Helmer, J. M. C. Plane, J. Qian, C. S. Gardner, *J. Geophys. Res.* **103**, 10,913 (1998); 4 Gg yr^{-1} is obtained from their flux after correcting the weight fraction of Fe in their calculation. An additional uncertainty in meteoritic fluxes is introduced because the actual mass available for incorporation into stratospheric aerosols may be different from the incident mass because the oxygen bound into minerals and oxides is released during ablation and then accumulated during later oxidation of iron, silicon, and other elements.

34. D. J. Hofmann, J. M. Rosen, J. M. Kiernan, J. Laby, *J. Atmos. Sci.* **33**, 1783 (1976); M. Chin and D. D. Davis, *J. Geophys. Res.* **100**, 8993 (1995); D. K. Weisenstein *et al.*, *ibid.* **102**, 13,019 (1997); M. J. Mills, thesis, University of Colorado, Boulder (1996). Condensible material converted (from 100 to 160 Gg yr^{-1} of sulfur, from oxidation of OCS and SO_2), with the assumption that 70% of the aerosol by weight is H_2SO_4.
35. Y. Chen, S. M. Kreidenweis, L. M. McInnes, D. C. Rogers, P. J. DeMott, *Geophys. Res. Lett.* **25**, 1391 (1998); J. Heintzenberg, K. Okada, J. Ström, *Atmos. Res.* **41**, 81 (1996).
36. U. M. Biermann *et al.*, *Geophys. Res. Lett.* **23**, 1693 (1996).
37. G. S. Kent, M. P. McCormick, S. K. Schaffner, *J. Geophys. Res.* **96**, 5249 (1991).
38. R. Dixon, in preparation.
39. J. J. Zieman *et al.*, *J. Geophys. Res.* **100**, 25,975 (1995); J. G. Watson *et al.*, *Aer. Sci. Technol.* **21**, 1 (1994). Range shown is from "primary geological samples."
40. The assistance of the pilots, ground crews, and staff of the WB-57F is gratefully acknowledged. This work was funded by NOAA base funding, the NASA Atmospheric Effects of Aviation Project, and the NASA Upper Atmospheric Research Program. B. Gary assisted with the MTP data analysis.

16 September 1998; accepted 2 November 1998

Structure of a Covalently Trapped Catalytic Complex of HIV-1 Reverse Transcriptase: Implications for Drug Resistance

Huifang Huang,* Rajiv Chopra,* Gregory L. Verdine,† Stephen C. Harrison†

A combinatorial disulfide cross-linking strategy was used to prepare a stalled complex of human immunodeficiency virus–type 1 (HIV-1) reverse transcriptase with a DNA template:primer and a deoxynucleoside triphosphate (dNTP), and the crystal structure of the complex was determined at a resolution of 3.2 angstroms. The presence of a dideoxynucleotide at the 3′-primer terminus allows capture of a state in which the substrates are poised for attack on the dNTP. Conformational changes that accompany formation of the catalytic complex produce distinct clusters of the residues that are altered in viruses resistant to nucleoside analog drugs. The positioning of these residues in the neighborhood of the dNTP helps to resolve some long-standing puzzles about the molecular basis of resistance. The resistance mutations are likely to influence binding or reactivity of the inhibitors, relative to normal dNTPs, and the clustering of the mutations correlates with the chemical structure of the drug.

The reverse transcriptase (RT) of HIV-1 is an important target of antiviral therapy in the treatment of acquired immunodeficiency syndrome (AIDS) (*1*). RT has two distinct enzymatic activities, an RNA- or DNA-dependent DNA polymerase and a ribonuclease (RNase) H, but current agents are directed only against the polymerase. Five of the seven inhibitors currently licensed in the United States are chain-terminating nucleoside analogs [for example, 3′-azido-2′,3′-dideoxythymidine (AZT), 2′,3′-dideoxyinosine (ddI), and 2′-deoxy-3′-thiacytidine (3TC)]. The other two inhibitors are members of a chemically diverse group of nonnucleoside RT inhibitors (NNRTIs).

HIV-1 RT is a dimer of two related chains, a 66-kD subunit (p66) and a 51-kD subunit (p51) derived from p66 by proteolytic cleavage (*2*). The two chains have in common four domains [referred to as "fingers," "palm," "thumb," and "connection" (*3*)], and p66 also has a COOH-terminal RNase H. The p66 subunit has both the polymerase and RNase H active sites. The palm contains residues critical for polymerase catalytic activity, and its folded structure resembles that of a corresponding catalytic domain in other DNA and RNA polymerases.

Emergence of resistance in patients treated with RT inhibitors is a major limitation of antiviral therapy (*4, 5*). All NNRTIs bind near the polymerase active site, in a hydrophobic pocket created by displacement of the polypeptide segment connecting palm and thumb. Viral mutations conferring resistance to these drugs can be understood readily in terms of alterations in their common binding site. In contrast, the positions of altered residues in viruses resistant to nucleoside analogs do not follow so clear a pattern (*5*). Lack of a structure for a catalytic complex of RT with template:primer and dNTP substrates has hindered understanding how site-specific changes confer resistance to particular drugs.

Previously determined RT structures include a number of NNRTI complexes and the unliganded enzyme (*3, 6–10*). The only published structure of RT with a bound template:primer is an RT•template:primer•Fab ternary complex (*6*), which shows that the primer terminus lies near three catalytically essential aspartic acid residues in the palm and that the duplex extends along the enzyme surface toward the RNase H. The position of fingers and thumb define a deep cleft, with the polymerase active site at its base. This feature, also present in crystals of RT with bound NNRTIs, has led to a model in which the 5′ extension of the template passes through the cleft, interacting with residues in the fingers and palm that are mutated in drug-resistant strains (*3, 6*). Other DNA polymerases contain a similar cleft, but recent structures of catalytic complexes show that it closes down when substrates bind and that it does not serve as a channel for polynucleotide chains (*11, 12*).

We have devised a way of isolating stalled, covalently tethered complexes of RT with template:primer, and we have crystallized one such species. In the complex described here at 3.2 Å resolution, the primer terminus is a dideoxynucleotide and thus unable to attack an incoming dNTP. The crystal of this catalytic complex contains bound dTTP in precisely the expected position for attack by the (missing) 3′ OH. The fingers domain bends, relative to other RT structures, so that various residues near the fingertips form part of the dNTP-binding site. This conformational adjustment defines the complete catalytic site and leads to a revised interpretation of the mechanism by which various mutations confer resistance to nucleoside analog drugs.

Trapping of a Catalytic HIV-1 RT–Substrate Complex

The difficulty in obtaining good crystals of an RT•template:primer•dNTP complex stems from RT's relatively modest specificity in binding to polymerizable versus nonpolymerizable sites in DNA, even in the presence of a dNTP. To overcome this problem, we chose

H. Huang and G. L. Verdine are in the Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. R. Chopra and S. C. Harrison are at the Howard Hughes Medical Institute and Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: verdine@glviris.harvard.edu and schadmin@crystal.harvard.edu

to cross-link the enzyme and template:primer with a disulfide bond (*13*). A unique cross-linked species was produced by using DNA with a single tethered thiol (*14*) and RT with a site-specific cysteine mutation (*15*). Modeling based on the earlier RT•template:primer•Fab complex (*6*) had led to the proposal that amino acid side chains positioned along one face of α helix H in p66 (Q258, G262, W266) (*16*) would track in the minor groove of the template:primer duplex (*17*). A cysteine residue introduced at one of these positions should therefore reach toward the "floor" of the minor groove, several base pairs from the polymerase active site; a thiol-bearing tether protruding from the minor groove could form a disulfide cross-link with the engineered cysteine, trapping the catalytic RT-DNA complex (Fig. 1). We placed a thiol tether in the minor groove by attaching it to the N^2 position of a guanine base (Fig. 1) (*18*). Rather than select a single configuration of reacting thiol partners, we systematically varied their relative positions (Fig. 2A). As shown in Fig. 2, incubation of a modified template:primer with cysteine-engineered RT proteins in the presence of certain dNTP and ddNTP combinations results in defined cycles of polymerization that move the enzyme forward a fixed distance on its substrate. If the DNA template is now positioned so that the thiol tether and the engineered cysteine are sufficiently close, a disulfide cross-link can form, preventing dissociation of the product.

In the top panel of Fig. 2C are the results of an experiment in which RT(Q258C) was advanced by one to four base pairs along the thiol-containing template:primer (*19*). After one cycle of polymerization (+1 lane), no cross-linking was observed, nor was any observed after two cycles (+2). After three cycles (+3), however, about half of the p66 was found to be cross-linked to DNA. One further cycle of polymerization (+4) resulted in a complete loss of cross-linking. We interpret these results to indicate that after three cycles of polymerization, Cys^{258} and the DNA-thiol are optimally positioned to react; fewer or further cycles simply move the reacting partners away from the optimal configuration.

When the engineered cysteine residue in RT was transposed by one turn along helix H (Fig. 2C, second panel), cross-linking was again positionally selective, but the highest efficiency of cross-linking now occurred after two cycles of polymerization. Transposition

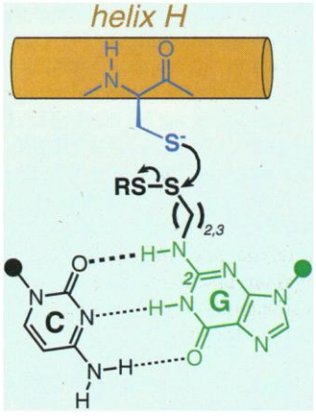

Fig. 1. RT-DNA tethering reaction. Chemistry of disulfide bond formation between the side chain of an engineered cysteine residue (blue) in helix H (gold) of RT to a thiol group in the minor groove of DNA (activated as the mixed disulfide), which is tethered to N^2 of a dG (green) in the template:primer.

Table 1. Statistics for data collection and refinement (*20*).

Feature	Native	Iodine
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell dimensions (Å)	78.5 150.0 279.3	79.0 150.8 280.9
Resolution (Å)	18–3.2 (3.3–3.2)	18–4.4 (4.5–4.4)
Mosaicity (°)	0.3	0.4
Unique reflections	54,281 (4,802)	15,424 (1,540)
Completeness (%)	93.8 (84.3)	62.5 (64.4)
Average *I*/s (*I*)*	11.2 (2.7)	6.2 (3.3)
Redundancy	4.2 (1.7)	2.2 (2.0)
R_{sym} (%)†	10.3 (37.2)	13.5 (33.8)
	Refinement	
Reflections (free)	51,049 (7,173)	
$R_{crystal}$‡ (R_{free})§ (%)	22.4 (29.8)	
Nonhydrogen protein atoms (Ca)	21,468 (1937)	

*I/s (I) is the mean reflection intensity divided by the estimated error. †$R_{sym} = |I_i - \langle I \rangle|/I_i$, where I_i is the intensity of an individual reflection and $\langle I \rangle$ is the average intensity over symmetry equivalents. ‡$R_{crystal} = ||F_o| - |F_c||/|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively. §R_{free} is equivalent to $R_{crystal}$ but calculated for a randomly chosen set of reflections that were omitted from the refinement process.

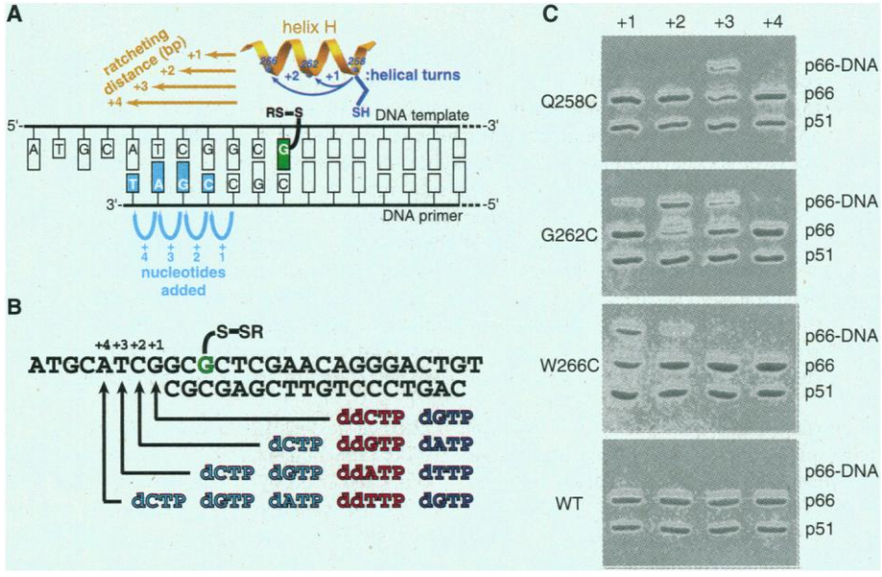

Fig. 2. Biochemical analysis of RT-DNA cross-linking. (**A**) Positional variations explored in the combinatorial search for efficient disulfide cross-linking between RT and DNA. The relative positions of the mutated cysteine and the tethered thiol group in DNA were varied (i) by introducing cysteine into successive turns of helix H (blue) and (ii) by having the polymerization reaction terminate after a defined number of cycles of the reaction (cyan), thereby ratcheting RT forward by a defined distance (gold) on the DNA substrate. This approach created 12 combinations of relative positions of the reacting thiol groups on RT and DNA. (**B**) The sequence of the template:primer and the constituents of the dNTP and ddNTP mixtures used in this study. Polymerization was initiated and continued by successive extensions with dNTPs (cyan) and terminated by incorporation of ddNTP (red). An additional dNTP (blue) then occupied the active site of RT, pairing with the corresponding base of the template. (**C**) SDS-PAGE analysis of the reactions between the three engineered cysteine mutants and the thiol-bearing template:primer, under the four polymerization conditions illustrated in (B). In the nonreducing PAGE used here with protein staining, disulfide cross-linking between the p66 subunit of RT and DNA results in the appearance of a new band having retarded mobility (p66-DNA), accompanied by reduced intensity of the p66 band itself.

of the cysteine by one additional turn along helix H resulted in a further one–base pair decrease in the optimal number of polymerization cycles (Fig. 2C, third panel). Wild-type RT failed to cross-link the thiol-tethered template:primer under identical reaction conditions (Fig. 2C, fourth panel). Of the three cysteine-engineered mutants studied here, RT(Q258C) exhibited the greatest degree of positional selectivity in disulfide cross-linking to DNA, and it was therefore chosen for structural analysis (*20*) (Table 1). The covalent trapping strategy described here may be useful with other protein-DNA complexes.

DNA Conformation and Protein Contacts

The DNA duplex binds along a groove, about 60 Å in length, stretching from the polymerase active site to the RNase H active site (Fig. 3A). The DNA has an A-like structure in the region of the polymerase active site, but with a distinctly wider major groove and smaller base pair tilt than the standard A conformation. Between base pair n-5 (the site of the covalent cross-link to the thumb) and n-11 (for numbering scheme, see Fig. 3B), the structure becomes gradually more B-like, but the major groove remains unusually deep for B DNA and the base pairs retain some tilt with respect to the helix axis. The A- to B-form transition is accompanied by an overall bend of about 40°, centered near base pair n-7. These features were seen in the DNA of the previously studied Rt•template:primer•Fab complex (*6*).

The 5′-template overhang bends away from the duplex and extends across the face of the fingers subdomain (Fig. 4A). Nucleotides n+2 and n+3 pack against the otherwise solvent-exposed residues Trp^{24}, Pro^{25}, Phe^{61}, and Ile^{63}. The course of the template overhang resembles that seen in the crystal structure of bacteriophage T7 DNA polymerase in template:primer complexes (*12*). It does not conform to the original suggestion that the template strand might pass through the cleft between fingers and thumb. Indeed, base n+1, the template nucleotide paired with an incoming dNTP, is wedged beneath the side chain of Leu^{74} and against the backbone of Gly^{152}, and these contacts clearly prevent the polynucleotide chain from continuing in a helical path (Fig. 4A).

The DNA duplex contacts mainly the p66 subunit of the RT heterodimer (Fig. 3B). The only direct interactions with bases occur in the minor groove, at positions n to n-3. There are van der Waals contacts with Pro^{157} and Met^{184} (base pair n) and with Ile^{94} (base pairs n-2 and n-3), and hydrogen bonds between Tyr^{183} and G(n-1). The prevalence of van der Waals contacts, even with hydrogen-bond acceptor atoms, recalls the minor-groove interaction made by TATA binding protein (TBP) (*21*) and differs from the hydrogen-bond networks seen in corresponding regions of other DNA polymerase complexes (*11, 12, 22*). Between nucleotides n and n-3, the sugar-phosphate backbone of the template strand contacts residues in the fingers and palm. These are largely van der Waals interactions, as might be expected for a contact along which the DNA backbone must move.

Between nucleotides n-1 and n-4 of the primer strand and n-5 and n-8 of the template strand, the minor groove faces the thumb (Fig. 3). The primer strand contacts the loop between palm and thumb [Met^{230} and Gly^{231}: the "primer grip" (*6*)], and both strands interact with residues in helices H and I. The backbone of the template lies in a positively charged groove between the COOH-terminus of helix I and strand β18 of the connection domain. The covalent cross-link joins residue 258 in helix H to G(n-5) of the template. Electron density for the cross-link is relatively weak, suggesting that the three-carbon linker adopts multiple conformations. The distance between the terminal carbons of the tether, 4.8 Å, is less than the maximum distance of about 7.5 Å (*23*), showing that the tether fits comfortably into the space between RT and DNA. Moreover, the segment of helix H and the DNA base pairs in the immediate vicinity of the cross-link exhibit normal structural parameters. Finally, the global features of the cross-linked complex, especially the B- to A-form transition of the template:primer and the presence of a metal-ligated dNTP in the active site, appear in various polymerase structures (*6, 11, 12*). We therefore conclude that the cross-link exerts little, if any, effect on the structure of the RT•template:primer•dNTP complex, but merely acts to prevent dissociation of the components.

Near the RNase H active site, the position of the template backbone appears to be close to the expected catalytic configuration for an RNA strand, even though a DNA duplex is not the correct substrate for the ribonuclease activity. In particular, there is a Mg ion bound by Asp^{443}, Asp^{549}, and phosphate n-17, which lies between Gln^{478} (mutated from Glu to inactivate the RNase H) and His^{539} [essential for activity in homologous enzymes (*24*)]. Arginine-448 inserts into the minor groove in the same region.

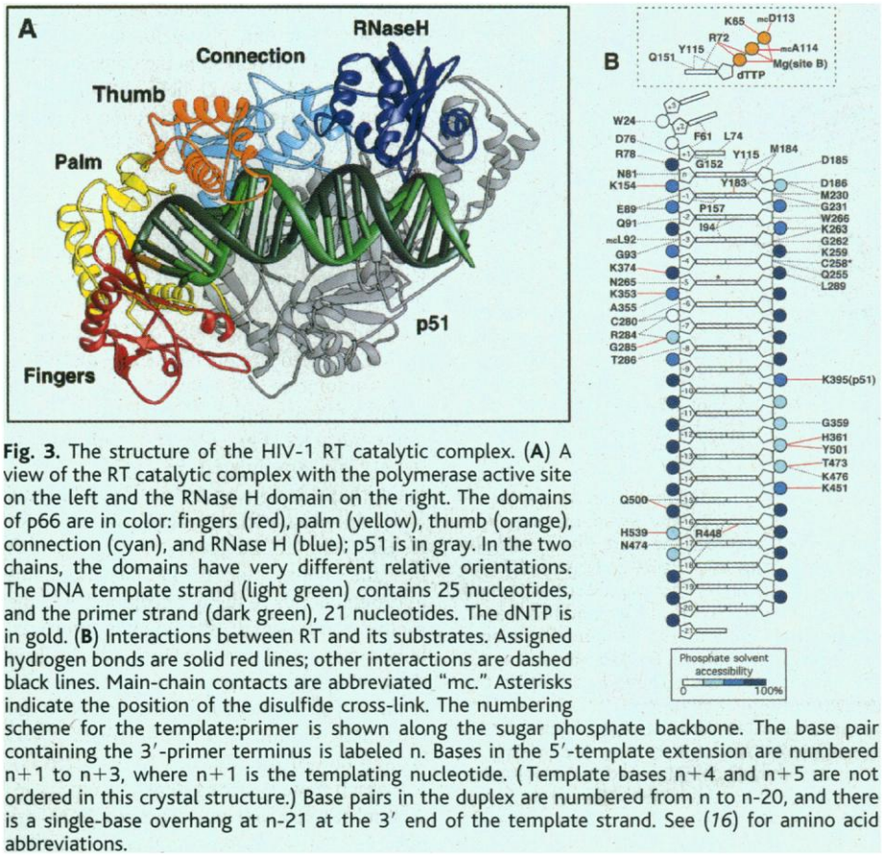

Fig. 3. The structure of the HIV-1 RT catalytic complex. (**A**) A view of the RT catalytic complex with the polymerase active site on the left and the RNase H domain on the right. The domains of p66 are in color: fingers (red), palm (yellow), thumb (orange), connection (cyan), and RNase H (blue); p51 is in gray. In the two chains, the domains have very different relative orientations. The DNA template strand (light green) contains 25 nucleotides, and the primer strand (dark green), 21 nucleotides. The dNTP is in gold. (**B**) Interactions between RT and its substrates. Assigned hydrogen bonds are solid red lines; other interactions are dashed black lines. Main-chain contacts are abbreviated "mc." Asterisks indicate the position of the disulfide cross-link. The numbering scheme for the template:primer is shown along the sugar phosphate backbone. The base pair containing the 3′-primer terminus is labeled n. Bases in the 5′-template extension are numbered n+1 to n+3, where n+1 is the templating nucleotide. (Template bases n+4 and n+5 are not ordered in this crystal structure.) Base pairs in the duplex are numbered from n to n-20, and there is a single-base overhang at n-21 at the 3′ end of the template strand. See (*16*) for amino acid abbreviations.

Conformational Changes in p66

Binding of the template:primer and dNTP induces significant conformational changes in parts of p66 (Fig. 5). The outer part of the fingers domain shifts significantly with respect to its base, bringing the "fingertips" closer to the palm. The palm moves slightly

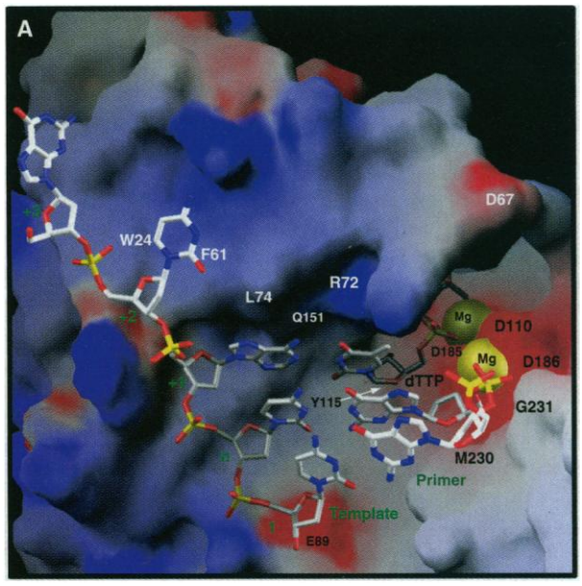

Fig. 4. The active site. (**A**) View into the dNTP pocket, formed by closure of the fingers domain (upper left, foreground) against the palm (center, background) and toward the thumb (lower right). The orientation corresponds to viewing from the right in Fig. 5A. Template and primer strands and dTTP are in stick representation; the protein is a surface rendering (red and blue indicate negative and positive electrostatic potential, respectively); Mg ions are yellow spheres. (**B**) Stereo view of polymerase active site. Template and primer strands, light and dark green, respectively; dTTP, gold; Mg ions, yellow spheres; assigned hydrogen bonds and metal ligand interactions, dashed lines.

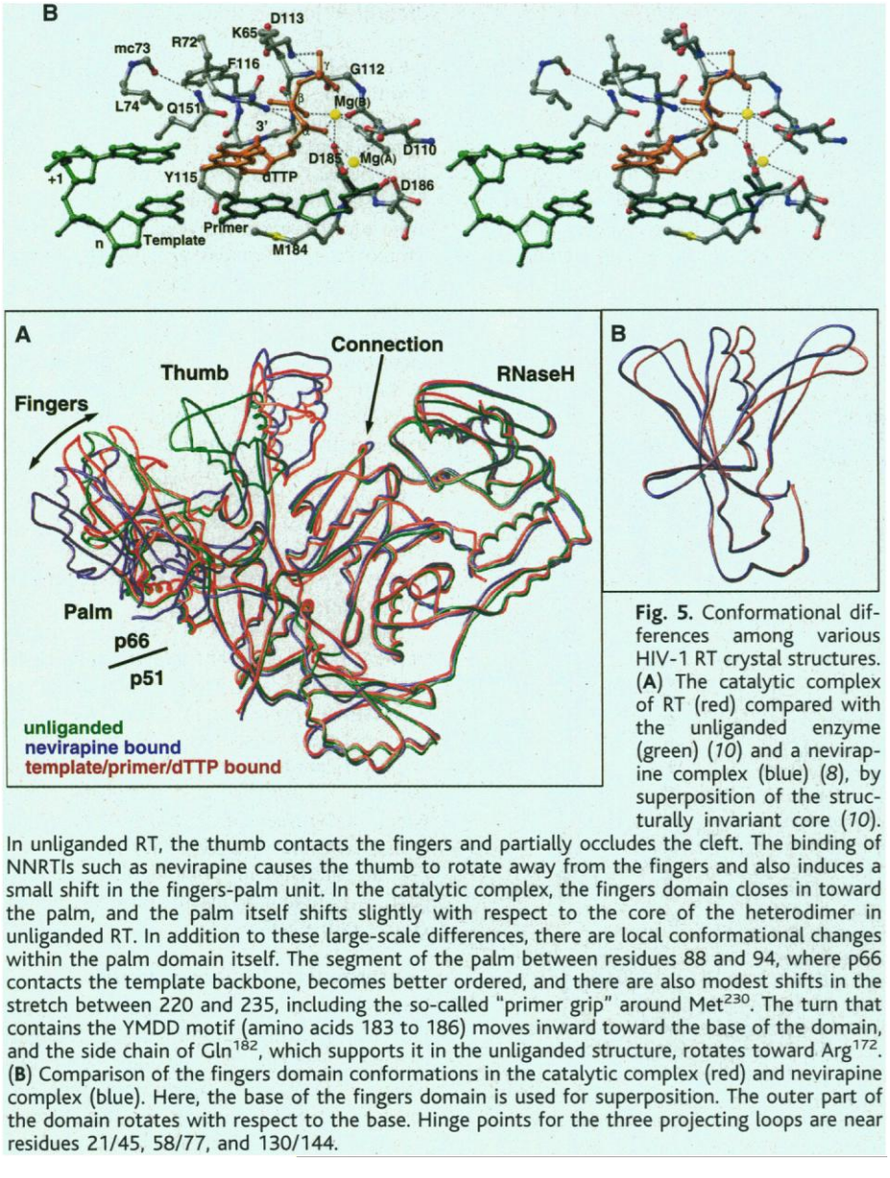

Fig. 5. Conformational differences among various HIV-1 RT crystal structures. (**A**) The catalytic complex of RT (red) compared with the unliganded enzyme (green) (*10*) and a nevirapine complex (blue) (*8*), by superposition of the structurally invariant core (*10*). In unliganded RT, the thumb contacts the fingers and partially occludes the cleft. The binding of NNRTIs such as nevirapine causes the thumb to rotate away from the fingers and also induces a small shift in the fingers-palm unit. In the catalytic complex, the fingers domain closes in toward the palm, and the palm itself shifts slightly with respect to the core of the heterodimer in unliganded RT. In addition to these large-scale differences, there are local conformational changes within the palm domain itself. The segment of the palm between residues 88 and 94, where p66 contacts the template backbone, becomes better ordered, and there are also modest shifts in the stretch between 220 and 235, including the so-called "primer grip" around Met^{230}. The turn that contains the YMDD motif (amino acids 183 to 186) moves inward toward the base of the domain, and the side chain of Gln^{182}, which supports it in the unliganded structure, rotates toward Arg^{172}. (**B**) Comparison of the fingers domain conformations in the catalytic complex (red) and nevirapine complex (blue). Here, the base of the fingers domain is used for superposition. The outer part of the domain rotates with respect to the base. Hinge points for the three projecting loops are near residues 21/45, 58/77, and 130/144.

(about 1 Å) toward the core, but most of the domain is essentially unchanged (Fig. 5A). The conformations of tyrosines 181 and 188, which rotate in response to NNRTI binding, are unperturbed with respect to unliganded RT (*10*). The position of the thumb agrees well with the one seen in the published complex with DNA and Fab (*6*).

The closure of the outer part of the fingers domain is similar to a bending inward of fingertips toward the palm by about 20° (Fig. 5B). This bending results in a relatively uniform distortion of the hydrophobic core of the fingers domain and in a displacement of the outermost parts of the domain by more than 5 Å. As a result, a number of residues are brought into contact with the incoming nucleoside triphosphate. Indeed, the large gap between fingers and palm seen in previous RT crystal structures is significantly closed off by the closure of the fingers, and what was a cleft in the NNRTI complexes is now almost a blind pocket for the dNTP. The outer loops of the fingers are relatively disordered in the "open" RT structures, and in most places they assume better defined positions in the catalytic complex. These loops contain sites of certain mutations that confer resistance to nucleoside analog inhibitors of RT.

Closure of the fingers is also a noteworthy feature of dNTP binding in members of both the DNA polymerase I (Pol I) and terminal transferase families (*11, 12, 25*). These enzymes have domains labeled fingers, thumb, and palm, but only the palm in the Pol I family has a fold related to the corresponding domain in RT. The O helix of the Pol I fingers is analogous in function and similar in position to the β3-β4 loop of RT.

The dNTP Site

The incoming nucleoside triphosphate—dTTP in this structure—pairs with the templating base. It remains unattacked because of the 2′,3′-dideoxyribonucleotide at the primer terminus. The base of the dTTP stacks on the terminus of the primer strand almost as in a continuous DNA strand, and the side chains of Arg^{72} and Gln^{151} pack against its outer surface (Fig. 4). The triphosphate moiety is coordinated by Lys^{65}, Arg^{72}, main-chain -NH groups of residues 113 and 114, and two metal ions (Mg). In our model, the guanidinium group of Arg^{72}, which lies flat against the dNTP base, donates hydrogen bonds to the α-phosphate, and the ε-amino group of Lys^{65} donates hydrogen bonds to the γ-phosphate (Fig. 4B). Both of these side chains move into position as a result of the fingers closure described in the preceding section. The triphosphate wraps around one of the Mg ions (designated B), for which there is strong density in the map. A nonbridging oxygen from each of the phosphates (pro-R_p for α and β) contributes

to the octahedral coordination of this metal; a backbone carbonyl (111) and side chains from aspartates 110 and 185 complete its shell. Density is weaker at the site for metal ion A, which would contact the 3′-OH of the primer terminus; the lower occupancy is probably due to the missing 3′-OH ligand. The likely octahedral coordination shell of this ion would include, in addition, all three aspartic acid side chains, the pro-R_p oxygen on the α-phosphate of the dNTP, and probably a water molecule. Overall, the conformation and metal ligation of the nucleoside triphosphate very closely resemble those of the dNTP in the T7 DNA polymerase replication complex, which has been studied at substantially higher resolution (*12*). A 3′-OH of the primer terminus would be in position for in-line attack on the α-phosphate, catalyzed by a two-metal mechanism, as expected from studies of numerous polymerases (*25, 26*).

The 3′-OH of the dTTP projects into a small pocket, lined by the side chains of Asp^{113}, Tyr^{115}, Phe^{116}, and Gln^{151}, and the peptide backbone between 113 and 115. The 3′-OH accepts a hydrogen bond from the main-chain -NH of Tyr^{115} and may donate one to the pro-S_p β-phosphate oxygen. Because most of the nucleoside analog drugs have modifications at the 3′ position, this "3′ pocket" is of particular significance for understanding structure-activity relations and for analyzing resistance mutations. The pocket can probably accommodate two or three water molecules in addition to the 3′-OH, and it clearly has room for the azido group of AZTTP. The side chain of Tyr^{115} is also important for modifications at the 2′ position of the dNTP. As proposed previously from the structure of a fingers-palm fragment of Moloney murine leukemia virus RT (*27*) and demonstrated by mutation of the homologous residue (a phenylalanine) to one with a smaller side chain (*28*), Tyr^{115} prevents incorporation of a ribonucleoside by interfering with a 2′-OH.

Steps in the Polymerase Reaction Cycle

HIV-1 RT binds its substrates in an ordered mechanism—the template:primer first, followed by the dNTP. Kinetic analysis indicates that after dNTP binding an apparent conformational change in the enzyme takes place, leading to attack by the 3′-OH of the primer terminus on the dNTP α-phosphorus and formation of a product complex (*29*). The postulated conformational change is the rate-limiting step in the forward reaction. We propose that our structure is likely to be a "snapshot" of the stage just after the rate-limiting step, in which the 3′-OH of the primer terminus is poised to attack the α-phosphorus.

The structure of unliganded HIV-1 RT probably represents another defined stage in the normal catalytic pathway, but only during initiation or reinitiation and not during processive synthesis. The position of contact between fingers and 5′ template extension in the present structure is near the contact between fingers and thumb in unliganded HIV-1 RT (*10*). Release of the fingers-thumb interaction that occurs in the unliganded molecule exposes Trp^{24} and Phe^{61} (contacts for the template extension, Fig. 4A), as well as the face of helix H. Movement of the thumb away from the fingers is not sufficient to cause rearrangement of the fingertips, however, as shown by their open position in both the previously described DNA complex [without dNTP (*6*)] and the NNRTI complexes (*3, 8, 9*). The nucleoside triphosphate itself, which interacts directly with two fingertip residues (Lys^{65} and Arg^{72}), may induce the closure.

The closing down of the fingertips traps both the template strand, by the interaction of Leu^{74} with the templating base, and the dNTP (Fig. 4A). It is likely that after the reaction the fingertips will bend back to their open position, leading to release of pyrophosphate and allowing the enzyme to bind the next dNTP. The reported structure of HIV-1 RT in complex with template:primer and Fab (*6*) is likely to be close to an "open" state of the enzyme after translocation to the next register but before

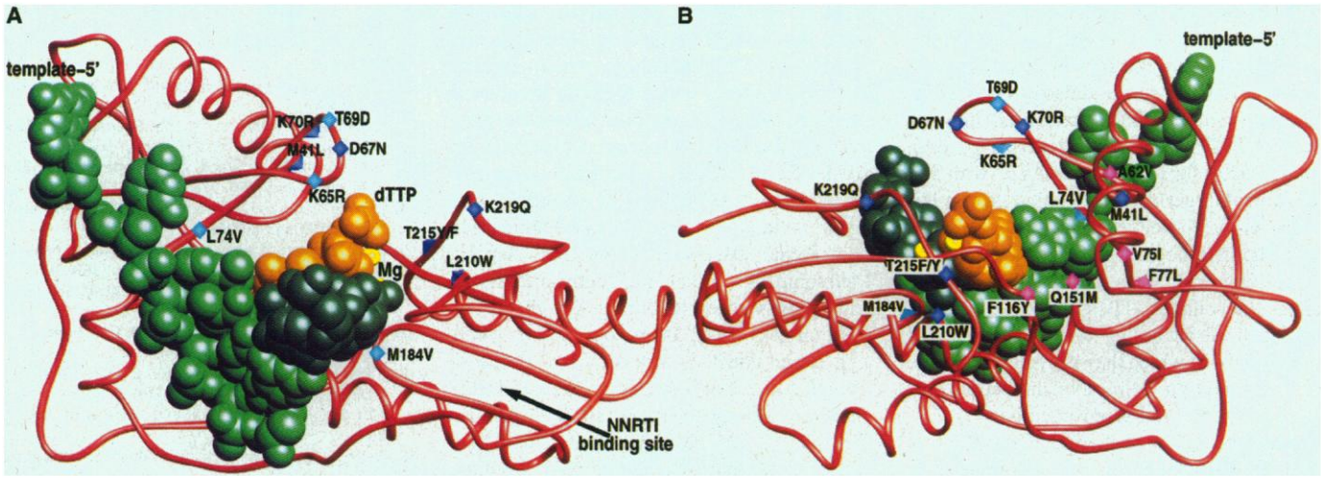

Fig. 6. Sites of mutations conferring resistance to various nucleoside analog drugs. (**A**) "Front" view, corresponding to the orientation in Fig. 4. The polypeptide backbone of the fingers and palm domains (residues 1 to 235) is shown as a red worm, and locations of resistance mutations are indicated by colored squares. The substrates are shown in Corey-Pauling-Koltun representation, with colors as in Fig. 3. The color code for mutations is as follows: light blue for resistance to ddI, ddC, and 3TC; blue for resistance to AZT; and violet for cross resistance to AZT and ddI or ddC. The location of the NNRTI binding site is shown by an arrow. Side chains of the residues at which mutations affect dideoxynucleotide sensitivity project forward: L74 bears on the templating base, and V at this position will also shift Q151 and R72 and hence the dNTP itself; M184 contacts the backbone and base at the primer terminus, and mutation to I or V will also generate a contact to the sugar ring of the dNTP; K65 contacts the γ-phosphate; and T69D (resistance to ddC) can probably best be explained by assuming a conformational effect on the fingers loop, transmitted to the dNTP by contacts from other fingers residues. (**B**) "Back" view, from the direction opposite to the one in (A). Side chains of AZT resistance mutations project toward this surface. One of the earliest mutations that appears in patients on AZT monotherapy is K70R. The Lys^{70} residue projects directly outward in the current model, but mutation to arginine (with five hydrogen-bond donors in fixed orientations on the guanidinium group) could readily induce side-chain reorientation, with contacts to Asp^{113} or the γ-phosphate. Subsequent appearance of T215Y/F confers higher levels of resistance. This mutation, likely to affect the rear of the 3′ pocket, is frequently "tuned" by appearance of others: K210W (which probably stabilizes the alteration at 215), M41L, and D67N and K219Q (which likely affect the interaction of fingers and palm and hence the formation of the 3′ pocket during the polymerization cycle) (*16*). Figure 4A was prepared with GRASP (*50*), and Figs. 3, 4B, 5, and 6, with RIBBONS (*51*).

binding of the dNTP. Flexing of the fingertips from open to closed must occur in synchrony with successive steps in the overall mechanism, as also suggested for enzymes of the Pol I family (*12*).

Mutations Conferring Resistance to Nucleoside Analog Inhibitors

The nucleoside analog inhibitors in current clinical use are chain terminators, modified at the 3′ position of the sugar ring. One (AZT) has a nonhydroxyl substituent at the 3′ position. The others (ddI, ddC, d4T, 3TC) have no 3′ substituent, and two have additional alterations in the sugar ring. Patterns of viral resistance have been studied in clinical isolates from early monotherapy trials as well as from more recent combination trials (*30–32*). Certain mutations are sufficient by themselves for drug resistance; others appear to "fine tune" or reinforce the effects of a primary event (*33*) (Fig. 6). We can now correlate these data with the crystal structure of a catalytic complex.

Two broad conclusions can be drawn from examination of the present structure. First, the point mutations sufficient for resistance [K65R, K70R, L74V, Q151M, M184I/V, and T215Y/F (*16*)] are all in the neighborhood of the incoming nucleotide, whereas many of the reinforcing changes are at a greater distance. The primary mutations are likely to affect directly the position, stability, or reactivity of the bound analog itself or additionally the templating base or primer terminus. Second, the sites of mutations conferring resistance to the dideoxy class of inhibitors (including 3TC) nearly all impinge on the dNTP from the "front" (Fig. 6A), whereas sites for resistance to AZT (with a 3′ substituent larger than a hydroxyl group) impinge on the 3′ pocket from the "rear" (Fig. 6B). The multinucleotide resistance mutation at position 151 is in the middle, directly across the bound dNTP. Thus, key mutations are localized according to the chemistry of the inhibitor to which they provide resistance. This observation helps explain patterns of cross-resistance and resensitization.

The orientation of the incoming dNTP is critically affected by a set of protein side-chain contacts as well as by the two metal ions. Small changes in the contacting residues can markedly influence the rate of nucleotide incorporation (or the rate of pyrophosphorolysis, which has been proposed to contribute to AZT resistance). The dNTP can also be influenced by residues that position the templating base or the primer terminus. The L74V mutation, which reduces sensitivity to ddI and ddC (*31*), affects a residue that locks the templating base tightly in place (Fig. 4). Leucine-74 also contacts the side chains of Arg^{72} and Gln^{151}, which in turn stack on the base of the dNTP. The mutation is therefore likely to influence dNTP binding directly, as well as through an effect on the template. The K65R mutation confers cross resistance to ddC, ddI, and 3TC (*34*). The lysine side chain contacts the γ-phosphate of the dNTP and stacks over Arg^{72} (Fig. 4). Substitution of Lys^{65} by arginine will thus induce packing rearrangements in precisely the same region as L74V. The M184V mutant, which appears rapidly in monotherapy trials with 3TC and exhibits some cross resistance to ddI and ddC (*35*), can influence both the dNTP and the primer terminus. The methionine side chain contacts the sugar and base of the 3′ nucleotide in the primer, but introduction of a β-branched side chain (isoleucine or valine) also creates a contact with the dNTP sugar ring. Modeling the correct (-) enantiomer of 3TC into our structure shows that interference with an isoleucine or valine at position 184 is enhanced (with respect to a conventional dNTP) by the configuration of the oxathiolane ring, accounting for the strong effect of these mutations on 3TC inhibition.

Understanding the molecular basis of AZT resistance is complicated by the absence of any measurable effect of resistance mutations on inhibition by AZT in vitro (*36*). Clustering of these mutations in the neighborhood of residues that create the 3′ pocket (Fig. 6B) argues for a direct effect on how the pocket accommodates the azido group of the drug, rather than for an indirect effect of some intracellular or viral cofactor. The important T215Y/F mutation, which confers high levels of resistance to AZT, involves a two-base change. Thus, presence of the phenyl ring of tyrosine or phenylalanine must be crucial for resistance, rather than elimination of the threonine. The bulky aromatic side chain is likely to contact the segment of polypeptide backbone containing Asp^{113}. Methionine-41, which frequently mutates to leucine in viruses harboring the change at position 215 (*37*), bears on Phe^{116}. Both Asp^{113} and Phe^{116} in turn contribute to the 3′ pocket, and changes in this pocket will influence its capacity to accept an azido group. The ddI and 3TC resistance mutations L74V and M184I/V partially restore AZT sensitivity to T215Y/F mutants, but the effect is not reciprocal (*31, 38*). The way front-face changes influence the position or stability of bound AZTTP may compensate for rear-face changes that reduce its efficiency of incorporation. Because they lack a 3′ substituent, the dideoxy analogs are likely to be insensitive to changes in the 3′ pocket.

Several resistance mutations map to the extreme tip of the β3-β4 loop, which is poorly ordered in our structure. These include T69D, associated with ddC insensitivity (*39*), and K70R, a primary site of AZT resistance (*32*). Studies of mutant complexes will be required to see what structural alterations such changes provide (see also caption to Fig. 6).

The Q151M mutation arises in a subset of patients on dual therapy trials (AZT and ddI or ddC) (*30, 40*). It affects both a contact to the dNTP and the character of the 3′ pocket. The mutation conserves the presence of a long side chain that can stack on the base of the dNTP; the hydrogen-bonding functions of the glutamine are evidently not essential for RT activity. The Q151M mutation is often followed by A62V, V75I, F77L, and F116Y. The first three changes probably modify the hydrophobic core of the fingers to adjust the orientation of the methionine. F116Y appears to restore a hydrogen-bond donor to the carbonyl of residue 73, just where it has been lost by the glutamine-to-methionine substitution. The residues in the polymerase of hepatitis B virus homologous to 116 and 151 are indeed tyrosine and methionine, respectively.

In summary, the pattern of nucleoside analog resistance mutations in HIV-1 RT fits logically with the structure of the closed complex. The clustering of mutations correlates well with cross resistance. Most of the current drugs in clinical use have no group to occupy the 3′ pocket, and cross resistance is common (*41*). Only AZT falls in the class with 3′ groups larger than -OH. Our analysis prompts further study of analogs with a sterically acceptable 3′ substituent.

References and Notes

1. H. Mitsuya, R. Yarchoan, S. Broder, *Science* **249**, 1533 (1990); D. Havlir and D. Richman, *Ann. Int. Med.* **124**, 984 (1996).
2. F. Di Marzo Veronese *et al.*, *Science* **231**, 1289 (1986).
3. L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice, T. A. Steitz, *ibid.* **256**, 1783 (1992).
4. J. W. Erickson and S. K. Burt, *Annu. Rev. Pharmacol. Toxicol.* **36**, 545 (1996).
5. D. D. Richman, in *Antiviral Chemotherapy* J. Mills, Ed. (Plenum, New York, 1996), vol. 4, pp. 383–395.
6. A. Jacobo-Molina *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6320 (1993).
7. J. Jager, S. J. Smerdon, J. Wang, D. C. Boisvert, T. A. Steitz, *Structure* **2**, 869 (1994).
8. J. Ren *et al.*, *Nature Struct. Biol.* **2**, 293 (1995).
9. J. Ding *et al.*, *Structure* **3**, 365 (1995); K. Das *et al.*, *J. Mol. Biol.* **264**, 1085 (1996); Y. Hsiou *et al.*, *ibid.*, p. 853.
10. D. W. Rodgers *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1222 (1995).
11. M. R. Sawaya, R. Prasad, S. H. Wilson, J. Kraut, H. Pelletier, *Biochemistry* **36**, 11205 (1997).
12. S. Doublie, S. Tabor, A. M. Long, C. D. Richardson, T. Ellenberger, *Nature* **391**, 251 (1998).
13. Disulfide cross-linking was chosen for the following reasons: (i) thiol-disulfide interchanges proceed readily under physiological conditions; (ii) disulfide bond formation is reversible and selective, and as such unlikely to yield an irrelevant cross-linked species through kinetic trapping, especially in the presence of a thiol-reducing agent; and (iii) efficient methods exist for site-specific attachment of tethered thiols to DNA (*14*) and for mutating other residues in a protein to cysteine.
14. D. Erlanson, J. Glover, G. L. Verdine, *J. Am. Chem. Soc.* **119**, 6927 (1997); D. Erlanson, S. Wolfe, L. Chen, G. Verdine, *Tetrahedron* **53**, 12041 (1997).
15. HIV-1 RT was expressed in *Escherichia coli* strain

BL21(DE3) as follows. An expression vector containing both p66 and p51 coding sequences was constructed, with independent promoters and attenuating sequences and opposite directions of transcription. The COOH-terminus of p51 also contained a six-histidine tag. Cysteine mutations were introduced into helix H of p66 to create tethering sites for a modified template:primer. In addition to a helix H mutation (Q258C, G262C, or W266C) [in which the first residue at the indicated position is mutated to the second residue (*16*)], each of the variant proteins used in this study had (i) the C280S mutation, introduced in prior crystallographic studies on RT (*6*), (ii) the E478Q mutation, introduced to eliminate RNase H catalytic activity (*42*) for purposes of future studies on RNA:DNA hybrids, and (iii) a P1K mutation to enhance expression levels. The three variant RT proteins retain one of the two cysteine residues present in wild-type p66 (C38) and both wild-type cysteine residues in p51 (C38 and C280). Cell pellets were lysed with a French press, and the enzyme was purified by Ni–nitrilotriacetic acid affinity chromatography, followed by cation exchange chromatography (sulfopropyl-Sepharose fast flow) and a final step of size-exclusion chromatography (Superdex 200).

16. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
17. K. Bebenek *et al.*, *Nature Struct. Biol.* **4**, 194 (1997).
18. We introduced the thiol tether into the template: primer through postsynthetic modification using the convertible nucleoside approach (*14*). Cystamine deprotection gave the C2 tether protected as the mixed 2-aminoethyl disulfide, which was used for biochemical studies; 3-aminopropane disulfide deprotection gave the C3 tether, which was used for crystallization. Template:primers containing the C2 and C3 tethers show virtually identical cross-linking specificity but exhibit different kinetics, with C3 being faster than C2 (*43*).
19. Each RT mutant (4 μM) was mixed with the thiol-tethered template:primer (10 μM) and a dNTP and ddNTP cocktail (100 μM each, refer to Fig. 2B) in the presence of 2 mM β-mercaptoethanol, which was added to increase the specificity of cross-linking. After a 2-hour incubation at 25°C, the cross-linking reaction was quenched by the addition of a thiol-capping reagent, methyl methanethiolsulfonate (20 mM), and the products were analyzed by SDS–polyacrylamide gel electrophoresis (SDS-PAGE) under nonreducing conditions.
20. The in situ polymerization–cross-linking procedure was carried out on a preparative scale, thereby generating a covalent complex comprising RT (Q258C) disulfide-bonded to a 27-nucleotide oligomer template:21-nucleotide oligomer primer. This complex was purified by anion-exchange chromatography on Mono-Q and then subjected to crystallization trials. The initial cross-linking reaction and subsequent crystallization trials were carried out in the presence of dTTP, with the aim of obtaining a structure containing the incoming nucleoside triphosphate bound in the enzyme active site (refer to Fig. 2). Crystals of RT•template:primer•dTTP complex were obtained by using hanging-drop vapor diffusion. Orthorhombic crystals ($P2_12_12_1$; a = 78.8 Å, b = 150.7 Å, c = 280.9 Å) grew as clustered plates in 1 to 2 weeks at 4°C. Drops were prepared by mixing a solution containing the purified RT-DNA complex [complex at 8 mg/ml , 10 mM MES (pH 6.5), 2 mM $MgCl_2$, 200 mM NaCl, 0.02 mM EDTA, 2.5 mM dTTP] mixed with an equal volume of the reservoir solution [14% (w/v) polyethylene glycol (PEG) 6000, 50 mM MES (pH 6.5), 1 M NaCl, 10 mM $MgCl_2$, 0.1 mM EDTA]. Nonreducing PAGE analysis of the crystals clearly indicated that RT(Q258C) and the template strand were quantitatively cross-linked (*43*), thus confirming the presence of the disulfide linker in the structure. Crystals were stabilized in a solution containing all the reservoir components, but with the concentration of PEG 6000 elevated to 16%, with 1 mM dTTP added. The crystals were briefly soaked (3 to 5 min) in a cryoprotecting solution containing all components of the stabilizing solution plus 25% glycerol, then mounted in a nylon loop and flash-frozen in liquid nitrogen.

Diffraction data were collected at the X25 beamline at National Synchrotron Light Source (NSLS) and the F1 beamline at Cornell High-Energy Synchrotron Source (CHESS) Laboratory (Table 1). Data were integrated, scanned, and reduced with the HKL software (*44*) and the CCP4 suite of crystallographic programs (*45*). The high-resolution structure of HIV-1 RT with a bound NNRTI (*8*) was used for molecular replacement calculations with programs from the CCP4 package and AMoRe (*46*). The search model was modified by removing the fingers, palm, and thumb of p66. With data from 12 to 3.5 Å, two solutions were evident, one significantly stronger than the other. The RNase H domains of the two molecules in the asymmetric unit form a continuous β sheet, as in previous RT crystal forms. Early stages of rebuilding [with the program O (*47*)] and adding back missing domains and nucleic acids exploited the two-fold noncrystallographic symmetry by iterative real-space averaging [program RAVE (*48*)]. Subsequent grouped-B refinement with X-PLOR (*49*) showed that the second molecule is less well ordered than the first, and in later stages we carried out all rebuilding in the density for the first molecule, using the coordinates thus obtained to generate both molecules for the next refinement round. Density for the A-form segment of the DNA duplex was evident in early stages, and the position of the DNA was confirmed by obtaining diffraction data from a complex with an iodinated primer (iodines at positions n-10 and n-17). Additional base pairs and the dNTP were added as the refinement proceeded. Phases were improved by many rounds of rebuilding and refinement, with noncrystallographic symmetry (NCS) restraints in all refinement protocols. NCS was defined by structural domains. Once all the atoms in the final model had been built, the thermal parameters were refined individually, with tight restraints. This procedure was necessary because of the significant differences in ordering of the two molecules in the asymmetric unit. The final structure (Table 1) has good geometry (root mean square standard deviations of bond lengths and bond angles are 0.014 Å and 2.7°, respectively).

21. J. L. Kim, D. B. Nikolov, S. K. Burley, *Nature* **365**, 520 (1993); Y. Kim, J. H. Geiger, S. Hahn, P. B. Sigler, *ibid.*, p. 512.
22. J. R. Kiefer, C. Mao, J. C. Braman, L. S. Beese, *ibid.* **391**, 304 (1998).
23. The distance between the terminal carbons was measured for an energy-minimized model -$\underline{C}H_2CH_2CH_2$-$SS\underline{C}H_2$- chain in the extended conformation, with a disulfide torsion angle of 90°. A tether with a distance of ~5 Å between the terminal carbons can be modeled by allowing one gauche torsional interaction in the tether, which requires little investment of energy.
24. Z. Hostomsky, Z. Hostomska, D. A. Matthews, in *Nucleases*, S. M. Linn, Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1993), pp. 341–376.
25. C. A. Brautigam and T. A. Steitz, *Curr. Opin. Struct. Biol.* **8**, 54 (1998).
26. T. A. Steitz, *ibid.* **3**, 31 (1993).
27. M. M. Georgiadis *et al.*, *Structure* **3**, 879 (1995).
28. G. Gao, M. Orlova, M. M. Georgiadis, W. A. Hendrickson, S. P. Goff, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 407 (1997).
29. R. A. Spence, W. M. Kati, K. S. Anderson, K. A. Johnson, *Science* **267**, 988 (1995); K. Rittinger, G. Divita, R. S. Goody, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8046 (1995).
30. R. W. Shafer *et al.*, *J. Infect. Dis.* **169**, 722 (1994).
31. M. H. St. Clair *et al.*, *Science* **253**, 1557 (1991).
32. B. A. Larder, G. Darby, D. D. Richman, *ibid.* **243**, 1731 (1989).
33. C. A. B. Boucher *et al.*, *J. Infect. Dis.* **165**, 105 (1992); G. J. Moyle, *J. Antimicrob. Chemother.* **40**, 765 (1997).
34. Z. Gu *et al.*, *Antimicrob. Agents Chemother.* **38**, 275 (1994); D. Zhang *et al.*, *ibid.*, p. 282.
35. M. Tisdale, S. D. Kemp, N. R. Parry, B. A. Larder, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 5653 (1993); R. Schuurman *et al.*, *J. Infect. Dis.* **171**, 1411 (1995).
36. R. Krebs, U. Immendorfer, S. H. Thrall, B. M. Wohrl, R. S. Goody, *Biochemistry* **36**, 10292 (1997); S. S. Carroll *et al.*, *ibid.* **33**, 2113 (1994); S. F. Lacey *et al.*, *J. Biol. Chem.* **267**, 15789 (1992).
37. P. Kellam, C. A. B. Boucher, B. A. Larder, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1934 (1992).
38. B. A. Larder, S. D. Kemp, P. R. Harrigan, *Science* **269**, 696 (1995).
39. J. E. Fitzgibbon *et al.*, *Antimicrob. Agents Chemother.* **36**, 153 (1992).
40. T. Shirasaka *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 562 (1993).
41. Mutations conferring resistance to abacavir, a carbocyclic dideoxy analog currently in clinical trials, overlap the set resistant to ddC, ddI, and 3TC [M. Tisdale, T. Alnadaf, D. Cousens, *Anitmicrob. Agents Chemother.* **41**, 1094 (1997)].
42. O. Schatz, F. Cromme, F. Gruninger-Leitch, S. F. J. Le Grice, *FEBS Lett.* **257**, 311 (1989).
43. H. Huang, R. Chopra, G. L. Verdine, S. C. Harrison, data not shown.
44. Z. Otwinowski and W. Minor, *Methods Enzymol.* **276**, 307 (1997).
45. CCP4, *Acta Crystallogr.* **D50**, 760 (1994).
46. J. Navaza, *ibid.* **A50**, 157 (1994).
47. T. A. Jones, J.-Y. Zhou, S. W. Cowan, M. Kjeldgaard, *ibid.* **A47**, 110 (1991).
48. T. A. Jones, in *Molecular Replacement*, E. J. Dodson, S. Gover, W. Wolf, Eds. (SERC Daresbury Laboratory, Warrington, UK, 1992), pp. 91–105.
49. A. T. Brünger, *X-PLOR version 3.1. A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
50. A. Nicholls, K. A. Sharp, B. Honig, *Proteins Struct. Funct. Genet.* **11**, 281 (1991).
51. M. Carson, *J. Mol. Graphics* **5**, 103 (1987).

52. We especially acknowledge B. Harris, who helped initiate this project; we thank the staffs of beamlines X25 at NSLS and F1 at CHESS; M. Lawrence, M. Jacobs, K. Leong, D. Erlanson, and R. Butcher and other members of the Harrison-Wiley and Verdine research groups; S. Doublie, T. Ellenberger, R. D'Aquila, and R. Colgrove, for discussion and advice; and E. Arnold, for coordinates of a refined complex before publication. The work was supported by NIH grants GM-39589 (to S.C.H.) and GM-44853 (to G.L.V.), NIH postdoctoral award GM-18621 (to H.H.), and by the Landon T. Clay Research Endowment (Harvard University). S.C.H. is an Investigator in the Howard Hughes Medical Institute. The coordinates of the complex have been deposited in the Brookhaven Protein Databank (1RTD).

3 August 1998; accepted 23 October 1998