ISSN 0006-2979, Biochemistry (Moscow), 2009, Vol. 74, No. 10, pp. 1132-1141. © Pleiades Publishing, Ltd., 2009.
Published in Russian in Biokhimiya, 2009, Vol. 74, No. 10, pp. 1388-1399.

Structure–Function Analysis of Mutant RNA-Dependent RNA Polymerase Complexes with VPg

Chaojiang Gu[1,2#], Tao Zeng[3,4#], Yong Li[1], Zhenghui Xu[1], Zhongxi Mo[3]*, and Congyi Zheng[1]*

[1]*State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan 430072, China; E-mail: cctcc202@whu.edu.cn*
[2]*Division of Molecular Virology, St. Luke's-Roosevelt Hospital Center, Columbia University, New York, NY 10019, USA*
[3]*School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China; E-mail: zhxmo@whu.edu.cn*
[4]*School of Computer Science, Wuhan University, Wuhan 430072, China*

Received August 25, 2008
Revision received April 21, 2009

Abstract—The replication of the foot-and-mouth disease virus (FMDV) genome is critically dependent upon the activity of a virally encoded RNA-dependent RNA polymerase (RdRp). In this study, four mutant RdRps of FMDV were isolated from viral quasi-species treated with ribavirin, of which two were single mutants (L123F and T381A) and two were double mutants (T291I/T381I and L123F/F244L). The mutant proteins were expressed in *Escherichia coli* and purified by His-bind resin chromatography. In combination with real-time RT-PCR, an *in vitro* RNA replication system that uses genome RNA/VPg as template–primers was used to determine polymerase activity. Mutant L123F exhibited a 0.6-fold decrease ($p < 0.001$) in polymerase activity relative to wild-type RdRp, whereas the activity of L123F/F244L and T381A was undetectable. Surprisingly, the activity of T291I/T381I yielded a 0.7-fold increase ($p < 0.001$) as compared to wild-type. In order to study the structure–function relationship of RdRp, all structures of the RdRp–RNA template–primer complex were obtained through homology modeling and molecular docking. The VPg1 orientation in the RdRp–VPg1 complexes was determined and analyzed with mathematical methods. Our results reveal that the orientation of VPg after binding to the polymerase determines the FMDV RdRp catalytic activity, which provides a basis for the rational design of novel antiviral agents.

DOI: 10.1134/S0006297909100095

Key words: foot-and-mouth disease virus, RNA-dependent RNA polymerase activity, homogenous modeling, molecular docking, VPg orientation

Foot-and-mouth disease (FMD) is a highly contagious disease caused by a most feared pathogen—foot-and-mouth disease virus (FMDV). FMDV belongs to the aphthovirus genus of the picornavirus family [1] and has seven antigenically distinct serotypes—O, A, C, South African Territories (SAT) 1, SAT2, SAT3, and Asia 1 [2]. The FMDV genome consists of a positive-sense single-stranded RNA molecule of approximately 8500 nucleotides linked to VPg(3B) at its 5′ end and polyadenylated at the 3′ terminus [3]. After attachment of the virus and entry into the cell, RNA is released into the cytoplasm, where VPg is removed from the 5′ terminus of the genomic RNA by a host enzyme [4]. Following the removal of VPg, the viral RNA acts as an mRNA for viral protein synthesis; this is necessary for viral RNA replication and assembly.

Genome replication in all positive-strand RNA viruses is critically dependent upon the activity of a virally encoded RNA-dependent RNA polymerase denoted as RdRp [5]. This enzyme is responsible for synthesizing negative-sense RNA, which is complementary to the positive-sense RNA, as well as the newly synthesized positive-sense RNA genomes that can be translated into viral proteins or encapsidated into new viral particles. Picornaviruses use a protein of 20-24 amino acids, termed VPg, to initiate viral RNA synthesis [6]. The FMDV genome encodes three different VPgs (VPg1, VPg2, VPg3) in tandem, and all three are active as primers for both positive- and negative-strand synthesis [7]. During replication initiation, the first step is the linkage of a UMP to the Tyr3

Abbreviations: FMD(V), foot-and-mouth disease (virus); RdRp, RNA-dependent RNA polymerase; RMSD, root mean square deviation; WT, wild-type.

[#] These authors contributed equally to this work.

* To whom correspondence should be addressed.

hydroxyl group of the VPg protein, which is catalyzed by the RNA polymerase that is also responsible for the elongation of the RNA strands. Thus, these polymerases are structurally adapted for using a protein as a primer [8]. Moreover, RdRp requires the uridinylated form of 3B/VPg peptide (VPgpU or VPgpUpU) to act as the primer for both positive- and negative-strand synthesis [7]. Ferrer-Orta et al. reported the structure and interactions of the FMDV 3D polymerase with a template–primer RNA by X-ray crystallography and discussed the functional role of several amino acid side chains [9]. More recently, they have also analyzed the structure of two complexes between the FMDV RdRp and its protein-primer VPg1 in its non-uridinylated and uridinylated forms [8]. Again, poty- and picornaviruses share similar genome organizations and polyprotein processing strategies. An NTP binding experiment with oxidized [α-^{32}P]UTP revealed that a potato virus A VPg contains an NTP-binding site [10]. All these previous findings suggest that VPg plays an important role in the initiation of RNA synthesis.

Ribavirin is a broad-spectrum agent that inhibits replication of various RNA-containing viruses. Its mechanism of action involves nucleoside induction of mutations resulting in error catastrophe during virus replication. The goal of this work was to study the effect of mutations of picornavirus RdRp, obtained from ribavirin-treated viral quasi-species, on initiation of RNA synthesis. The complex structures of these RdRp mutants of type O FMDV with VPg1, designed by homology modeling and molecular docking, suggested that two of the mutant RdRp forms might have lower affinity to VPg primer. The predicted differences in initiation efficacy of four mutant RdRps correlated with RdRp activity determined by a novel *in vitro* RNA synthesis system together with an accurate real-time RT-PCR method. These results contribute to better understanding of the initiation mechanism of viral genome replication.

MATERIALS AND METHODS

Ribavirin treatment and cloning of RdRp mutants. The FMDVs of serotype O used for this study were obtained from Lanzhou Veterinary Research Institute, Chinese Academy of Agriculture Sciences, Lanzhou. The baby hamster kidney 21 cells (BHK-21) were provided by the China Center for Type Culture Collection (CCTCC). BHK-21 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS) at 37°C with 5% CO_2. Cellular monolayers were infected with virus stock in 2% FBS DMEM at a multiplicity of infection (MOI) of 100 PFU per cell. Ribavirin treatment and cloning of RdRp mutants was described previously [11].

Protein expression and purification. The RdRp proteins were expressed and purified as described previously [9]. They were further purified on a poly(U)-Sepharose 4B (Amersham Biosciences) ion-exchange column. A typical yield was 28 mg of purified protein per liter of cell culture, and then these purified proteins were concentrated using a spin concentrator (30K MWCO; Pall). The purity of these proteins as assessed by SDS-PAGE was >95%. The purified proteins were measured by the Bradford method with bovine serum albumin as a standard. After being adjusted to 0.1 mg/ml, the proteins were stored in aliquots at –80°C until use.

VPg1 synthesis (synthetic peptides). The VPg used as a protein-primer was prepared by solid-phase synthesis of the FMDV OGbF15 VPg1 sequence (GPYVGPLERQKPLKVRAKLPQQE). It was purified by Sephadex G25 chromatography and HPLC, analyzed by mass spectrometry, and finally dissolved in diethyl pyrocarbonate (DEPC)-treated water to a concentration of 10 mg/ml.

Preparation of RNA template. Viral genomic RNA was collected from cytoplasmic lysates using PolyATtract mRNA Isolation Systems (Promega, USA).

***In vitro* RdRp activity assays.** The FMDV 3D polymerase system was used for *in vitro* RdRp activity assays. RNase contamination was avoided by treatment of all solutions and reaction tubes with DEPC. Purified RdRp was dialyzed against 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 10% glycerol. The optimized assay conditions were as follows: 20 mM Tris-HCl, pH 8.0, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 2 mM dithiothreitol, 50 mM NaCl, 50 μg/ml BSA, 10 U of RNasin, 0.25 mM of each NTP, 0.3 μg RNA template, and about 0.1 μg purified RdRp in a 50 μl reaction volume. The reaction mixtures were incubated at 30°C for 2 h and terminated by adding 5 μl 100 mM EDTA in phosphate-buffered saline (pH 7.4). The reaction products were extracted with phenol–chloroform–isoamyl alcohol and precipitated by ethanol. Finally, the precipitated products were resuspended in 1/10 volume DEPC-treated water. The resultant products (5 μl) were used for enzymatic activity determination by a novel real-time RT-PCR.

Strand-specific real-time RT-PCR assay. The TaqMan one-step quantitative RT-PCR assay was performed [12] using a primer pair FMDV-1 5′-GAACACATTCTTTACACCAGGAT-3′ and FMDV-2 5′-CATATCTTTGCCAATCAACATCAG-3′ and a TaqMan probe 5′-FAM-ACAACCTACCGCCGAGCCAATTC-TAMRA-3′. The RT-PCR was performed with the Platinum® Quantitative RT-PCR ThermoScript™ one-step Master Mix Reagents Kit (Invitrogen, USA) following the manufacturer's protocol. Reverse transcription and amplification were carried out using the Rotor-Gene 2000 Real-Time thermal cycler (Corbett Research, Australia). To compensate for variations inherent in sample preparation and reverse transcription, the *in vitro* synthesized RNA templates were used as a dual control along with the copy number of viral RNA molecules, which is more reliable than reported RT-PCR employing a DNA-

based standard. Optimal parameters were: cDNA synthesis at 50°C for 30 min, the thermal profile for PCR was 95°C for 5 min followed by 40 cycles of 94°C for 30 sec and 60°C for 90 sec. Each sample had two to three replicates, and all reactions were repeated two or three times independently to ensure the reproducibility of the results.

Homology modeling. The 3D model of type O FMDV RdRp protein and its four mutants was built using the type C FMDV RdRp [PDB:1WNE] as a template. This template shares about 96.8% identity with target sequence (using the ALIGN software), indicating the high model reliability. In order to obtain the best structural models for RdRp mutants (L123F, L123F/F244L, T381A, and T291I/T381I), the structures were modeled by four different homology-modeling servers: SWISS-MODEL (http://swissmodel.expasy.org) [13]; CPHmodels (http://www.cbs.dtu.dk/services/CPHmodels/) [14]; 3D-JIGSAW (http://www.bmm.icnet.uk/~3djigsaw/) [15]; ESyPred3D (http://www.fundp.ac.be/urbm/bioinfo/esypred/) [16].

Molecular docking. After evaluation of the structural models, the final 3D structures of RdRp mutants were constructed using the online CPHmodels server. The three-dimensional structure of VPg1 (a small molecule) was generated through *de novo* structure prediction by which high-resolution structure prediction (<1.5 Å) can be achieved for small protein domains (<85 residues) [17]. The A10 oligonucleotide was further docked with the FMDV 3D-VPg complex as a template. The structures of RdRp mutants and their ligands were prepared for docking using AutoDock Tools 1.1 [18]. All "heteroatoms", including water molecules and ions, were removed from the original structural data. The positions of polar hydrogen atoms and charges were assigned using the Kollman algorithm [19]. Atomic solvation parameters and fragmental volumes were determined using the addsol program. Gasteiger charges were computed, and energy was minimized with default parameters. AutoDock Tools software was also used to designate the rotatable bonds and generate a grid parameter file and a docking parameter file using default parameters. Finally, docking was performed with AutoDock 3.05 [20], which can predict the bound conformations of a small, flexible ligand to a non-flexible macromolecular target of known structure.

RESULTS

Screening for FMDV RdRp mutants. BHK-21 cells were infected with FMDV and then treated with ribavirin. The mutated RdRp genes were expressed by plasmid pET-30a. Two single mutants (L123F, T381A) and two double mutants (L123F/F244L, T291I/T381I) were isolated from the treated virus quasi-species. Each mutant protein was purified by His-bind resin chromatography using original resin to avoid cross-contamination. All of these four mutant polymerases are stable and soluble proteins that can be expressed at levels comparable to that of wild-type. Each of them contains more than 1 mg purified proteins, which were measured by the Bradford method with bovine serum albumin as a standard (Fig. 1).

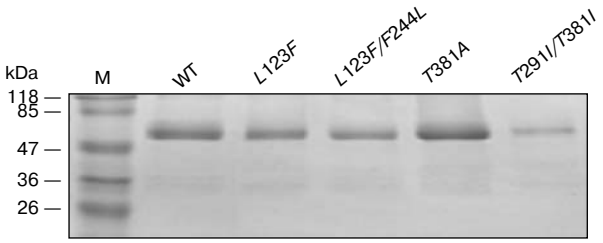

Fig. 1. Preparation of mutant proteins of RdRp in this study. The purified proteins were separated by SDS-PAGE (10% polyacrylamide) and were stained with Coomassie blue. The names of the mutant proteins are derived from the respective amino acid, the position along the RdRp polypeptide, and the final amino acid at that position.

Structure of both wild-type and mutant FMDV 3D polymerase. In order to understand better the mechanism of initiation of viral genome replication, the structure of the FMDV 3D polymerase complex is required. We obtained the amino acid sequence of the wide-type FMDV 3D polymerase and the four mutants (each 470 amino acid residues in length); however, it is impractical to determine the 3D structures of all five polymerases using physicochemical methods, for it is very time consuming and expensive. Therefore, it is reasonable to predict the 3D structures of these polymerases using methods based on theoretical models.

Among all theoretical approaches currently available, comparative modeling is the most reliable method that can generate a 3D model of a protein directly from its amino acid sequence. However, it requires at least one experimentally solved 3D structure (template) sharing at least 50% sequence identity with the target sequence [13]. Since the sequences of wild-type (WT) and mutant RdRps have over 96% identity to our used template [PDB:1WNE], we confirmed that homologous modeling would be the best choice for modeling RdRps. There are many methods for predicting protein 3D structure, but each of them has specific advantages, limitations, and applications. On the assumption that modeling structure can reflect the structural changes of the mutants, we need to evaluate all potential methods to search for appropriate modeling structures. Therefore, the following two criteria extracted from previous study were employed to evaluate the effectiveness of candidates, which are quantitative and qualitative estimations on the influence of the mutations, respectively.

Criterion 1. It has been demonstrated that the resolution of any homologous modeling methods cannot be smaller than 0.5 Å [13]. So, if the difference between the

template and predicted mutant is smaller than 0.5 Å, the functional difference predicted from the structural difference is not reliable.

Criterion 2. An important point in structural prediction concerns local effect of a site mutation (influence on the residues near the mutation in the sequence of the peptide chain) and distant effect (influence to residues near in space but distant in sequence). A method that reflects distant effect is preferred.

As it is known, in a PDB file [21] the 3D structure of a protein is represented by ordered coordinates of constituent atoms. So we note the 3D structure of a protein P as $P = (\lambda_1, \lambda_2, ..., \lambda_n)$, where $\lambda_i = (x_i, y_i, z_i) \in R^3$ $(i = 1,2, ..., n)$ is the spatial coordinates of the Cα atom in the i-th residue. For later discussion and analysis, we define the structural difference between two proteins $P_1 = (\lambda_1^1, \lambda_2^1, ..., \lambda_n^1)$ and $P_2 = (\lambda_1^2, \lambda_2^2, ..., \lambda_n^2)$, denoted by $\rho(P_1, P_2)$, as follows,

$$\rho(P_1, P_2) = \|P_1 - P_2\|_1 = \sum_{j=1}^{n} \left|\lambda_j^1 - \lambda_j^2\right|,$$

where

$$\left|\lambda_j^1 - \lambda_j^2\right| = \sqrt{(x_j^1 - x_j^2)^2 + (y_j^1 - y_j^2)^2 + (z_j^1 - z_j^2)^2},$$

$j = 1, 2, ..., n$.

In this study, the root mean square deviation (RMSD) was used to measure the structure difference between two proteins, and the structure alignment was used to minimize RMSD of the structural difference between the two proteins. Note that for different aims the coordinate system used to represent the spatial structure of protein P_1 and protein P_2 might be different. So, before making a comparison between the two protein structures, it is necessary to transform the spatial coordinates so that they are in the same coordinate system. In this study, CE software [22] was employed to perform the transformation of coordinate systems and the structure alignment.

In our study, four famous homologous structure modeling softwares were considered including Swiss-Model, CPHmodels, 3D-JIGSAW, and ESyPred3D [13-16], and each of them was evaluated for its ability to predict the 3D structure of the wild type (WT) and four mutant RdRps. The RMSD value for each of the predicted mutants relative to WT was calculated (Table 1).

The data in Table 1 indicate that Swiss-Model and 3D-JIGSAW do not fit the first criterion proposed above. More specifically, Swiss-Model emphasizes the local effect of site mutation while it ignores distant effect. For instance, Swiss-Model predicted drastic changes in L123F/F244L at mutated residues 123 and 244 and non-mutated residue 341, compared with WT, but no obvious change at any other locations (Fig. 2b). The 3D-JIGSAW prediction for the L123F/F244L mutant was a trivial shift along the whole chain with moderate change at mutated residue 244 (Fig. 2b). 3D-JIGSAW also predicted that T381A is identical to WT. Thus, this data suggests that 3D-JIGSAW is insensitive to site mutation.

We further analyzed the performances of CPHmodels and ESyPred3D. ESyPred3D predicted a significant difference between WT and the L123F/F244L mutant along the entire length of the chain (Fig. 2b). Its predictions for L123F, T381A, and T291I/T381I mutants show similar entire chain differences (Fig. 2, a, c, and d). Most other point mutations only changed limited structures. In addition, ESyPred3D predicted substantial differences along residues 418-464 between the L123F/F244L mutant and WT, while the other three methods predicted them with no obvious difference. This indicated that this method could not be used to model site mutations in such cases. The differences predicted by CPHmodels met the two criteria and requirements of our study: the structure difference between each two RdRps (whether wild-type, or mutant) is reliable in quantity and significant local and distant effects are also observed. Thus, CPHmodels was finally selected to build 3D structures of the RdRps.

VPg structure bound to RdRp. Because VPg information in the RdRp–VPg complex is very limited in the current Protein Data Base, uncovering the change in binding behavior of VPg in a wild-type complex or a mutant complex is the first step to understand mechanism of initiation of FMDV RNA replication. VPg1 protein used as primer is so small (23 amino acid residues) that its 3D

Table 1. RMSD values of structure difference between the wild-type and the mutant RdRps

Method	ρ(*WT*, *L123F*)	ρ(*WT*, *L123F/F244L*)	ρ(*WT*, *T381A*)	ρ(*WT*, *T291I/T381I*)
SWISS-MODEL	0.061	0.083	0.067	0.060
CPHModels	0.640	0.710	0.734	0.566
3D-JIGSAW	0.001	0.003	0.000	0.010
ESyPred3D	0.975	0.661	0.647	0.883

Note: The modeling constructions were performed with CPHmodels software.

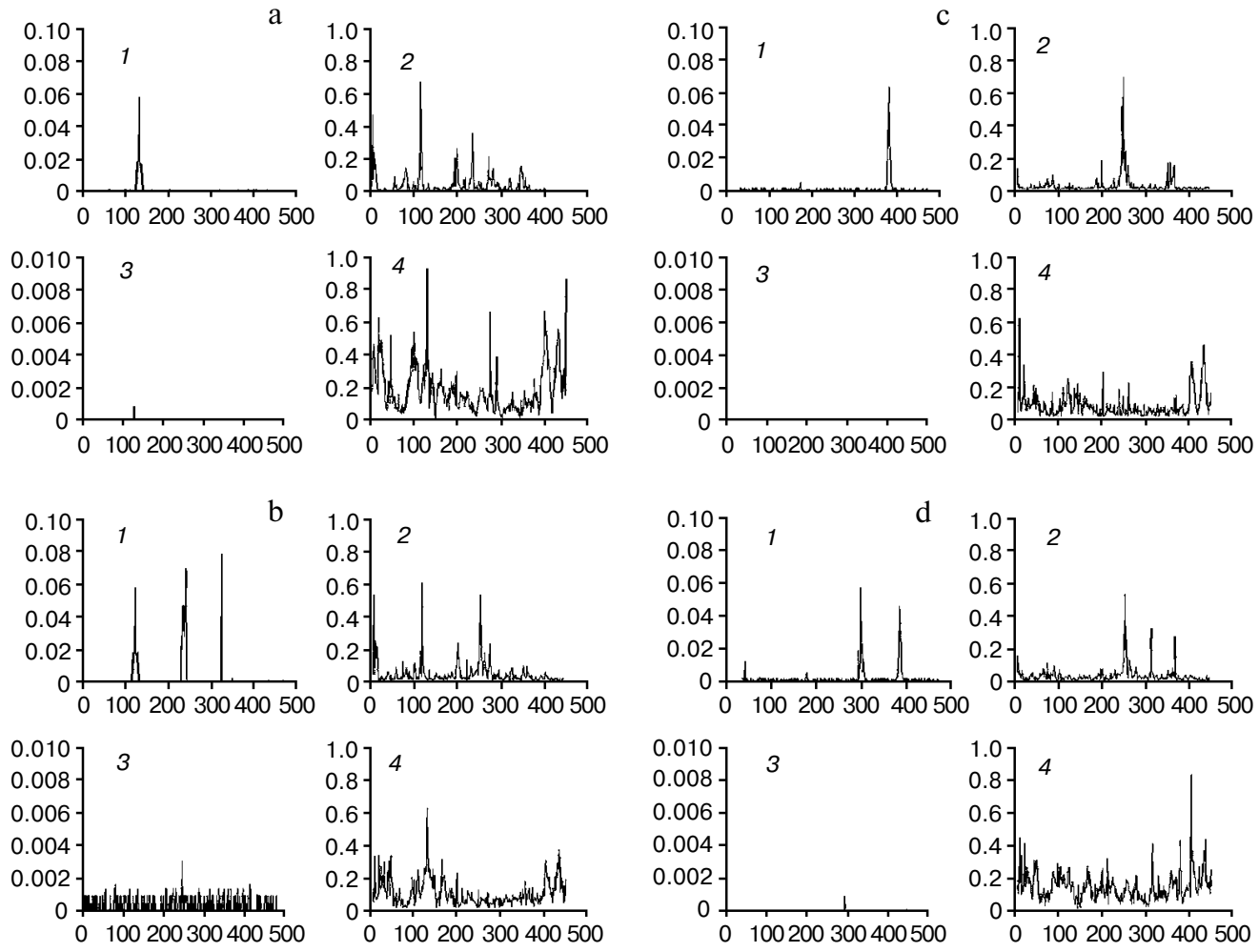

Fig. 2. Comparison of the WT RdRp structure with the mutant RdRp structure. The structures of RdRp were produced by Swiss-Model (*1*), CPHmodels (*2*), 3D-JIGSAW (*3*), and ESyPred3D (*4*) software used for homology modeling, respectively. The X-axis represents the position of amino acid residues in the peptide chain. The Y-axis represents the scale of differences. a) Differences between WT and L123F predictions with four methods. b) Differences between WT and L123F/F244L predictions with four methods. c) Differences between WT and T381A predictions with four methods. d) Differences between WT and T291I/T381I predictions with four methods.

structure cannot be built by homologous modeling. So in this study we selected *de novo* structure prediction [17] to build the 3D structure of VPg1.

In computational drug design, molecular docking could offer a most possible static binding affinity of the complex, and molecular dynamic simulation could supply the dynamic properties of the complex binding affinity. However, neither method identifies all of the contacts of key residues necessary for the binding interface because the affinity of the side chains of residues are ambiguous when predicted by those methods. So, attention here was not given to the influence of the mutation on the side chain of active residues, but on the influence of the mutation on VPg binding affinity.

According to Ferrer-Orta et al. [8], the direction of RNA replication by RdRp is the orientation of VPg upon the primer–enzyme complex. The N-terminal portion of VPg is located close to the NTP entry cavity and projects the side chain of residue Tyr3 into the active site. The peptide chain snakes through the large RNA binding cleft towards the thumb domain of the 3D protein. Thus, VPg binding behavior differences in wild-type and mutant complexes would be reflected by the varied projections of VPg in every complex, and these movements might be related to FMDV RdRp catalytic activity.

First of all, we used AutoDock 3.05 [20] to dock the VPg1 ligand onto the target Asp245 of the WT and the four mutant RdRps, respectively, and we used Genetic Algorithm/Local Search to find the optimal docking structure. We employed minimal docking energy to evaluate the functional capacity with the assumption that changes in energy during docking reflect changes in function. The best possible RdRp–VPg1 complex structure was made after performing 10 minimal energy calcula-

tions (Fig. 3; see color insert). We calculated RMSD values of the structural differences between each combination of the complexes (Table 2A) to display the similarities and differences among the five complexes.

We took the projection of VPg1 from Gly1 to Gln23 as the orientation of VPg1 in each of the complexes. It should be noted that the docking complexes can be observed from different points of view, so the coordinate system to represent the spatial structure of each complex can differ. Hence, it was necessary to transform the structures into the same coordinate system by structural alignment. In addition, since the VPg1 molecule is not symmetrical, it could spin in different binding affinities. Under such a condition, the change of binding behavior of VPg1 in the different complexes cannot be observed using the previously mentioned simple projections of VPg1. For these reasons, we utilized quaternion to further measure binding differences of VPg1 in each of the complexes. Here we cite an example (L123F mutant), which illustrates the use of this method to study VPg1 affinity based on orientation. Let P(L123F, k) be the coordinates of the k-th Cα atom in the L123F mutant chain, and P(WT, k) be the coordinates of the k-th Cα atom in the WT chain. As mentioned previously, the structural alignment of the WT complex and the L123F mutant complex using CE software would offer the following coordinate system transformation:

$$P(L123F,k)=\begin{bmatrix}0.112975 & -0.992095 & 0.055282\\ 0.008138 & 0.056560 & 0.998366\\ -0.993565 & -0.112340 & 0.014463\end{bmatrix}P(WT,k)+\begin{bmatrix}41.028812\\ 0.407715\\ 45.803864\end{bmatrix}=$$
$$=R\cdot P(WT,k)+T,$$

$$R=\begin{bmatrix}0.112975 & -0.992095 & 0.055282\\ 0.008138 & 0.056560 & 0.998366\\ -0.993565 & -0.112340 & 0.014463\end{bmatrix},\qquad T=\begin{bmatrix}41.028812\\ 0.407715\\ 45.803864\end{bmatrix},$$

where R is the rotation matrix and T is the translation matrix. From the viewpoint of geometry, scale operation or translation operation does not affect the orientation of VPg1. So, only the rotation matrix R is considered to be related to the change in orientation and the spin of VPg1. The rotation matrix R can be transformed into the corresponding quaternion [23]:

Table 2A. RMSD values between RdRp–VPg complexes*

Complex	*L*123*F*	*L*123*F*/*F*244*L*	*T*381*A*	*T*291*I*/*T*381*I*
WT	12.19	15.14	15.01	0.40
*L*123*F*		12.53	12.64	12.23
*L*123*F*/*F*244*L*			0.69	15.13
*T*381*A*				15.13

* AutoDock 3.05 was used to dock the ligand VPg1 (23 amino acid residues) onto the target Asp245 of the WT and four mutant receptors, respectively. Genetic Algorithm/Local Search was used to find the optimal docking structure.

Table 2B. VPg orientation and rotation quaternion in the RdRp complex**

VPg orientation	
WT complex	[10.226, –6.823, –5.047]
*L*123*F* complex	[6.113, –9.964, –6.319]
*L*123*F*/*F*244*L* complex	[–6.446, –5.406, 10.286]
*T*381*A* complex	[–6.653, –6.034, 9.794]
*T*291*I*/*T*381*I* complex	[10.006, –6.762, –5.544]
Transformation quaternion	
WT to *L*123*F*	[–0.510380, 0.481955, 0.459600, 245.920730']
WT to *L*123*F*/*F*244*L*	[0.501057, 0.479773, 0.500530, 242.384251']
WT to *T*381*A*	[0.486525, 0.449696, 0.529503, 243.984610']
WT to *T*291*I*/*T*381*I*	[0.010253, 0.021523, –0.001905, 60.056750']
*L*123F to *L*123*F*/*F*244*L*	[0.516189, –0.474334, 0.520633, 238.331542']
*L*123*F*/*F*244*L* to *T*381*A*	[0.049009, –0.000561, –0.019811, 60.276791']

** Transformations were carried out using CE software.

$$R = [x, y, z, w] =$$

$$= [-0.510380, 0.481955, 0.459600, 245.920730'],$$

where $[x, y, z] = [-0.510380, 0.481955, 0.459600]$ represents the symmetrical axle, and $w = 245.920730$ is the angle of rotation. After calculation, we determined that the VPg1 orientation in the WT complex is [10.226, –6.823, –5.047], and that in the L123F mutant complex is [6.113, –9.964, –6.319]. Obviously, VPg1 in the L123F mutant complex points in a similar direction but exhibits a certain angle difference when compared to that in the WT complex. Table 2B contains the calculation results of the structures of the other complexes.

Data presented in Tables 2A and 2B suggest that VPg1 projects similarly in WT and the T291I/T381I mutant, pointing from the small cleft to the large cleft of the active cavity. In contrast, VPg1 in RdRps with the T381A, on L123F/F244L mutant projects from the large cleft to the small cleft. However, since the radius of the small cleft does not hold dsRNA, RNA replication cannot be performed in this direction. Thus, we predicted that the activity of the T291I/T381I mutant would be similar to WT, while that of L123F, L123F/F244L, and T381A would be reduced significantly or even lost completely. The VPg orientations in wild-type and mutant RdRps are shown in Fig. 3.

Biochemical analysis of structure-based mutants of polymerase. Much is known about structural motifs of RdRp [24]. Seven essential motifs have been located (Fig. 4). Motif A, B, C, and E come from the palm domain that stabilizes the position of the 3′ end of the primer, and motif F is in fingers. Several key residues of motifs C and E mediate several common contacts, and this affinity can

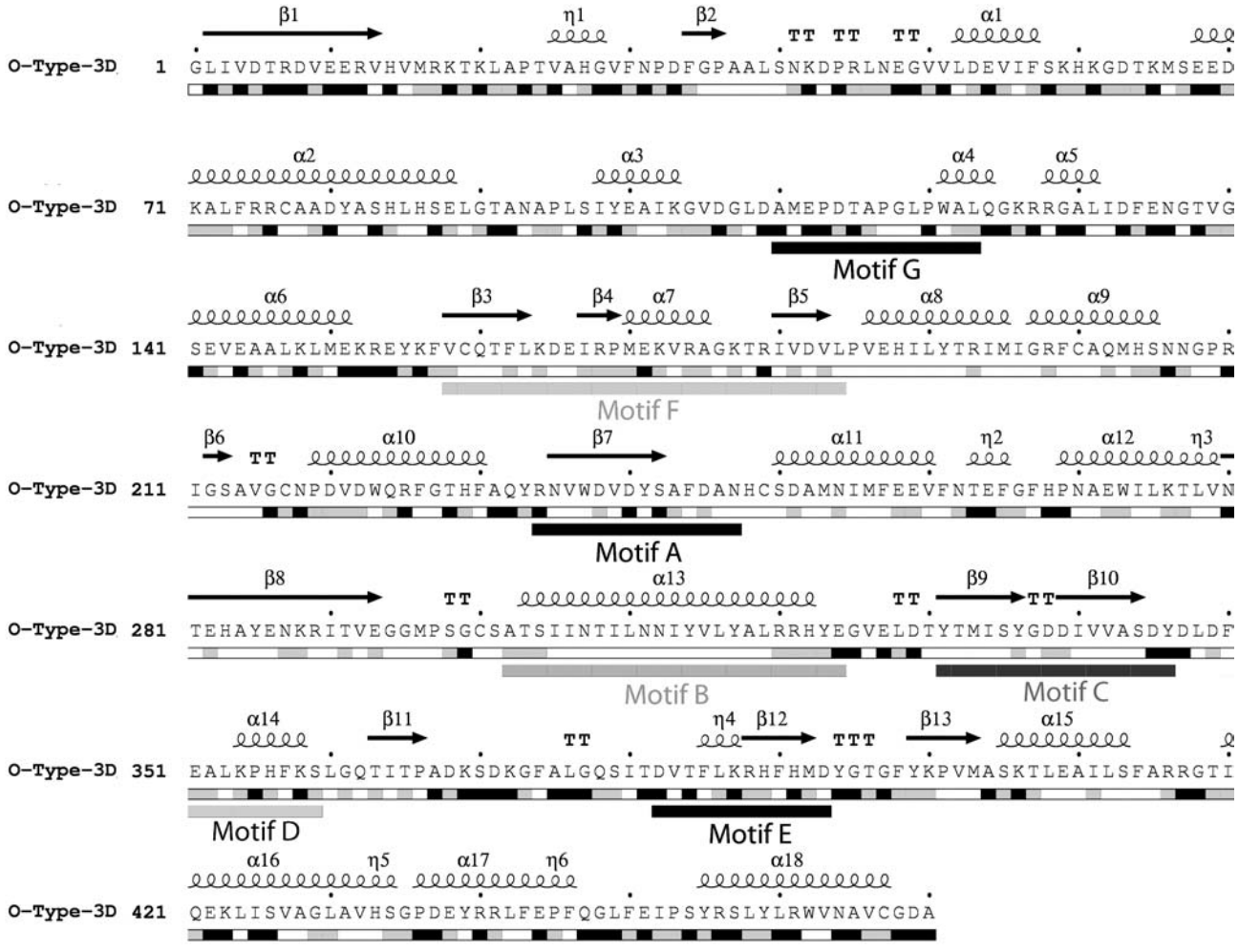

Fig. 4. Known location of motifs on type O FMDV RdRp. By alignment with template of RdRp of type C FMDV, seven known essential functional motifs are located. In detail, motif A is located at 234-247 bp, mainly comprising a β-sheet; motif B is located at 302-324 bp, mainly comprising an α-helix; motif C is located at 331-346 bp, comprising two short β-sheets; motif D is located at 350-359 bp, comprising a short α-helix; motif E is located at 382-393 bp, comprising α-helix and β-sheet; the largest motif F is located at 158-184 bp, grouping three β-sheets and one α-helix; and motif G is located at 110-123 bp, forming a loop structure.

form an optimal orientation for RNA elongation. In addition, essential residues of motifs A, B, and F are involved in distinct patterns of interactions that might help the positioning of the acceptor base of the template strand and could participate in the recognition and binding of the incoming rNTP [24]. However, the mechanism of the critical event of initiation of RNA synthesis during picornavirus replication is still largely unknown owing to the absence of complex structural information. Our docked structure shows how the VPg protein accesses the activity site cavity from the front of the molecule through the large RNA binding cleft, mimicking, at least in part, the RNA molecule. The N-terminal position of VPg projects into the active site where the hydroxyl moiety of residue Tyr3 is in good proximity to the catalytic aspartate Asp245 of motif A and Asp338 of motif C.

In this position, Tyr3 essentially mimics the 3'OH of the primer strand during RNA elongation. Site-directed mutations reported by Ferrer-Orta et al. [8] are predominantly located around the catalytic site. In their model, the conserved residues Asp338 and Asp339 of motif C, together with Asp245 in motif A, are in the right position to bind to divalent cations and to help 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. Observation from our model, most of the mutations can directly alter the catalytic activity of the enzyme, consistent with decreased catalytic activity from experiment. Our *in vitro* biological analysis of the catalytic activity of these mutants demonstrated that T291I/T381I has slightly increased activity, while the other mutants have lost part or all of their activity (Fig. 5). This experimental data agrees with our predictions. Careful analysis of the crystal structure in the neighborhood of the mutated residues might explain the observed loss of activity of the mutant proteins. Phe244 is in close proximity to the sugar moiety of the UMP substrate. The nitrogen atoms of the main chain of residues Phe244 (motif A) involve an interaction with an extensive network formed by the phosphate groups of NTP. In addition, residue Phe244 contributes to the highly hydrophobic RNA-binding cavity [24]. It seems that a mutation to Leu would have a major effect on RdRp activity. Thr381 lies on a loop near the entry of the NTP; the hydroxyl side chain of Thr381 is hydrogen bonded to the hydroxyl group of Ser379, providing some rigidity to that loop. So, mutation to Ala might affect the dynamics of that particular loop and consequently affect the binding of the VPg protein. This could alter the RdRp activity. Thr381 is also close to a group of residues (His389, His391, Tyr394) that is expected to affect the uridinylation of VPg [5]. Leu123 lies on a loop close to N-terminal Glu11 that is believed to play an important role in the RdRp activity. Therefore, mutation of Leu123 to Phe is expected to modulate the interaction with Glu11, which could affect the replication activity of the RdRp. From the available structures, it is not clear why the double mutant T291I–T381I showed increased replication activity of RdRp. A possible explanation for this could be that the substitution T291I conferred the right conformation of the VPg1 in the complexes and rescued the enzyme capacity at a level higher than that of the wild type enzyme.

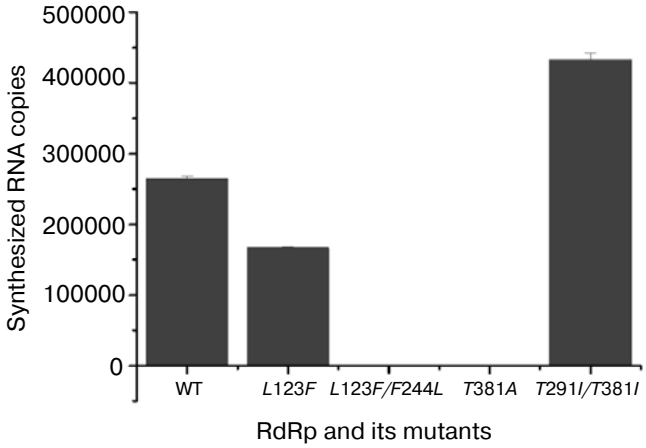

Fig. 5. Activity of RdRp mutants determined by *in vitro* RNA synthesis assays. In this assay, the viral genome RNAs utilized as RNA templates were added to a 50-μl reaction mixture containing about 0.1 μg purified RdRp. The reaction mixtures were incubated for 120 min at 30°C and terminated by adding 100 mM EDTA in phosphate-buffered saline (pH 7.4). The synthesized RNA was quantified by a strand-specific real-time RT-PCR as described previously. The data represent the mean of three separate experiments including the standard error (bar).

DISCUSSION

In our previous study, ribavirin was used as an antiviral reagent and FMDV as a model RNA virus. The ribavirin passage experiments indicated that ribavirin is a highly efficient RNA virus mutagen. The primary antiviral mechanism is attributable to lethal mutagenesis/error catastrophe [11]. Recently, it was shown that RTP (ribavirin triphosphate) is incorporated into the viral RNA genome by the poliovirus RNA polymerase $3D^{pol}$, which allows for base mismatches, thereby enhancing the mutation frequency and leading to lethal mutagenesis of the viral genome [25, 26]. Due to the importance of the highly structured and multifunctional RNA polymerase, the sequence of the viral encoded enzyme RdRp is generally highly conserved. Thus, mutations, especially around its active sites, likely result in a decline in activity or an enhancement in the replication error rate. Four mutant RNA-dependent RNA polymerases of FMDV containing either single replacements (L123F, T381A) or double replacements (L123F/F244L, T291I/T381I) were

obtained from a viral quasi-species treated with 1000 μM ribavirin after one round of infection.

When attempting to understand protein function or engaging in the drug design process, insights into the three-dimensional structure of proteins is of great importance. The experimental elucidation of the 3D-structure of proteins is, however, often hampered by difficulties in obtaining sufficient protein, diffracting crystals, and many other technical aspects. Therefore, it is not surprising that predictive methods have gained much interest. Among all current theoretical approaches, comparative modeling is the only method that can reliably generate a 3D model of a protein (target) from its amino acid sequence. Successful model building requires at least one experimentally solved template that has significant amino acid sequence similarity to the target sequence. The structural modeling for type O FMDV RdRp was performed using type C FMDV RdRp [PDB:1WNE] as a template. The best software for homologous modeling of RdRp mutants was identified by experiments. Through 3D structure evaluation, we have shown that CPHmodels is sensitive to site-mutation-caused structural change, provides sufficient resolution power, and is suitable for analyzing structural differences between multiple mutants. To investigate the effect of FMDV RdRp mutation on RNA polymerase, we studied the binding pose of primer VPg1 by docking it to the WT polymerase and its four mutants. AutoDock 3.05 software was used to dock primer molecules (VPg1) with RdRp. An estimation was made for the effect of site mutation on the function of FMDV RdRp based on the predicted mutant structure and conserved regions. The structure models of RdRp–primer complexes through molecular docking suggest that RNA replication diminishes when VPg is oriented towards the small cleft. This prediction is supported by our experimental results. Polymerase appears to use conformational change to couple correct nucleotide binding or base pairing to catalytic efficiency. Compounds that induce the "active" conformation of the polymerase, irrespective of the nature of the base pair, may decrease polymerase fidelity. A slight decrease in polymerase fidelity has profound effects on virus viability because the overall effect is amplified by multiple rounds of genome replication [27]. At the same time, mutations of critical amino acid residues in the polymerase actives sites could enhance the replication error rate or disable the activity of the polymerase. In our report, Phe244 is in close proximity to the sugar moiety of the UMP substrate, so a major effect on the RdRp activity upon mutation to Leu might have occurred. Thr381 lies on a loop close to the entry of the NTP; the hydroxyl side chain of Thr381 is hydrogen bonded to the hydroxyl group of Ser379, providing some rigidity to that loop. Therefore, a mutation to Ala might affect the dynamics of that particular loop and consequently affect the binding of the VPg protein and hence the RdRp activity. Thr381 is also not far from the group residues (His389, His391, Tyr394) that are expected to affect the uridinylation of VPg. Leu123 lies on a loop close to Glu11 at the N-terminal that is believed to play an important role in the RdRp activity.

In polymerization assays, the purified mutant polymerase with L123F showed a decreased catalytic capacity and the L123F/F244L and T381A mutants had no detectable activity in comparison to WT when using an RNA genome as a template in an *in vitro* replication system. The results indicate that the 3D polymerase including different domains that can alter the interaction of the enzyme with mutagenic nucleoside analogs is exquisitely sensitive to mutations and act as target protein during FMDV extinction caused by ribavirin. This suggests that the viral RdRp might be a good target for antiviral drug treatment. Consistent with the functional study, the results of homology modeling and molecular docking also show that the VPg1 orientation at the active site is different in the four mutants. Compared to the VPg1 orientation in wild-type RdRp, VPg1 orientation is altered greatly in the L123F mutant (activity is reduced), and is in the opposite direction in the L123F/F244L mutant (activity is lost) and the T381A mutant (activity is lost). However, it is noteworthy that VPg1 orientation in the active site of the T291I/T381I mutant is similar to that of the wild-type enzyme, but that the double mutant enzyme shows increased replication activity of RdRp. Hence, it is important to further investigate this double mutant to understand the observed phenomenon, which might be related to the T291I/T381I polymerase resistance for ribavirin-induced error catastrophe.

In summary, *in vitro* activity assays and bioinformatic analysis for wild-type and four mutant RdRp–VPg1 complexes of type O FMDV have been described. The main contributions of this study are the prediction of 3D structures of the wild-type RdRp–VPg1 complex and four mutant complexes by homology modeling and molecular docking and the determination of VPg1 orientation in these RdRp–VPg1 complexes by mathematical methods. The latter reveals the relationship between replication activity of RdRp and VPg1 orientation in the RdRp–VPg1 complex, providing a new theoretical insight into the mechanism of initiation. Our prediction results are supported by *in vitro* biological analysis of RdRps function and by biochemical analysis of structure of polymerase mutants.

We also want to point out that the conclusions of this study apply to the WT and mutant RdRp–VPg1 complexes of type O FMDV. The FMDV genome encodes three different VPgs (VPg1, VPg2, VPg3) in tandem, and all three are active as primers for RNA replication. Deletion of any individual copy of VPg has a deleterious effect on RNA replication [28]. Therefore, in following study it is necessary to also consider VPg2 and VPg3 binding to RdRp, besides VPg1 binding to RdRp, in order to build a complete profile of the role of VPg in RdRp genome

replication. We plan to obtain more RdRp mutants and RdRp–VPg complex mutants through amino acid replacement in some positions of RdRp and VPgs. Obviously, it would be difficult to study the 3D structures of so many RdRp and RdRp–VPg complexes by techniques such as X-ray crystallography or nuclear magnetic resonance spectroscopy (NMR), but such analyses could be done using the bioinformatic approaches used in this study. These bioinformatic approaches could be used to investigate the RdRps of other picornaviruses as well.

This work was supported in part by the Special Funds for Major State Basic Research of China (No. G1999011904, 2006CB504300) and the Key Project of Science Research of the Chinese Ministry of Education (No. 505009).

REFERENCES

1. Belsham, G. J. (1993) *Prog. Biophys. Mol. Biol.*, **60**, 241-260.
2. Pereira, H. G. (1981) in *Virus Diseases of Food Animals* (Gibbs, E. P. J., ed.) Vol. 2, Academic Press, New York, pp. 333-363.
3. Chinsangaram, J., Piccone, M. E., and Grubman, M. J. (1999) *J. Virol.*, **73**, 9891-9898.
4. Murray, K. E., and Barton, D. J. (2003) *J. Virol.*, **77**, 4739-4750.
5. O' Reilly, E. K., and Kao, C. C. (1998) *Virology*, **252**, 287-303.
6. Paul, A. V. (2002) in *Molecular Biology of Picornaviruses* (Semler, B. L., and Wimmer, E., eds.) ASM Press, Washington, DC, pp. 227-246.
7. Nayak, A., Goodfellow, I. G., and Belsham, G. J. (2005) *J. Virol.*, **79**, 7698-7706.
8. Ferrer-Orta, C., Arias, A., Agudo, R., Perez-Luque, R., Escarmis, C., Domingo, E., and Verdaguer, N. (2006) *EMBO J.*, **25**, 880-888.
9. Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmis, C., Domingo, E., and Verdaguer, N. (2004) *J. Biol. Chem.*, **279**, 47212-47221.
10. Puustinen, P., and Makinen, K. (2004) *J. Biol. Chem.*, **279**, 38103-38110.
11. Gu, C. J., Zheng, C. Y., Zhang, Q., Shi, L. L., Li, Y., and Qu, S. F. (2006) *J. Biochem. Mol. Biol.*, **39**, 9-15.
12. Gu, C. J., Zheng, C. Y., Shi, L. L., Zhang,Q., Li, Y., Lu, B., Xiong, Y., Qu, S. F., Shao, J. J., and Chang, H. Y. (2007) *Virus Genes*, **34**, 289-298.
13. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) *Nucleic Acids Res.*, **31**, 3381-3385.
14. Lund, O., Nielsen, M., Lundegaard, C., and Worning, P. (2002) in *Abstracts of CASP5 Conference*, California, p. A102.
15. Bates, P. A., Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2001) *Proteins Suppl.*, **5**, 39-46.
16. Lambert, C., Leonard, N., De, B., and Depiereux, E. (2002) *Bioinformatics*, **18**, 1250-1256.
17. Bradley, P., Misura, K. M., and Baker, D. (2005) *Science*, **39**, 1868-1871.
18. Sanner, M. F., Duncan, B. S., Carillo, C. J., and Olson, A. J. (1999) *Pac. Symp. Biocomput.*, 401-412.
19. Chiche, L., Gregoret, L. M., Cohen, F. E., and Kollman, P. A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 3240-3243.
20. Goodsell, D. S., Morris, G. M., and Olson, A. J. (1996) *J. Mol. Recognit.*, **9**, 1-5.
21. RCSB Protein Data Bank (www.rcsb.org).
22. Guda, C., Scheeff, E. D., Bourne, P. E., and Shindyalov, I. N. (2001) *Pac. Symp. Biocomput.*, **6**, 275-286.
23. Theobald, D. L. (2005) *Acta Crystallogr. A*, **61**, 478-480.
24. Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmis, C., Domingo, E., and Verdaguer, N. (2007) *Proc. Natl. Acad. Sci. USA*, **104**, 9463-9468.
25. Crotty, S., Maag, D., Arnold, J. J., Zhang, W. D., Lau, J. Y. N., Hong, Z., Andino, R., and Cameron, C. E. (2000) *Nature Med.*, **6**, 1375-1379.
26. Arnold, J. J., Vignuzzi, M., Stone, J. K., Andino, R., and Cameron, C. E. (2005) *J. Biol. Chem.*, **280**, 25706-25716.
27. Cameron, C. E., and Castro, C. (2001) *Curr. Opin. Infect. Dis.*, **14**, 757-764.
28. Falk, M. M., Sobrino, F., and Beck, E. (1992) *J. Virol.*, **66**, 2251-2260.