doi:10.1016/j.jmb.2005.12.044 J. Mol. Biol. (2006) 357, 665–675

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Nucleotide Channel of RNA-dependent RNA Polymerase used for Intermolecular Uridylylation of Protein Primer

Andres B. Tellez[1], Scott Crowder[2], Jeannie F. Spagnolo[2], Aaron A. Thompson[3], Olve B. Peersen[3], Douglas L. Brutlag[1] and Karla Kirkegaard[2]*

[1]*Departments of Biomedical Informatics Stanford University School of Medicine, Stanford CA 94305, USA*

[2]*Microbiology and Immunology Stanford University School of Medicine, Stanford CA 94305, USA*

[3]*Department of Biochemistry Colorado State University Fort Collins CO 80523, USA*

**Corresponding author*

Poliovirus VPg is a 22 amino acid residue peptide that serves as the protein primer for replication of the viral RNA genome. VPg is known to bind directly to the viral RNA-dependent RNA polymerase, 3D, for covalent uridylylation, yielding mono and di-uridylylated products, VPg-pU and VPg-pUpU, which are subsequently elongated. To model the docking of the VPg substrate to a putative VPg-binding site on the 3D polymerase molecule, we performed a variety of structure-based computations followed by experimental verification. First, potential VPg folded structures were identified, yielding a suite of predicted β-hairpin structures. These putative VPg structures were then docked to the region of the polymerase implicated by genetic experiments to bind VPg, using grid-based and fragment-based methods. Residues in VPg predicted to affect binding were identified through molecular dynamics simulations, and their effects on the 3D-VPg interaction were tested computationally and biochemically. Experiments with mutant VPg and mutant polymerase molecules confirmed the predicted binding site for VPg on the back side of the polymerase molecule during the uridylylation reaction, opposite to that predicted to bind elongating RNA primers.

Keywords: polymerase; protein priming; RNA replication; molecular dynamics; docking

Introduction

Viral RNA complexes are higher-order structures that comprise myriad protein–protein, protein–RNA and protein–lipid interactions. Both positive and negative strands of poliovirus RNA are initiated using VPg, a 22 amino acid residue peptide, as a primer.[1,2] Protein-mediated RNA priming occurs *via* a protein–protein interaction with VPg binding to the polymerase and subsequently being uridylylated at its third residue, tyrosine, in a reaction that requires UTP and an AA dinucleotide-containing template.[1,3] It is likely that uridylylation of VPg is *via* the same two metal ion mechanism and active site used for RNA-dependent RNA polymerization,[4] because the uridylylation reaction is sensitive to mutations in the same catalytic aspartate residues as those used for RNA-dependent RNA elongation.[5]

E-mail address of the corresponding author: karlak@stanford.edu

In the RNA replication complex of poliovirus, the polymerase is tethered to membranes *via* direct binding to protein 3AB, a precursor of VPg (also termed 3B).[6] The 87 amino acid residue 3A moiety of protein 3AB contains a hydrophobic domain near the C terminus that is required for its membrane association; however, the seven most C-terminal residues adjacent to the VPg region are hydrophilic.[6] Previous work has identified a 3AB binding site on the polymerase,[5] which is, interestingly, on the side opposite where the RNA template is believed to bind, based on analogy with other polymerases,[7–9] and the co-crystallization of the highly related foot-and-mouth-disease RNA-dependent RNA polymerase with an RNA template and primer.[10] Protein 3AB is a competitive inhibitor of VPg uridylylation,[11] suggesting that VPg and protein 3AB share the same binding site on 3D polymerase. This raises the possibility that VPg is uridylyated from the "back" of the polymerase by accessing the active site through the nucleotide channel. The structure of the 22 residue VPg

peptide is undetermined and no structure of the VPg–3D polymerase complex has been solved. To generate hypotheses about the structure of the VPg–polymerase complex, we began with computational modeling of the interaction.

Computational biology provides several different approaches for the docking of one macromolecule onto another, but the verification of these predicted structures usually requires genetic and biochemical experiments. Grid-based docking algorithms typically assume that at least one of the proteins in the interaction is a rigid body. These programs, including Zdock and FlexiDock (Materials and Methods),[12] create a three-dimensional grid in the putative binding site and allow the ligand structure to occupy positions along the grid points. On the other hand, fragment-based docking methods such as FlexX and Extenza,[13,14] build the ligand in the putative binding site one residue at a time, searching for favorable energetic interactions as the ligand increases in size. Programs that model protein motion can accommodate only a limited number of degrees of freedom because large conformational changes in proteins and flexible substrates create an intractable number of possible complexes. Here, we used genetic experiments to guide the docking of computationally folded VPg into its putative binding site on a suite of different potential polymerase conformations, followed by experimental verification of the predicted complex.

Results

Computational studies predict a β-hairpin structure for poliovirus VPg

The structure of the 22 amino acid residue poliovirus-encoded protein VPg (Figure 1) has not been solved. We sought to determine the most probable secondary structure of this protein primer for poliovirus RNA replication by computational modeling before attempting to dock it onto the known structure of poliovirus polymerase. All predicted folds for poliovirus VPg were found to exhibit β-hairpin structures, such as the example shown in Figure 1(a). In all such structures, the N and C-terminal residues were brought into close apposition by a turn induced by a triple-proline motif (P7, P12 and P14), bending the peptide into two anti-parallel β-strands. The potential structural importance of the triple-proline motif is supported by alignment of multiple VPg sequences (A. Palmenberg, personal communication), which shows a high level of conservation at these positions among enteroviruses. On the basis of these predictions, a β-hairpin structure such as that shown in Figure 1(a) became our working model for a probable structure of VPg in complex with the viral polymerase.

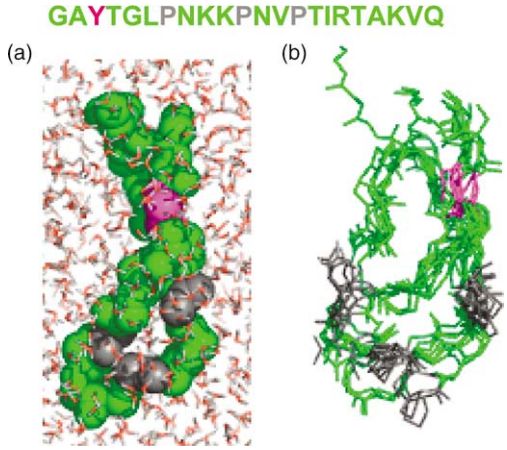

Figure 1. (a) A typical main-chain fold of VPg predicted by computation, surrounded by water. (b) Superposition of ten VPg structures taken from a 10 ns simulation, sampled at 1 ns intervals. The α-carbon chain of Tyr3 is magenta; Pro residues 7, 11 and 14 are gray.

Molecular dynamics simulations using solvent water were employed to test the stability of the β-hairpin structure computationally. In Figure 1(b), snapshots of the main chain from a representative simulation were taken every nanosecond for 10 ns and superimposed. The β-strands flanking the Pro residues showed considerable flexibility in these simulations, as did the precise position of the residues in the β-hairpin turn. However, the presence of the turn and the overall β-hairpin fold remained stable features of the predicted VPg structure (Figure 1(b)).

Computational docking predicts specific residues in VPg and polymerase involved in complex formation

To develop a testable hypothesis for the molecular interaction between 3D polymerase and VPg, several folded VPg structures were computationally docked on different potential conformations of poliovirus polymerase using both fragment-based and grid-based docking algorithms. Various folded VPg conformations are shown in Figure 1(b); putative polymerase conformations were obtained by molecular dynamics simulations[15] of the full-length three-dimensional structure.[16] The docked VPg ligand was constrained to be within 7 Å (from C^{α} to C^{α}) of residues Phe377, Arg379, Val391 and Glu382 (Figure 2(a)); these residues were chosen because they are necessary for either 3AB binding, VPg uridylylation or both.[5] The position of Tyr3 of VPg, to which uridine residues are attached covalently, was not constrained relative to the active site residues of the polymerase.

A typical docked VPg structure, making contact with the putative binding residues Phe377, Arg379, Val391 and Glu382 on the polymerase, is shown in Figure 2(b). All putative VPg–polymerase complexes shared a set of common contacts. Specifically, VPg residues Lys9 and Lys10 at the β-turn were found close to or directly contacting

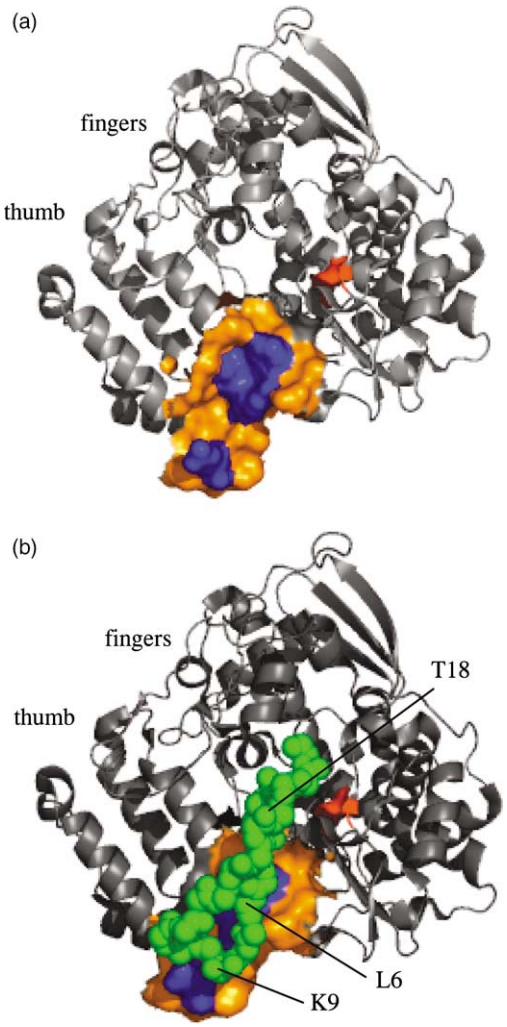

Figure 2. (a) The putative 3AB binding site[5,32] (Glu382, Phe377, Arg379 and Val391) on poliovirus 3D polymerase that was used to filter predicted VPg–polymerase complexes is shown in blue. Active-site residue Asp328 is shown in red. Residues within seven Å of the 3AB residues used to guide the VPg docking are shown in orange. (b) An example of a docked VPg structure is shown with the main chain rendered in green spheres.

polymerase residues Glu382 and Asp 381, respectively. Additionally, hydrophobic residues Gly5, Leu6 and Pro7 on one face of the β-hairpin VPg structure were always found in close proximity to a hydrophobic pocket on the polymerase composed of Phe377, Arg379 and Val391. Both charged and hydrophobic residues in VPg make complementary contacts on the polymerase and are ubiquitous in the docked structures. The ubiquity of the polymerase contacts at Glu382, Phe377, Arg379 and Val391 was not surprising, of course, because the VPg docking was originally constrained to be proximal to these residues. However, the uniformity of the contacts with VPg and the proximity of the Tyr3 hydroxyl group to the catalytic Asp residues of the polymerase argued that the predicted complexes were plausible candidates for the actual structure. The particular complex shown in Figure 2(b) is one of several in which the hydroxyl group of Tyr3 of VPg that is uridylylated (Figure 1) is in close proximity to a metal ion in the polymerase active site.[17]

Mutations in Leu6, Lys9 and Lys10 lead to release of VPg in computational simulations

To probe the role of VPg residues in the VPg–polymerase interaction computationally, the relative stabilities of the complexes between 3D polymerase and various VPg peptides were monitored by molecular dynamics simulations. Specifically, the simulations used wild-type and mutant VPg molecules in which Leu6, Lys9 and Lys10 were computationally replaced with Ala (L6A, K9A and K10A). The screening dielectric constant of H_2O was held constant, despite the charge distribution changes in the VPg molecules, by the particle-mesh Ewald algorithm for coulombic interaction and the SPC water molecule model included in the simulation package (Materials and Methods). Wild-type and mutant VPg molecules were docked onto 3D polymerase in 12 different initial conformations, followed by a 2 ns molecular dynamics simulation. Six distances between residues on VPg and residues on the polymerase (Figure 3(a)) were monitored throughout the course of molecular dynamics simulations[15] to compare the dynamics of the different VPgs at corresponding starting positions. A representative sample of such simulations is shown in Figure 3(b), in which one of the intermolecular distances, between the α-carbon atoms on VPg residue Lys10 and polymerase residue Glu382, was plotted as a function of time. The simulations all started with the wild-type and mutant VPg molecules in the same position relative to the polymerase active site; the slight differences in starting position shown in Figure 3(b) result from stochastic placement of water molecules followed by subsequent energy minimization as well as variation between intial velocities assigned to the atoms in the different VPg molecules. The wild-type VPg–polymerase complex was quite stable over the 2 ns time-frame of the simulation; all six intermolecular distances decreased with time (Figure 3(c)) indicating that, in seeking a more energetically favorable state, the proteins moved closer together. In contrast, the L6A, K9A and K10A mutant VPg molecules all created apparent instabilities in the VPg–polymerase complex; as can be seen in Figure 3(c), most of the intermolecular distances in these complexes increased with time of simulation.

Experimental alanine substitution of VPg residues Leu6, Lys9 and Lys10 decreases the rate of VPg uridylylation

To test the importance of VPg residues Leu6, Lys9 and Lys10 in VPg binding to poliovirus polymerase experimentally, two different biochemical assays were employed. First, the enzymatic uridylylation

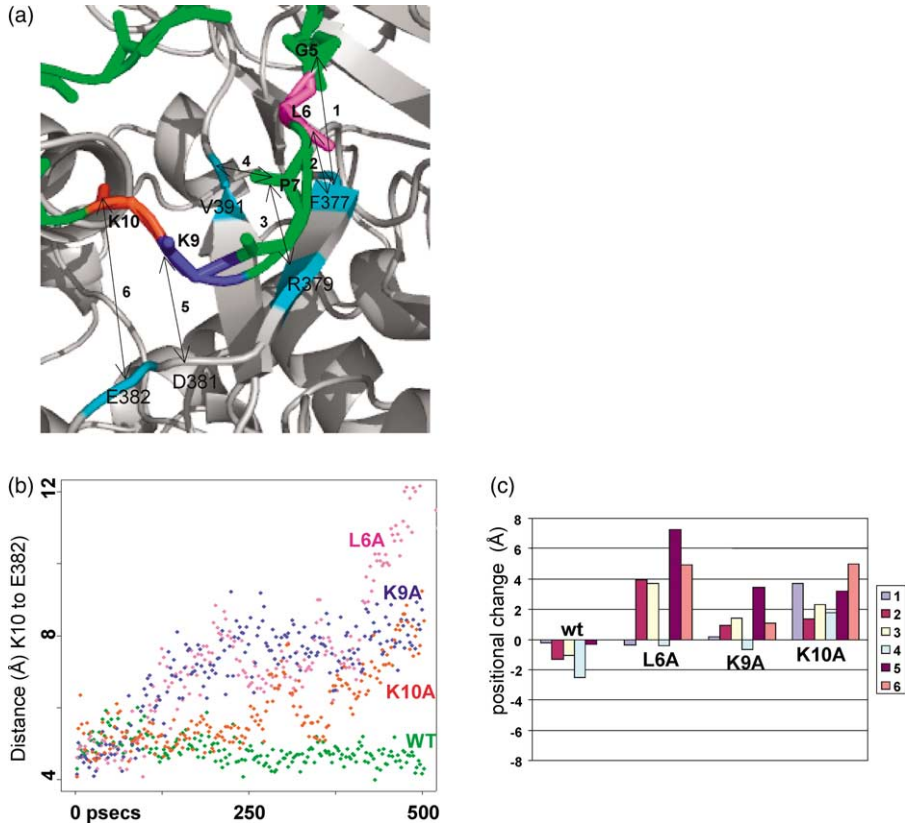

Figure 3. (a) A typical docked VPg structure showing interactions and contacts found in all predicted VPg–3D polymerase complexes. The polymerase residues within the 3AB binding site are shown in blue; the distances between those residues and putative contact residues on VPg, shown in bold, are numbered. (b) The distance between the C^{α} atoms of VPg residue Lys9 and polymerase residue Asp381 was plotted as a function of time for wild-type and mutant VPg molecules. (c) The change in position of VPg residues, relative to polymerase, has been plotted with each of the intermolecular distances shown in (a) for wild-type and mutant VPg molecules.

of mutant and wild-type VPg peptides by 3D polymerase was measured by monitoring the incorporation of labeled UMP using [α-^{32}P]UTP as a substrate; the 40 min time point used for each reaction was determined in time courses to reflect the initial rate (data not shown). As shown in Figure 4(a)–(c), mutant VPg peptides L6A, K9A and K10A reduced the rate of VPg uridylylation relative

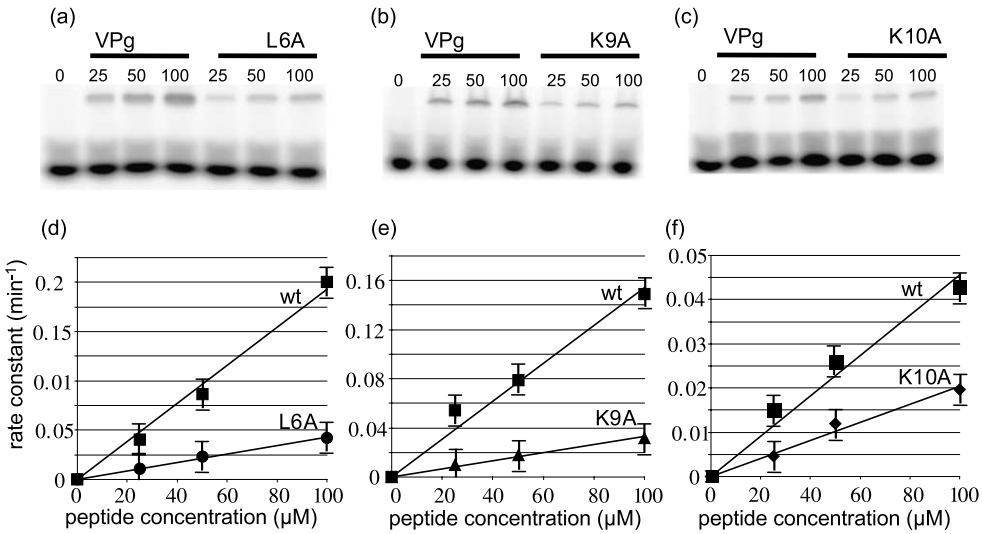

Figure 4. Activity assays to measure uridylylation rates of wild-type and mutant VPg molecules. Wild-type VPg uridylylation is compared with (a) VPg-L6A (b) VPg-K9A and (c) K10A-VPg at the concentrations indicated. Below, the initial rates of the uridylylation reactions for wild-type VPg and (d) VPg-L6A, (e) VPg-K9A and (f) VPg-K10A are shown. The slope of the line is k_{cat}/K_m.

to wild-type VPg (Figure 4(d)–(f)). The rate of VPg-pU formation divided by the polymerase concentration is graphed in Figure 4(d) as a function of VPg concentration. For each VPg tested, the slope was calculated to yield the second-order rate constant for the reaction, k_{cat}/K_m. All three mutant VPg peptides showed reduced k_{cat}/K_m values relative to wild-type VPg (Figure 4(d)). Specifically, the k_{cat}/K_m values for L6A, K9A and K10A are 4.8-fold, 5.0-fold and 2.8-fold, respectively, lower than wild-type VPg.

Experimental alanine or aspartate substitution of VPg residue Thr18 slightly increases the rate of VPg uridylylation

Alanine substitution mutations were chosen because alanine is equally soluble in hydrophobic and hydrophilic environments.[18,19] However, substitution of a charged or polar residue by alanine could cause long-range electrostatic effects that could also influence the binding affinity of the mutant VPg molecules for 3D polymerase. To test whether a large change in electrostatic distribution would negatively impact VPg–polymerase interactions, Thr18, a residue on the VPg surface not predicted to make direct contact with the polymerase, was changed to either Ala or Asp. Figure 5 shows the binding data for VPg-T18A and T18D; if anything, slightly increased affinity of these mutant VPgs for polymerase was observed. The k_{cat}/K_m values for VPg-T18A (Figure 5(a) and (b)) and T18D (Figure 5(c) and (d)) were measured as 1.7-fold and 1.3-fold higher, respectively, than wild-type VPg. The small increase observed when Thr18 is substituted with Ala or Asp shows that changing the dielectric constant does not reduce the rate of VPg uridylylation by poliovirus 3D polymerase *per se*. Therefore, it is most likely that the reduced rates of uridylylation of the L6A, K9A and K10A mutant Vpgs (Figure 4) result from the alteration of direct polymerase contacts. However, the observed reductions in k_{cat}/K_m could have resulted from changes in k_{cat}, the intrinsic efficiency of the reaction with no change in polymerase affinity for bound mutant VPg. Therefore, experiments were designed to test whether the Leu6, Lys9 and Lys10 mutations affected the direct binding of VPg-containing sequences to the polymerase.

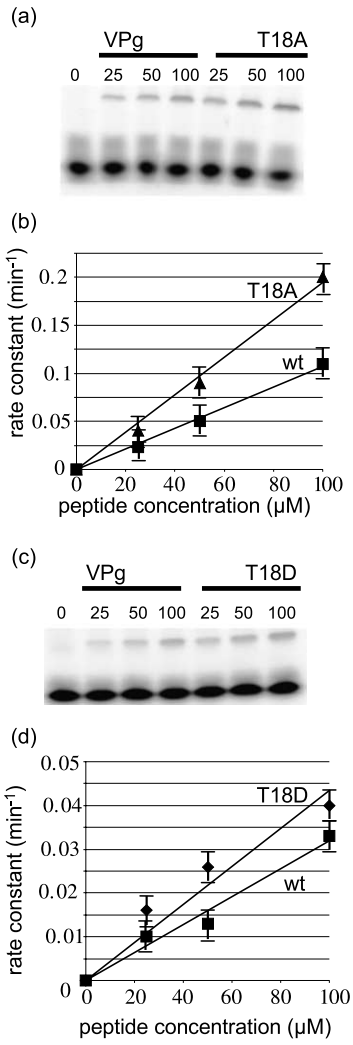

Figure 5. Uridylylation rates of wild-type and mutant VPg molecules that contain mutations not predicted to affect VPg–polymerase interactions. Wild-type and mutant VPg uridylylation rates are compared for various concentrations of (a) and (b) VPg-T18A and (c) and (d) VPg-T18D. Initial rates of the uridylylation reactions for wild-type VPg and (b) VPg-T18A, (d) VPg-T18D are shown and the slope of the line is k_{cat}/K_m.

VPg residues predicted to contact 3D polymerase are necessary for competition of X_7-VPg, which contains seven additional N-terminal amino acid residues derived from 3A

Although the 22 amino acid residue peptide VPg is the substrate for uridylylation by poliovirus polymerase, it is thought to bind to the polymerase as a larger precursor, protein 3AB, which displays higher affinities for polymerase. It is not yet known why the affinity of protein 3AB for the polymerase is greater than that of VPg: at least some of the protein–protein interactions are likely to be the same in the polymerase–VPg and polymerase–protein 3AB complexes for several reasons. First, Val391, Glu382, R379 and Phe377 of the polymerase were identified *via* their roles in 3AB binding, but mutations in these residues reduce the rate of VPg uridylylation.[5] Second, protein 3AB, which is not uridylylated under most assay conditions, is a competitive inhibitor of VPg uridylyation.[11] Finally, the only mutations in protein 3AB that were found to reduce the apparent affinity for the polymerase in two-hybrid assays were VPg (protein 3B) sequences.[20] It is possible that the 3A portion of protein 3AB acts as a stabilizing scaffold, or chaperone, to aid VPg binding to the polymerase.

To ask whether the L6A, K9A and K10A mutations reduced the binding of VPg-containing proteins to polymerase, we synthesized extended versions of

mutant and wild-type VPg to use as competitive inhibitors of wild-type VPg uridylylation. Specifically, wild-type, L6A, K9A and K10A VPg sequences were all synthesized with the seven amino acid residues from the C terminus of protein 3A that are predicted to be cytosolic appended to their N termini; these molecules were termed X_7-VPg peptides. Figure 6(a) shows the inhibition of wild-type VPg uridylylation observed with increasing amounts of X_7-VPg; the observed K_i was approximately 20 μM. Competition of wild-type VPg uridylylation was observed for the X_7-VPg L6A, X_7-VPg K9A and X_7-VPg K10A peptides as well, but higher concentrations were required (Figure 6(b) and (c)). The results of these competition experiments are quantified in Figure 6(c), which indicated apparent K_i values of 20 μM for the wild-type X_7-VPg peptide, and higher K_i values of 70 μM, 45 μM and 42 μM, respectively, for the L6A, K9A and K10A X_7-VPg variants. The changes in K_i values for the mutant X_7-VPg peptides argue that they bind to the polymerase with reduced affinity. Therefore, the simplest explanation of the increased k_{cat}/K_m for the mutant VPg molecules is that the L6A, K9A and K10A mutations reduced the binding affinity of VPg for the viral polymerase. This result is consistent with the computational model, which posited that Leu6, Lys9 and Lys10 directly contact a binding surface on the polymerase.

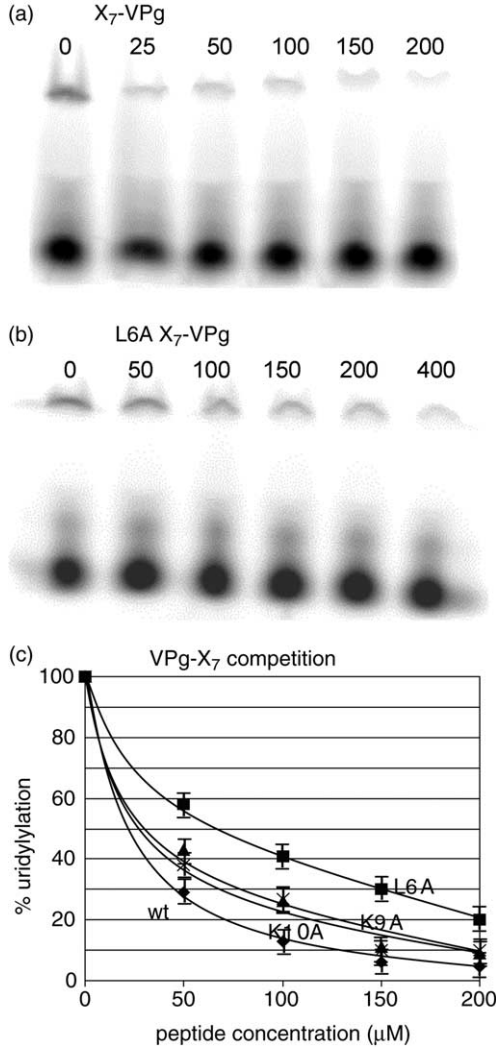

Figure 6. Measuring competition of wild-type VPg uridylyation by wild-type and K9A mutant X_7-VPg. VPg uridylylation reactions were performed in the presence of 1 μM polymerase and 50 μM wild-type VPg substrate. (a) VPg uridylylation in the presence of increasing concentrations of wild-type X_7-VPg. (b) VPg uridylylation in the presence of increasing concentrations of X_7-VPg-L6A. (c) The percentage uridylylation is plotted as a function of inhibitor concentration; 100% uridylylation is the amount of VPg uridylylated in the absence of inhibitor. The inhibitor concentration that corresponds to a 50% reduction of the enzymatic activity is K_i.

Double mutant analysis supports a model in which VPg Lys9 interacts with polymerase Glu382

The computational model shown in Figure 2(b) was developed from previous data that suggested the binding surface on the polymerase included 3D polymerase Phe377, Arg379, Glu382 and Val391. To test the specific hypothesis that VPg residue Lys9 contacts Glu382 on 3D polymerase directly *via* a charge–charge interaction (Figure 2(b)), the uridylylation rates of both wild-type and mutant VPgs by both wild-type and E382A mutant polymerases were monitored. If Lys9 and Glu382 form a direct contact, then the combination of K9A VPg and 3D polymerase E382A mutations should be less disruptive than would be observed if their effects on polymerase–VPg interaction were independent.[18,19,21] On the other hand, if these residues are not in direct contact, the effects of mutations should be additive.

To increase the sensitivity of this double-mutant experiment and to test effects on binding more directly, we measured the competition of wild-type VPg uridylylation by wild-type and K9A mutant X_7-VPg. As shown in Figure 7, the apparent K_i of wild-type X_7-VPg was 20 μM when used as a competitor for wild-type VPg uridylylation by wild-type polymerase, as compared to approximately 45 μM when used as a competitor for wild-type VPg uridylylation with E382A polymerase. When the competitive inhibition assay was repeated using X_7-VPg K9A as the inhibitor, the apparent K_i was 45–50 μM with both wild-type and E382A polymerases (Figure 7(c)). The equivalent K_i values of X_7-VPg with 3D polymerase E382A, X_7-VPg K9A with wild-type 3D polymerase and X_7-VPg K9A with 3D polymerase E382A are consistent with the hypothesis that 3D polymerase Glu382 and VPg Lys9 form a direct contact and do not contact other residues.

Inhibition of VPg uridylylation by wild-type and L6A X_7-VPg with wild-type and F377A polymerase is consistent with direct contact between VPg residue Leu6 and polymerase residue Phe377

Our computational model further suggested that VPg Leu6 contacts a hydrophobic surface on the

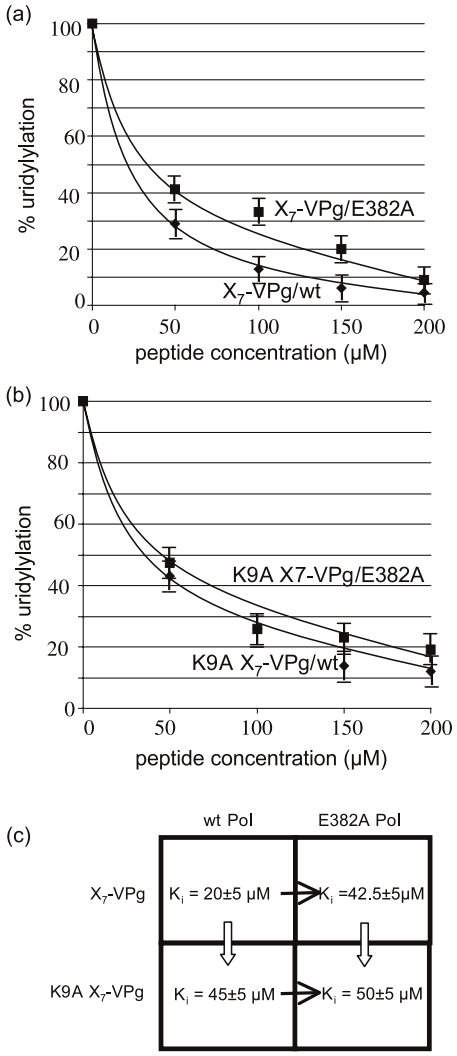

Figure 7. Activity and competitive inhibition assays to probe the VPg Lys9–polymerase Glu382 contact using combinations of mutant and wild-type VPg and polymerase molecules. (a) Competition of wild-type VPg uridylylation by wildtype X_7-VPg is plotted for both wild-type and E382A polymerase. (b) Competition of wild-type VPg uridylylation by the K9A X_7-VPg inhibitor is plotted for wild-type and E382A polymerase. (c) Visualization of the combinations of mutant and wild-type VPg and polymerase molecules shows the change in K_i for each possible combination of polymerase and inhibitor. In the table, the polymerases are the column headings and the row headings are the competitive inhibitors. Each entry shows the K_i that results from the combination of the indicated polymerase and competitive inhibitor. The matching arrows compare the change in K_i for corresponding legs of the double mutant cycle.

polymerase that includes Phe377. The VPg-Leu6–polymerase-Phe377 interaction was probed by measuring the uridylylation of wild-type VPg by wild-type and F377A polymerases in the presence of either wild-type or mutant X_7-VPg-L6A competitors. Competition by wild-type X_7-VPg of uridylylation by wild-type or F377A 3D polymerase resulted in K_i values of 20 μM and 150 μM, respectively, a 7.5-fold effect of the polymerase mutation (Figure 8(a) and (c)). However, competition by mutant X_7-VPg L6A of uridylylation by wild-type or F377A 3D polymerase resulted in K_i values of 70 μM and 210 μM, respectively: only a threefold effect of the polymerase mutation. As before, the largest change in K_i occurred upon the mutation of either Leu6 in VPg or Phe377 in the polymerase; once the first mutation of either residue was in place, the additional change from introducing the second mutation was reduced. We suspect that the Leu6 contact with the hydrophobic pocket on the polymerase that contains Phe377 may

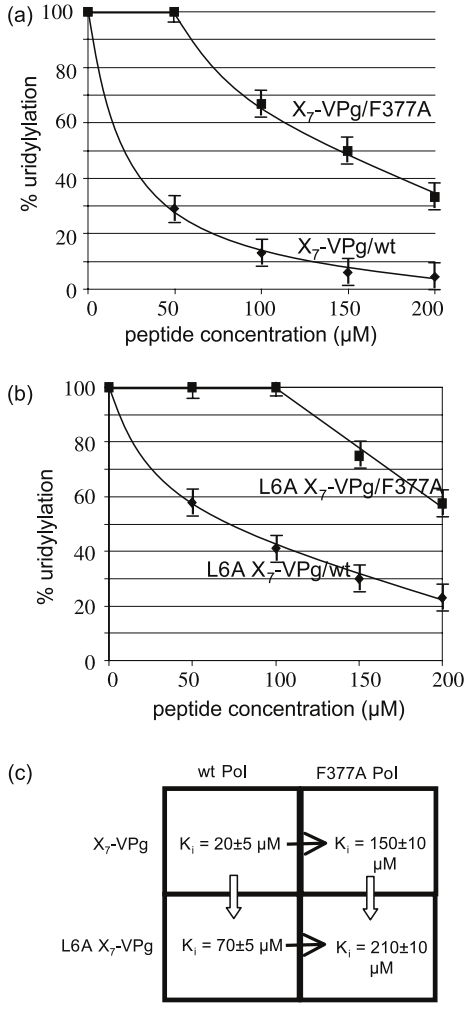

Figure 8. Competitive inhibition assays to probe the VPg-Leu6–polymerase Phe377 contact using a double-mutant cycle. (a) Competition of wild-type VPg uridylylation by wild-type X_7-VPg is plotted for both wild-type and F377A polymerase. (b) Competition of wild-type VPg uridylylation by the L6A X_7-VPg inhibitor is plotted for wild-type and F377A polymerase. (c) Visualization of the double mutant cycle. The change in K_i for each possible combination of polymerase and inhibitor are shown. The polymerases are the column headings and the row headings are the inhibitors. Each entry shows the K_i that results from the combination of the indicated polymerase and competitive inhibitor. The matching arrows compare the change in K_i for corresponding legs of the double-mutant cycle.

involve other polymerase residues as well, making the double-mutant combination more defective than would be expected if Phe377 were the only Leu6 contact, as was observed for the defective charge–charge contact between Lys9 and Glu382. Nonetheless, the initial mutation of the Leu6 residue of VPg rendered mutation of the Phe377 of 3D polymerase less disruptive, consistent with the hypothesis that the contact between Leu6 of VPg and Phe377 of 3D polymerase is direct.

Discussion

Several protein-folding algorithms were employed to create plausible folds and side-chain configurations for VPg. The docking of putative VPg folds, in turn, led to multiple modeled polymerase structures using several docking algorithms. Residues in VPg predicted to have large effects on binding, through salt-bridges or hydrophobic interactions, were identified through inspection and molecular dynamics simulations. On the basis of the posited interactions, several mutant peptides predicted to display weaker affinities for polymerase were designed. The effects of these mutations on VPg binding and uridylylation were then tested individually and in combination with mutant polymerase; the results are consistent with a model using the novel binding of a substrate to the "back" of the polymerase.

It was proposed recently that VPg uridylyation by rhinovirus and foot-and-mouth disease 3D polymerases employs the same binding site as RNA-dependent RNA polymerization, on the "front" of the polymerase.[10,22] In this scenario, the residues on the "back" of the polymerase (Phe377, Arg379, Glu382 and Val391 in the poliovirus 3D polymerase) were postulated not to be part of the actual VPg surface, but instead to define a site of allosteric regulation. If VPg uridylylation and RNA elongation share the same binding site, then mutations that affect structural integrity of the VPg site would be likely to affect nucleotide polymerase activity as well. However, mutation of residues Phe377, Arg379, Glu382 and Val391 in the poliovirus polymerase reduce the rate of VPg binding and uridylylation but not the rate of nucleic acid-primed polymerase activity.[15] Instead, these data and the results presented here are consistent with a model in which the polymerase uses a VPg binding site on the backside of the "thumb" region for uridylylation. As depicted in Figure 9, the VPg binding residues in polymerase (Figure 9(a)) and the polymerase-binding residues in the VPg molecules (Figure 9(b)) are partially conserved for the picornaviruses whose polymerase structures have been solved. Superposition of the known picornavirus polymerase structures (Figure 9(c)) shows that the residues identified in the sequence alignment all correspond to the same surface on the back of the thumb domain. The 2C-cre RNA and any other template RNAs presumably bind to the front of the

(a) 3D polymerase sequences

PV	377	FFRADEKYPFLIHPV	391
HRV 14	376	YFKPDQQFPFLVHPV	390
HRV 16	376	GFRQDDKYKFLIHPT	390
HRV 1B	376	GFKQDEKHTFLIHPT	390
FMDV	388	HFHMDYG-TGFYKPV	402

(b) VPg sequences

PV	3	YTGLPNKKP	11
HRV 14	3	YSGEPKPKT	11
HRV 16	3	YSGEPKPKT	11
HRV 1B	3	YSGEPKEKT	11
FMDV	3	YSGPLERQP	11

(c)

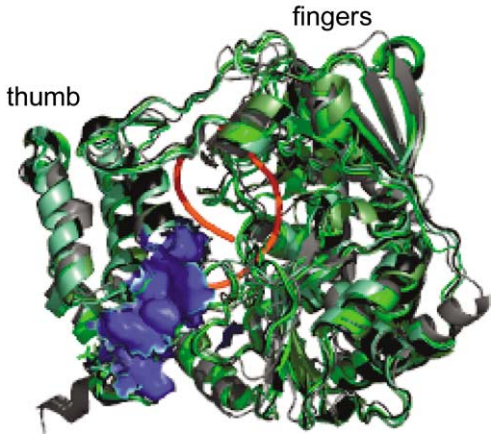

Figure 9. Sequence alignments of VPg-binding residues among picornavirus polymerases (a) whose structures have been solved[10,16,22,33] and (b) polymerase-binding residues in the cognate VPg sequences. (c) The crystal structures of the polymerases of foot-and-mouth disease virus (gray), poliovirus Mahoney type 1 (green), human rhinovirus (HRV) 1B (black), HRV 16 (olive) and HRV 14(light green) are superimposed. The VPg-binding residues identified in the sequence alignment in (a) are shown in blue. In (b), the VPg residues predicted to contact the polymerase directly are shown in blue. The template RNA from the foot-and-mouth disease virus polymerase co-crystal is shown in red.[10]

polymerases, in a position similar to that visualized for the template RNA in the co-crystal between foot-and-mouth-disease 3D polymerase and an RNA template primer.[10] Subsequent to VPg uridylylation on the back of the polymerase, it is likely that the protein primer translocates to a site on the front of a polymerase molecule for further elongation.

If, as we envisage, VPg uridylylation occurs on the back side of the polymerase and the uridylylated primer then moves to the cleft of the palm for further RNA elongation, the question arises as to how this translocation occurs. As Appleby *et al.* observed, it seems unlikely that the uridylylated protein primer spelunks through the nucleotide tunnel.[22] We consider it more likely that the newly uridylylated VPg is transferred from its binding site on the back of one polymerase to the front of the active site on another enzyme (Figure 10).

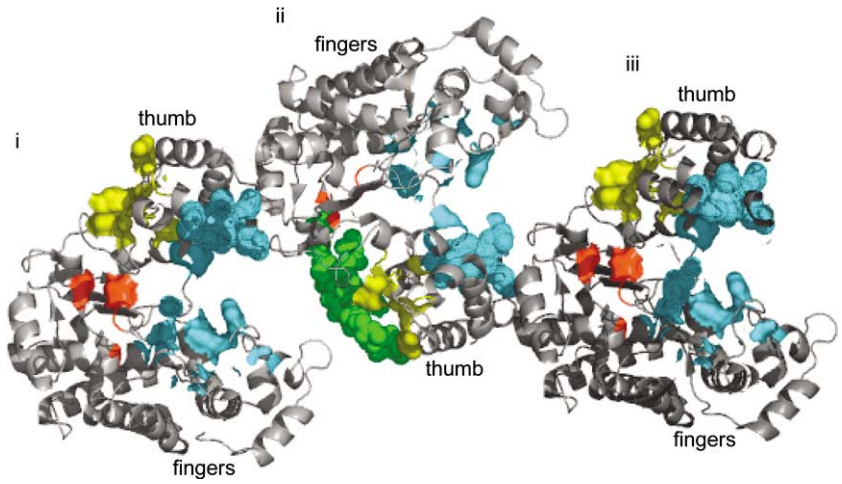

Figure 10. A model for intermolecular VPg uridylylation and RNA elongation. Three polymerase molecules are aligned *via* Interface I[17,23,25]; the VPg main chain is shown in green. Putative RNA-binding residues are teal, the active site is red and the posited VPg binding site is yellow. In this scenario, the back of the "palm" of the polymerase molecule (ii) that uridylylates VPg contacts the "thumb" of the polymerase (i) that will elongate the uridylylated protein primer.

We further propose that this molecular hand-off is facilitated by the favorable geometry of polymerase–polymerase interactions along interface I, a site of polymerase–polymerase interactions observed originally in a partial crystal structure of poliovirus polymerase[17] and disrupted by mutation in the protein crystallized to yield the full-length structure.[16] Mutations predicted to destabilize interface I diminish poliovirus growth,[23,24] reduce the RNA-binding affinity of poliovirus 3D polymerase,[23] and disrupt elongation under conditions in which RNA concentration is rate-limiting.[25] A rate-limiting step between VPg uridylylation and further elongation is consistent with the observation of accumulated VPg-pUpU within infected cells.[26]

The interplay of experimental and computational biology is a powerful approach in the investigation of protein–protein interactions. Computation can provide insights and hypotheses that may be difficult to predict or understand from first principles. In this case, VPg structure prediction was followed by computational docking, which helped select residues for alanine-scanning mutagenesis. Subsequent biochemical data were found to be consistent with the computational model, suggesting that the predicted conformations reflect true protein–protein interactions in the VPg–polymerase complex, in which the VPg–polymerase complexes at two different steps in the initiation of RNA replication, uridylyation of the protein primer and elongation of the protein primer, utilize distinct polymerase surfaces.

Materials and Methods

Peptide folding and simulation

The 22 amino acid residue sequence of VPg was folded computationally using the Rosetta algorithm,[27] as implemented in the Hidden Markov Model Structure Prediction server.[28] The algorithm finds sequence alignments between the protein of interest and similar proteins whose structures can be found in the protein data base,[29] then uses the folds from these similar proteins to create a consensus model for the protein of interest. Similar predicted structures for poliovirus VPg were found using two independent programs: Sybyl, which uses energy-based modeling, and Insight, which employs statistical analysis and sequence alignment to generate models for the structures of proteins of interest†. The amino acid side-chains were added to each main-chain fold with the SCWRL rotamer library.[30] To test the stability of the β-hairpin structure computationally, molecular dynamics simulations were performed using the Rosetta folds as initial structures.[15] These simulations calculate the force on each atom using Newton's second law (Force = Mass × Acceleration) and then integrate to find the motion of the atoms in the system for a very short time-interval. Integrating across many such time-intervals, the range of dynamic motion of the molecular system can be explored.[31]

Polymerase simulation and conformers

Docking calculations were performed initially using a polymerase structure created by threading the polymerase amino acid sequence and partial structure[17] onto the crystal structure of the homologous rabbit haemorrhagic disease virus polymerase using the Deep View threading program‡. Subsequently, the three-dimensional structure of the full-length poliovirus polymerase was obtained as described.[16] All conformations were also created by molecular dynamics simulation of this full-length structure using the GROMOS97 force–field and explicit solvent molecules.

Docking putative VPg folds to polymerase

The FlexX fragment-based algorithm was used to dock VPg to the 3D polymerase.[14] The VPg peptide was split into overlapping fragments of eight residues and each fragment was docked in turn to the 3D polymerase to create a computationally tractable calculation. Docking was performed also with the FlexiDock grid-based

† Tripos SYBYL, Tripos Inc., 1699 South Hanley Rd, St. Louis, MO.

‡ Guex, N., Peitsch, M., Schwede, T. & Diemand, A. (1995). DeepView/Swiss-PdbViewer, GlaxoSmithKline.

algorithm†. To guide the docking, program parameters were set to use any surface that contained a contact within a 7 Å radius of residues Phe377, Arg379, Glu382 and Val391, identified in the polymerase 3AB binding site.[5] The docking of all putative VPg folds to the multiple polymerase conformations modeled resulted in hundreds of potential VPg–3D polymerase complexes. To filter these results, complexes were selected in which all four polymerase residues known to be involved in 3AB binding were within 3 Å of at least one VPg residue.

Computational mutagenesis and simulated effects of alanine scanning

To assay the importance of the putative VPg–polymerase contacts, alanine was substituted *in silico* for individual residues of VPg, and the resulting VPg–polymerase complexes were simulated using molecular dynamics. Specifically, folded wild-type and mutant VPg molecules were placed in the vicinity of the active site visually using Deep View, and the dynamics of each VPg–polymerase pair were measured. Initially, wild-type VPg and polymerase pairs were simulated using 12 different conformations where VPg was placed within 7 Å of the active site. All simulations ran for 2 ns each, using the GROMOS-97 force–field and explicit solvent molecules. Computational residue substitutions were made to produce L6A, K9A and K10A mutant VPg molecules. The dynamics of each of the alternate VPg molecules were simulated from the 12 starting positions used for the wild-type VPg and polymerase simulations. Each simulation was performed for 2 ns under force–field and solvent conditions identical with those used for the wild-type VPg–polymerase experiments.

Polymerase purification and uridylylation assay

Wild-type VPg as well as VPg-L6A, VPg-K9A, and VPg-K10A were synthesized by Global Peptides, Fort Collins, CO. The wild-type and E382A poliovirus polymerases were purified as described.[5] Uridylylation assays were performed with 1 μM polymerase in 50 mM Hepes (pH 7.5), 5 mM DTT, 0.5 mM $MnCl_2$, 30 mM NaCl, 10% (v/v) glycerol, 140 μg/ml of poly(A), 50 μM UTP, 60 μCi/ml of [α-^{32}P]UTP; (3000 Ci/mmol; PerkinElmer Life Sciences) and the indicated concentration of VPg. Reactions were incubated for 30 min at 30 °C, halted by addition of EDTA to 2 mM, and then heated for 4 min at 90 °C. Separation of the ^{32}P-labeled, uridylylated VPg was accomplished using a Tris–Tricine gel (3 M Tris (pH 9.55), 0.3% (w/v) SDS). Visualization was achieved with a Kodak Phosphor screen and the Molecular Dynamics Storm 860 Scanner. The Image Quant program was used for quantification of the VPg signal.

X_7-VPg competition assays

The binding affinities of alternate X_7-VPg molecules were measured *via* competitive inhibition assays. X_7-VPg is identical with VPg, except that it contains an additional seven N-terminal residues (KLFAGHQ) derived from the C terminus of the VPg precursor, protein 3AB. Competition assays were performed as described above, using 50 μM VPg and the specified concentration of each X_7-VPg as the competitive inhibitor.

Acknowledgements

We thank Peter Sarnow and Oliver C. Richards for critical reading of the manuscript, Dan Herschlag for advice on kinetic measurements, Khanh Soliven for help with the uridylylation assay, Scott Hobson and Steve Schultz for introducing us to the properties of X_7-VPg and Ollie Richards and Eckard Wimmer for ongoing discussions. This work was supported by NIH grant AI-42119 (to K.K.), the Stanford Bioinformatics Training Grant, and an American Cancer Society Postdoctoral Fellowhip (to J.F.S.).

References

1. Paul, A. V., van Boom, J. H., Filippov, D. & Wimmer, E. (1998). Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature*, **393**, 280–284.
2. Rueckert, R. R. (1995). Picornaviridae and their replication. In *Virology* (Fields, B. N. *et al*., eds), pp. 609–654, Raven Press, New York.
3. Paul, A. V., Yin, J., Mugavero, J., Rieder, E., Liu, Y. & Wimmer, E. (2003). A "slide-back" mechanism for the initiation of protein-primed RNA synthesis by the RNA polymerase of poliovirus. *J. Biol. Chem.* **278**, 43951–43960.
4. Beese, L. S., Derbyshire, V. & Steitz, T. A. (1993). Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science*, **260**, 352–355.
5. Lyle, J. M., Clewell, A., Richmond, K., Richards, O. C., Hope, D. A., Schultz, S. C. & Kirkegaard, K. (2002). Similar structural basis for membrane localization and protein priming by an RNA-dependent RNA polymerase. *J. Biol. Chem.* **277**, 16324–16331.
6. Towner, J. S., Ho, T. V. & Semler, B. L. (1996). Determinants of membrane association for poliovirus protein 3AB. *J. Biol. Chem.* **271**, 26810–26818.
7. Cheetham, G. M. T. & Steitz, T. A. (1999). Structure of a transcribing T7 RNA polymerase initiation complex. *Science*, **286**, 2305–2309.
8. Huang, H., Chopra, R., Verdine, G. L. & Harrison, S. C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science*, **282**, 1669–1675.
9. Jager, J. & Pata, J. D. (1999). Getting a grip: polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* **9**, 21–28.
10. Ferrer-Orta, C., Arias, A., Perez-Luque, R., Escarmis, C., Domingo, E. & Verdaguer, N. (2004). Structure of foot-and-mouth disease virus RNA-dependent RNA polymerase and its complex with a template-primer RNA. *J. Biol. Chem.* **279**, 47212–47221.
11. Boerner, J. E., Lyle, J. M., Daijogo, S., Semler, B. L., Schultz, S. C., Kirkegaard, K. & Richards, O. C. (2005). Allosteric effects of ligands and mutations on poliovirus RNA-dependent RNA polymerase. *J. Virol.* **79**, 7803–7811.
12. Chen, R. & Weng, Z. (2002). Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins: Struct. Funct. Genet.* **47**, 281–294.
13. Budin, N., Majeux, N. & Caflisch, A. (2001). Fragment-based flexible ligand docking by evolutionary optimization. *Biol. Chem.* **382**, 1365–1372.

14. Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470–489.
15. Lindahl, E., Hess, B., van der Spoel, D. & van Drunen, R. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **7**, 302–317.
16. Thompson, A. A. & Peersen, O. B. (2004). Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO J.* **23**, 3462–3471.
17. Hansen, J. L., Long, A. M. & Schultz, S. C. (1997). Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure*, **5**, 1109–1122.
18. Jin, L. & Wells, J. A. (1994). Dissecting the energetics of an antibody-antigen interface by alanine shaving and molecular grafting. *Protein Sci.* **3**, 2351–2357.
19. Wells, J. A. (1996). Binding in the growth hormone receptor complex. *Proc. Natl Acad. Sci. USA*, **93**, 1–6.
20. Xiang, W., Cuconati, A., Paul, A. V., Cao, X. & Wimmer, E. (1995). Molecular dissection of the multifunctional poliovirus RNA-binding protein 3AB. *Rna*, **1**, 892–904.
21. Fersht, A. (1984). *Enzyme Structure and Mechanism*, W.H. Freeman Co., New York.
22. Appleby, T. C., Luecke, H., Shim, J. H., Wu, J. Z., Cheney, I. W., Zhong, W. *et al*. (2005). Crystal structure of complete rhinovirus RNA polymerase suggests front loading of protein primer. *J. Virol.* **79**, 277–288.
23. Hobson, S. D., Rosenblum, E. S., Richards, O. C., Richmond, K., Kirkegaard, K. & Schultz, S. C. (2001). Oligomeric structures of poliovirus polymerase are important for function. *EMBO J.* **20**, 1153–1163.
24. Pathak, H. B., Ghosh, S. K., Roberts, A. W., Sharma, S. D., Yoder, J. D., Arnold, J. J. *et al*. (2002). Structure–function relationships of the RNA-dependent RNA polymerase from poliovirus (3Dpol). A surface of the primary oligomerization domain functions in capsid precursor processing and VPg uridylylation. *J. Biol. Chem.* **277**, 31551–31562.
25. Lyle, J. M., Bullitt, E., Bienz, K. & Kirkegaard, K. (2002). Visualization and functional analysis of RNA-dependent RNA polymerase lattices. *Science*, **296**, 2218–2222.
26. Crawford, N. M. & Baltimore, D. (1983). Genome-linked protein VPg of poliovirus is present as free VPg and VPg-pUpU in poliovirus-infected cells. *Proc. Natl Acad. Sci. USA*, **80**, 7452–7455.
27. Simmons, K. T., Kooperberg, C., Huang, E. & Baker, D. (1997). Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *J. Mol. Biol.* **268**, 209–225.
28. Bystroff, C. & Shao, Y. (2002). Fully automated ab initio protein structure prediction using I-SITES, HMMSTR and ROSETTA. *Bioinformatics*, **18**, S54–S61.
29. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. *et al*. (2000). The Protein Data Bank. *Nucl. Acids Res.* **28**, 235–242.
30. Canutescu, A. A., Shelenkov, A. A. & Dunbrack, R. L., Jr (2003). A graph theory algorithm for protein side-chain prediction. *Protein Sci.* **12**, 2001–2014.
31. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. (1995). GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56.
32. Hope, D. A., Diamond, S. E. & Kirkegaard, K. (1997). Genetic dissection of interaction between poliovirus 3D polymerase and viral protein 3AB. *J. Virol.* **71**, 9490–9498.
33. Love, R. A., Maegley, K. A., Yu, X., Ferre, R. A., Lingardo, L. K., Diehl, W. *et al*. (2004). The crystal structure of the RNA-dependent RNA polymerase from human rhinovirus: a dual function target for common cold antiviral therapy. *Structure*, **12**, 1533–1544.

Edited by J. Doudna

(Received 25 July 2005; received in revised form 12 December 2005; accepted 12 December 2005)
Available online 5 January 2006