of a post-incorporation state (POST1 or 'state 4') generated by incubating CTP and either Mg^{2+} or Mn^{2+} with crystals of a PV polymerase complex formed with primer-template duplex RNA (Gong & Peersen, 2010). These structures revealed a conformation at the active site that resembled the PRE1 state in NV pol, but unlike in the NV pol complexes the bond between the α and β phosphoryl groups of CTP was now hydrolyzed and the bond between the α-phosphoryl group of CTP and the 3′-OH group of the terminal nucleotide of the primer had now formed. Because the PRE1 state of PV polymerase has yet to be crystallized, an understanding of the conformational changes occurring between the PRE1 and POST1 states in a single RdRP system has yet to be achieved. Assuming that the structure that we are now describing of the BACK1 state in NV pol is a good mimic of the true BACK1 state, our analysis reveals for the first time some surprising conformational changes associated with nucleotidyl transfer in NV pol.

3.2. Structure of the C-terminal tail

As seen previously in the PRE1 complex, RNA binding displaces the C-terminal tail from the active-site cleft of NV RdRP (Zamyatkin *et al.*, 2008). In the BACK1 complex, residues 506–516 are not defined by electron density, suggesting that the tip of the C-terminal tail is unstructured. Compared with the PRE1 structure, in which residues 489–516 are not defined by electron density, however, the BACK1

structure shows that 16 additional residues (489–505) of the C-terminal tail are well ordered and clearly defined by electron density (Fig. 3). The structure of this portion of the tail is very similar to that seen in the APO1 structure up to Ser502. The conformation of residues 503–505 of BACK1 differs from the conformation seen in APO1, because in APO1 these residues lie in the RNA-binding cleft of the enzyme and would sterically clash with the position occupied by the 3′-end of the template strand in the BACK1 complex (Fig. 3). As a result, residues 503–505 reach around the 3′-end of the template strand in a manner that may differ if the template strand were longer, since a longer template strand may sterically clash with the structure formed by residues 503–505 in the BACK1 complex seen here with a short 8-mer template strand. This observation also suggests that the length of the template strand extending from the RNA-binding cleft may affect the conformation of the C-terminal tail domain. This is an important inference, because much longer template strands can be expected in the natural state during the replication of an intact viral genomic or subgenomic RNA template strand.

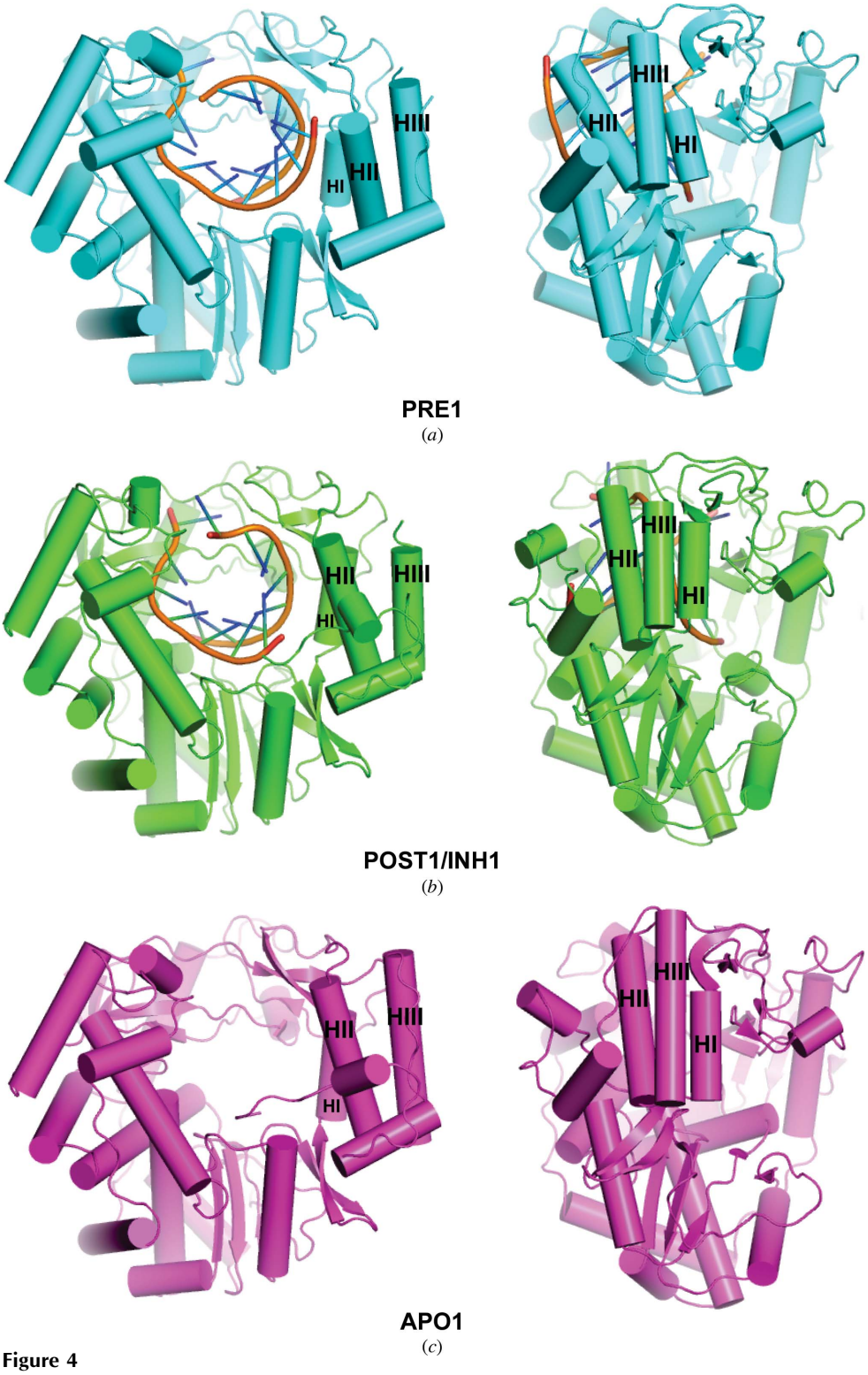

Figure 4
Overall ribbon representations of the (*a*) PRE1 (PDB entry 3bso), (*b*) BACK1 and (*c*) APO1 (PDB entry 1sh0) complexes. The left panels show a 'front' view down the RNA-binding groove and the right panels show 'side' views showing the rotation of helix II towards the 'front' in the PRE1 complex. Helix II in the thumb domain is marked.

3.3. Conformational changes in the thumb domain

Perhaps the most surprising aspect of the BACK1 complex is the conformation of the thumb domain. The conformation of the thumb domain in the BACK1 complex more closely resembles the conformation seen in the APO1 state than that in the PRE1 complex. As mentioned previously, the largest difference seen between APO1 and PRE1 is the orientation of helix II, the centrally located helix in the thumb domain (residues 435–449). In PRE1, helix II appears to rotate by 25° from the conformation seen in APO1 (Figs. 4 and 5). In contrast, when the palm and fingers domains of APO1 and BACK1 are superimposed, the orientation of helix II differs by only 5° in the two structures.

In the PRE1 conformation, the N-terminal end of helix II is tilted towards the 'front' of the RNA-binding cleft, away from the

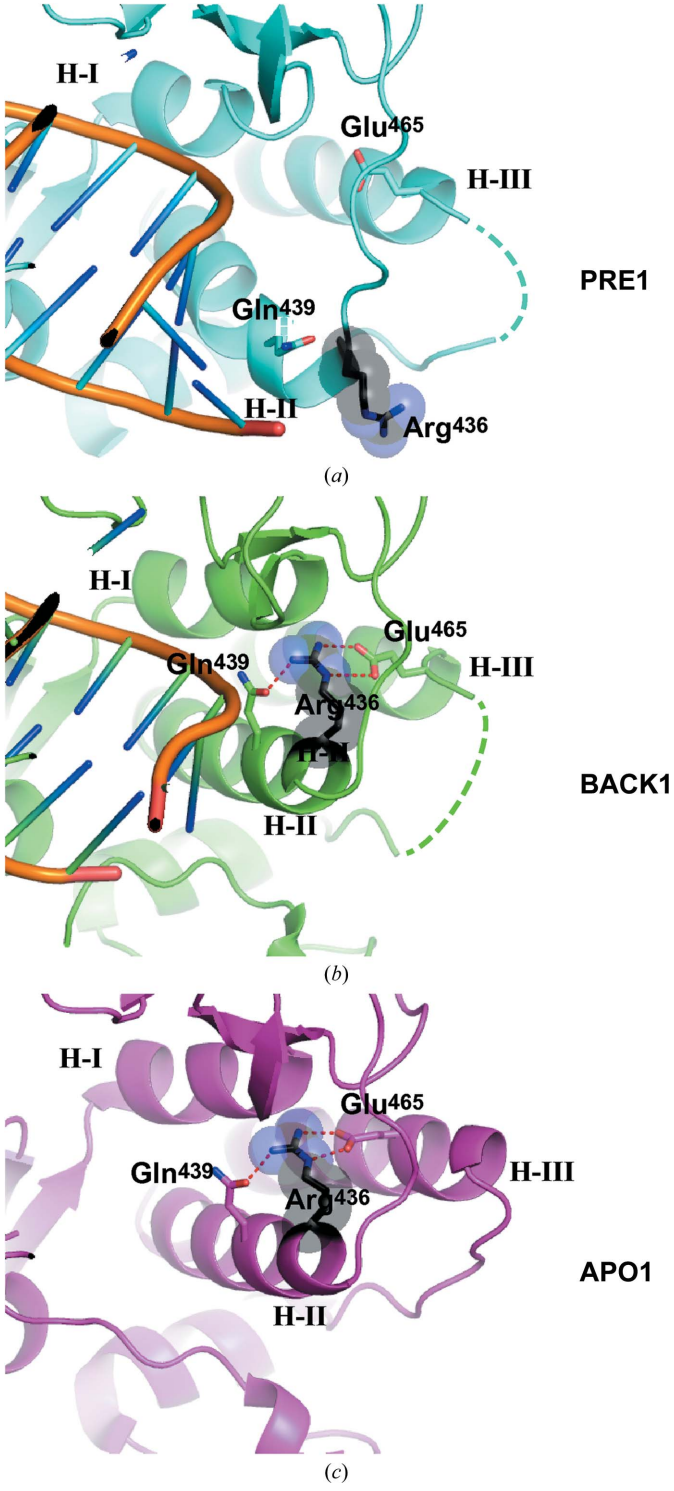

Figure 5
Conformational changes in the interface between helices I, II and III of the thumb domain. Helices are marked HI, HII and HIII. In the PRE1 complex (PDB entry 3bso) (*a*), Arg436 is fully solvent-exposed and helix II is rotated away from helices I and III. In the BACK1 (*b*) and APO1 (PDB entry 1sh0) (*c*) structures, Arg436 lies buried between the central helices of the thumb domain and forms hydrogen bonds (red dashed lines) to Gln439 in helix II and Glu465 in helix III. Hydrogen bonds between the terminal guanidino N atoms of Arg436 and the main-chain carbonyl O atom of Gln414 in helix I are not shown. Residues in the loop connecting HII and HIII that are not well ordered enough to be modeled in the PRE1 and BACK1 structures are indicated by dashed lines.

active site. The result of the 25° rotation relative to APO1 places the main-chain atoms of some residues at the N-terminal end of helix II ∼6–7 Å closer to the 'front' of the molecule. The largest motion is seen in the guanidino group of Arg436, which moves ∼16 Å owing to a lever-arm effect from a centrally buried position in the APO1 and BACK1 states to a surface-exposed position in the PRE1 state (Fig. 5). When it is buried, the guanidino group lies between helices I, II and III in the thumb domain, forming electrostatic interactions with the carboxylate side chain of Glu465 from helix III, as well as hydrogen bonds to the main-chain carbonyl O atom of Gln414 from helix I and the side-chain carbonyl O atom of Gln439 from helix II, which in turn donates a hydrogen bond to the side-chain amide group of Gln414. Because of the centrally buried location of Arg436, this residue appears to be a key trigger residue that possibly relays small changes in conformation among the three central helices in the thumb domain.

It should be noted that Arg436 is highly conserved in all caliciviral polymerases (Fig. 6), and a very similar network of interactions is seen in the crystal structures of other caliciviral polymerases, including the more distantly related lagovirus polymerase from *Rabbit hemorrhagic disease virus* (RHDV; Ng *et al.*, 2002), human norovirus (Ng *et al.*, 2004; Högbom *et al.*, 2009), human sapovirus (Fullerton *et al.*, 2007) and murine norovirus (Lee *et al.*, 2011; Alam *et al.*, 2012). Based on the level of sequence similarity at all of the key positions forming interactions in the calicivirus structures (with the slight exception of Glu465 in the vesiviruses), the same set of interactions and a similar conformational change is likely to occur in all caliciviral polymerases. In the crystal structures of RdRPs from the more distantly related picornaviruses and flaviviruses, however, the interface between helices I, II and III does not contain a residue like Arg436, and conformational changes of the sort seen on comparing the PRE1 NV polymerase structure with the APO1 and BACK1 structures have not yet been seen in any of the picornaviral or flaviviral RdRP structures (Choi, 2012; Gong *et al.*, 2013). If similar conformational changes occur in RdRPs from picornaviruses and flaviviruses, they presumably take place through a different mechanism.

One of the effects of the conformational change from the closed PRE1 structure to the open BACK1 structure involves a slight widening of the active-site cleft and the loss of interactions between the primer RNA strand and the side chain of Gln414 (Fig. 7). As noted above, the buried position of the Arg436 side chain in the 'open' BACK1 state allows the guanidino group to form a hydrogen bond to the side chain of Gln439, which rotates the side chain into position to hydrogen bond to the side chain of Gln414 in helix I. Formation of this hydrogen bond in turn disrupts a bidentate hydrogen-bond network between the side chain of Gln414 and the 2′-OH of a nucleotide in the RNA primer. Because the polymerase binds the RNA duplex weakly and only forms a few specific interactions with the RNA, the disruption of a specific binding interaction in combination with the movement of helix I away from the RNA primer may combine to loosen the grip of the enzyme on the RNA duplex as a means to promote RNA

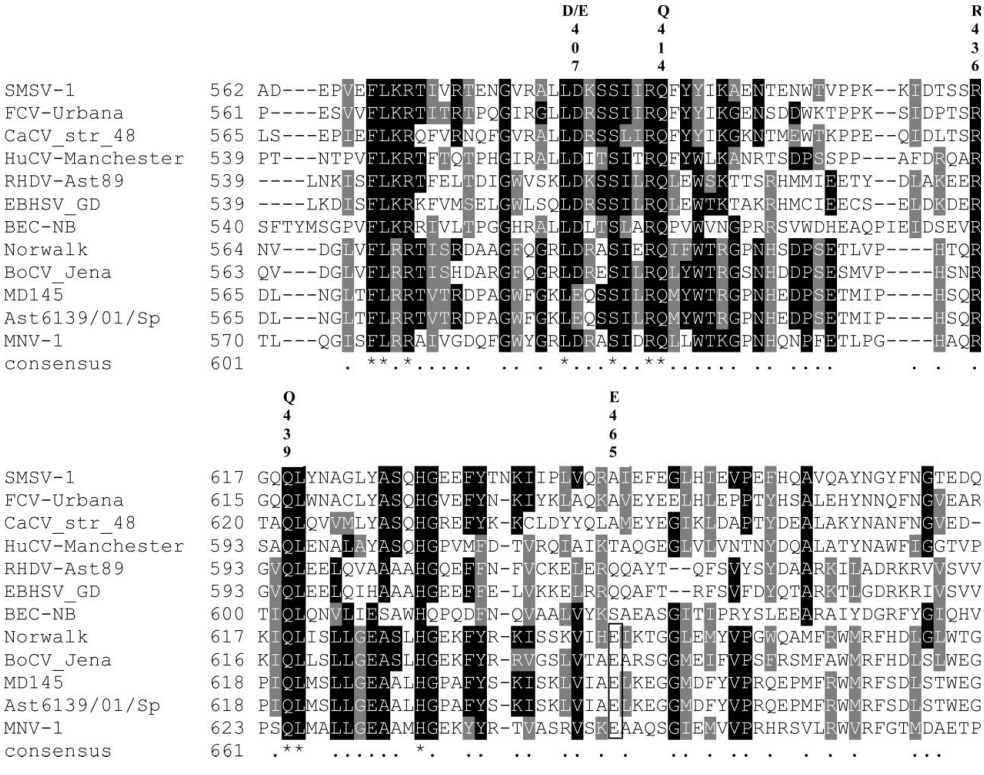

Figure 6
Amino-acid sequence alignment of the nonstructural protein 7 (NS7) polymerase sequences from representative members of the admitted genera (*Norovirus*, *Sapovirus*, *Lagovirus*, *Vesivirus* and *Nebovirus*) within the *Caliciviridae* family. Accession numbers are indicated in parentheses: SMSV-1 (AF181081), FCV-Urbana (L40021), CaCV_str_48 (AB070225), HuCV-Manchester (X86560), RHDV-Ast89 (Z49271), EBHSV_GD (Z69620), BEC-NB (Q8JN60), Norwalk (M87661), BoCV_Jena (AJ011099), MD145 (AY032605), Ast6139/01/Sp (AJ583672) and MNV-1 (AY228235). For clarity, the N- and C-terminal sequences not including relevant residues have been deleted. The numbers heading each sequence line indicate amino-acid residues starting from the N-terminus of the NS6-7 precursor. The positions of NV (Ast6139/01/Sp) RdRP residues (Gln414; Ala436; Gln439 and Glu4654) are indicated.

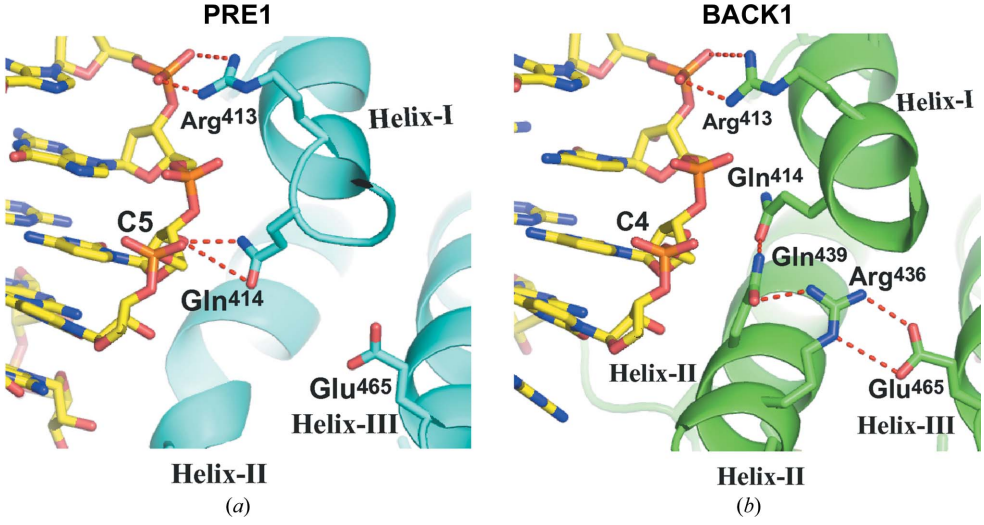

Figure 7
Views of the rearrangement of the hydrogen-bonding network around Gln414. (*a*) In the PRE1 complex (PDB entry 3bso), the side chain of Gln414 forms bidentate hydrogen-bonding interactions with the 2′-OH of C5 in the primer RNA strand. (*b*) In the BACK1 complex, the side chain of Gln414 accepts a hydrogen bond from the side chain of Gln439, which in turn accepts a hydrogen bonds from Arg436. Helices I and II both move away slightly from the RNA primer, which may loosen the grip of the polymerase on the RNA in preparation for the subsequent RNA-translocation step.

translocation. Replacing the highly conserved Gln414 side chain with a shorter side chain incapable of forming hydrogen bonds to either the RNA primer or Gln439 may lend insight into the hypothesized role of this residue in coupling conformational changes in the thumb domain to the process of RNA translocation.

3.4. Coupling of the active-site structure to conformational changes in the thumb domain

The relatively large changes in the conformation of the thumb domain appear to be linked to changes in the active site in BACK1 *versus* PRE1, but the mechanisms underlying the linkage between the structural changes in these two regions are not clear. Superposition of the BACK1 structure onto that of PRE1 reveals only subtle changes in the active site, as described above, and some of these changes appear to be linked to changes near the N-terminus of helix I, which in turn affect helices II and III (Fig. 8). Among the larger changes observed in the region between the active site and helix I are (i) changes in the solvent structure around the divalent metal ions that are owing in part to the absence of the terminal phosphate groups in BACK1, (ii) a change in the ordering and position of the mobile 370s loop, which contains the highly conserved motif D residue Lys374 responsible for general acid catalysis (Castro *et al.*, 2009; Yang *et al.*, 2012) and (iii) a change in the conformation and the position of Glu407, a solvent-exposed residue at the N-terminal end of helix I in the thumb domain (Fig. 8). In homologues of NV polymerase, the residue at position 407 nearly always contains an acidic side chain (Fig. 6). Although it is not clear from the PRE1 and BACK1 structures why

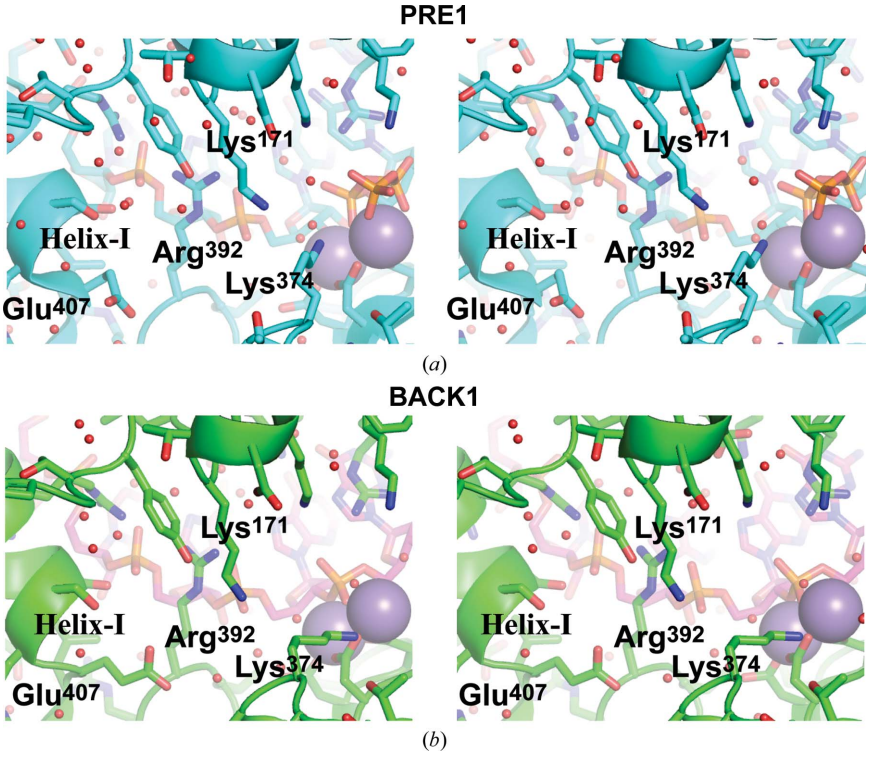

Figure 8
Stereoscopic views of the interfacial region between the active site and helix I of the thumb domain. (*a*) PRE1 (PDB entry 3bso) and (*b*) BACK1 structures showing some of the amino-acid residues, ordered solvent molecules (small red spheres), bound RNA and nucleoside triphosphate, and divalent metal ions (large gray spheres) that may couple changes in conformation at the active site with conformational changes in the thumb domain. Helix I in the thumb domain is also labeled.

an acidic residue at this position is important, the location of this side chain at the interface between the active site and helix I of the thumb domain and the changes in solvent structure around the active site suggest that this residue may help to transmit changes at the active site into changes in the conformation of the thumb domain. Further structural and functional studies of site-directed mutants targeting this residue may help to clarify the role of this residue and others in the interface between the active site and helix I.

Difference distance matrix plots (Richards & Kundrot, 1988) also highlight some of the differences in conformation between the different states of NV polymerase (Fig. 9). A positive difference in distance between any pair of residues indicates that the residues are farther apart in conformation *A versus* conformation *B*. When the APO1 structure (conformation *A*) is compared with either the PRE1 (Fig. 9*a*) or BACK1 (Fig. 9*b*) structures, the largest positive differences are located between the 370s loop near the active site and the rest of the protein, indicating a more open loop conformation in APO1, as discussed previously (Zamyatkin *et al.*, 2008). When comparing APO1 and PRE1 (Fig. 9*a*), the largest negative differences (indicating regions of the protein that are closer together or in a tighter, more closed conformation in APO1) are seen for helices I and II in the thumb domain, also as discussed previously (Zamyatkin *et al.*, 2008) and above (Fig. 4). However, when comparing APO1 and BACK1 (Fig. 9*b*), the differences in the thumb domain are clearly less evident, with the largest changes seen in the slightly more closed conformation adopted by helix II in APO1. When comparing BACK1 with PRE1 (Fig. 9*c*), the key differences are seen in a more closed conformation for helix I and a slightly more open conformation for helix I. As discussed below, the structural differences seen amongst the APO1, PRE1 and BACK1 conformations are likely to facilitate different functions of the enzyme in the various steps of the enzymatic cycle.

4. Discussion

The BACK1 complex reported in this study appears to be the result of NV polymerase binding to a primer-template duplex in the manner expected for the state that occurs after the incorporation of a nucleotide in the normal enzymatic cycle. The key differences lying at the heart of the numerous conformational differences between the PRE1 and BACK1 states are (i) the presence of a covalent bond between the 3′-O atom of the penultimate residue in the primer strand of RNA and the α-phosphate of the terminal residue and (ii) the absence of the two terminal phosphate residues in the terminal nucleotide residue, which occupies a nearly identical position to that occupied by the nucleoside triphosphate in the BACK1 complex. Our structures indicate that these two changes in covalent bonding make only small changes to the active site, which clearly remains in a 'closed' conformation in both PRE1 and BACK1. However, the small perturbations in structure appear to be coupled to and perhaps amplified in some way to changes in the thumb domain. As NV polymerase moves from the closed, pre-incorporation complex (PRE1) to the closed, post-incorporation complex (POST1), the post-incorporation state contains a closed active site but the open thumb conformation seen in BACK1 is an intermediate state bridging the PRE1 state with a closed active site and a closed thumb conformation and a hypothetical POST2 state with an open active site and an open thumb conformation. The similarities in the structure of the thumb domain in the BACK1 and APO1 states suggest that the change in conformation of the thumb domain may facilitate looser interactions with the RNA. The loosening of interactions between protein and RNA following the nucleotidyl-transfer step may lead to either the release or the translocation of the RNA duplex, a step which is clearly required for the enzyme to regenerate an open nucleotide-binding site following the incorporation of a nucleotide into the nascent RNA primer strand.

As the conformational changes described here are dynamic in nature, further insight into the mechanisms underlying

these conformational changes may be expected to come from spectroscopic studies (Yang *et al.*, 2012) and molecular-dynamics simulations (Moustafa *et al.*, 2011; Shen, Moustafa *et al.*, 2012) in addition to crystallography. Studies on mutants trapped in or favoring one conformation over the other may help to provide further insight into this issue. Structural differences between the PRE1 and BACK1 states indicate that mutations affecting the highly conserved interhelical interface involving Arg436, Gln414, Gln439 and Glu465 (Figs. 5 and 6) will affect conformational changes in the thumb domain. The residues at the interface between the base of helix I and the active site, especially Glu407, are also likely to play key roles in coupling small conformational changes in the active site with more dramatic and functionally important conformational changes in the thumb domain.

As a practical application that may arise from a deeper understanding of the mechanisms underlying these conformational changes, we are currently searching for small-molecule compounds to stabilize or trap the enzyme in states in which the active site and/or thumb domain adopt either open or closed conformations. Such compounds are expected to interfere with the conformational changes expected to be essential for normal enzymatic activity in NV polymerase. As a result, such compounds could act as allosteric inhibitors and hence as lead compounds for novel antiviral therapeutics in a manner reminiscent of pharmacologically efficacious inhibitors for other viral polymerases, such as HIV reverse transcriptase (Sarafianos *et al.*, 2009) and *Hepatitis C virus* NS5B (Powdrill *et al.*, 2010). A deeper understanding of the roles of conformational changes in the mechanism of this important enzyme should prove invaluable in discovering novel allosteric inhibitors for the polymerase enzymes from NV and a wide range of related RNA viruses.

We thank Dr Isabelle Barrette-Ng for helpful discussions. X-ray diffraction data were measured on beamline 08ID-1 of the Canadian Light Source (CLS), which is supported by the Natural Sciences and Engineering Research Council of Canada, the National Research Council Canada, the Canadian Institutes of Health Research, the Province of Saskatchewan, Western Economic Diversification Canada and the University of Saskatchewan. This work was supported by NSERC Discovery Grant 262089, CIHR Operating Grant MOP-67209 and an AHFMR Senior Scholar Award to KN. Work at F. Parra's laboratory has been supported by grant BIO2009–07535 from the Spanish Ministerio de Ciencia e Innovación and the European Union FEDER program.

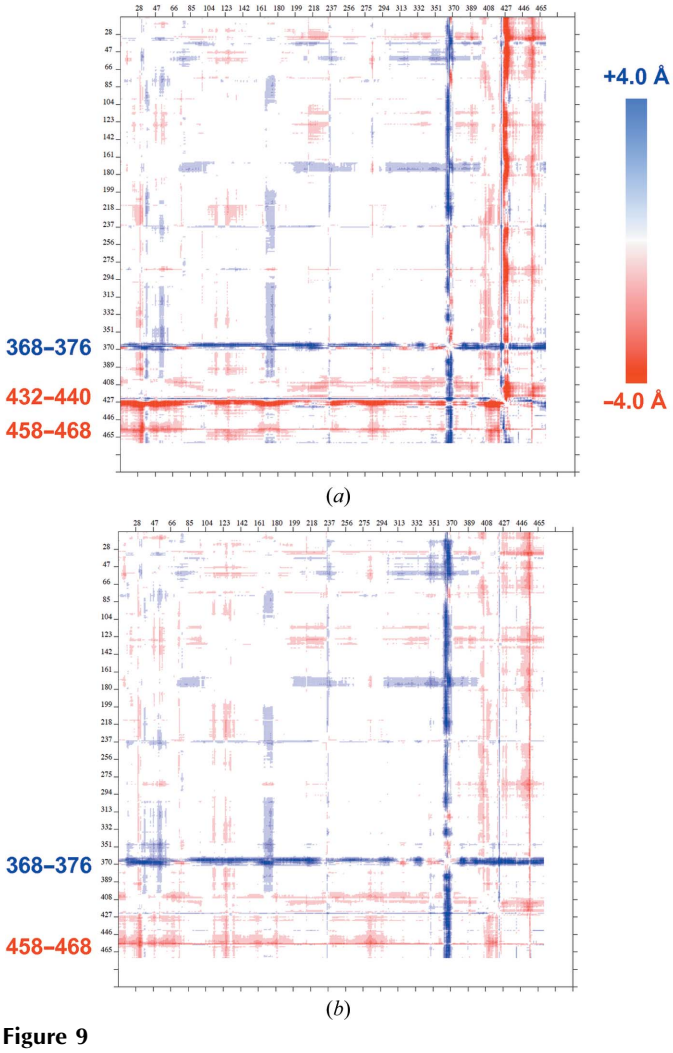

Figure 9
Difference distance matrix plots calculated to compare the differences in interatomic distances for the C^{α} atoms in (*a*) APO1 *versus* PRE1, (*b*) APO1 *versus* BACK1 and (*c*) BACK1 *versus* PRE1. Blue dots indicate differences in interatomic distances that are larger in the first structure *versus* the second structure. Red dots indicate differences in interatomic distances that are smaller in the first structure *versus* the second structure. The darkest blue dots indicate distances that are greater than or equal to +4.0 Å and the darkest red dots indicate distances that are more negative than −4.0 Å. Owing to slight differences in the number of ordered residues at the ends of different models and the amount of disorder in the loop connecting helices II and III in the thumb domain, only residues 6–466 and 474–488 were included in the difference distance matrix calculations for all three structures. The ranges of residues showing the largest differences in interatomic distances for each pair of structures are indicated to the left of each plot in blue and red.

References

Alam, I., Lee, J.-H., Cho, K. J., Han, K. R., Yang, J. M., Chung, M. S. & Kim, K. H. (2012). *Virology*, **426**, 143–151.
Arnold, J. J. & Cameron, C. E. (2004). *Biochemistry*, **43**, 5126–5137.
Beard, W. A. & Wilson, S. H. (2006). *Chem. Rev.* **106**, 361–382.
Blanton, L. H., Adams, S. M., Beard, R. S., Wei, G., Bulens, S. N., Widdowson, M. A., Glass, R. I. & Monroe, S. S. (2006). *J. Infect. Dis.* **193**, 413–421.
Caillet-Saguy, C., Lim, S. P., Shi, P.-Y., Lescar, J. & Bressanelli, S. (2014). *Antiviral Res.* **105**, 8–16.
Cameron, C. E., Moustafa, I. M. & Arnold, J. J. (2009). *Curr. Opin. Struct. Biol.* **19**, 768–774.
Castro, C., Smidansky, E. D., Arnold, J. J., Maksimchuk, K. R., Moustafa, I., Uchida, A., Götte, M., Konigsberg, W. & Cameron, C. E. (2009). *Nature Struct. Mol. Biol.* **16**, 212–218.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cheung, A. C. M. & Cramer, P. (2011). *Nature (London)*, **471**, 249–253.
Choi, K. H. (2012). *Adv. Exp. Med. Biol.* **726**, 267–304.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Doublié, S., Sawaya, M. R. & Ellenberger, T. (1999). *Structure*, **7**, R31–R35.
Eltahla, A. A., Lim, K. L., Eden, J.-S., Kelly, A. G., Mackenzie, J. M. & White, P. A. (2014). *Antimicrob. Agents Chemother.* **58**, 3115–3123.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Ferrer-Orta, C., Arias, A., Escarmís, C. & Verdaguer, N. (2006). *Curr. Opin. Struct. Biol.* **16**, 27–34.
Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmís, C., Domingo, E. & Verdaguer, N. (2004). *J. Biol. Chem.* **279**, 47212–47221.
Ferrer-Orta, C., Arias, A., Pérez-Luque, R., Escarmís, C., Domingo, E. & Verdaguer, N. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 9463–9468.
Fullerton, S. W., Blaschke, M., Coutard, B., Gebhardt, J., Gorbalenya, A., Canard, B., Tucker, P. A. & Rohayem, J. (2007). *J. Virol.* **81**, 1858–1871.
Garriga, D., Ferrer-Orta, C., Querol-Audí, J., Oliva, B. & Verdaguer, N. (2013). *J. Mol. Biol.* **425**, 2279–2287.
Gentile, I., Coppola, N., Buonomo, A. R., Zappulo, E. & Borgia, G. (2014). *Expert Opin. Investig. Drugs*, **23**, 1211–1223.
Gong, P., Kortus, M. G., Nix, J. C., Davis, R. E. & Peersen, O. B. (2013). *PLoS One*, **8**, e60272.
Gong, P. & Peersen, O. B. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 22505–22510.
Högbom, M., Jäger, K., Robel, I., Unge, T. & Rohayem, J. (2009). *J. Gen. Virol.* **90**, 281–291.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kleywegt, G. J. (1996). *Acta Cryst.* D**52**, 842–857.
Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A. & Benkovic, S. J. (1987). *Biochemistry*, **26**, 8410–8417.
Lee, J.-H., Alam, I., Han, K. R., Cho, S., Shin, S., Kang, S., Yang, J. M. & Kim, K. H. (2011). *J. Gen. Virol.* **92**, 1607–1616.
Moustafa, I. M., Shen, H., Morton, B., Colina, C. M. & Cameron, C. E. (2011). *J. Mol. Biol.* **410**, 159–181.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Ng, K. K.-S., Arnold, J. J. & Cameron, C. E. (2008). *Curr. Top. Microbiol. Immunol.* **320**, 137–156.
Ng, K. K.-S., Cherney, M. M., Vazquez, A. L., Machin, A., Alonso, J. M., Parra, F. & James, M. N. G. (2002). *J. Biol. Chem.* **277**, 1381–1387.
Ng, K. K.-S., Pendás-Franco, N., Rojo, J., Boga, J. A., Machín, A., Alonso, J. M. & Parra, F. (2004). *J. Biol. Chem.* **279**, 16638–16645.
Ortín, J. & Parra, F. (2006). *Annu. Rev. Microbiol.* **60**, 305–326.
Powdrill, M. H., Bernatchez, J. A. & Götte, M. (2010). *Viruses*, **2**, 2169–2195.
Read, R. J. (2001). *Acta Cryst.* D**57**, 1373–1382.
Richards, F. M. & Kundrot, C. E. (1988). *Proteins*, **3**, 71–84.
Rothwell, P. J. & Waksman, G. (2005). *Adv. Protein Chem.* **71**, 401–440.
Sarafianos, S. G., Marchand, B., Das, K., Himmel, D. M., Parniak, M. A., Hughes, S. H. & Arnold, E. (2009). *J. Mol. Biol.* **385**, 693–713.
Shen, H., Moustafa, I. M., Cameron, C. E. & Colina, C. M. (2012). *J. Phys. Chem. B*, **116**, 14515–14524.
Shen, H., Sun, H. & Li, G. (2012). *PLoS Comput. Biol.* **8**, e1002851.
Sholders, A. J. & Peersen, O. B. (2014). *J. Mol. Biol.* **426**, 1407–1419.
Showalter, A. K., Lamarche, B. J., Bakhtina, M., Su, M.-I., Tang, K.-H. & Tsai, M.-D. (2006). *Chem. Rev.* **106**, 340–360.
Steitz, T. A. (2006). *EMBO J.* **25**, 3458–3468.
Tan, C. W., Lai, J. K. F., Sam, I.-C. & Chan, Y. F. (2014). *J. Biomed. Sci.* **21**, 14.
Verdaguer, N. & Ferrer-Orta, C. (2012). *Structure*, **20**, 1448–1450.
Wang, D., Bushnell, D. A., Huang, X., Westover, K. D., Levitt, M. & Kornberg, R. D. (2009). *Science*, **324**, 1203–1206.
Weixlbaumer, A., Leon, K., Landick, R. & Darst, S. A. (2013). *Cell*, **152**, 431–441.
Yang, X., Smidansky, E. D., Maksimchuk, K. R., Lum, D., Welch, J. L., Arnold, J. J., Cameron, C. E. & Boehr, D. D. (2012). *Structure*, **20**, 1519–1527.
Zamyatkin, D. F., Parra, F., Alonso, J. M., Harki, D. A., Peterson, B. R., Grochulski, P. & Ng, K. K.-S. (2008). *J. Biol. Chem.* **283**, 7705–7712.
Zamyatkin, D. F., Parra, F., Machín, A., Grochulski, P. & Ng, K. K.-S. (2009). *J. Mol. Biol.* **390**, 10–16.