Structure of the RNA-dependent RNA polymerase of poliovirus

Jeffrey L Hansen[1], Alexander M Long[2] and Steve C Schultz[1]*

Background: The central player in the replication of RNA viruses is the viral RNA-dependent RNA polymerase. The 53 kDa poliovirus polymerase, together with other viral and possibly host proteins, carries out viral RNA replication in the host cell cytoplasm. RNA-dependent RNA polymerases comprise a distinct category of polymerases that have limited sequence similarity to reverse transcriptases (RNA-dependent DNA polymerases) and perhaps also to DNA-dependent polymerases. Previously reported structures of RNA-dependent DNA polymerases, DNA-dependent DNA polymerases and a DNA-dependent RNA polymerase show that structural and evolutionary relationships exist between the different polymerase categories.

Results: We have determined the structure of the RNA-dependent RNA polymerase of poliovirus at 2.6 Å resolution by X-ray crystallography. It has the same overall shape as other polymerases, commonly described by analogy to a right hand. The structures of the 'fingers' and 'thumb' subdomains of poliovirus polymerase differ from those of other polymerases, but the palm subdomain contains a core structure very similar to that of other polymerases. This conserved core structure is composed of four of the amino acid sequence motifs described for RNA-dependent polymerases. Structure-based alignments of these motifs has enabled us to modify and extend previous sequence and structural alignments so as to relate sequence conservation to function. Extensive regions of polymerase–polymerase interactions observed in the crystals suggest an unusual higher order structure that we believe is important for polymerase function.

Conclusions: As a first example of a structure of an RNA-dependent RNA polymerase, the poliovirus polymerase structure provides for a better understanding of polymerase structure, function and evolution. In addition, it has yielded insights into an unusual higher order structure that may be critical for poliovirus polymerase function.

Addresses: [1]Campus Box 215, Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309, USA and [2]Vertex Pharmaceuticals, 130 Waverly Street, Cambridge, MA 02139, USA.

*Corresponding author.
E-mail: Schultz@MMOL.colorado.edu

Introduction

Poliovirus, a small positive-strand RNA virus, is an important prototype for other picornaviruses such as rhinovirus, hepatitis A virus, coxsackie virus, echoviruses, foot and mouth disease virus, and encephalomyocarditis virus. The 7 500 nucleotide single-strand RNA genome of poliovirus contains one long open reading frame which is translated into a 247 kDa polyprotein. The C-terminal 461 amino acid polypeptide is the RNA-dependent RNA polymerase (also referred to as 3Dpol) [1,2]. The polymerase, together with other viral proteins and, possibly, host proteins, carries out viral RNA replication in the cytoplasm of infected cells to generate new template, messenger, and viral RNAs from the original infecting RNA (reviewed in [3,4]).

Although poliovirus RNA replication occurs in large membrane-associated replication complexes in cells [5,6], the polymerase alone is active in the absence of any other proteins *in vitro* [7–10]. Many aspects of poliovirus polymerase activity, including RNA binding, NTP binding, polymerization of nucleotides, RNA strand displacement, and interactions with other viral proteins have been investigated biochemically as well as genetically (reviewed in [3,4]). RNA-binding and polymerization activities are highly cooperative with respect to polymerase concentration, suggesting that polymerase–polymerase interactions are important for function [11]. Polymerase–polymerase interactions have also been observed by chemical cross-linking in solution [11], and in the yeast two-hybrid genetic interaction assay [12].

Amino acid sequence similarities indicate that poliovirus polymerase is related structurally and evolutionarily to other RNA-dependent RNA polymerases [13–15], and perhaps also to RNA-dependent DNA polymerases [16–18] and, more distantly, to DNA-dependent polymerases [19]; Figure 1 illustrates these relationships. The relationships between RNA-dependent RNA and RNA-dependent DNA polymerases have been recently questioned, however, due to the very limited nature of the sequence similarities [20]. Crystal structures have been reported for polymerases from three of the four categories. All of these

Figure 1
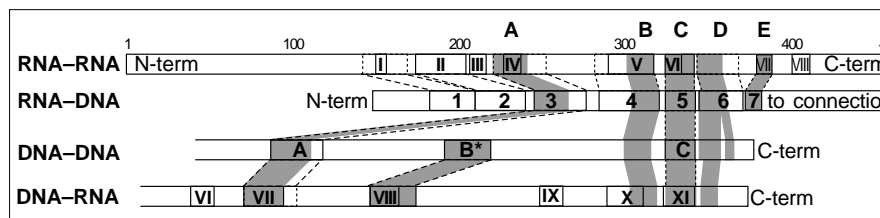
Amino acid sequence alignments for the four categories of nucleic acid polymerase. The numbers at the top of the figure refer to residue numbers of poliovirus polymerase. Sequence motifs within each category are enclosed with solid lines. These motifs are those described by Koonin [14] for positive-strand RNA viruses (RNA–RNA); Xiong and Eickbush [18] for RNA-dependent DNA polymerases (RNA–DNA); DeLarue et al. [19] for Pol I and Polα type DNA-dependent DNA polymerases (DNA–DNA); and Masters et al. [61] for the small single-subunit DNA-dependent RNA polymerases (DNA–RNA). Sequence-based alignments between the polymerase categories are designated by dashed lines; the sequence-based alignments are those of Xiong and Eickbush [18] for RNA-dependent polymerases; DeLarue et al. [19] for DNA-dependent polymerases and for all four categories of polymerases. The structure/sequence motifs described in this paper are indicated by shading and are labeled A, B, C, D, and E which are the designations of Poch [17]. These alignments expand on sequence and structural alignments described previously [23,44]. Note that DNA-dependent polymerases contain a motif B (labeled B*) that is different from the motif B as designated by Poch for RNA-dependent polymerases; motif B as labeled here refers to a structure/sequence motif that occurs in all four categories of polymerase and corresponds to the C-terminal portion of motif B of Poch [17].

structures resemble a right hand, an analogy that was first used to describe the structure of the large (Klenow) fragment of DNA polymerase I [21]. The 'fingers' and 'thumb' subdomains of the two DNA-dependent polymerases, Klenow and T7 RNA polymerase, are similar to each other [22] but are very different from the corresponding subdomains of reverse transcriptases [23]. The 'palm' subdomains contain a core structure that is similar in each of these three categories. The core structure of the palm subdomain contains four of the amino acid sequence motifs described for RNA-dependent polymerases [23]. Intriguingly, as has been noted previously [24–26], the conserved core structure of polymerase palm subdomains exhibits a protein fold that is strikingly similar to that of the RNA-recognition motif (RRM) found in proteins involved in splicing, several ribosomal proteins, and of a variety of other proteins [27]. Rat DNA polymerase β is an exception to the polymerases discussed here, in that it is structurally distinct [28–30] and apparently belongs to a different evolutionary superfamily of nucleotidyl transferases [31].

We have determined the structure of the RNA-dependent RNA polymerase of poliovirus. Comparison of this structure with structures from each of the other three categories provides additional insights into polymerase structure and evolution. The structure of poliovirus polymerase also provides new insights into its activities in solution and its functions in viral replication.

Results and discussion

Crystallization, data collection, and structure determination

Poliovirus polymerase was expressed in *Escherichia coli* using an inducible T7-based expression system, and purified as described in the Materials and methods section. Typical yields were 10–20 mg of > 99% pure poliovirus polymerase from 15 g cells (wet weight). Crystals with two different morphologies (needle shaped and hexagonally shaped) were initially grown by gradually reducing NaCl concentrations from 0.5 M to 0.1 M. Larger, higher quality hexagonally shaped crystals of poliovirus polymerase grew from solutions of 3–5 mg/ml polymerase, 1.2 M $CaCl_2$ in 2–6 weeks. SDS–PAGE of washed crystals verified that they contained exclusively full-size poliovirus polymerase (data not shown). Crystals were harvested into solutions containing 1.0–1.2 M $CaCl_2$ without loss of diffraction.

The crystals of poliovirus polymerase are trigonal, space group $P3_221$ with cell dimensions a = b = 88.1 Å, c = 158.5 Å. The crystals contain one molecule per asymmetric unit with V_M = 3.3 Å³/dalton. Diffraction is anisotropic with reflections to 2.4 Å resolution along a* and b* and to 2.8 Å resolution along c*; therefore, we are describing the structure as a 2.6 Å resolution structure. Native and derivative data were collected using a Rigaku R-AXIS IIC. Data used for the initial structure determination were collected from crystals at 20°C. Data used for refinement of the structure were collected from crystals cooled to –25°C, which reduced decay of high resolution reflection intensities during data collection; crystals used in the –25°C collections were soaked in solutions that contained 20% glycerol or 2 M glucose to prevent freezing. These conditions for data collection gave rise to small changes in cell dimensions (a = b = 87.3, c = 158.5 for 20°C and a = b = 88.1, c = 158.5 for –25°C). Although slight shifts within the unit cell occurred as a result of these different conditions, no significant changes in the overall structure were observed. Although much effort was directed toward quick freezing the crystals, a reduction in the quality of diffraction at high resolution was observed in all attempts at quick freezing. The data were reduced to reflection intensities using either the Molecular Structure Corporation (MSC) software or DENZO and SCALEPACK [32]. Statistics for the –25°C native data used in refinement of the structure are listed in Table 1.

The structure was determined using multiple isomorphous heavy atom replacement (MIR) methods. Statistics for the

Table 1

Statistics for data (≥ 0σ) collected from crystals of poliovirus polymerase cooled to –25°C*.

Resolution	Average I	Average error	R factor (2σ data)	Completeness (%)
99.00–5.91	7015.9	340.0	0.034	97.0
5.91–4.69	4395.0	225.3	0.039	99.0
4.69–4.10	4377.3	238.9	0.044	98.4
4.10–3.72	3055.8	192.3	0.054	98.4
3.72–3.46	2136.9	178.0	0.069	98.7
3.46–3.25	1347.0	169.1	0.087	98.6
3.25–3.09	835.2	164.5	0.113	98.8
3.09–2.96	596.8	164.8	0.132	98.5
2.96–2.84	460.4	164.9	0.149	98.6
2.84–2.74[†]	343.6	157.9	0.165	98.3
2.74–2.66[†]	266.9	152.2	0.177	97.9
2.66–2.58[†]	216.2	140.4	0.193	97.1
2.58–2.51[†]	165.6	132.8	0.208	95.0
2.51–2.45[†]	151.8	128.0	0.207	85.3
2.45–2.40[†]	134.5	122.1	0.198	66.0

Data were reduced using DENZO. [†]Because diffraction is anisotropic, with reflections to 2.8 Å resolution along a and b* and to 2.4 Å resolution along c*, we are referring to this as a 2.6 Å resolution structure. R factor = $\Sigma |I - <I>|/\Sigma I$.

Table 2

Statistics for heavy atom derivatives used in the structure determination.

Crystal	Resolution	R_{sym}* (all data)	R_{cross}[†]	Occupancies		Phasing[‡] power	R_{Cullis}§
				Site 1	Site 2		
Native	2.6 Å	8.5%	–	–	–	–	–
$MeHgC1_a$	2.7 Å	9.0%	17.6%	0.120	0.049	1.8	0.52
$MeHgC1_b$	2.7 Å	8.9%	18.3%	0.111	0.022	1.9	0.49
$MeHgC1_c$	2.7 Å	9.1%	18.6%	0.177	0.040	1.5	0.65
$HgC1_{2a}$	2.7 Å	10.1%	21.0%	0.120	0.083	1.3	0.67
$HgC1_{2b}$	2.8 Å	8.8%	20.2%	0.108	0.084	1.4	0.58
$TbC1_3$	3.2 Å	8.0%	13.2%	0.084	–	0.6	0.89

*$R_{sym} = \Sigma |I - <I>|/\Sigma I$. [†]$R_{cross} = \Sigma |F_{PH} - F_P|/\Sigma |F_{PH}|$. [‡]Phasing power (centric) = $\Sigma |F_{Hcalc}|/\Sigma |F_{PH}| - |F_P + F_{Hcalc}||$.
§R_{Cullis}(centric) = $\Sigma |F_{PH} - |F_P + F_{Hcalc}||/\Sigma |F_{PH} - F_P|$.

heavy atom derivatives used in calculating phases (using MLPHARE [33]) are listed in Table 2. The two observed mercury sites correspond to two of the five cysteine residues (Cys212 and Cys418) of poliovirus polymerase; mercury sites at Cys290 and Cys96 were subsequently observed in difference Fourier maps. These four mercury sites were also useful in fitting and verifying the structure. The MIR phases were modified by solvent flattening and

histogram matching using SQUASH [34] or dm [35]. The structure was fit into these electron density maps using O [36] and refined using X-PLOR [37]. The current model contains residues 12–37, 67–97, 181–266, 291–461, and 24 water molecules; the remaining regions of the protein are disordered in the crystals. The structure is currently refined to an R factor of 21.8% and an R_{free} [38] of 27.0%. Root mean square (rms) deviations from ideal bond lengths are 0.014 Å and from ideal bond angles are 2.18°. Figure 2 shows an (F_o–F_c) simulated annealed omit electron-density map that includes the highly conserved YGDD (Tyr–Gly–Asp–Asp) motif of RNA-dependent RNA polymerases. The mean B factor for all non-hydrogen atoms is 30.8 Å² with a standard deviation of 11.6 Å² excluding residues flanking disordered regions which generally have B factors of >60 Å². Ramachandran analysis (calculated with PROCHECK) shows that 86.2% (238) of the residues

Figure 2

Stereoview of a simulated annealed (F_o–F_c) omit electron-density map (1.5σ) of the active site of poliovirus polymerase. The highly conserved aspartate of motif A and the highly conserved YGDD (Tyr–Gly–Asp–Asp) sequence of motif C are labeled. Also included is a modeled calcium ion bound between Asp233 and Asp329 of motifs A and C, respectively.
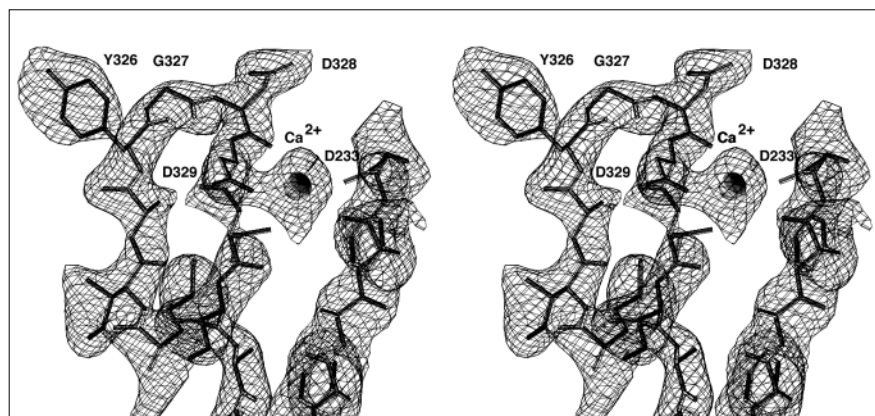
are in most favored regions; 12.3% (34) of the residues are in allowed regions; 1.4% (4) of the residues are in generous regions; and 0.0% of the residues are in disallowed regions. Nearly all of the residues in the allowed and generous regions are located in peptide segments near the disordered portions of the protein. Although backbone electron density is clear in these areas, these residues were often difficult to fit precisely. Residues 12–22 were especially difficult to fit. Two residues in the generously allowed regions (Lys359 and Lys375) are near well-ordered regions of the protein, but occur in surface loops and are not as well ordered as neighboring residues.
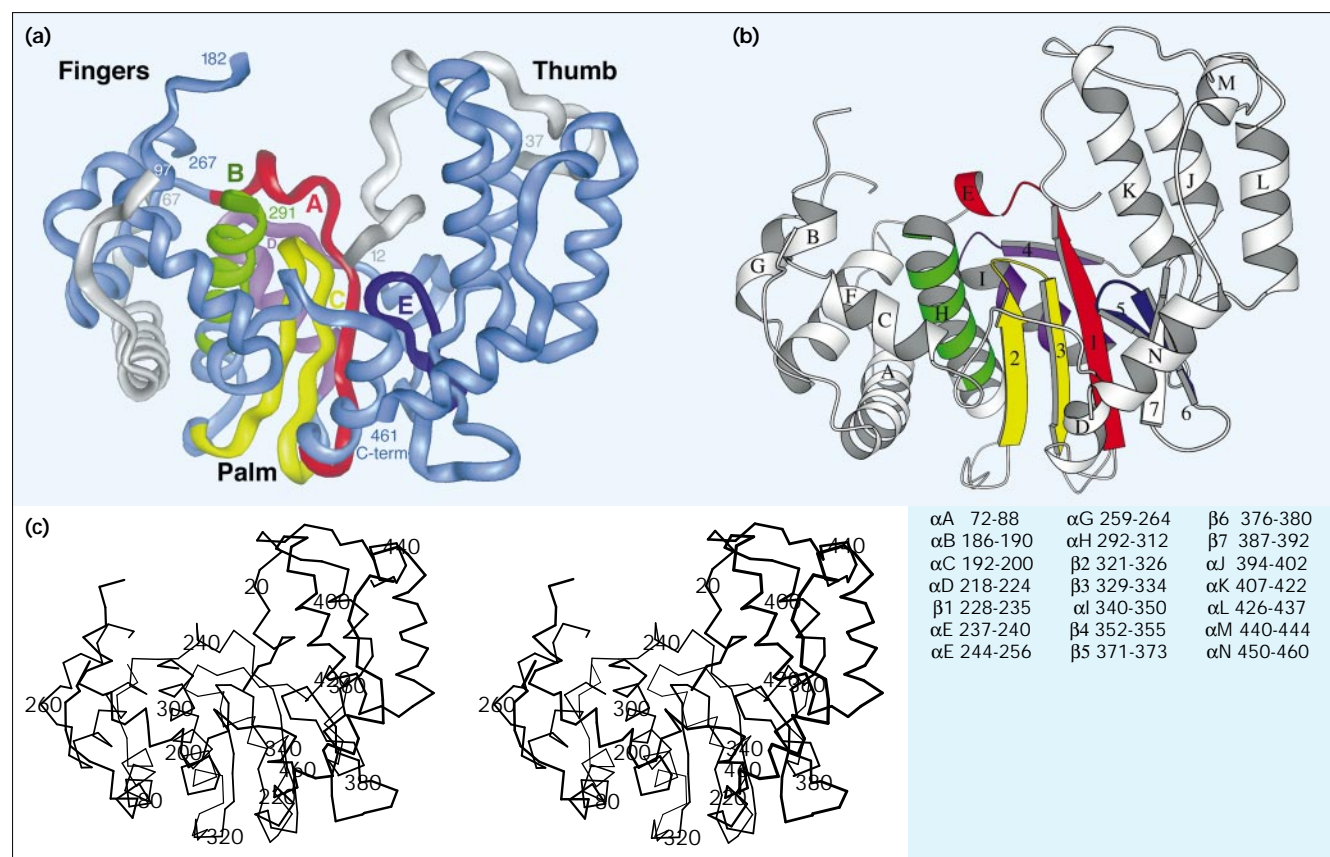
General features of the poliovirus polymerase structure

The poliovirus polymerase contains recognizable palm, thumb, and fingers subdomains (Figure 3). The palm subdomain contains five of the amino acid sequence motifs of RNA-dependent polymerases, referred to as A, B, C, D, and E [17]. The thumb subdomain is composed mostly of residues C-terminal of the palm subdomain and is largely

α-helical. In addition, a polypeptide strand from the very N-terminal region of the protein interacts with the top of the thumb (shown in white in Figure 3a); this unusual feature will be discussed later. The fingers subdomain of poliovirus polymerase is composed of two polypeptide segments, a larger segment that precedes motif A and a smaller segment composed of residues between motifs A and B of the palm subdomain. Although the top of the fingers subdomain is disordered in the crystals, the bottom portions are ordered and clearly resolvable.

In addition to palm, thumb, and fingers subdomains, the poliovirus polymerase contains structural elements composed of residues N-terminal of the fingers subdomain that have no counterpart in the other categories of polymerases. The ordered portions of the N-terminal regions are shown in white in Figure 3a. Residues 12–37 constitute the N-terminal strand of the thumb subdomain and residues 67–97 form an α helix (αA) beneath the fingers subdomain. The residues that join these two regions

Figure 3

αA 72-88	αG 259-264	β6 376-380
αB 186-190	αH 292-312	β7 387-392
αC 192-200	β2 321-326	αJ 394-402
αD 218-224	β3 329-334	αK 407-422
β1 228-235	αI 340-350	αL 426-437
αE 237-240	β4 352-355	αM 440-444
αE 244-256	β5 371-373	αN 450-460

Structure of the RNA-dependent RNA polymerase of poliovirus. **(a)** Ribbon representation; the thumb, fingers, and palm subdomains are labeled. The structure/sequence motifs of the palm subdomain are labeled and colored as follows: A in red, B in green, C in yellow, D in light purple, and E in dark purple; in white are the two N-terminal regions (residues 12–37 and 67–97) that are ordered in the crystals. Residues flanking disordered regions are also numbered. **(b)** Schematic of the structure with structure/sequence motifs, colored as described for (a). Residues contained within α helices and β strands are listed at the bottom of the figure. **(c)** Stereoview carbon trace.

(residues 38–66) are disordered in the crystals. This region of disorder is unfortunate in that the possible ways of connecting the N-terminal regions are the subject of great interest, as will be discussed below.

The palm subdomain and polymerase motifs
In poliovirus polymerase, the polypeptide regions that correspond to four of the amino acid sequence motifs of RNA-dependent polymerases (A–D in Figure 1) fold into a structure that forms the core of the palm subdomain (Figure 3). This core structure consists of two α helices that pack beneath a four-stranded antiparallel β sheet. The strands of the antiparallel β sheet are composed of residues from motifs A, C, and part of D, while the α helices are composed of residues from motif B and the remainder of motif D. This same core structure is present in the palm subdomains of all four categories of polymerases. Figure 4a shows structures of polymerases from each of the four categories with the structure/sequence motifs highlighted. A fifth motif, motif E, which occurs in RNA-dependent but not DNA-dependent polymerases, packs between the palm and thumb subdomains. In correlating individual residues of the amino acid sequence motifs with their positions in the structures, some significant modifications of previous amino acid sequence alignments are necessary. Revised structure-based sequence alignments for Klenow, T7 RNA polymerase, HIV-1 reverse transcriptase, and poliovirus polymerase are shown in Figure 4b.

Motif A of poliovirus polymerase (red in Figures 3 and 4) forms one of the four β strands (β1) of the core structure followed by a short helical turn (αE) at the C-terminal end of the motif. RNA-dependent DNA polymerases have a similar short helical turn near the C-terminal end of motif A [25,26], whereas the corresponding residues of DNA-dependent polymerases constitute the beginning of a longer α helix [21,22]. Near the end of the β strand of motif A just preceding the helix is the completely conserved aspartate that has been aligned in all previous sequence and structure comparisons; this residue is expected to coordinate catalytically essential metal ions [19,39–41]. Following the completely conserved aspartate, all reported sequence alignments have introduced a single amino acid gap into motif A of RNA-dependent RNA polymerases relative to the other three categories of polymerases [17–19]. The structure of poliovirus polymerase shows that this gap in the alignment is not correct.
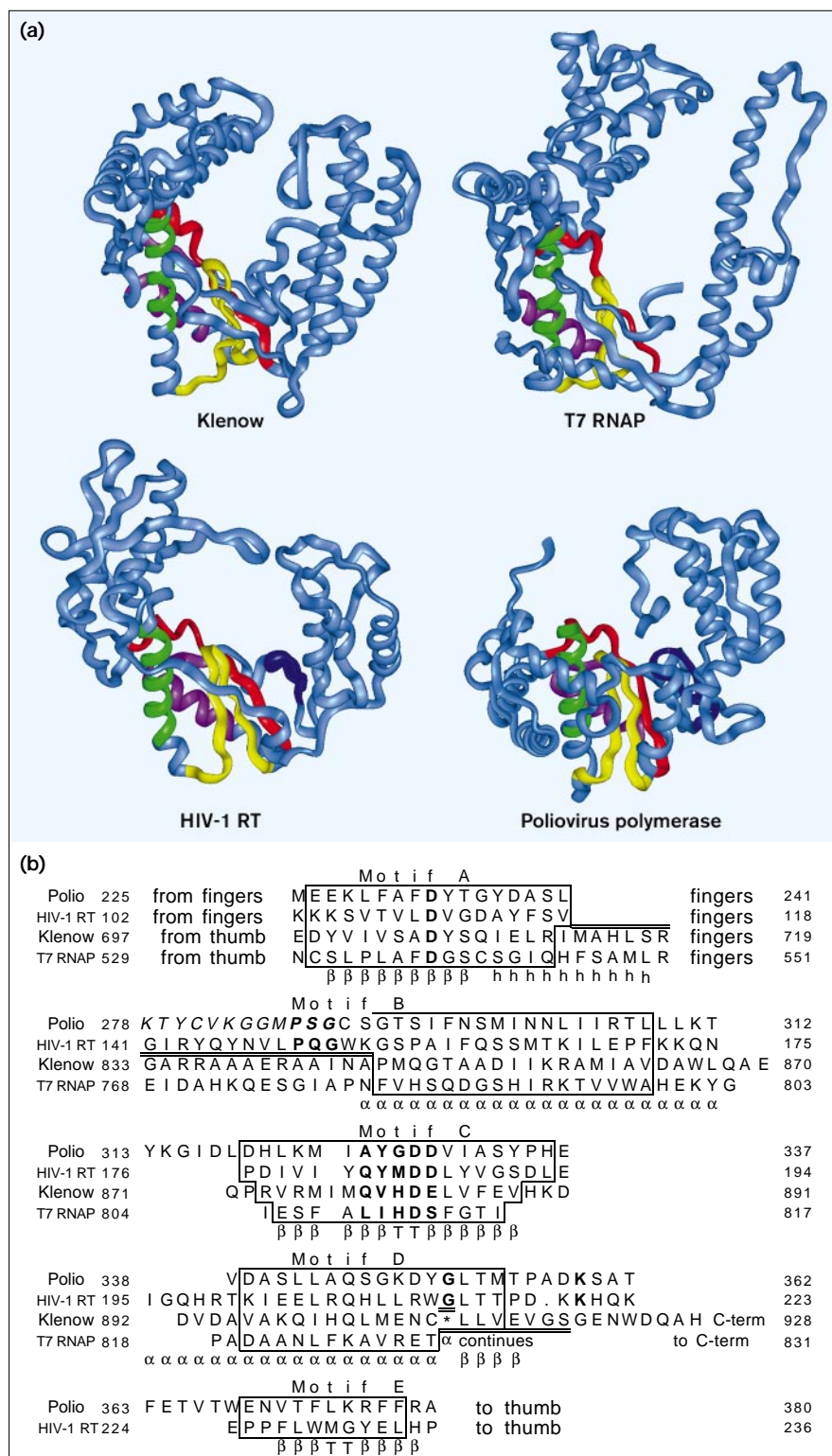
Modifying the sequence alignments to exclude the gap in motif A of RNA-dependent RNA polymerases has important consequences. First, Ser240 of poliovirus polymerase now corresponds with Ser117 of HIV-1 reverse transcriptase (Figure 4b). In both structures, the sidechain of this serine caps the end of the helical turn of motif A by hydrogen bonding to a backbone carbonyl. Second, in our modified

sequence alignment, Tyr115 of HIV-1 reverse transcriptase aligns with Asp238 of poliovirus polymerase. Sequence comparisons show that this position is almost completely conserved as phenylalanine or tyrosine in reverse transcriptases and is almost completely conserved as aspartate in RNA-dependent RNA polymerases of positive-strand RNA viruses. We believe that this difference derives from one of the primary functional differences between RNA-dependent DNA polymerases and RNA-dependent RNA polymerases: discrimination between NTPs and dNTPs. Indeed, Georgiadis *et al.* [25] have suggested that, in reverse transcriptases, a conserved aromatic sidechain at this position discriminates against NTPs in favor of dNTPs by sterically interfering with the 2′ hydroxyl of nucleoside triphosphates. Mutation of this residue (Phe155) in Moloney murine leukemia virus (MMLV) reverse transcriptase indeed correlates with altered discrimination between dNTPs and NTPs [42]. An aspartate at this position in RNA-dependent RNA polymerases, as observed in our structure, could favor NTPs over dNTPs, perhaps by interacting directly with the 2′ hydroxyl group of an incoming NTP.

Motif B of poliovirus polymerase (green in Figures 3 and 4) forms one of two α helices that pack beneath the four-stranded antiparallel β sheet of the polymerase core structure. The N-terminal portion of motif B, which includes the highly conserved PXG (Pro–X–Glu, where X is any residue) sequence of RNA-dependent polymerases, is disordered in the poliovirus polymerase crystals. The C-terminal portion of motif B, however, is ordered and is part of a long α helix (αH). The helix containing motif B is tilted differently in each of the four categories of polymerase; as a result, the C-terminal end of this helix occupies rather different positions in the different polymerases. Importantly, however, a portion of this helix is similarly positioned in all four categories of polymerase: it is in this region that all four motifs come together to form the 'heart' of the core structure of the polymerase palm subdomains. Amino acid sidechains from all four of the motifs contribute to this heart region of the palm subdomain.

Within motif B (as within motif A), previous sequence alignments had introduced a single amino acid gap in RNA-dependent RNA polymerases with respect to the other categories of polymerase [17–19]. Although a compelling sequence alignment arises by introducing this gap, it is inconsistent with structural alignments. As within motif A, removing this gap has important consequences. For example, it is now apparent that the most conserved asparagine in RNA-dependent RNA polymerases (Asn297 in poliovirus polymerase) aligns structurally with a highly conserved aromatic residue in RNA-dependent DNA polymerases (Phe160 of HIV-1 reverse transcriptase). Interestingly, in poliovirus polymerase, Asn297 hydrogen bonds

Figure 4

(a)

Klenow   T7 RNAP

HIV-1 RT   Poliovirus polymerase

(b)

Motif A

Polio	225	from fingers	M E E K L F A F **D** Y T G Y D A S L	fingers	241
HIV-1 RT	102	from fingers	K K K S V T V L **D** V G D A Y F S V	fingers	118
Klenow	697	from thumb	E D Y V I V S A **D** Y S Q I E L R I M A H L S R	fingers	719
T7 RNAP	529	from thumb	N C S L P L A F **D** G S C S G I Q H F S A M L R	fingers	551

β β β β β β β β   h h h h h h h h h h

Motif B

Polio	278	*K T Y C V K G G M* **P S G** C S G T S I F N S M I N N L I I R T L L L K T		312
HIV-1 RT	141	G I R Y Q Y N V L **P Q G** W K G S P A I F Q S S M T K I L E P F K K Q N		175
Klenow	833	G A R R A A A E R A A I N A P M Q G T A A D I I K R A M I A V D A W L Q A E		870
T7 RNAP	768	E I D A H K Q E S G I A P N F V H S Q D G S H I R K T V V W A H E K Y G		803

α α α α α α α α α α α α α α α α α α α α α α

Motif C

Polio	313	Y K G I D L D H L K M I A **A Y G D D** V I A S Y P H E		337
HIV-1 RT	176	P D I V I Y Q **Y M D D** L Y V G S D L E		194
Klenow	871	Q P R V R M I M **Q V H D E** L V F E V H K D		891
T7 RNAP	804	I E S F A **L I H D S** F G T I		817

β β β   β β β T T β β β β β β

Motif D

Polio	338	V D A S L L A Q S G K D Y **G L** T M T P A D **K** S A T		362
HIV-1 RT	195	I G Q H R T K I E E L R Q H L L R W **G L** T T P D . K **K** H Q K		223
Klenow	892	D V D A V A K Q I H Q L M E N C * L L V E V G S G E N W D Q A H C-term		928
T7 RNAP	818	P A D A A N L F K A V R E T α continues to C-term		831

α α α α α α α α α α α α α α α α α α α   β β β β

Motif E

Polio	363	F E T V T W E N V T F L K R F F R A to thumb		380
HIV-1 RT	224	E P P F L W M G Y E L H P to thumb		236

β β β T T β β β β

Polymerase structure/sequence motifs in representative structures from each of the four categories of polymerases. **(a)** Structures of the polymerase domains of the large (Klenow) fragment of *E. coli* DNA polymerase I [21], a DNA-dependent DNA polymerase; T7 RNA polymerase (T7 RNAP) [22], a DNA-dependent RNA polymerase; HIV-1 reverse transcriptase (HIV-1 RT) [23], an RNA-dependent DNA polymerase; and poliovirus polymerase, an RNA-dependent RNA polymerase. Each is positioned with the thumb subdomain to the right and the fingers subdomain to the left. The structure/sequence motifs are color coded: A in red, B in green, C in yellow, D in purple, and E in dark purple. **(b)** Structure-based sequence alignments for poliovirus polymerase, HIV-1 reverse transcriptase, the large (Klenow) fragment of DNA polymerase I, and T7 RNA polymerase. Also shown is motif E for the RNA-dependent polymerases. Motif B refers to the structure/sequence motif shown in green in part (a) of this figure and not motif B* of DNA-dependent polymerases; α-helical and β-strand regions of the motifs are indicated on the bottom line. The alignments described here were defined by difference distance calculations and comparisons of secondary and tertiary structures that will be described elsewhere. For T7 RNA polymerase, alignment of motif A is based on the assumption that the aspartate in this motif will align with the completely conserved aspartate in other polymerases. This assumption shifts the sequence assignments for this structure by one residue. In Klenow, motif D contains an insert with the sequence TRLDVP which is designated by an asterisk in the alignment. The extents of structural agreements are boxed with a solid line; the most conserved residues are shown in bold; doubled lines indicate regions in which the structures are different in the different polymerase categories; italic text indicates regions that cannot be compared due to disorder in the crystals. These alignments differ significantly from previous sequence [17–19] and structure alignments [44].

with the conserved Asp238 of motif A, perhaps positioning this residue to discriminate between NTPs and dNTPs, as discussed previously. In HIV-1 reverse transcriptase,

Tyr115 interacts with Phe160 such that the hydrophobic Phe160–Tyr115 interaction in HIV-1 reverse transcriptase corresponds to the Asn297–Asp238 interaction in poliovirus

polymerase. This co-variation in sequence probably arises from the important functional difference of using NTPs rather than dNTPs.

Motif C of poliovirus polymerase (yellow in Figures 3 and 4) forms a β-turn-β structure, which is part of the antiparallel β sheet of the polymerase core. The turn region of motif C contains two aspartates (Asp328 and Asp329) that are highly conserved in RNA-dependent polymerases; the first is also conserved in DNA-dependent polymerases. In poliovirus polymerase, the two aspartates occupy positions 3 and 4 of a type II′ β turn, which positions both aspartate sidechains on the outside face of the core structure. This structure is very similar in all of the four categories of polymerase and is precisely positioned in the heart of the core by interactions with residues from each of the four motifs.

The two adjacent aspartates of motif C are quite close to the conserved aspartate of motif A, and these clustered aspartates are proposed to coordinate catalytically essential metals [19,39–41]. Indeed, for poliovirus polymerase, mutating the conserved aspartate of motif A (Asp233) or the first conserved aspartate of motif C (Asp328) results in an inactive polymerase [43]. Changing the second aspartate of motif C (Asp329) to asparagine results in a change in metal specificity [43]. In the crystal structure of poliovirus polymerase, strong electron density is observed between the aspartate of motif A and the second aspartate of motif C (Figure 2). We have modeled this as a calcium ion, because this was the only divalent metal ion present in the crystallizations. In the structure of MMLV reverse transcriptase [25], a metal was observed instead to coordinate between the two adjacent aspartates of motif C. Given that calcium does not support poliovirus polymerase activity and that nucleotides and primed-template nucleic acids were not present in either structure, the functional relevance of either of these metal coordination schemes is unclear.

Motif D of poliovirus polymerase (light purple in Figures 3 and 4) forms an α helix-turn-β strand structure (αI and β4). The α helix packs beneath the β sheet of the core structure. The β strand of motif D makes limited antiparallel β sheet interactions with the outside of motif A to complete the four-stranded antiparallel β sheet of the core structure. The turn region packs against the base of the fingers subdomain. The parts of motif D that precede and follow the turn (which includes the C-terminal end of the helix and the N-terminal end of the β strand) are similarly positioned in each of the four categories of polymerase, in the heart of the core structure where all four motifs come together. Even in Klenow [21], which has a larger turn region in motif D than the other polymerases, the helix and the β strand are positioned in a manner similar to that of other polymerases. Accordingly, we have

introduced an insertion into our structure-based sequence alignment to account for the larger turn region of motif D in Klenow (Figure 4b), as have Steitz *et al.* [44] in aligning Klenow and HIV-1 reverse transcriptase.

A fifth region of homology, motif E, is present in RNA-dependent but not DNA-dependent polymerases [17,18]. Motif E (dark purple in Figures 3 and 4) is positioned between the palm and thumb subdomains and is not integral to the conserved core structure. Motif E forms a short β-turn-β structure (β5 and β6) in both poliovirus polymerase and HIV-1 reverse transcriptase that interacts extensively with the face of the β sheet of the core structure. These interactions are distinctly hydrophobic and account for the conservation of several hydrophobic residues in motifs A, C, and D of RNA-dependent polymerases. The structure of motif E varies significantly even in different crystal structures of HIV-1 reverse transcriptase [41,45,46]; these variations appear to correlate with the position of the thumb subdomain. A β strand of the thumb subdomain (β7) interacts with motif E, yielding a short three-stranded antiparallel β sheet; this interaction may contribute to the positioning of the thumb subdomain. The conformation of motif E in poliovirus polymerase is most similar to the conformation of the unliganded form of HIV-1 reverse transcriptase [46] and is least similar to the neverapine-bound form of HIV-1 reverse transcriptase [41].

The palm subdomain of poliovirus polymerase also contains a polypeptide segment N-terminal to motif A that associates with the front edge of the core structure. This region consists of an α helix (αC) that passes in front of motif B, a loop that associates along the front edge of motif C, and then a short α helix (αD) immediately preceding motif A. All of the currently solved polymerase structures contain similarly located polypeptide segments. In HIV-1 reverse transcriptase, this segment meanders across the front of the palm as an irregularly structured strand with no significant similarities to poliovirus polymerase. Although amino acid sequence homologies have been proposed for these regions of poliovirus polymerase and HIV-1 reverse transcriptase (Figure 1; motif II of poliovirus polymerase and motif 2 of HIV-1 reverse transcriptase) [18], no structural similarities are observed in these regions.

Structural similarities between polymerase palm subdomains and the RRM

The conserved core structure of polymerase palm subdomains is strikingly similar in topology of secondary structural elements and in overall packing to the structures of the U1A RRM [47], ribosomal proteins L7/L12 [48] and S6 [24], the anticodon-binding domain of phenylalanyl-tRNA synthetase [49], the phosphocarrier protein Hpr [50, 51], the enzyme acyl phosphatase [52], the signal transducing protein PII [27], the regulatory subunit of aspartate

transcarbamylase [53], nucleotide diphosphate kinase [54], and procarboxypeptidase B [55]. Several of these structures are shown in Figure 5. All of the structure/sequence motifs of the polymerase palm subdomains correspond to structural elements of the RRM fold. Motif A of the polymerases corresponds to the highly conserved RNP-2 sequence motif of proteins involved in splicing, and part of motif C corresponds to the highly conserved RNP-1 sequence motif. The only three significant structural differences between the polymerase core structure and the RRM correlate with functional differences. Firstly, residues between motifs A and B of polymerases form part of the fingers subdomain. Because no fingers subdomain exists in other RRM proteins, the polypeptide chain proceeds directly to the α helix that corresponds to motif B of the polymerases. Secondly, the helical region of motif A of the polymerases is important for substrate binding and catalysis; this structure is not present in other RRM proteins. Thirdly, the turn region of motif C of polymerases is important for catalysis and the analogous region of the RRM is important for RNA recognition [56]. Accordingly, these turn regions have quite different structures in various RRM proteins.

Whether the similarities between the core structures of polymerase palm subdomains and the RRM arose via convergent or divergent pathways is unknown. Convergence from independent origins would suggest that, for some reason (e.g. folding, stability, kinetics of formation or function), the RRM fold is a very favorable structure. Identical connectivities would further require that the RRM fold favor this particular sequential arrangement of secondary structural elements. Divergence from a common ancestor is also an exciting possibility. Perhaps the RRM fold developed early in the ancient ribonucleoprotein (RNP) world and was a central player in the transition from the RNA

Figure 5
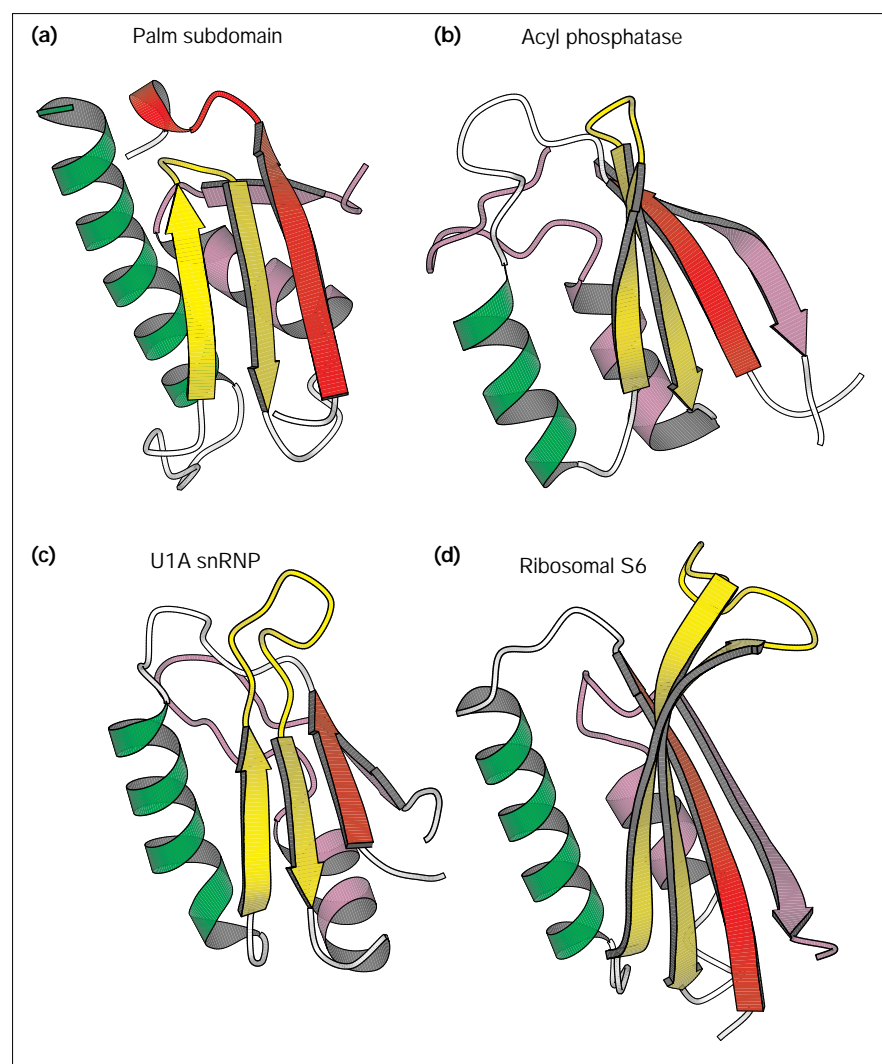
The RRM-like folds of (a) poliovirus polymerase palm domain, (b) acyl phosphatase [52], (c) the U1A-RNP domain [47] and (d) ribosomal protein S6 [24].

world to the RNP world and from the RNP world to the DNA world. The development of the RRM fold might have been a defining event in the evolution of life that ultimately contributed to each of the three processes of the central dogma of biology: replication, transcription, and translation.

The fingers subdomain

The fingers subdomain of poliovirus polymerase is composed of two polypeptide segments, one N-terminal of the palm subdomain (residues 97–194) and a second between motifs A and B (residues 240–285). This arrangement is similar to that of RNA-dependent DNA polymerases, in which the fingers are composed of a short polypeptide segment (approximately 30 amino acids) between motifs A and B and a larger polypeptide segment (80–100 amino acids) N-terminal of motif A. In contrast, the fingers subdomains of both categories of DNA-dependent polymerases are composed almost entirely of residues between motifs A and B. Unfortunately, significant portions of the fingers subdomain of poliovirus polymerase are disordered in the crystals (residues 98–180 and 267–290). These disordered segments are expected to form the top portion of the fingers subdomain. Similarities in their location in the primary sequence, as well as proposed sequence homologies (Figure 1) [18], might suggest that the top portions of the fingers subdomains of poliovirus polymerase and reverse transcriptases are structurally similar. Without additional information, however, these possible relationships remain unclear. The ordered portions of the poliovirus polymerase fingers subdomain constitute the lower portions of the fingers (αB, αF, and αG in Figure 3); these regions are very different from the analogous regions of HIV-l reverse transcriptase.

The thumb subdomain

The thumb subdomain of poliovirus polymerase (Figures 3 and 6) is composed primarily of the C-terminal-most 80 amino acid residues. This organization is the same as in RNA-dependent DNA polymerases [41], but contrasts with the N-terminal thumb subdomains of DNA-dependent polymerases [21]. The thumb subdomain of poliovirus polymerase begins with a β strand (β7) that interacts with the edge of the β strands of motif E (β5 and β6) to form a short three-stranded antiparallel β sheet. The thumb subdomain of HIV-1 reverse transcriptase also begins with a β strand that similarly interacts with motif E; these interactions may be important for positioning the thumb subdomain. The remainder of the thumb is composed of a series of five α helices. The first three of these form a three-helix bundle, the fourth (αM) is positioned at the top of the thumb subdomain, and the fifth (αN) is positioned along the front edge of the β strand (β7) of the thumb subdomain. Although the thumb subdomain of HIV-1 reverse transcriptase is also largely α-helical, the arrangement of these helices is different from that of poliovirus polymerase (Figure 6). The only similarities in the thumb subdomains of poliovirus polymerase and HIV-1 RT are first that both begin with a β strand that interacts with motif E, and second an α helix

Figure 6

Comparison of the thumb subdomains of poliovirus polymerase and HIV-1 reverse transcriptase (RT). Although both thumb subdomains are largely α-helical, the arrangement of helices is different in these two structures. The 'descending' helices (helix K of poliovirus polymerase and helix H of HIV-1 RT), which are similarly located facing the active-site cleft, are darkened.
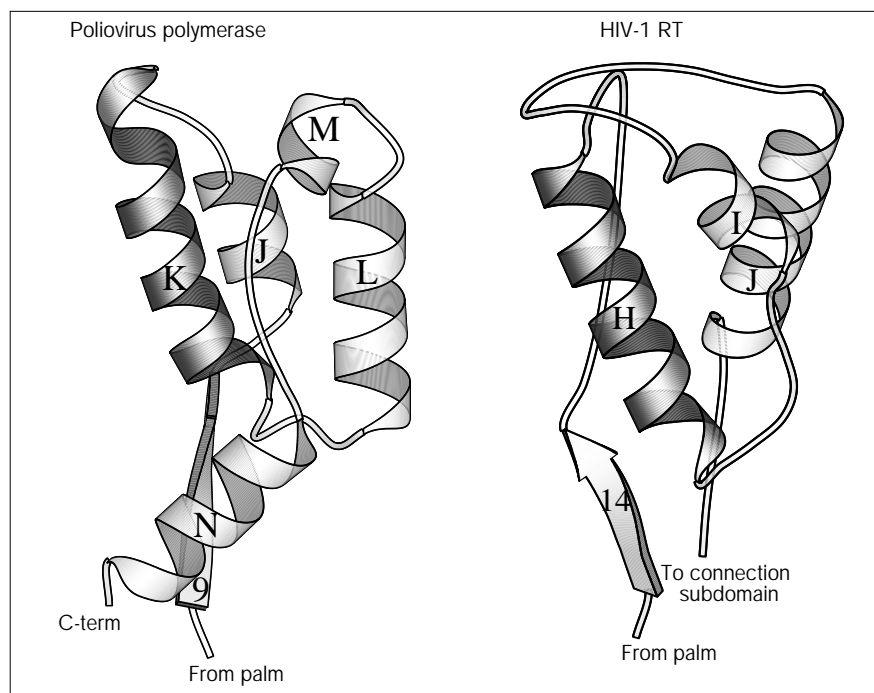
(αK of poliovirus polymerase and αH of HIV-1 reverse transcriptase) is similarly positioned along the active site cleft. These helices are highlighted in Figure 6. In the co-crystal structure of HIV-1 reverse transcriptase complexed with duplex DNA [45], α helix H and the residues immediately following it are positioned in the minor groove of the DNA. Interestingly, Klenow and T7 RNA polymerases also contain an α helix along their active site cleft, and in the Klenow–DNA [57] and Taq poly-merase–DNA [58] complexes this helix is also positioned in the minor groove of the DNA.

N-terminal regions

Two segments of the N-terminal polypeptide region of poliovirus polymerase are ordered in the poliovirus poly-merase crystals (shown in white in Figure 3a). The first ordered region (residues 12–37) composes part of the thumb subdomain. These residues extend as a single polypeptide strand from the active-site cleft up across the top of the thumb subdomain. This unusual interaction intimately links the N-terminal residues of the poly-merase with the C-terminal thumb region. Residues 25–35 are tightly wedged between the helices of the thumb with Phe30, Val33 and Phe34 packed into a hydrophobic core near the top of the thumb subdomain. The N-terminal residues are essential for polymerase activity; deletion of Trp5 [59] or of residues 1–6 (SCS,

unpublished observations) completely inactivates the polymerase. The position of residues 12–25 in the active site cleft is consistent with an essential role in polymer-ization activity, but specifically how this region might be involved is not known.

The second ordered segment of the N-terminal region (residues 67–97) is on the opposite side of the polymerase, at the bottom of the fingers subdomain (Figure 3). These residues form a long α helix that is positioned along the base of the fingers subdomain. If the two N-terminal seg-ments are part of the same molecule, residues 38–66 would have to reach >45 Å across the active site cleft of poliovirus polymerase. Alternatively, residues 12–37 might derive from an adjacent molecule, which would have important implications for the oligomerization of poliovirus polymerase.

Does poliovirus polymerase function as an oligomer?

Poliovirus polymerase exhibits cooperativity in RNA binding and polymerization activity with respect to poly-merase concentration [11], which indicates that interac-tions between polymerase molecules are important for these activities. Polymerase–polymerase interactions have also been observed directly by chemical cross-linking [11] and in the yeast two-hybrid assay [12]. We therefore eval-uated the packing of polymerase molecules in our crystals

Figure 7
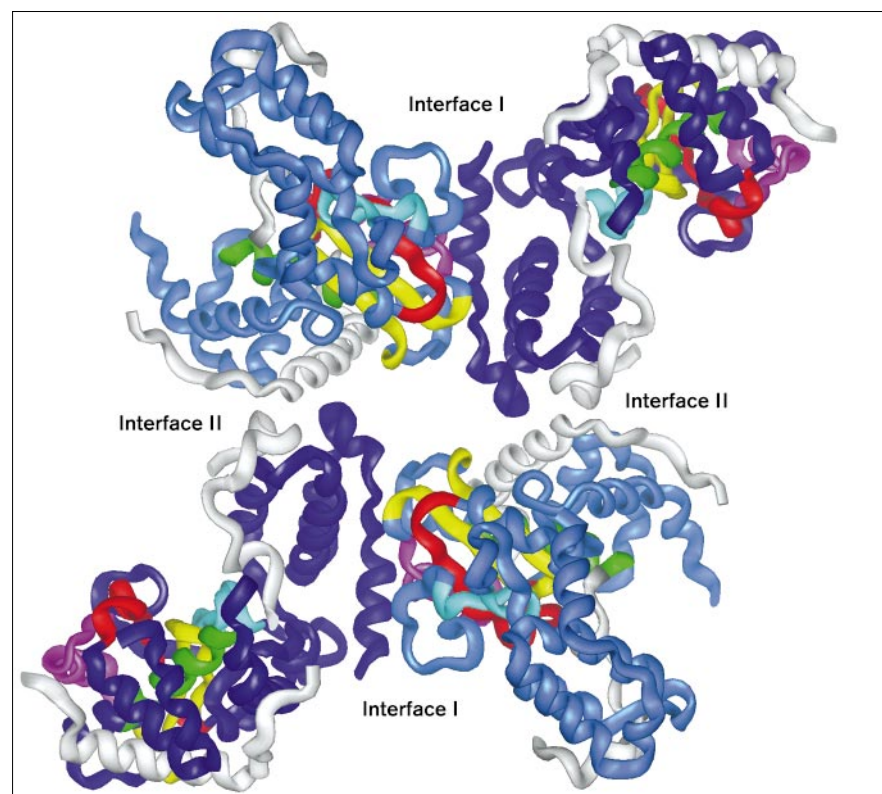
Interactions between polymerase molecules in the crystals of poliovirus polymerase. The structure/sequence motifs and N-terminal regions are color-coded as in Figure 3a. The two different types of interactions are labeled interface I and interface II.

for potentially relevant polymerase–polymerase interactions and found that two significant interfaces between polymerase molecules are present (Figure 7).

Interface I derives from extensive interactions between the front of the thumb subdomain of one molecule and the back of the palm subdomain of the adjacent molecule (Figure 7). The surface on the front of the thumb subdomain is composed largely of residues in and immediately preceding the C-terminal helix (αN) of the polymerase (residues 446–461). In addition, three sidechains (Gln411, Asp412, and Arg415) are contributed from helix L. The second surface is composed of residues in the non-conserved loops between motifs B and C and between motifs C and D as well as three residues from the helix of motif D (αI). Interface I is large, with at least 23 amino acid sidechains involved in the interaction, giving a total buried surface area of 1480 $Å^2$. The sidechain interactions include a variety of hydrophobic, ionic, and hydrogen bonding interactions. For example, the hydrophobic sidechain of Leu446 extends from the surface of one molecule into a hydrophobic pocket of the adjacent molecule; the sidechain of Arg456 is interacting with Asp339 and Ser341 of the adjacent molecule. The extent and specificity of interactions at interface I suggest that this interface is not just a consequence of crystal packing, but that it is designed to provide for functionally important polymerase–polymerase interactions.

Interface I is directional, or head-to-tail, such that the oligomeric unit is not of defined size but could extend indefinitely in both directions. Adjacent polymerase molecules pack along a crystallographic 2_1 screw axis, such that each molecule is rotated 180° and translated 44 Å relative to the adjacent molecule.

When RNA is modeled into our structure, based on the location of double-stranded DNA in the HIV-1 reverse transcriptase–DNA co-crystal structure [45], the RNA falls roughly along the oligomeric fiber formed from the interactions at interface I. Our working hypothesis is, therefore, that oligomerization via interface I is important for nucleic acid interactions. Indeed, if one considers two poliovirus polymerase molecules associated via interface I, this interaction extends the nucleic acid-binding surface in a manner similar to that of the connection and RNase H subdomains and the p51 subunit in the HIV-1 reverse transcriptase heterodimer [23,45]. We are now testing the hypothesis that oligomerization of poliovirus polymerase via interface I is important for nucleic acid binding.

Interface II involves the two N-terminal regions of the polymerase. The N-terminal strand (residues 12–37) that interacts with the thumb subdomain and the helix at the base of the fingers subdomain (residues 67–97) contact each other directly at interface II. The specific nature of

interface II depends on how these two regions are connected by residues 38–66, which are disordered in the crystals. Connecting these two regions within a molecule would require that residues 38–66 reach >45 Å across the active site cleft of the polymerase, which would be an unusual structural feature. Alternatively, these segments might belong to different molecules and connect across interface II (note the proximity of the white regions of adjacent molecules across this interface in Figure 7). This possibility is attractive, in that the connecting distance is shorter (<30 Å) and the connected regions are on the same surface, so no other parts of the polymerase intervene. Since the very N-terminal residues of the polymerase are essential for catalytic activity, a *trans* contribution of the N-terminal strand by a neighboring molecule would mean that interactions between polymerase molecules at interface II are required for polymerase activity.

Polymerase interactions at interface II yield a two-fold symmetric association of the filaments formed by interactions along interface I (Figures 7 and 8). The interface I filaments cross at a 120° angle such that two given filaments interact and then diverge from each other. Figure 8 illustrates this arrangement of molecules. An interesting scenario arises if interface II involves an intermolecular association in which the N-terminal strand over the thumb is contributed by the neighboring polymerase associated via interface II. Given that the N-terminus is essential for activity, the interaction between filaments via interface II would be required for polymerase activity and would create one active polymerase center in each of the associated filaments. Further interactions with other filaments could occur at analogous sites along each filament. What emerges is a network of polymerase filaments in which layers of parallel interface I filaments associate with other layers at 120° angles via the interface II interactions. The interface II associations would correspond to active polymerase sites within the polymerase network. This network could pack more or less tightly according to the frequency of interface II interactions to form, perhaps, a small, tightly packed core of polymerase activity or perhaps a large, loosely connected network of polymerase filaments. Whereas it seems likely that interface I is important for polymerase function, the importance of interface II depends on whether the N-terminal strand over the thumb represents an intramolecular or an intermolecular interaction.

Poliovirus is known to replicate its RNA in large membrane-associated replication complexes that contain the polymerase [5,6] as well as several other viral proteins and possibly host proteins (reviewed in [3]). Our structural models for oligomerization of poliovirus polymerase may be providing a first glimpse of the poliovirus replication machinery.

Figure 8

Model for the oligomerization of poliovirus polymerase. Two 'fibers' formed by interactions along interface I, each containing four molecules of poliovirus polymerase (labeled 1, 2, 3, 4 and 1′, 2′, 3′, 4′), are shown interacting via interface II. These fibers cross at 120° angles such that 1 and 1′ are coming out of the page and 4 and 4′ are going into the page. Each fiber is highlighted with an axis shown in black. A twofold rotation axis that relates the two fibers passes through the center of the figure perpendicular to the page. The regions shown in white are the N-terminal strands of the thumb subdomains that may be contributed *in trans* by a molecule in the other fiber (i.e. 2→3′ and 2′→3). Only one set of interface II interactions is shown. Additional interactions along the interface I fibers would potentially lead to a complex network of interacting polymerase filaments.
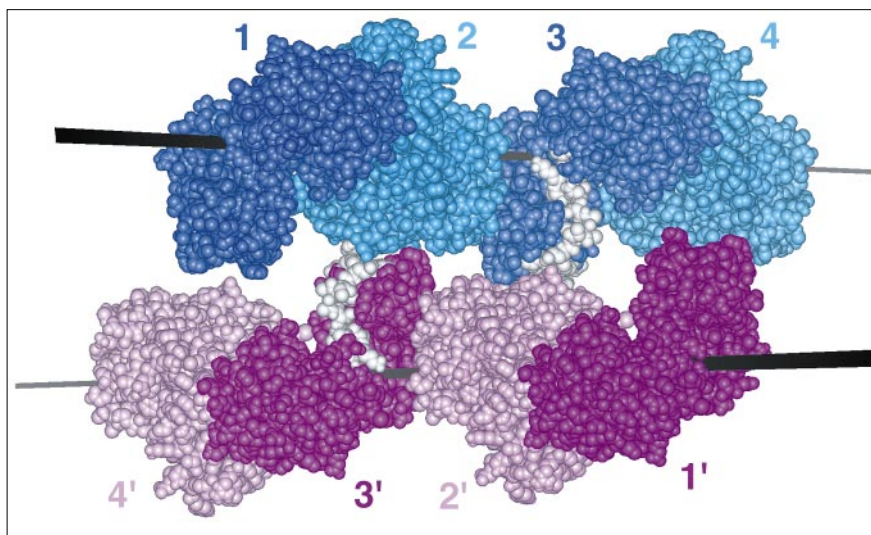
Biological implications

RNA viruses encode an RNA-dependent RNA polymerase that enables them to replicate their RNA genome directly in the cytoplasm of host cells without the need for DNA. We report here the structure of the RNA-dependent RNA polymerase of poliovirus. This is the first solved structure of a viral RNA-dependent RNA polymerase and provides a first example of a structure from the final category typically used to describe polymerases. Comparison of the poliovirus polymerase structure with those of other solved polymerase structures reveals that it has the same overall 'right hand' shape. Although the detailed structures of the fingers and thumb subdomains are different from those of other polymerases, the palm subdomain contains a core structure similar to those of other polymerases, composed of four of the amino acid sequence motifs described for RNA-dependent polymerases [17, 18]. Interestingly, the conserved core structure of polymerase palm subdomains has a very common protein fold that is also found in proteins involved in translation and splicing. Potential evolutionarily relationships between such proteins has important implications for protein evolution and perhaps for the origins of life.

Polymerase–polymerase interactions appear to be important for poliovirus polymerase function [11]. We observe two regions of polymerase–polymerase interactions in the crystal that extend well beyond those typically observed in crystal packing. The first interface (I) involves more than 23 amino acid sidechains on two surfaces of the protein and is directional, such that it defines an extended fiber of polymerase molecules with a distinct directionality. When RNA is modeled into our structure, it falls roughly along the fibers formed by interface I interactions. **The second interface (II) involves interactions between the N-terminal polypeptide segments, but residues connecting the N-terminal-most polypeptide segment to the rest of the protein are disordered in the crystals. We believe that the N-terminal polypeptide segment may derive *in trans* from another polymerase molecule, such that the N terminus of one polymerase contributes to the active-site cleft of a second. Since the N-terminal residues of poliovirus polymerase are required for activity ([59]; SCS, unpublished observations), interactions at interface II would then be required for polymerase activity. Interface II interactions give rise to an association of interface I filaments such that an unusual model arises in which fibers of polymerase molecules form and interact, potentially forming a network of polymerase molecules. Active sites would correspond to the positions at which the fibers interact via interface II. This unusual higher order structure may be a first glimpse at the structure of the poliovirus replication complex.**

Materials and methods

Purification of poliovirus polymerase from E. coli

Poliovirus polymerase was purified from *E. coli* BL21(DE3)pLysS using inducible T7-based expression systems. The plasmids contain the entire coding region for the 3D polymerase with an additional Met added to the N terminus as required for expression. *E. coli* harboring pT5T-3D (T Jarvis and K Kirkegaard, personal communication) or pKKT7E-3D (SCS, unpublished results) were grown in 2XYT media to early log phase, cooled to room temperature, grown to OD_{600} = 0.5, and then induced with 0.5 mM IPTG for 12–16 hours. Poliovirus polymerase was purified as follows: cells were lysed by sonication; polymerase was precipitated with ammonium sulfate at 40% of saturation; the protein was loaded onto an S-sepharose column in 50 mM NaCl, 25 mM HEPES (pH 8.5), 0.02% sodium azide, 0.1 mM EDTA, 2 mM DTT and, after washing with approximately 6–8 column volumes of this same buffer, the protein was eluted with 0.2 M NaCl in 25 mM HEPES (pH 8.5), 15% glycerol, 0.02% sodium azide, 0.1 mM EDTA, 2 mM DTT, and 0.5% n-octyl-β-D-

Table 3

Refinement statistics for RNA-dependent RNA polymerase of poliovirus (2σ data; 2568 scatters).

Resolution	Number of reflections	R value	R_{free}
10.00–4.65	3480	0.210	0.263
4.65–3.76	3308	0.167	0.235
3.76–3.30	3008	0.202	0.269
3.30–3.01	2588	0.252	0.278
3.01–2.80	2145	0.281	0.316
2.80–2.64	1632	0.321	0.374
2.64–2.51	1180	0.349	0.350
2.51–2.40	752	0.369	0.335
Total	18,093	0.218	0.270

The rms deviations from ideal bond length = 0.014 Å; rms deviations from ideal bond angles = 2.18°. 10% of the reflection intensities were used for R_{free} analysis.

glucopyranoside; the protein was diluted to reduce the NaCl concentration to 0.15 M, loaded onto a Q-sepharose FPLC column and eluted with a linear gradient of NaCl from 0.1 M to 0.35 M NaCl in 25 mM HEPES (pH 8.5), 15% glycerol, 0.02% sodium azide, 0.1 mM EDTA, 2 mM DTT, and 0.5% n-octyl-β-D-glucopyranoside. Upon elution from this column, the protein was >99% pure as estimated from SDS–PAGE. Typically, 10–30 mg of pure polymerase is obtained per liter of *E. coli* culture. The N-terminal Met was removed in *E. coli* as determined by N-terminal sequence analysis (data not shown).

Crystallization and data collection

Crystals were grown by hanging drop vapor diffusion of approximately 3-5 mg/ml purified poliovirus polymerase dissolved in 0.6 M $CaCl_2$, 50 mM PIPES (pH 6.0–7.0), 3% glycerol, 10 mM DTT, and 0.2 mM EDTA against a well solution containing 1.2 M $CaCl_2$. The crystals grew as hexagonal blocks in 2–6 weeks at 15°–22°C. The crystals are trigonal, space group $P3_221$ with a = b = 88.1 Å, c = 158.5 Å. Diffraction was anisotropic with diffraction along c* to 2.4 Å resolution and diffraction along a* and b* to 2.8 Å resolution. The crystals were harvested into 1.0–1.2 M $CaCl_2$, 50 mM PIPES (pH 6.5), 10 mM DTT, and 0.2 mM EDTA.

Native and derivative data used for the initial structure determination were collected at 20°C. Native data used for refinement of the structure were collected from crystals cooled to −25°C. These crystals were soaked in the harvesting solution described above with 20% glycerol or 2 M D-Glucose added to prevent freezing. The data were reduced to reflection intensities using either the Molecular Structure Corporation (MSC) software or DENZO and SCALEPACK [32].

Structure determination and refinement

The structure was solved by multiple isomorphous heavy-atom replacement methods. Heavy-atom derivatives were prepared by soaking crystals in solutions containing 1 mM of the heavy atom compound, 1.2 M $CaCl_2$, and 50 mM PIPES (pH 6.5) for 12–14 hours. Native and derivative data were scaled and analyzed using the CCP4 program suite [60]. Heavy-atom positions were identified by visual inspection of difference Patterson maps and using difference Fourier maps. Heavy-atom positions and occupancies were refined and phases were calculated using MLPHARE [33] (Table 3). The MIR phases were modified by solvent flattening and histogram matching using either SQUASH [34] or dm [35] and the structure was fit using O [36]. The structure was refined using X-PLOR [37].

References
1. Kitamura, N., *et al.*, & Wimmer, E. (1981). Primary structure, gene organization and polypeptide expression of poliovirus RNA. *Nature* **291**, 547–553.
2. Racaniello, V.R. & Baltimore, D. (1981). Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. *Proc. Natl. Acad. Sci. USA* **78**, 4887–4891.
3. Richards, O.C. & Ehrenfeld, E. (1990). Poliovirus RNA replication. *Curr. Top. Microbiol. Immunol.* **161**, 89–119.
4. Wimmer, E., Hellen, C.U.T. & Cao, X. (1993). Genetics of poliovirus. *Ann. Rev. Genet.* **27**, 353–436.
5. Baltimore, D., Franklin, R.M., H.J.E. & Tamm, I. (1963). Poliovirus-induced RNA polymerase and the effects of virus-specific inhibitors on its production. *Proc. Natl. Acad. Sci. USA* **49**, 843–849.
6. Caliguiri, L.A. & Tamm, I. (1970). The role of cytoplasmic membranes in poliovirus biosynthesis. *Virology* **42**, 100–110.
7. Lundquist, R.E., Ehrenfeld, E. & Maizel, J.V. (1974). Isolation of a viral polypeptide associated with poliovirus RNA polymerase. *Proc. Natl. Acad. Sci. USA* **71**, 4773–4777.
8. Flanegan, J.B. & Van Dyke, T.A. (1979). Isolation of a soluble and template-dependent poliovirus RNA polymerase that copies virion RNA *in vitro. J. Virol.* **32**, 155–161.
9. Morrow, C.D., Warren, B. & Lentz, M.R. (1987). Expression of enzymatically active poliovirus RNA-dependent RNA polymerase in *Escherichia coli. Proc. Natl. Acad. Sci. USA* **84**, 6050–6054.
10. Rothstein, M.A., Richards, O.C., Amin, C. & Ehrenfeld, E. (1988). Enzymatic activity of poliovirus RNA polymerase synthesized in *Escherichia coli* from viral cDNA. *Virology* **164**, 301–308.
11. Pata, J.D., Schultz, S.C. & Kirkegaard, K. (1995). Functional oligomerization of poliovirus RNA-dependent RNA polymerase. *RNA* **1**, 466–477.
12. Hope, D.A., Diamond, S.E. & Kirkegaard, K.K. (1997). Genetic dissection of interactions between poliovirus 3D pol and viral protein 3AB. *J. Virol.*, In press.
13. Argos, P., Kamer, G., Nicklin, M.J.H. & Wimmer, E. (1984). Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. *Nucleic Acids Res.* **12**, 7251–7267.
14. Koonin, E.V. (1991). The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J. Gen. Virol.* **72**, 2197–2206.
15. Bruenn, J.A. (1991). Relationships among the positive-strand and double-strand RNA viruses as viewed through their RNA-dependent RNA polymerases. *Nucleic Acids Res.* **19**, 217–226.
16. Kamer, G. & Argos, P. (1984). Primary structural comparison of RNA-dependent polymerases from plant, animal, and bacterial viruses. *Nucleic Acids Res.* **12**, 7269–7282.
17. Poch, O., Sauvaget, I., Delarue, M. & Tordo, N. (1989). Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* **8**, 3867–3874.
18. Xiong, Y. & Eickbush, T.H. (1990). Origin and evolution of retro-elements based upon their reverse transcriptase sequences. *EMBO J.* **9**, 3353–3362.
19. DeLarue, M., Poch, O., Tordo, N., Moras, D. & Argos, P. (1990). An attempt to unify the structure of polymerases. *Protein Eng.* **3**, 461–467.
20. Zanotto, P.M., Gibbs, M.J., Gould, E.A. & Holmes, E.C. (1996). A re-evaluation of the higher taxonomy of viruses based on RNA polymerases. *J. Virol.* **70**, 6083–6096.
21. Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. & Steitz, T.A. (1985). Structure of the large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* **313**, 762–766.

22. Sousa, R., Chung, Y.J., Rose, J.P. & Wang, B.-C. (1993). Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution. *Nature* **364**, 593–599.
23. Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A. & Steitz, T.A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1783–1790.
24. Lindahl, M., *et al.*, & Amons, R. (1994). Crystal structure of the ribosomal protein S6 from *Thermus thermophilus. EMBO J.* **13**, 1249–1254.
25. Georgiadis, M.M., Jessen, S.M., Ogata, C.M., Telesnitsky, A., Goff, S.P. & Hendrickson, W.A. (1995). Mechanistic implications from the structure of a catalytic fragment of moloney murine leukemia virus reverse transcriptase. *Structure* **3**, 879–892.
26. Unge, T., *et al.*, & Standberg, B. (1994). 2.2 Å resolution structure of the N-terminal half of HIV-1 reverse transcriptase. *Structure* **2**, 953–961.
27. Cheah, E., Carr, P.D., Suffolk, P.M., Vasudevan, S.G., Dixon, N.E. & Ollis, D.L. (1994). Structure of the *Escherichia coli* signal transducing protein PII. *Structure* **2**, 981–990.
28. Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H. & Kraut, J. (1994). Structure of ternary complexes of rat DNA polymerase b, a DNA template-primer, and ddCTP. *Science* **264**, 1891–1903.
29. Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H. & Kraut, J. (1994). Crystal structure of rat DNA polymerase β: evidence for a common polymerase mechanism. *Science* **264**, 1930–1935.
30. Davies, J.F.J., Almassy, R.J., Hostomska, Z., Ferre, R.A. & Hostomsky, Z. (1994). 2.3 Å crystal structure of the catalytic domain of DNA polymerase β. *Cell* **76**, 1123–1133.
31. Holm, L. & Sander, C. (1995). DNA polymerase β belongs to an ancient nucleotidytransferase superfamily. *Trends Biochem. Sci.* **237**, 345–347.
32. Otwinowski, Z. & Minor, W. (1996). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**,
33. Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In *Isomorphous Replacement and Anomalous Scattering*, (Wolf, W., Evans, P.R., & Leslie, A.G.W., eds), pp. 80–86, Daresbury Laboratory, Daresbury, UK.
34. Zhang, K.Y.J. & Maing, P. (1990). The use of Sayre's equation with solvent flattening and histogram matching for phase extension and refinement of protein structures. *Acta Cryst. A* **46**, 377–381.
35. Cowtan, K.D. (1994). 'DM': an automated procedure for phase improvement by density modification. In *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography.* **31**, 34–38.
36. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
37. Brünger, A.T. (1993). *X-PLOR Manual Version 3.1: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
38. Brunger, A.T. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Cryst. D* **49**, 24–36.
39. Argos, P. (1988). A sequence motif in many polymerases. *Nucleic Acids Res.* **16**, 9909–9916.
40. Beese, L.S. & Steitz, T.A. (1991). Structural basis for the 3′–5′ exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J.* **10**, 25–33.
41. Kohlstaedt, L.A., Friedman, J.M., Rice, P.A. & Steitz, T.A. (1992). The structure of HIV-1 reverse transcriptase snaps into focus. *J. NIH Res.* **4**, 78–83.
42. Gao, G., Olova, M., Georgiadis, M.M., Hendrickson, W.A. & Goff, S.P. (1997). Conferring RNA polymerase activity to a DNA polymerase: a single residue in reverse transcriptase controls substrate selection. *Proc. Natl. Acad. Sci. USA* **94**, 407–411.
43. Jablonski, S.A., Luo, M. & Morrow, C.D. (1991). Enzymatic activity of poliovirus RNA polymerase mutants with single amino acid changes in the conserved YGDD amino acid motif. *J. Virol.* **65**, 4565–4572.
44. Steitz, T.A., *et al.*, & Rice, P.A. (1993). Two DNA polymerases: HIV reverse transcriptase and the Klenow fragment of *Escherichia coli* DNA polymerase I. *Cold Spring Harbor Symposia on Quantitative Biology* LVIII. 495–504.
45. Jacobo-Molina, A., *et al.*, & Arnold, E. (1993). Crystal structure of human immunodeficiency virus type I reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. USA* **90**, 6320–6324.
46. Rodgers, D.W., *et al.*, & Harrison, S.C. (1995). The structure of unliganded reverse transcriptase from the human immunodeficiency virus type I. *Proc. Natl. Acad. Sci. USA* **92**, 1222–1226.
47. Nagai, K., Oubridge, C., Jessen, T.H., Li, J. & Evans, P.R. (1990). Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* **348**, 515–520.
48. Leijonmarck, M. & Liljas, A. (1987). Structure of the C-terminal domain of the ribosomal protein L7/L12 from *Escherichia coli* at 1.7 Å. *J. Mol. Biol.* **195**, 555–580.
49. Goldgur, Y., *et al.*, & Safro, M. (1997). The crystal structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus* complexed with cognate tRNA[Phe]. *Structure* **5**, 59–68.
50. Herzberg, O., Reddy, P., Sutrina, S., Saier, M.H., Reizet, J & Kapifia, G. (1992). Structure of the histidine-containing phosphocarrier protein HPr from *Bacillus Subtilis* at 2.0 Å resolution. *Proc. Natl. Acad. Sci. USA* **89**, 2499–2503.
51. Wittekind, M., Rajagopal, P., Branchini, B.R., Reizer, R., Saier, M.H. & Klevit, R.E. (1992). Solution structure of the phosphocarrier protein HPr from *Bacillus subtilis* by two-dimensional NMR spectroscopy. *Protein Science* **1**, 1363–1376.
52. Pastore, A., Saudek, V., Ramponi, G., & Williams, R.J.P. (1992). Three-dimensional structure of acyl phosphatase. *J. Mol. Biol.* **224**, 427–440.
53. Stevens, R.C., Gouaux, J.E. & Lipscomb, W.N. (1990). Structural consequences of effector binding to the T state of aspartate carbamoyltranferase: crystal structure of the unliganded and ATP and CTP-complexed enzymes at 2.6 Å resolution. *Biochemistry* **29**, 7691–7701.
54. Dumas, C., *et al.*, & Janin, J. (1992). X-ray structure of nucleoside diphosphate kinase. *EMBO J.* **11**, 3203–3208.
55. Coll, M., Guasch, A., Aviles, R. & Huber, R. (1991). Three-dimensional structure of porcine procarboxypeptidase B: a structural basis of its inactivity. *EMBO J.* **10**, 1–9.
56. Oubridge, C., Ito, N., Evans, P.R., Teo, C.-H. & Nagai, K. (1994). Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**, 432–438.
57. Beese, L.S., Derbyshire, V. & Steitz, T.A. (1993). Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* **260**, 352–355.
58. Eom, S.H., Wang, J. & Steitz, T.A. (1996). Structure of *taq* polymerase with DNA at the polymerase active site. *Nature* **382**, 278–281.
59. Plotch, S.J., Palant, O. & Gluzman, Y. (1989). Purification and properties of poliovirus RNA Polymerase expressed in *Escherichia coli. J. Virol.* **63**, 216–225.
60. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.
61. Masters, B.S., Stohl, L.L. & Clayton, S.D. (1987). Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell* **51**, 89–99.