Structure of HIV-1 reverse transcriptase in a complex with the non-nucleoside inhibitor α-APA R 95845 at 2.8 Å resolution

J Ding[1], K Das[1], C Tantillo[1], W Zhang[1], AD Clark, Jr[1], S Jessen[1], X Lu[1], Y Hsiou[1], A Jacobo-Molina[1†], K Andries[2], R Pauwels[3], H Moereels[2], L Koymans[2], PAJ Janssen[2], RH Smith, Jr[4,5], M Kroeger Koepke[4], CJ Michejda[4], SH Hughes[4] and E Arnold[1*]

[1]Center for Advanced Biotechnology and Medicine (CABM) and Rutgers University Chemistry Department, 679 Hoes Lane, Piscataway, NJ 08854-5638, USA, [2]Janssen Research Foundation, Turnhoutseweg 30, B-2340 Beerse, Belgium, [3]TIBOTEC, Institute for Antiviral Research, Drie Eikenstraat 661, B-2650 Edegem, Belgium, [4]ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, P.O. Box B, Frederick, MD 21702-1201, USA and [5]Department of Chemistry, Western Maryland College, Westminster, MD 21157, USA

Background: HIV-1 reverse transcriptase (RT) is a multifunctional enzyme that copies the RNA genome of HIV-1 into DNA. It is a heterodimer composed of a 66 kDa (p66) and a 51 kDa (p51) subunit. HIV-1 RT is a crucial target for structure-based drug design, and potent inhibitors have been identified, whose efficacy, however, is limited by drug resistance.

Results: The crystal structure of HIV-1 RT in complex with the non-nucleoside inhibitor α-anilinophenylacetamide (α-APA) R 95845 has been determined at 2.8 Å resolution. The inhibitor binds in a hydrophobic pocket near the polymerase active site. The pocket contains five aromatic amino acid residues and the interactions of the side chains of these residues with the aromatic rings of non-nucleoside inhibitors appear to be important for inhibitor binding. Most of the amino acid residues where mutations have been correlated with high levels of resistance to non-nucleoside inhibitors of HIV-1 RT are located close to α-APA. The overall fold of HIV-1 RT in complex with α-APA is similar to that found when in complex with nevirapine, another non-nucleoside inhibitor, but there are significant conformational changes relative to an HIV-1 RT/DNA/Fab complex.

Conclusions: The non-nucleoside inhibitor-binding pocket has a flexible structure whose mobility may be required for effective polymerization, and may be part of a hinge that permits relative movements of two subdomains of the p66 subunit denoted the 'palm' and 'thumb'. An understanding of the structure of the inhibitor-binding pocket, of the interactions between HIV-1 RT and α-APA, and of the locations of mutations that confer resistance to inhibitors provides a basis for structure-based design of chemotherapeutic agents for the treatment of AIDS.

Structure 15 April 1995, **3**:365–379

Key words: AIDS, drug resistance, mechanism of non-nucleoside inhibition, polymerase structure, protein–drug interactions

Introduction

Human immunodeficiency virus type 1 (HIV-1) is the causative agent of AIDS. HIV-1 reverse transcriptase (RT) is responsible for the transformation of the single-stranded RNA genome of HIV-1 into the linear double-stranded (ds) DNA that becomes permanently integrated into the chromosomes of the host cells. RT is the therapeutic target of many of the agents that inhibit replication of HIV-1 and has been the subject of extensive scientific studies (reviewed in [1–3]). Currently, two functionally distinct classes of HIV-1 RT inhibitors have been discovered and are being used clinically or are in clinical trials [3,4]. The first of these, the nucleoside analog inhibitors, are analogs of normal deoxynucleoside triphosphates (dNTPs). The only drugs now used for the treatment of AIDS patients, 3′-azido-3′-deoxythymidine (AZT), 2′,3′-dideoxyinosine (ddI), 2′,3′-dideoxycytidine (ddC), and 2′,3′-didehydro-2′,3′-dideoxythymidine (d4T) are nucleoside analogs. However, nucleoside analogs are not specific for HIV-1 RT, and are incorporated into cellular DNA by the host DNA polymerases and can therefore cause serious side effects. Moreover, treatment with these analogs leads to the emergence of drug-resistant viral strains that contain mutations in HIV-1 RT (reviewed in [3–5]). The other class of HIV-1 RT inhibitors, known as non-nucleoside inhibitors, include tetrahydroimidazo(4,5,1-1-*jk*)(1,4)-benzodiazepin-2(1*H*)-one and -thione (TIBO) derivatives [6], dipyridodiazepinones [7], pyridinones [8], bis(heteroaryl)piperazines (BHAPs) [9], and 2′,5′-bis-O-(tertbutyldimethylsilyl)-3′-spiro-5″-(4″-amino-1″,2″-oxathiole-2″,2″-dioxide)pyrimidine (TSAO) derivatives [10] (reviewed in [3,11]). The α-anilinophenylacetamide (α-APA) compounds (Fig. 1) were developed at Janssen Research Foundation [12] and represent an additional class of highly potent and specific non-nucleoside inhibitors. The dichlorinated

*Corresponding author. †Present address: Centro de Biotecnologia y Departamento de Quimica, ITESM, Sucursal de Correos "J", Monterrey, N.L. 64849, México.

 ISSN 0969-2126

compound α-APA R 90385 has an IC_{50} (50% inhibitory concentration in MT-4 cells) of 5 nM and a selectivity index (50% cytotoxic concentration/50% inhibitory concentration) of 84 000. In addition, these compounds are chemically simple and economical to synthesize, exhibiting potential for bulk production and rapid development of new analogs [12]. R 89439 (the racemic mixture of R 90385 and its enantiomer), which has been named Loviride, is currently being tested in clinical trials in AIDS patients and has been shown to induce a prolonged rise in the CD4 count in asymptomatic HIV-1 positive patients (M De Brabander, *et al.*, & PA Janssen, abstract PB0242, p. 203, The Tenth International Conference on AIDS, Japan, August, 1994; S Staszewski, *et al.*, & R Van Den Broeck, abstract 513B, p. 59, The Tenth International Conference on AIDS, Japan, August, 1994). In contrast to nucleoside analogs, non-nucleoside inhibitors are highly specific for HIV-1 RT and do not inhibit either HIV-2 RT or normal cellular polymerases, thus exhibiting lower cytotoxicity and fewer side effects. However, non-nucleoside inhibitors also induce the emergence of resistant strains of virus [3,5,12,13]. Multidrug combination therapy has been tried, but drug-resistant viral strains still emerge [14].

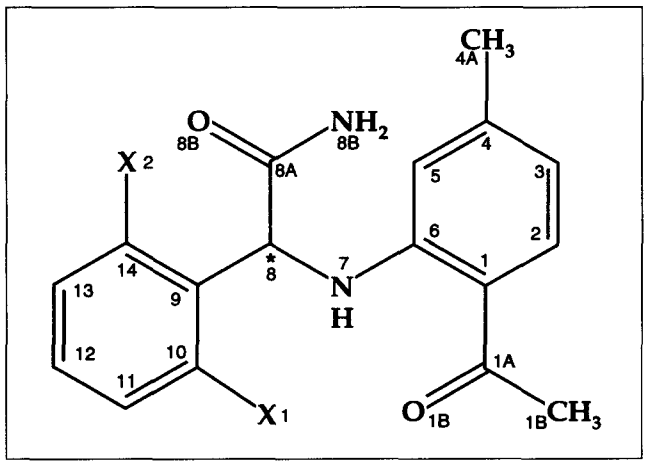

Fig. 1. Chemical structure of α-anilinophenylacetamide, α-APA. In the analog R 95845, X=Br and in R 90385, X=Cl. The asterisk denotes the location of the chiral carbon atom which has the *S* configuration in both R 95845 and R 90385.

HIV-1 RT is a heterodimer that contains two related polypeptides of 66 kDa (p66) and 51 kDa (p51). The polymerase domain of HIV-1 RT comprises the N-terminal 440 amino acid residues of the p66 subunit. The ribonuclease H (RNase H) domain comprises the remaining 120 amino acid residues at the C terminus of p66. The p51 subunit, which is proteolytically derived from p66 or from some larger precursor, has the same amino acid sequence as the polymerase domain of p66. The polymerase domains of both p66 and p51 contain four subdomains, denoted fingers, palm, thumb and connection [15]. The overall fold of individual subdomains is similar in both subunits. However, the relative spatial arrangement of the subdomains in the two subunits is surprisingly different. In the p66 subunit, the fingers, palm and thumb subdomains form a large cleft that accommodates the template-primer dsDNA [16]. The trio of aspartic acid residues (110, 185 and 186) that is strictly conserved in retroviral RTs and forms part of the polymerase active site, lies on the 'floor' of this binding cleft.

The three-dimensional structure of HIV-1 RT has been determined by X-ray crystallography in a complex with the non-nucleoside inhibitor nevirapine [15,17] and in a ternary complex with a 19-mer/18-mer dsDNA template-primer and the Fab of a monoclonal antibody (HIV-1 RT/DNA/Fab) [16] (reviewed in [18]). The structure of HIV-1 RT in the absence of inhibitor or template-primer was also described (R Raag, *et al.*, & E Arnold, abstract B03, p. 44, American Crystallographic Association Meeting, Atlanta, June, 1994; [19,20]). The structure of a truncated version of HIV-1 RT corresponding to the fingers and palm subdomains has been reported at 2.2 Å resolution [21].

Here we report the crystal structure of HIV-1 RT in a complex with the non-nucleoside inhibitor α-APA R 95845 (the dibrominated form of α-APA R 90385) at 2.8 Å resolution and the interactions between the bound inhibitor and the amino acid residues forming the inhibitor-binding pocket. There are significant differences in the structures of HIV-1 RT/α-APA and HIV-1 RT/DNA/Fab complexes. The potential functional implications of these differences are discussed.

Results and discussion

Summary of the structure determination and refinement

In the following discussion, RT will be used to refer to HIV-1 RT unless otherwise specified. A summary of the crystallographic data is given in Table 1. The initial phasing was determined using the molecular replacement (MR) method as implemented in the program X-PLOR [22]. The starting model was a polyalanine model of RT, built on the basis of the coordinates from the RT/DNA/Fab complex ([16]; J Ding *et al.*, unpublished data), using the partial Cα model of the RT/nevirapine complex [15] as a guide. The MR phasing was supplemented by multiple isomorphous replacement (MIR) phasing, solvent flattening, phase extensions, and iterative combination of partial model and MIR phases. Electron-density map averaging from multiple crystal forms of RT/non-nucleoside inhibitor complexes yielded the most reliable maps for model building. The location of α-APA was determined using the difference Fourier method. The orientation and conformation of α-APA was interpreted on the basis of the difference Fourier (F_o–F_c) maps and the difference Fourier electron density between the RT/R 95845 complex and the RT/R 90385 complex, where R 95845 and R 90385 are dibrominated and dichlorinated analogs of α-APA, respectively (Fig. 2a). Molecular modeling was carried out iteratively using the graphics program 'O' [23]. The structure refinement was performed using conventional

Table 1. Summary of crystallographic data.

Dataset	Native (R 95845)	Hg-dUTP	TAMM	Native (R 90385)
No. of crystals	15	7	14	11
No. of images	136	47	65	84
Resolution (Å)	2.6	3.0	3.0	2.6
Unique reflections[a]	37 099	13 524	21 615	34 168
Observed reflections	113 074	37 642	48 594	91 341
Completeness (%)	78	41	67	71
R_{merge}[b]	0.10	0.14	0.11	0.11
R_{deriv}[c]		0.37	0.34	
Phasing power[d]		1.84	1.64	
$R_{centric}$[e]		0.76	0.69	
FOM[f]		0.32	0.34	

Abbreviations: Hg-dUTP, 5-mercurideoxyundine 5′-triphosphate; TAMM, tetrakis(acetoxymerucuri)methane. [a]The reflections listed here have I≥2σ(I); in multiple isomorphous replacement (MIR) phase calculations and structure refinement, we used reflections with F≥3σ(F). [b]$R_{merge} = \Sigma|I_{obs} - \langle I \rangle|/\Sigma\langle I \rangle$. [c]$R_{deriv} = \Sigma||I_{PH} - I_P|/\Sigma I_P$. [d]Phasing power $= \sqrt{[\Sigma|F_H|^2/\Sigma(|F_{PH,obs}| - |F_{PH,calc}|)^2]}$. [e]$R_{centric} = \Sigma||F_{PH} \pm F_P| - F_H|/\Sigma|F_{PH} - F_P|$. [f]The overall MIR figure of merit (FOM) is 0.40 for 21 810 reflections to 3.0 Å resolution [F≥3σ(F)] and 0.81 after iterative solvent flattening (see the Materials and methods section).

positional refinement and the simulated annealing protocol within X-PLOR [22]. The statistics of model refinement are summarized in Table 2.

Overall structure of the HIV-1 RT p66/p51 heterodimer

Fig. 2b shows a representative portion of a $2mF_o-F_c$ (SIGMAA-weighted [24,25]) difference Fourier map at 2.8 Å resolution in the region around the non-nucleoside inhibitor-binding pocket. Overall, the electron density for most of the amino acid residues and the inhibitor was clearly defined. The structure of the RT/α-APA complex is shown in Fig. 3. Most of the p66/p51 heterodimer in the RT/α-APA complex has the same secondary-structure assignment as was given for the RT/DNA/Fab complex (Fig. 2 of [16]). As in the RT/nevirapine and RT/DNA/Fab structures, the polymerase domain of the p66 subunit resembles a right hand and the fingers, palm, and thumb subdomains form a cleft that can accommodate the template-primer substrate [15,16]. The thumb subdomain of p66 in the RT/α-APA structure is not folded down into the DNA-binding cleft as it is in the RT structures lacking bound inhibitor or DNA (R Raag, *et al.*, & E Arnold, abstract B03, p. 44, American Crystallographic Association Meeting, Atlanta, June, 1994; [20]); instead, it is in an upright position. This position is similar to that of the p66 thumb in the RT/nevirapine and RT/DNA/Fab structures, suggesting that both non-nucleoside inhibitors and nucleic acid template-primers influence the position of the thumb. As in the RT/DNA/Fab and the RT/nevirapine structures, the p66/p51 heterodimer interface is primarily formed by interactions in three regions: first, between the p66 palm (residues 85–90, 93–96, 99–100 and 161–162) and the p51 fingers (residues 52–57, 131, 136–137, 140 and 143) subdomains; second, between the p66 connection (residues 381–384, 402 and 405–410) and the p51 connection (residues 331, 364–365, 392–394, 397, 401, 405 and 417–419) and fingers (27–28 and 136–137) subdomains;

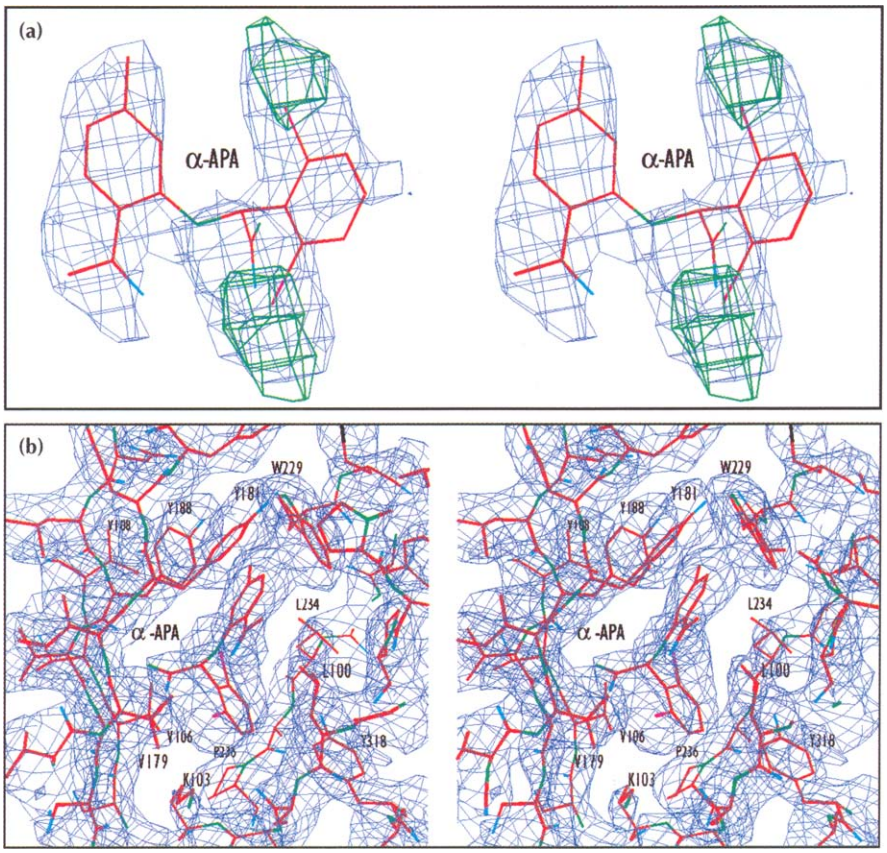

Fig. 2. (a) Stereoview of difference Fourier maps showing the fit of α-APA into the electron density. The α-APA coordinates correspond to the current refined model. The green map is the difference Fourier map at 3.5 Å resolution between the HIV-1 RT/α-APA R 95845 (dibrominated) complex and the HIV-1 RT/α-APA R 90385 (dichlorinated) complex contoured at a 10σ level which revealed the positions of two bromine atoms. The blue map is the mF_o-F_c difference Fourier map at 3.0 Å resolution calculated from a model before α-APA was included in the refinement; the phases were computed from the back-transformation of electron density averaged from three HIV-1 RT/inhibitor complexes; the map is contoured at 1.5σ. **(b)** Stereoview of a portion of a $2mF_o-F_c$ difference Fourier map at 2.8 Å resolution in the region around the non-nucleoside inhibitor-binding pocket showing the α-APA inhibitor and some of the nearby amino acid residues. The phases were computed from the current atomic model (R=0.255) and the map is contoured at 1.2σ.

Table 2. Statistics of model refinement.

Resolution range	10.0–2.8 Å
No. of reflections	31 444 (working set: 30 186; test set: 1258)
Completeness	81.8% (working set: 78.5%; test set: 3.3%)
R-factor[a]	0.255
Free R-factor	0.360
Number of:	
protein residues	985 (p66: 1–558; p51: 1–427)
protein non-H atoms	7792 (p66: 4368; p51: 3424)
inhibitor non-H atoms	23
Averaged B-factor	39 Å^2 (p66: 43 Å^2, p51: 34 Å^2)
Rms bond lengths	0.014 Å
Rms bond angles	0.81°
Rms dihedral angles	25.24°
Rms improper angles	1.64°

[a] $R = \Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$.

and third, between the RNase H (residues 432–436, 439–441, 459, 500, 532, 536 and 541–546) and the p51 thumb (residues 255, 258–259 and 280–290) and connection (residues 421–422) subdomains (Fig. 3) [26,27].

Location of the α-APA inhibitor and nature of the non-nucleoside inhibitor-binding pocket

The α-APA inhibitor is located in the core of a highly hydrophobic pocket, referred to as the non-nucleoside inhibitor-binding pocket, composed primarily of amino acids from secondary-structure elements of the p66 palm subdomain: the β5b–β6 loop (Pro95, Leu100, Lys101 and Lys103), β6 (Val106 and Val108), the β9–β10 hairpin (Val179, Tyr181, Tyr188 and Gly190), and the β12–β13 hairpin (Phe227, Trp229, Leu234 and Pro236), as well as β15 (Tyr318) from the p66 thumb subdomain (Fig. 4). The pocket lies near, but is distinct from, the polymerase active site which contains the highly conserved Tyr-Met-Asp-Asp (YMDD) motif at positions 183–186. A possible solvent-accessible entrance to the pocket is located at the p66/p51 heterodimer interface and the residues at the rim of the putative entrance are Leu100, Lys101, Lys103, Val179 and Tyr181 of p66, and Glu138 of the β7–β8 loop of the p51 fingers (Fig. 4). Most of the amino acid residues that form the binding pocket are hydrophobic and five of them have aromatic side chains. The only hydrophilic residues around the pocket (Lys101 and Lys103 of p66, and Glu138 of p51) are located at the putative entrance. The potential roles of these residues are not yet clear but all confer resistance to non-nucleoside inhibitors when mutated (see below). It is conceivable that their flexible and polar side chains may act to recognize molecules that are trying to enter the pocket and/or to prevent the bound inhibitor escaping from the pocket.

Only one non-nucleoside inhibitor-binding site is found in the p66/p51 heterodimer [15]. Relative to the p66 subunit, the thumb subdomain of p51 is rotated away from the DNA-binding cleft and the connection subdomain is folded up onto the palm between the fingers and thumb. The region of p51 corresponding to the β12–β13–β14 sheet of p66, which forms an important part of the binding pocket, is completely unraveled and extended in the RT/α-APA structure. The movement of the thumb also correlates with the movement of β15.

Interactions of the α-APA inhibitor with specific amino acid residues

Table 3 lists the interactions between α-APA and the surrounding amino acid residues. Because of the hydrophobic nature of the non-nucleoside inhibitor-binding pocket and the inhibitor, most of the close contacts between α-APA and the surrounding residues are hydrophobic interactions. There are 43 protein–inhibitor interactions with a distance of <3.6 Å, two of which are potential hydrogen-bonding interactions between the amide group of α-APA and the carbonyl groups of amino acid residues Tyr188 (Tyr188 O...N8B α-APA, 2.8 Å) and Val189 (Val189 O...N8B α-APA, 3.2 Å) (Table 3).

The non-nucleoside inhibitors are a diverse class of compounds with different chemical structures, but they are all hydrophobic and contain aromatic groups. We proposed that interactions between the aromatic rings of inhibitors and the aromatic side chains of Tyr181, Tyr188 and other residues in the binding pocket would be important for the binding of most non-nucleoside inhibitors [18]. Analysis of the inhibitor–protein interactions shows that the two

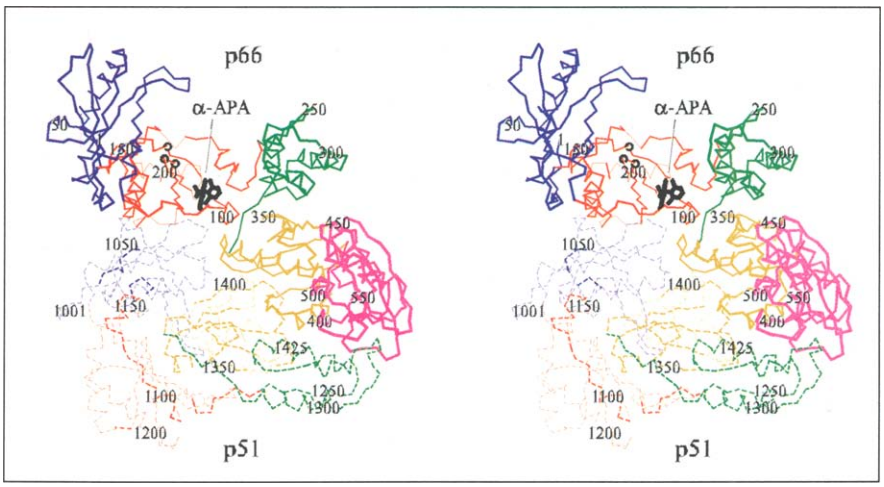

Fig. 3. Stereo Cα trace of the HIV-1 RT p66/p51 heterodimer in complex with α-APA showing the overall structure, the heterodimer interface, the location of the polymerase active site, and the location of the non-nucleoside inhibitor-binding pocket. The p66 and p51 subunits are drawn in solid and dashed lines,respectively. The bound inhibitor is shown as a ball-and-stick model and the three aspartic acids at the polymerase active site are shown as open circles. Every 50th residue is labeled: p66, 1–558 and p51, 1001–1427. The subdomains are colored as: fingers, blue; palm, red; thumb, green; connection, yellow; and RNase H, magenta. (Figure generated using MOLSCRIPT [64].)

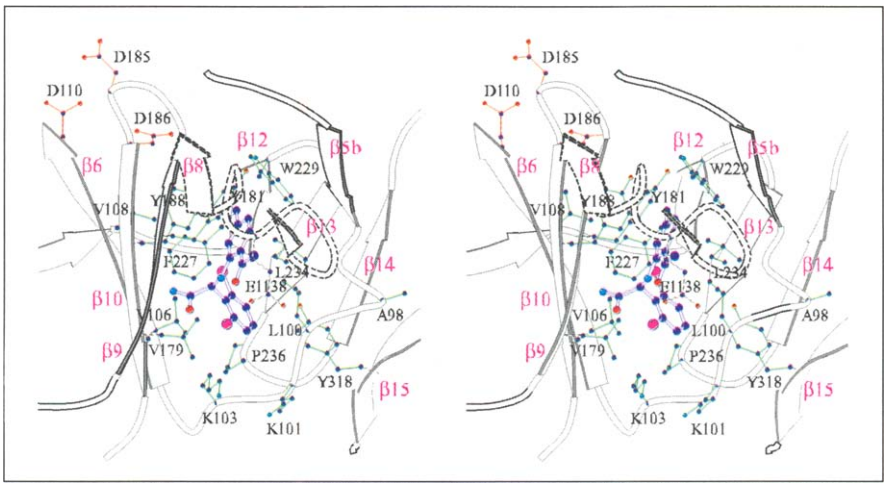

Fig. 4. Close-up of the non-nucleoside inhibitor-binding pocket in the structure of the HIV-1 RT/α-APA complex looking through a putative entrance to the pocket, showing interactions between α-APA and nearby amino acid residues. α-APA is shown in purple as a ball-and-stick model with carbons purple, nitrogens cyan, oxygens red and bromines magenta. The β7–β8 portion of p51 has a dashed outline. The side chains are shown for the amino acid residues that make close contacts with α-APA (in green), and for the three essential aspartic acid residues D110, D185 and D186 (in red) at the polymerase active site. Dashed lines indicate connections between the side chains and the β-strands.

aromatic rings of α-APA have potentially significant aromatic–aromatic interactions with the aromatic amino acid residues in the pocket (Table 4). The centroid-to-centroid distances are within the energetically favorable range (3.9–6.8 Å) commonly observed in protein structures [28]. As expected, the aromatic–aromatic interactions in the current structure (listed in Table 4) are either of the favorable edge-to-face (tilted T) or of the favorable offset stacked (parallel displaced) types [29] (Fig. 4). Interactions among the aromatic side chains of the amino acid residues forming the pocket could influence the structure of the pocket and could help explain the phenotypic consequences of some of the mutations that confer resistance to non-nucleoside inhibitors (see below).

Non-nucleoside inhibitor-resistant mutations are located close to the bound inhibitor: implication for mechanisms of resistance to non-nucleoside inhibitors

Most of the non-nucleoside inhibitor-resistant mutation sites that have been reported (for example, Ala98, Leu100, Lys101, Lys103, Val106, Val108, Glu138, Val179, Tyr181, Tyr188, Gly190 and Pro236) map to amino acid residues that are located close to the bound α-APA inhibitor (Fig. 4). The shortest distances between the α-APA R 95845 and the residues whose mutation confers resistance to non-nucleoside inhibitors are also given in Table 3. With the exception of Ala98 and Val108, each of the commonly identified non-nucleoside inhibitor-resistant mutations in the p66 subunit has a contact at a distance of ~4 Å or less.

The nature of the mutations that cause resistance varies and the potential mechanism of resistance will depend on the specific amino acid change. Pauwels *et al.* [12] reported that the mutation Leu100→Ile in HIV-1 RT caused an approximately three-fold drop in sensitivity of HIV-1 to α-APA R 89439. This moderate resistance may be due to a mild steric conflict with the inhibitor, as the side chain of Leu100 makes a close contact with the bound α-APA (Table 3). Changes at position 181, in which the tyrosine is mutated to cysteine or isoleucine, have much more dramatic consequences [12]. A viral variant that contains the Tyr181→Cys mutation, which has been identified as causing high-level resistance to nearly all non-nucleoside inhibitors [5,30], has >4000-fold resistance to R 89439 in cell culture. In addition, recombinant HIV-1 RTs that contain Tyr181→Cys or Tyr181→Ile mutations show a dramatic loss of sensitivity to inhibition by α-APA R 89439 [12]. The loss of the aromatic side chain at residue 181 could cause a significant reduction in the favorable aromatic ring interactions. For example, Tyr181 has aromatic–aromatic interactions with α-APA, the side chains of Tyr188 and Trp229 (both of which are also involved in the aromatic cluster with α-APA), and the side chain of Tyr183, hence, mutation of Tyr181 may influence the conformations of the inhibitor and these amino acid residues. In addition, because Tyr181 is located at what may serve as an entrance to the pocket (Fig. 4), the mutations Tyr181→Cys and Tyr181→Ile eliminate aromatic–aromatic interactions with non-nucleoside inhibitors that could help 'steer' the compounds into the pocket. The importance of the interaction between Tyr188 and α-APA is supported by the observation that recombinant HIV-1 RT containing a Tyr188→Leu mutation is >1000-fold more resistant to R 89439 than is wild-type HIV-1 RT [12]. The aromatic ring of Tyr188 has stabilizing interactions with aromatic moieties of α-APA, Tyr181 and Phe227 (Table 4). The importance of having aromatic side chains at both Tyr181 and Tyr188 for binding of non-nucleoside inhibitors is also supported by site-directed mutagenesis studies showing that HIV-1 RT with tryptophan or phenylalanine at these positions shows little or no resistance to non-nucleoside inhibitors [31]. Most of the other non-nucleoside inhibitor-resistant mutation sites are also located close enough to the bound inhibitor that changes in these residues are likely to affect the shape of the inhibitor-binding pocket. Some of the additional mechanisms may include increased side-chain bulk leading to steric conflict (for example, Val108→Ile, Val179→Asp, Val179→Glu, Gly190→Glu and Pro236→Leu), loss of contact (for example, Lys103→Asn and Val106→Ala), or the change of local charge or electrostatic potential (for example, Lys101→Glu, Lys103→Asn, Val179→Asp, Val179→Glu and Gly190→Glu). Ala98 and Val108 are located 8.7 Å and 5.9 Å away from the α-APA inhibitor, respectively.

Table 3. Interactions or shortest distances (in Å) between α-APA R 95845 and amino acid residues in the non-nucleoside inhibitor-binding pocket, at the polymerase active site, and at the non-nucleoside inhibitor-resistant mutation sites.

Residue	Atom	α-APA	Distance	Mutation	Location	Reference
Pro95	Cβ	C2	4.5[a]		β5b	
Pro97	Cα	C2	6.5[a]		β5b–β6	
Ala98	N	C2	8.7[a]	Ala→Gly	β5b–β6	[30]
Leu100	Cδ1	C1	3.6	Leu→Ile	β5b–β6	[58–60]
Lys101	O	C13	3.5	Lys→Glu	β5b–β6	[33]
Lys103	Cγ	Br2	3.8[a]	Lys→Asn	β5b–β6	[13,60]
Val106	Cγ1	N8B	3.2	Val→Ala	β6	[59]
	Cγ1	Br1	3.6			
Val108	Cγ2	Br1	5.9[a]	Val→Ile	β6	[58]
Val179	O	O8B	3.4	Val→Asp	β9	[30]
	Cβ	O8B	3.6	Val→Glu		[30]
	Cγ1	O1B	3.6			
Tyr 181	Cβ	C1	3.6	Tyr→Cys	β9	[13,61]
	Cβ	C1A	3.6			
	Cγ	C1	3.4			
	Cγ	C1A	3.6			
	Cγ	C2	3.3			
	Cδ1	C1B	3.5			
	Cδ2	C1	3.5			
	Cδ2	C2	3.1			
	Cδ2	C3	3.1			
	Cδ2	C4	3.6			
	Cε2	C2	3.3			
	Cε2	C3	3.1			
Tyr188	C	N8B	3.1	Tyr→His	β10	[60]
	O	N8B	2.8[b]			
	O	C8A	3.5			
	Cβ	N8B	3.3			
	Cδ2	C4A	3.6			
	Cε2	C4A	3.3			
Val189	C	N8B	3.2		β10	
	O	N8B	3.2[b]			
Gly190	N	N8B	3.3	Gly→Glu	β10	[5,39,62]
	Cα	N8B	3.5			
	Cα	O8B	3.6			
Phe227	Cδ2	Br1	4.2[a]		β12	
Trp229	Cβ	C4A	3.6		β12	
	Cγ	C4A	3.1			
	Cδ2	C3	3.6			
	Cδ2	C4A	3.0			
	Cε2	C3	3.4			
	Cε2	C4A	3.5			
	Cε3	C3	3.6			
	Cε3	C4A	3.3			
	Cζ2	C3	3.2			
	Cζ3	C3	3.4			
	Cη2	C3	3.2			
Leu234	Cδ1	Br1	4.0[a]		β13	
Pro236	Cα	C12	4.3[a]	Pro→Leu	β13–β14	[63]
Tyr318	Cε2	C12	3.0		β15	
	Cε2	C13	3.6			
Glu138 (p51)	Cγ	C1B	4.0[a]	Glu→Lys	β7–β8	[60]
Asp110[c]	N	N8B	11.2[a]		β6	
Asp185[c]	C	C4A	10.4[a]		β9–β10	
Asp186[c]	Cβ	C4A	7.0[a]		β9–β10	

[a]The shortest distances of any atom pair is given for these residues. [b]These atoms have potential hydrogen-bonding interactions. [c]These three aspartic acid residues are at the polymerase active site.

Mutations at these sites probably affect the conformation of nearby residues that interact directly with the inhibitor.

As p51 is a proteolytic product of p66, any mutations that give rise to drug resistance are present in both subunits. It was predicted, based on the structure of HIV-1 RT, that most of the mutations that confer resistance to non-nucleoside inhibitors will act through the change in the p66 subunit. The exception is the mutation Glu138→Lys which was predicted to act through the change in p51 [18]. Subunit-specific mutagenesis has been used to confirm the prediction: resistance to TSAO and TIBO compounds occurs only with the Glu138→Lys change in the p51 subunit and not in the p66 subunit, whereas most other mutations confer resistance only when present in the p66 subunit [32,33]. In the p66 subunit, Glu138 of the β7–β8 connecting loop is located on the outer surface of the heterodimer, far away from the non-nucleoside inhibitor-binding pocket. However, Glu138 of p51 is located at the putative entrance to the non-nucleoside binding pocket. Although Glu138 of p51 has no direct contact with the inhibitor in this structure, its side chain approaches the inhibitor (Glu138 Cγ...C1B α-APA, 4.0 Å) and the side chain of Tyr181 (Glu138 O...Cε1 Tyr181, 3.8 Å) quite closely. These interactions may affect the precise position of α-APA and the side-chain orientation of Tyr181.

Table 4. Centroid-to-centroid distances for pairs of aromatic rings that have potential aromatic–aromatic interactions in the non-nucleoside inhibitor-binding pocket.

Aromatic ring 1	Aromatic ring 2	Centroid-to-centroid distance (Å)	Shortest C–C distance (Å)	Interaction type
α-APA ring 1[a]	Tyr181	3.9	3.1	Offset stacked
	Tyr188	5.4	3.7	Edge-to-face
	Trp229	4.4[b]	3.2	Edge-to-face
α-APA ring 2[a]	Tyr318	5.5	3.0	Offset stacked
Tyr181	Tyr183	6.4	4.2	Offset stacked
	Tyr188	6.8	5.0[c]	Edge-to-face
	Trp229	5.8[b]	4.0	Edge-to-face
Tyr183	Trp229	6.6[b]	5.2[c]	Edge-to-face
Tyr188	Phe227	6.3	4.6	Edge-to-face
Trp229	Tyr232	5.6	4.2	Offset stacked
Tyr232	Trp239	5.8[b]	3.7	Edge-to-face

[a]Ring 1 of α-APA is composed of C1–C6; ring 2 is composed of C9–C14. [b]Only the six-membered rings of Trp229 and Trp239 were included in the centroid-to-centroid distance calculation. [c]Although the centroid-to-centroid distances for these aromatic pairs are <7 Å [28], the shortest C–C distance is >4.8 Å [29] and these interactions will be weaker.

Non-nucleoside inhibitors vary widely in chemical structure and some of the compounds, such as TSAO and BHAP derivatives, are substantially larger than the α-APA compounds. It is likely that different non-nucleoside inhibitors contact different portions of the binding pocket and have the potential for interacting with different sets of RT amino acid residues, which would account for the varying spectra of inhibitory activity and resistance.

Analysis of amino acid residues in the non-nucleoside inhibitor-binding pocket of HIV-2 RT and SIV RT

The non-nucleoside inhibitors are highly specific for HIV-1 RT and have little or no inhibitory activity against a variety of other viral and cellular polymerases tested [6–9,34,35], including the closely related HIV-2 RT and simian immunodeficiency virus (SIV) RT. Sequence

alignment of HIV-1 RT and HIV-2 RT, which have an overall amino acid sequence identity of ~60%, demonstrates that the two enzymes have a number of amino acid differences in the region of the non-nucleoside inhibitor-binding pocket (Fig. 5). A series of studies using chimeric HIV-1/HIV-2 RT enzymes indicated that several regions of the RT sequence were involved in non-nucleoside inhibitor binding [36–38]. It was shown that the region 176–190, which forms part of the pocket, was especially important in conferring sensitivity to non-nucleoside inhibitors. Site-directed mutagenesis studies of HIV-1 RT and HIV-2 RT have underscored the important roles of Tyr181 and Tyr188 in binding of non-nucleoside inhibitors [36,37]. HIV-2 RT has isoleucine and leucine residues at positions 181 and 188, respectively. Therefore, the energetically favorable aromatic–aromatic interactions observed between these residues in HIV-1 RT and the inhibitor are lost. Additional differences in the pocket residues occur at positions 101, 106, 179, 190 and 227 (Fig. 5). Mutations at residues 179 and 190 confer resistance to some non-nucleoside inhibitors [39,40]. Residues 101 and 106 have been identified as drug-resistant mutation sites and were found to be crucial for enzymatic activity [37,39,40]. The amino acid differences at these positions may contribute to the lack of sensitivity of HIV-2 RT to non-nucleoside inhibitors.

SIVs are a family of primate viruses that are related to HIV-1 [41]. Sequence comparisons indicated that SIVmac and SIVsm [42] strains are very closely related to HIV-2, whereas SIVagm and SIVmnd strains are more divergent [41,43,44]. The amino acid sequences of the RTs of SIVmac and HIV-2 are quite similar in the region that forms the non-nucleoside inhibitor-binding pocket of HIV-1 RT (Fig. 5), which explains the failure of non-nucleoside inhibitors to block SIVmac251 replication [7,8,34]. The TIBO compound, R82150, does inhibit two SIVagm strains (SIVagm3 and SIVagmTYO-1) and weakly inhibits SIVmndGB1 [45]. This inhibition could be due to the presence of aromatic amino acid residues at positions 181 (tyrosine in SIVmnd) and 188 (tryptophan in SIVagm and phenylalanine in SIVmnd) (Fig. 5). A new lentivirus identified in wild chimpanzees, SIVcpz, is more closely related to HIV-1 [43]. The RTs of SIVcpz and HIV-1 exhibit 85% identity at the amino acid sequence level. In the non-nucleoside inhibitor-binding pocket, only one residue difference exists, at position 179 (valine in HIV-1 RT; threonine in SIVcpz RT) (Fig. 5). This suggests that the RT of SIVcpz should have a hydrophobic pocket that is quite similar to that of HIV-1 RT. We would predict that the SIVcpz RT will be sensitive to at least some non-nucleoside inhibitors.

Comparisons of HIV-1 RT/α-APA structure with HIV-1 RT/DNA/Fab structure

The RT/α-APA structure is more similar to the RT/nevirapine structure [15] than to the RT/DNA/Fab [16], the unliganded RT [20] and the RT/Fab (R. Raag, *et al.*, & E Arnold, abstract B03, p. 44, American Crystallographic Association Meeting, Atlanta, June, 1994)

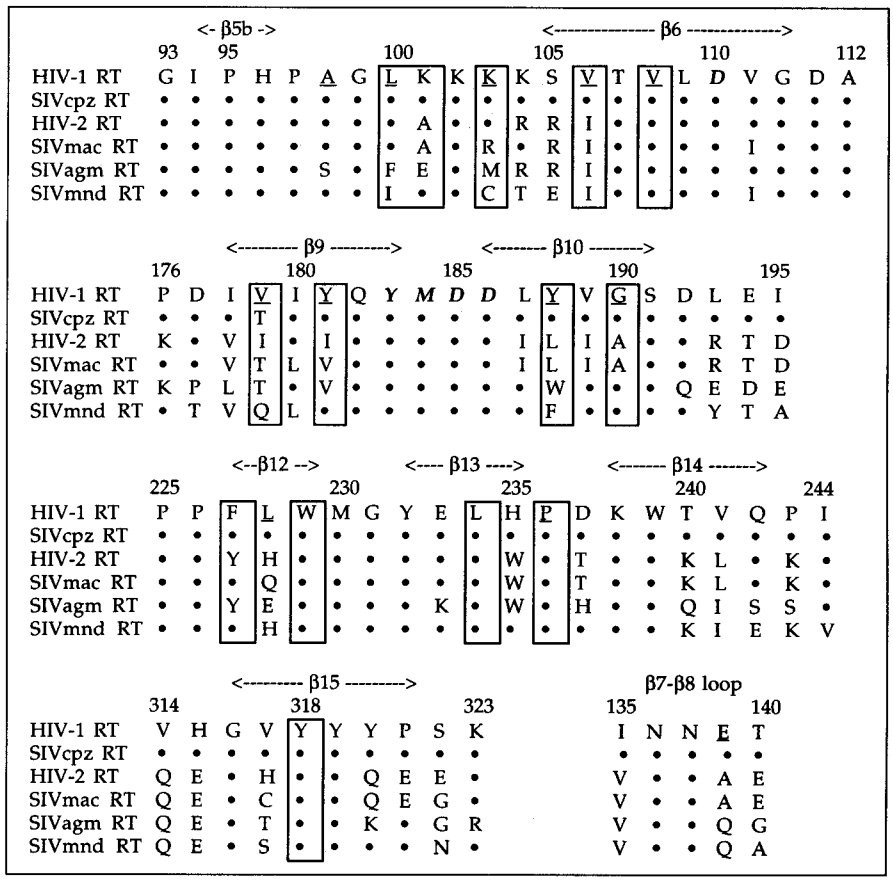

Fig. 5. Sequence alignment of amino acid residues from HIV-1, HIV-2, and simian immunodeficiency virus (SIV) RTs in the regions that form the non-nucleoside inhibitor-binding pocket. Residues forming part of the pocket are boxed. Residues whose mutations confer drug resistance are underlined. The YMDD motif and Asp110, which are essential for polymerase catalytic activity, are shown in bold italicized text. The secondary-structure assignment for HIV-1 RT is indicated [16]. The sequences displayed are: HIV-1 RT (strain LAI) [65]; SIVcpz RT [43]; HIV-2 RT (strain ROD) [66]; SIVmac RT (strain 251) [67]; SIVagm RT (strain 3) [68]; and SIVmnd RT (strain GB-1) [69].

structures. Superposition of the Cα coordinates of the RT/α-APA and the RT/nevirapine structures revealed few differences in the overall trace and the folding of secondary-structure elements. Superposition of the Cα coordinates of unliganded RT and the RT/DNA/Fab complex showed that the individual subdomains in both structures have comparable positions except for the p66 thumb. Conversely, superposition of the Cα backbones of RT in the RT/α-APA and RT/DNA/Fab complexes revealed not only numerous local conformational changes in the vicinity of the non-nucleoside inhibitor-binding pocket, but also significant global conformational changes in the p66/p51 heterodimer. Comparisons of unliganded RT *versus* the RT/DNA/Fab complex, and of unliganded RT *versus* the RT/nevirapine complex have been described [20]. Jäger *et al.* [19] have presented an interesting analysis of the relationship between the RT/nevirapine and RT/DNA/Fab complexes and described a swivel motion between the p66 fingers and palm subdomains and the other subdomains. The current comparison of the RT/α-APA and RT/DNA/Fab structures has revealed, in addition, that groups of subdomains appear to move in a coordinated fashion relative to the p66 fingers and palm subdomains.

Global conformational differences

Superposition analysis of the p66/p51 heterodimer in the RT/α-APA and RT/DNA/Fab structures revealed that RT could be divided into four large superimposable regions that correspond to groups of subdomains (R1–R4) (Table 5; Fig. 6). Smaller fragments not included in these regions are those that adopt different conformations in the two complexes and the relatively disordered regions. The results of the superposition analysis are summarized in Table 5.

The heterodimer interface seems not to be an inherent boundary between the superimposable regions. R2 and R3 contain subdomains from both p66 and p51 and the relative movements between subdomains that contact each other across the heterodimer interface are small compared with those between subdomains of the same subunit. Interestingly, both R2 and R3 in the RT/DNA/Fab complex are positioned closer to the DNA-binding cleft than in the RT/α-APA complex. Thus, relative to the DNA-binding cleft in the RT/DNA/Fab complex, the binding cleft in the RT/α-APA complex is expanded by ~5 Å at the tip of the thumb subdomain. This may reflect rearrangements of the heterodimer upon DNA binding, owing to the interactions between the DNA and the polymerase and RNase H domains. In addition, interactions between α-APA and structural elements at the base of the thumb may widen the cleft (Fig. 7) by forcing a separation between regions R1 and R2. Modification of the cleft width by binding of non-nucleoside inhibitors may affect binding or translocation of the dsDNA template-primer [46].

In unliganded HIV-1 RT structures, the p66 thumb subdomain folds into the DNA-binding cleft (R Raag, *et al.*, & E Arnold, abstract B03, p. 44, American Crystallographic Association Meeting, Atlanta, June, 1994; [20]). This movement is independent of the movements of the rest of the heterodimer and, therefore, is different from the movement of the p66 thumb observed here. In the present comparison, the p66 thumb and connection subdomains, and the p51 fingers subdomain (R2) constitute one movable entity. Thus, the widening of the DNA-binding cleft observed in the RT/α-APA structure cannot be accounted for by the independent 'up and down' motions of the thumb, but instead must result from relative movements of regions R1 and R2.

Table 5. Analysis of the global conformational changes of the HIV-1 RT p66/p51 heterodimer between the HIV-1 RT/α-APA and the HIV-1 RT/DNA/Fab structures.

Region[a]	Subdomains included	Number of atoms	Rms deviation (Å)	Rotation (°)	Translation (Å)
R1	p66 fingers p66 palm	191	1.2	0.0	0.0
R2	p66 thumb p66 connection p51 fingers	234	1.3	5.5	1.6
R3	RNase H p51 thumb	178	1.0	19.0	0.9
R4	p51 palm p51 connection	175	1.1	17.8	1.0

[a]Residues included in each region are listed as follows: R1 p66 (1–25, 27–50, 52–65, 77–89, 95–111, 113–134, 140–184, 186–193, 196–218); R2 p66 (252–285, 292–312, 321–356, 362–402, 408–427), p51 (20–42, 47–50, 55–63, 70–80, 115–122, 124–150); R3 p66 (428–555), p51 (252–271, 280–283, 285–296, 301–314); R4 p51 (93–97, 103–114, 151–159, 161–219, 232–357, 362–417, 419–425).

The independent movement of region R3 may be related to its role in crystal packing. In all reported crystal structures of HIV-1 RT [15,16,19,20], including the RT/α-APA complex, and the isolated RNase H domain [47,48] structures, the five-stranded β-sheet of one RNase H subdomain interacts with the five-stranded β-sheet of another two-fold-related RNase H domain to form an intermolecular 10-stranded β-sheet. Because this crystal-packing interaction has been observed in every crystalline form of HIV-1 RT (C2, $P3_212$ and F222) and HIV-1 RNase H reported so far, it may represent a favorable crystal-packing motif. It is not known whether this two-fold crystal-packing interaction is of biological significance.

It is unclear whether some or all of these conformational rearrangements result from differences in crystal packing in different crystalline forms or from template-primer and/or inhibitor binding, or a combination of causes. Additional structures of HIV-1 RT, for example, in complex with both nucleic acid and a non-nucleoside inhibitor, may help identify the potential causes of the observed conformational changes. Regardless of the cause, however, it appears that HIV-1 RT is a flexible enzyme that contains several regions capable of moving

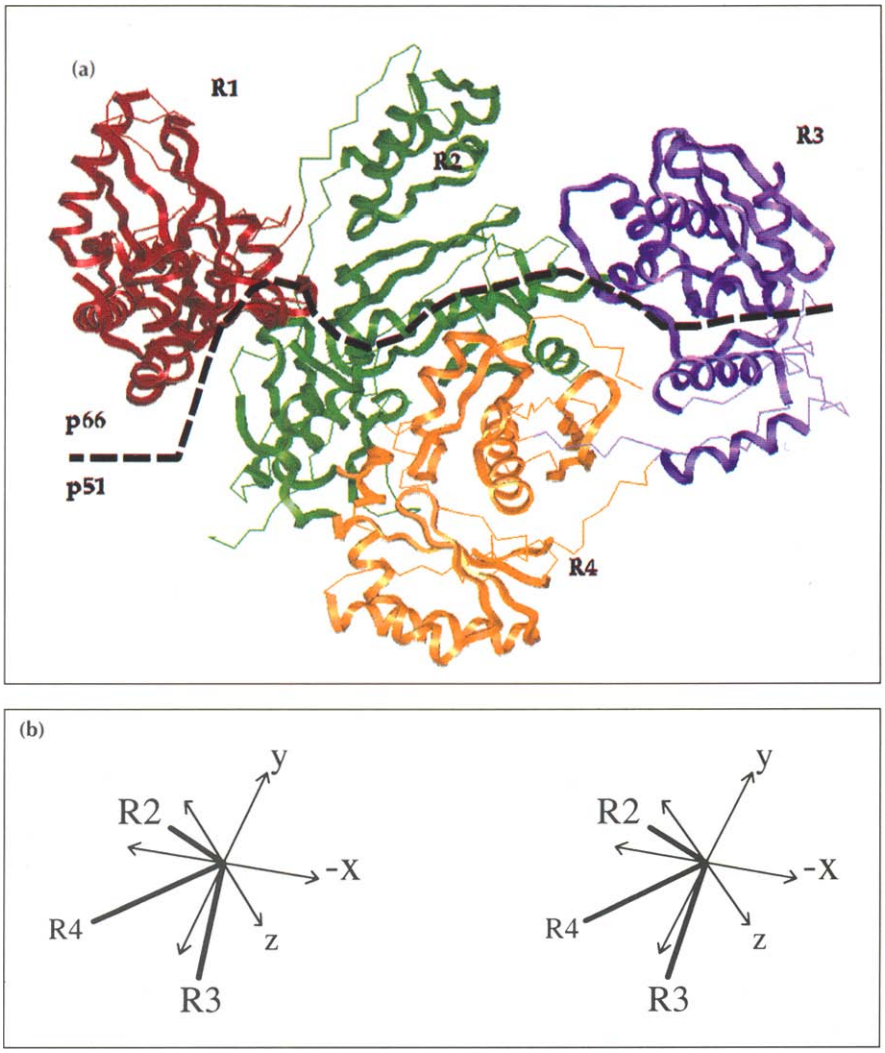

Fig. 6. Global conformational differences between the RT/α-APA and the RT/DNA/Fab complexes. **(a)** Colored ribbons of the p66/p51 heterodimer in the RT/α-APA structure showing the superimposable regions between the RT/α-APA and the RT/DNA/Fab complexes. Regions that superimpose well (RI–R4) are labeled in correspondence with Table 5. Regions that do not superimpose well are depicted as Cα traces. The black dashed line corresponds to the heterodimer interface between the p66 and p51 subunits. To obtain the best view of the regions observed to move concertedly, the view shown is rotated by ~30° along both a vertical axis and the axis perpendicular to the plane of the page relative to the orientation shown in Fig. 3. **(b)** Vector diagram showing the direction of movement of the superimposable regions between the RT/α-APA and the RT/DNA/Fab complexes. The view is the same as in (a). The vectors correspond to average Cα displacements between superimposable regions in the two structures when the heterodimers are aligned on the basis of the R1 superposition (the p66 fingers and palm subdomains). The vectors represent the magnitude and direction of movement for the transformation of each region from the RT/α-APA complex to its counterpart in the RT/DNA/Fab complex. The coordinate axes are shown in thinner lines and are labeled x, y and z.

relative to one another. This is also reflected in the asymmetry of the heterodimer, whose subdomains are folded similarly but are packed quite differently in the p66 and p51 subunits. The biological significance of this flexibility is unclear. However, we suggest that flexibility may be required for the enzyme's multiple functions. If template-primer binding induces conformational changes throughout the heterodimer, this could provide a means of allosteric communication between separated regions of the enzyme, for example, between the polymerase and RNase H active sites, during the potentially coupled processes of dNTP incorporation and RNA-template degradation. Inhibition by non-nucleoside compounds may also be explained in terms of allosteric control of dNTP incorporation, through interactions with nearby amino acid residues that may induce conformational changes at the polymerase active site.

Local conformational differences near the non-nucleoside inhibitor-binding pocket

There are marked differences in the side-chain orientations of Tyr181 and Tyr188 in the presence and absence of non-nucleoside inhibitor. In the RT/DNA/Fab and unliganded RT structures, the side chains of Tyr181 and Tyr188 in p66 both point away from the polymerase active site and towards the hydrophobic core. In the RT/α-APA and RT/nevirapine structures, these two residues have their side chains pointing towards the polymerase active site, creating space in the pocket to accommodate the inhibitor. Moreover, in the structure of the RT/α-APA complex, the side chain of Tyr181 is in a position that pushes Trp229 away from its position in the RT/DNA/Fab complex, which may be partly responsible for the conformational change of the β12–β13–β14 sheet (see below). The connecting loop of the β9–β10 hairpin in p66 is folded differently in the RT/α-APA and the RT/DNA/Fab structures (Fig. 7). When the core portion of the palm subdomains of the two structures (αE–β9–β10–αF) are superimposed, the overall rms fit of 60 Cα atoms is 0.99 Å. However, the Cα positions for the YMDD motif (183–186) differ by 0.9 Å, 2.2 Å, 3.2 Å and 0.9 Å, respectively. The conformations of the YMDD motif of the p51 subunit in these two structures are more similar to that of the p66 subunit in the RT/α-APA structure than to that of the RT/DNA/Fab p66 subunit. The YMDD motifs of both p66 and p51 in the structures of unliganded RT and the RT/nevirapine complex adopt conformations that are quite similar to that in the p66 subunit in the RT/α-APA structure (data not shown). It seems plausible

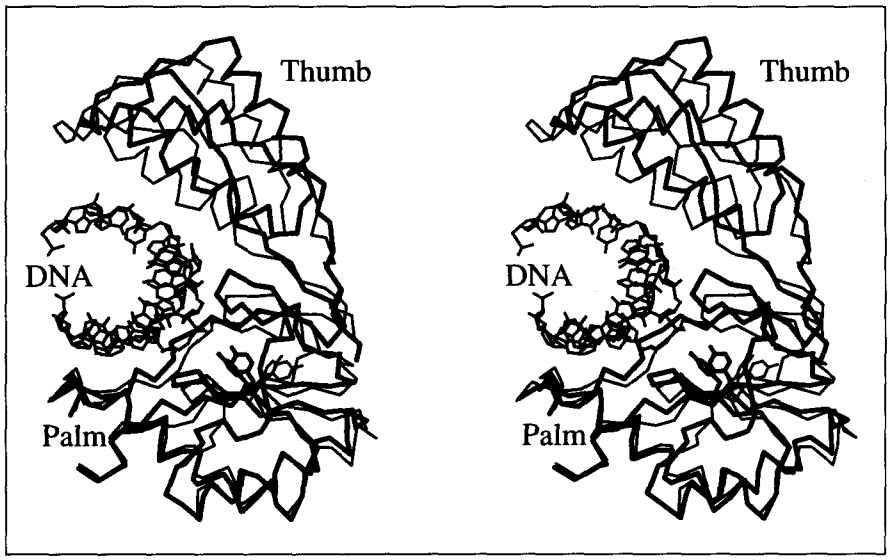

Fig. 7. Stereoview of the superposition of the RT/α-APA (thick lines) and the RT/DNA/Fab (thin lines) structures, showing only the p66 palm and thumb subdomains. The DNA shown corresponds to a portion of the model in the RT/DNA/Fab complex. The view is down the DNA axis from the RNase H domain. The superposition is based on the p66 fingers and palm subdomains (R1, Table 5). The thumb subdomain in the RT/α-APA complex is positioned further away from the DNA-binding cleft compared with the thumb in the RT/DNA/Fab complex. The side chains of Tyr181 and Tyr188 in the RT/α-APA complex (thick lines) rotate away from the positions in the RT/DNA/Fab complex (thin lines) and point towards the polymerase active site, thus, creating space for binding non-nucleoside inhibitors.

that the conformational difference of the YMDD motif observed in the p66 subunit of the RT/DNA/Fab complex relative to its conformation in other structures may be caused by binding of the template-primer. The precise conformation and mobility of this catalytically essential region may play an important role in the mechanism of non-nucleoside inhibition (see below).

Significant conformational changes in the β12–β13–β14 sheet of the p66 palm subdomain were observed in the RT/α-APA structure relative to the RT/DNA/Fab structure and may be a result of binding α-APA because there are several close contacts between α-APA and this β-sheet (Table 3). The β12–β13 hairpin has been denoted the 'primer grip' because of its close interactions with the 3′-terminus of the primer strand. This segment forms part of the non-nucleoside inhibitor-binding pocket. Site-directed mutagenesis studies indicated that modification of these residues can have significant effects on the polymerase activity ([49]; P Boyer, personal communication). In the RT/α-APA structure, movements of β11b, the β11b–β12 loop, and the β12–β13 hairpin relative to their positions in the RT/DNA/Fab structure result in expansion of the non-nucleoside inhibitor-binding pocket (Fig. 7). The β13–β14 loop, which contains Pro236, is pulled towards the bound inhibitor. This yields a differential twisting of the β14 strand out of the plane of the β-sheet. The binding of the non-nucleoside inhibitor deforms the primer grip of the p66 palm subdomain. Alternatively, these conformational changes may be associated with the coordinated movements of the p66 thumb and connection subdomains discussed in the previous section.

Implications of the HIV-1 RT/α-APA structure for mechanisms of non-nucleoside inhibition of HIV-1 RT

Numerous lines of evidence suggest that conformational mobility may be required during DNA polymerization by HIV-1 RT and other polymerases. The accumulated structural and biochemical evidence shows that α-APA, nevirapine and other non-nucleoside inhibitors bind at similar locations in HIV-1 RT. A preliminary comparison of the RT/α-APA, RT/TIBO (K Das, *et al.*, unpublished data) and RT/nevirapine [15,17] structures revealed striking similarity in the binding modes of these diverse non-nucleoside inhibitors and remarkable consistency of a butterfly-like shape adopted by the inhibitor molecules (J Ding *et al.*, unpublished data). In their description of the structure of a complex of HIV-1 RT with nevirapine, Kohlstaedt *et al.* [15] postulated that the non-nucleoside inhibitor may be working either by altering the precise geometry of the polymerase active site, or by restricting the mobility of the p66 thumb subdomain. Whatever the mechanism of inhibition, it is likely to be similar for the various non-nucleoside inhibitors.

One possible mechanism of inhibition of HIV-1 RT by non-nucleoside inhibitors could be that the conformational changes in the inhibitor-binding pocket distort the precise geometry and/or mobility of the nearby polymerase active site. During DNA polymerization, the polymerase catalytic site may need conformational flexibility in order to interact with the constantly changing template-primer substrates, distinguish different dNTP substrates, and permit translocation of the template-primer following nucleotide incorporation. Comparison of the RT/α-APA and RT/DNA/Fab structures suggests that binding of non-nucleoside inhibitors may keep the side chains of Tyr181 and Tyr188 rotated away from the pocket and towards the polymerase active site. This change, or some of the other conformational changes observed in the immediate vicinity of the binding pocket, could affect the mobility of some key elements of the nearby polymerase active site, including the YMDD motif, which is in a different conformation in the inhibitor-bound and DNA-bound structures (discussed above). As a consequence, the polymerase active site may no longer be able to deform its geometry to interact with the template-primer or to adjust to a conformation favorable for efficient catalysis. It is important to note that binding of non-nucleoside inhibitors to HIV-1 RT does not markedly decrease binding of template-primer or dNTP substrates [8,34,35].

In an alternative scenario for the mechanism of inhibition, the non-nucleoside inhibitor-binding pocket may function as a hinge between the palm and the thumb subdomains (Fig. 7) and the binding of a non-nucleoside inhibitor could distort the conformation of the primer grip of the palm subdomain and restrict the mobility of the thumb subdomain. The binding of the non-nucleoside inhibitor changes the conformation of the β12–β13–β14 sheet and the primer grip. Consequently, the distorted primer grip may not be able to interact with the primer strand and/or function properly in positioning the primer strand relative to the polymerase active site. The binding of the non-nucleoside inhibitor expands the DNA-binding cleft and may restrict the mobility of the thumb, which could prevent the thumb from interacting appropriately with the nucleic acid during the recognition of the template-primer [16] and, as a consequence, eliminate the polymerization activity.

Biological implications

One of the most important steps in HIV-1 replication is the synthesis of double-stranded (ds) DNA from the viral RNA genome by reverse transcriptase (RT). The only drugs currently used in the treatment of AIDS are the nucleoside analogs AZT, ddI, and ddC, which all inhibit RT. The clinical utility of these drugs is limited by two factors: serious side effects and the rapid emergence of drug-resistant mutations in HIV-1 RT. There are several classes of non-nucleoside inhibitors. They are highly specific for HIV-1 RT and, therefore, are more attractive candidates for clinical use than the nucleoside analogs. However, they too encounter drug-resistant viral strains.

The structure of HIV-1 RT in complex with a non-nucleoside inhibitor, α-anilinophenylacetamide (α-APA) R95845, provides several insights that may be important in the design of improved drugs for AIDS treatment. The inhibitor binds in a hydrophobic pocket lined by a number of aromatic side chains that are likely to stabilize the binding of a non-nucleoside inhibitor, explaining the observation that non-nucleoside inhibitors invariably contain at least one aromatic ring. The mutations that cause high-level resistance to non-nucleoside inhibitors are located close to the bound inhibitor, and presumably affect the stability of inhibitor binding.

Comparison of the structure of HIV-1 RT bound to α-APA with that of HIV-1 RT bound to dsDNA/Fab shows significant changes in the local environment of the inhibitor-binding pocket, including rotations of the side chains of Tyr181 and Tyr188 out of the pocket and movements of a hinge-like structure consisting of two β-sheets of the p66 palm subdomain. There are also changes in the overall structure of the protein: several rigid-body rearrangements of subdomains, portions of subdomains, or groups of subdomains were observed. These mobile portions of the structure may be required to move in a coordinated fashion during DNA polymerization.

We propose that the binding of non-nucleoside inhibitors may distort the geometry, or restrict the mobility, of the polymerase active site, which is close to the inhibitor-binding pocket, and that the effects on the three catalytically essential aspartic acid residues (Asp110, Asp185 and Asp186) may be especially important. Alternatively, the inhibitor-binding site may act as a hinge between the palm and the thumb subdomains, which may move relative to each other during polymerization, and this movement may be prevented by inhibitor binding. Because the hairpin structure formed by β-strands 12 and 13, which functions as a 'primer grip', is also part of the inhibitor-binding pocket, the proper positioning of the primer terminus at the active site may also be affected by inhibitor binding.

Materials and methods

Protein and inhibitor preparation and crystallization

The samples of HIV-1 RT used in crystallization were prepared as described in [50]. Crystals were grown using conditions [51] modified from those reported by Kohlstaedt *et al.* [15] for crystallization of a complex of HIV-1 RT with nevirapine. The mixing procedure was designed to enhance inhibitor solubility and availability upon mixing with concentrated HIV-1 RT solution. A 20 mM stock solution of the non-nucleoside α-APA inhibitor R95845 was prepared by dissolving it in dimethylsulfoxide (DMSO). A 15 mM working solution of α-APA containing 5% β-D-octylglucopyranoside (β-OG) was then prepared by adding the appropriate volume of 20% (w/v) β-OG to the 20 mM stock solution and mixing thoroughly. The purified HIV-1 RT enzyme was concentrated and transferred into a buffer of 10 mM Tris-HCl (pH 8.0), 75 mM NaCl using Centricon-30 microconcentrators (Amicon), yielding a solution with an HIV-1 RT concentration of 40–45 mg ml^{-1}. The concentrate was used immediately or divided into 2.5 mg aliquots and frozen at –20°C for future use. Thawed aliquot(s) of HIV-1 RT concentrate were kept on ice and then quickly mixed by vortexing with the working solution at a 2:1 molar ratio of inhibitor to HIV-1 RT and kept on ice for 15 min. Crystallization was performed using the hanging-drop vapor-diffusion method at 4°C using a crystallization solution composed of 50 mM Bis-Tris propane (pH 6.8), 100 mM $(NH_4)_2SO_4$, 10% glycerol and 12% (w/v) polyethyleneglycol (PEG) 8000. The hanging drops were prepared by adding equal volumes (5 μl) of the crystallization solution to the concentrated protein solution and mixing thoroughly to give a final protein concentration in the drop of 19–20 mg ml^{-1}. The initial precipitation observed in the drops upon mixing the HIV-1 RT with the crystallization solution dissipated after 6–10 h at 4°C. Crystals began to appear 24 h after drop preparation and reached a size of 1.5 mm×0.3 mm×0.2 mm after ~7 days at 4°C. The crystals often resembled knife-blades or arrowheads in appearance and belonged to the monoclinic space group C2 with cell dimensions of a=223.3 Å b=69.9 Å, c=106.5 Å, and

β=105.4°, which are close to these reported for the apparently isomorphous crystals of the HIV-1 RT/nevirapine complex [15]. One asymmetric unit contains one HIV-1 RT p66/p51 heterodimer and one α-APA inhibitor with a total molecular weight of 117 kDa, corresponding to a specific volume V_M of 3.42 Å^3 Da^{-1}, and a solvent content of 64% by volume, assuming that the standard partial specific volume for protein is 0.74 ml g^{-1}.

For preparation of the tetrakis(acetoxymercuri)methane (TAMM, Strem Chemicals) heavy-atom derivative, crystals were first rinsed for 5–10 s in crystallization solution and then soaked in crystallization solution plus 5 mM TAMM and 0.5 mM α-APA inhibitor R 95845 for 20–27 h followed by mounting. For the preparation of the 5-mercurideoxyuridine 5′-triphosphate (Hg-dUTP, Sigma) heavy-atom derivative, crystals were rinsed for 5–10 s in crystallization solution and then soaked in crystallization solution plus 2 mM dithiothreitol and 0.5 mM α-APA inhibitor R 95845 for 4–4.5 h. Crystals were subsequently rinsed twice (briefly) in crystallization solution and back-soaked for 2.5 h in crystallization solution plus 0.5 mM α-APA inhibitor R 95845. Crystals were finally soaked in crystallization solution plus 3 mM Hg-dUTP and 0.5 mM α-APA inhibitor R 95845 for 12–15 h and then mounted.

X-ray data collection

Crystals used for X-ray data collection had a typical size of 1.2 mm×0.3 mm×0.2 mm. The X-ray diffraction data were measured at the Cornell High Energy Synchrotron Source (CHESS) using the highly intense F1 beamline. The diffraction data were collected at −15°C using 1.5–2.0° oscillations from morphologically oriented crystals and recorded on Fuji storage phosphor image plates with a crystal-to-image plate distance of 30 cm. The X-ray wavelength used was λ=0.91 Å and the beam was limited using a 0.3 mm collimator. Exposure times were usually 6–12 s. The image plates were scanned using a BAS2000 Fuji storage phosphor scanner with a raster step of 100 μm. The digitized data were processed, merged, and scaled together from multiple crystals using a modified version of the Purdue oscillation film processing package, MTOPS, (G Kamer and E Arnold, unpublished program). Two native datasets were collected (Table 1): one dataset is for HIV-1 RT in complex with the dibrominated α-APA compound R 95845; the other is for HIV-1 RT in complex with R 90385, the dichlorinated analog of R 95845. Because the dibrominated native dataset is relatively more complete and has slightly higher quality, the structure determination and refinement reported here is principally based on this dataset. Two heavy-atom derivative datasets of the HIV-1 RT/α-APA R 95845 complex were also collected (Table 1) and were scaled to the dibrominated native dataset using mean local scaling by resolution ranges [52].

Structure determination

The crystal structure of the HIV-1 RT/α-APA complex was determined using the molecular replacement (MR) method as implemented in the program X-PLOR [22]. Since the structure of the HIV-1 RT/DNA/Fab complex was solved [16] and refined to 2.8 Å resolution (J Ding *et al.*, unpublished data), the initial experiments tried to use either full or polyalanine models of the HIV-1 RT p66/p51 heterodimer from the HIV-1 RT/DNA/Fab complex as starting models for molecular replacement. However, the rotation function search and subsequent Patterson correlation (PC) refinement [53] of 6000 highest rotation function search peaks revealed no outstanding solutions. Preliminary comparison of the RT/DNA/Fab complex and the RT/nevirapine complex had indicated that the relative orientations and positions of the RT subdomains in the two structures have some substantial differences [18], that may account for the failure of the rotation function search described above. Therefore, a new starting model was built that consisted of a partial polyalanine model of RT from the RT/DNA/Fab complex (PDB entry 1HMI) [16] (J Ding *et al.*, unpublished data), where the partial Cα model from the RT/nevirapine complex [15] was used as a guide for backbone position (PDB entry 1HVT). This model contained 937 amino acid residues and 4660 non-hydrogen atoms. Most of the connecting loops, which were missing in the original RT/nevirapine structure, were built based on the RT/DNA/Fab structure with the exception of residues 225–233 and 244–255 in p66 and 230–250 in p51. The rotation function search using diffraction data between 10.0 Å and 4.0 Å resolution and the subsequent PC refinement revealed a single solution at (ψ,θ,ϕ)= (0.0,0.0,0.0) which is 8σ higher than the next peak. The ensuing translation function search also yielded a single peak at (0.004,0.0,0.004), which is 11σ higher than the second peak. Rigid-body refinement of the partial polyalanine model in which the model was divided into five pieces yielded an R-factor of 0.44, and 20 cycles of positional refinement using conjugate-gradient minimization reduced the R-factor further to 0.41 for reflections between 10.0 Å and 3.0 Å resolution.

Isomorphous difference Fourier syntheses for the two heavy-atom derivatives using the MR phases revealed two sites of mercury binding at Cys38 in both the p66 and the p51 subunits, which are the same binding positions as in the RT/DNA/Fab complex (J Ding *et al.*, unpublished data) and confirmed that the MR solution was correct. These heavy-atom sites for the two isomorphous derivatives were also independently verified using both minimum and sum functions in the Patterson vector search program VMAP (R Williams and E Arnold, unpublished program). The MR phasing was then supplemented by MIR phasing, solvent flattening and phase extension. The phases were improved with the approach of iterative combination of the partial model and the initial MIR phases using the program PHASES (W Furey, and S Swaminathan, abstract PA33, p.73, American Crystallographic Association Meeting, New Orleans, April, 1990).

Model building and structure refinement

Iterative molecular modeling was performed using the graphics program 'O' [23]. Residue assignment and side chain placement were guided by difference Fourier maps with various coefficients, and omit electron-density maps (both conventional and simulated-annealing omit maps). The refinement of the structure was carried out using the molecular dynamics technique in the program X-PLOR [22]. At early stages, the refinement was carried out using the conventional positional refinement protocol to avoid the movement of main-chain atoms into the side-chain densities. The simulated-annealing protocol was employed when most of the side chains had been placed correctly. In the omitted regions of the model, the missing amino acid residues were gradually (usually one or two residues per cycle) inserted from both ends of the polypeptide chains if there was clear electron density for the new residues.

During the process of structure determination, several diffraction datasets were also collected for crystals of other HIV-1 RT/non-nucleoside inhibitor complexes that belong to the same space group, and initial structures were determined based on the HIV-1 RT/α-APA structure. The unit cell dimensions for these different complexes are related but have significant

differences, indicating incomplete isomorphism. Thus, the technique of averaging electron density from multiple crystal forms of different HIV-1 RT/inhibitor complexes was applied using the CCP4 package [25] and RAVE [54]. The structures used in the multiple crystal form map averaging procedure are primarily the HIV-1 RT/α-APA complex, the HIV-1 RT/TIBO R86183 complex (a=227.9 Å, b=70.3 Å, c=106.0 Å, β=105.5°) (K Das *et al.*, unpublished data), and the HIV-1 RT/BHAP U-90152 complex (a=227.5 Å, b=69.8 Å, c=106.1 Å, β=105.7°) (W Zhang *et al.*, unpublished data). Each crystal form was initially modeled by breaking the p66/p51 heterodimer into 13 segments according to individual subdomains which were positioned and oriented using rigid-body refinement. The map averaging was initially carried out at 3.5 Å resolution and then gradually extended to 3.0 Å resolution while increasing resolution in steps of 0.02 Å.

The multiple crystal form map averaging yielded the best quality maps, which resolved many ambiguities and showed dramatically improved electron density for many side chains in the regions where there was weak or no density in any previous maps. Therefore, in the later stages of model building, we used averaged maps of multiple crystal forms of HIV-1 RT/inhibitor complexes with difference Fourier maps as well as omit maps as references. Computation of the free R-factor [55] after each cycle of refinement was also performed to monitor the progress and the quality of the model. A subset of 1258 reflections, corresponding to 4% of the observed unique reflections, was randomly selected as the test dataset for evaluation of the free R-factor. At the final stages of structure refinement, the newly released full model of the HIV-1 RT/nevirapine complex (PDB entry 3HVT) and the coordinates of an unliganded HIV-1 RT structure [20] were used as references for model building. This was especially valuable in resolving the ambiguity of backbone tracing in some regions with poor electron density and in the placement of some side chains.

The location of the non-nucleoside inhibitor was initially revealed by difference Fourier calculations. After several cycles of refinement and successive model building, the difference Fourier electron-density maps clearly indicated significant positive density in the non-nucleoside inhibitor-binding pocket. However, the initial difference Fourier density was not of sufficient quality to uniquely determine the orientation and conformation of the inhibitor. This led to an initial misinterpretation of the inhibitor orientation which was subsequently corrected as the phases and map quality improved. As two separate datasets were collected for HIV-1 RT complexes: one with dibrominated α-APA inhibitor R95845 and the other with dichlorinated α-APA inhibitor R90385, it was anticipated that the difference Fourier calculation between these two complexes would reveal the two bromine positions, which would facilitate assignment of the precise position and conformation of the α-APA inhibitor. The difference Fourier maps between the HIV-1 RT/R95845 complex and the HIV-1 RT/R90385 complex revealed two significantly higher peaks which were located in the non-nucleoside inhibitor-binding pocket and were attributed to the two bromine atoms (Fig. 2a). As the refinement and modeling continued, the electron density for the whole α-APA compound became clearer. The α-APA inhibitor was included in the structure refinement only after the difference electron density of α-APA was well resolved and its orientation was unambiguous (Fig. 2a). The orientation of the primary amide group of α-APA was assigned based on plausible hydrogen bonds of the amide nitrogen atom with the carbonyl oxygen atoms of Tyr188 and Val189. Alternative orientations of this group are possible and cannot be ruled out by the current analysis.

The refinement of the current model, including 7815 non-hydrogen atoms, has converged to an R-factor of 0.255 for 30 186 reflections and a free R-factor of 0.360 for 1258 reflections between 10.0 Å and 2.8 Å resolution (81.8% complete to 2.8 Å). Fig. 2b shows a representative portion of a $2mF_o-F_c$ Fourier electron-density map at 2.8 Å resolution in the region around the non-nucleoside inhibitor-binding pocket. A summary of the current model and corresponding refinement statistics is given in Table 2. The loops β3–β4 (65–72), and β7–β8 (133–143), the region β14–αH (249–258), the αI–αJ loop (281–293), and the C-terminal 15 residues of p66, as well as β11b and β12 (218–230) of p51 are highly disordered with very poor electron density. The backbone trace in these regions is therefore only tentatively assigned. Due to the very weak or invisible electron density for the side chains, the following residues have been modeled as alanines during the refinement: p66 — 36, 66, 71, 72, 134–139, 218–223, 249, 263, 281–293, 297–302, 311, 312, 323, 356–358, 366, 451, 556–558; and p51 — 13, 22, 70, 113, 121, 197, 199, 218–220, 224–230, 253, 270, 278, 281, 305, 356–358, 361–363. At the current stage of refinement, water molecules have not been included. However, there are clear indications that, throughout the electron-density maps, some residual electron densities exist at a 2–3σ level that have reasonable stereochemistry to be water molecules. All atoms were refined with full occupancy. A single overall temperature factor was refined at the early stages of refinement and converged to a value of ~43 $Å^2$, which is a little higher than the B-factor deduced from the Wilson plot (B=37 $Å^2$). At the final stages of refinement, restrained individual isotropic B-factors were refined, which decreased both the R-factor and the free R-factor by ~1.5%. The regions with the highest B-factors are surface regions exposed to solvent, especially the fingers and thumb subdomains of p66 and the thumb subdomain of p51. The mean error in atomic positions was estimated by the method of Luzzati [56] to be 0.42 Å. The stereochemistry of the model has been examined using the programs X-PLOR [22] and PROCHECK [57]. The analysis of main-chain conformation shows that 79.4% of residues lie in the most favored regions and no residues are found in disallowed regions of the Ramachandran plot.

Analysis of conformational differences between the HIV-1 RT/α-APA complex and the HIV-1 RT/DNA/Fab complex

The analysis was performed using the least-squares routines lsq_explicit and lsq_improve in the program 'O' [23]. A least-squares superposition of all Cα atoms in the p66/p51 heterodimer in the RT/α-APA and RT/DNA/Fab structures yielded an rms deviation of 3.5 Å for 982 atoms. To achieve better superpositions, we broke the p66/p51 heterodimer into superimposable regions using the following strategy. First, superpositions were carried out for each individual subdomain. Portions outside each superimposed subdomain were then analyzed to locate fragments containing four or more sequential Cα atoms that were within a distance cutoff of 1.9 Å when superimposed. This analysis resulted in new transformations for groups of fragments that superimposed similarly throughout the p66/p51 heterodimer. Based on the new transformations, a second cycle of superposition analysis was performed with a distance cutoff of 2.5 Å to locate larger superimposable regions. This procedure identified four large superimposable regions in the size range 175–234 Cα atoms for the whole p66/p51 heterodimer (Fig. 6; Table 5). Each of these regions in the two

structures can be superimposed with an rms deviation of ≤1.3 Å. Table 5 gives the rotational and translational parameters which were used to superimpose the individual regions of the RT/α-APA structure onto their counterparts in the RT/DNA/Fab structure when the heterodimers were aligned on the basis of the R1 superposition (the p66 fingers and palm subdomains). The absolute value for the rotation of each region varies between 5.5° and 19.0° and the translation varies between 0.9 Å and 1.6 Å. Regions that could not be readily superimposed include portions of the inhibitor-binding pocket, N- and C-terminal portions of polypeptide chains, and the relatively disordered regions.

The coordinates of the HIV-1 RT p66/p51 heterodimer and the non-hydrogen atom coordinates of the α-APA inhibitor R 95845 have been deposited with the Brookhaven Protein Data Bank for immediate release (entry number 1HNI).

Acknowledgements: We thank P Van Daele for synthesizing the dibrominated derivative of α-APA R 95845 for these studies, D Rodgers and SC Harrison for providing the unliganded HIV-1 RT coordinates prior to general distribution, and P Boyer, P Clark, A Holmes, W Huber, G Kamer, L Kaven, J Marcotrigiano, R Nanni, D Oren, P Patel, R Raag, D Resnick, B Roy, and G Tarpley for helpful discussions and assistance, and S Ealick and D Bilderback (CHESS) and J Maizel, Jr. (Frederick Biomedical Supercomputing Center Cray-YMP). The work in EA's laboratory has been supported by Janssen Research Foundation, a Johnson & Johnson Focused Giving Award, NIH grants AI 27690 and AI 36144, and a grant from the American Foundation for AIDS Research to AJM. SJ was supported by an AIDS Research Fellowship from Deutsches Krebsforschungszentrum and WZ is supported by a CABM Graduate Fellowship. Research at FCRDC is sponsored in part by the National Cancer Institute, DHHS, under contract No. N01 CO-74101 and N01 CO-46000 with ABL, by NIGMS, and by a grant from the National Science Foundation (CHE-8910890, CHE-9215925) to RHS.

References

1. Jacobo-Molina, A. & Arnold, E. (1991). HIV reverse transcriptase structure-function relationships. *Biochemistry* **30**, 6351–6361.
2. Le Grice, S.F.J. (1993). Human immunodeficiency virus reverse transcriptase. In *Reverse Transcriptase.* (Skalka, A.M. and Goff, S.P., eds), pp. 163–191, Cold Spring Harbor Laboratory Press, Plainview, NY.
3. Tantillo, C., *et al.*, & Arnold, E. (1994). Locations of anti-AIDS drug binding sites and resistance mutations in the three-dimensional structure of HIV-1 reverse transcriptase: implications for mechanisms of drug inhibition and resistance. *J. Mol. Biol.* **243**, 369–387.
4. Larder, B.A. (1993). Inhibitors of HIV reverse transcriptase as antiviral agents and drug resistance. In *Reverse Transcriptase.* (Skalka, A.M. and Goff, S.P., eds), pp. 205–222, Cold Spring Harbor Laboratory Press, Plainview, NY.
5. Richman, D.D. (1993). Resistance of clinical isolates of human immunodeficiency virus to antiretroviral agents. *Antimicrob. Agents Chemother.* **37**, 1207–1213.
6. Pauwels, R., *et al.*, & Janssen, P.A.J. (1990). Potent and selective inhibition of HIV-1 replication *in vitro* by a novel series of TIBO derivatives. *Nature* **343**, 470–474.
7. Merluzzi, V.J., *et al.*, & Sullivan, J.L. (1990). Inhibition of HIV-1 replication by a non-nucleoside reverse transcriptase inhibitor. *Science* **250**, 1411–1413.
8. Goldman, M.E., *et al.*, & Stern, A.M. (1991). Pyridinone derivatives: specific human immunodeficiency virus type 1 reverse transcriptase inhibitors with antiviral activity. *Proc. Natl. Acad. Sci. USA* **88**, 6863–6867.
9. Romero, D.L., *et al.*, & Tarpley, W.G. (1991). Non-nucleoside reverse transcriptase inhibitors that potently and specifically block human immunodeficiency virus type 1 replication. *Proc. Natl. Acad. Sci. USA* **88**, 8806–8810.
10. Balzarini, J., *et al.*, & De Clercq, E. (1992). 2',5'-Bis-O-(tertbutyldimethylsilyl)-3'-spiro-5''-(4''-amino-1'',2''-oxathiole-2'',2''-dioxide) pyrimidine (TSAO) nucleoside analogs: highly selective inhibitors of human immunodeficiency virus type 1 that are targeted at the viral reverse transcriptase. *Proc. Natl. Acad. Sci. USA* **89**, 4392–4396.
11. Young, S.D. (1993). Non-nucleoside inhibitors of HIV-1 reverse transcriptase. *Perspect. Drug Discov. Design* **1**, 181–192.
12. Pauwels, R., *et al.*, & Janssen, P.A.J. (1993). Potent and highly selective human immunodeficiency virus type 1 (HIV-1) inhibition by a series of α-anilinophenylacetamide derivatives targeted at HIV-1 reverse transcriptase. *Proc. Natl. Acad. Sci. USA* **90**, 1711–1715.
13. Nunberg, J.H., *et al.*, & Goldman, M.E. (1991). Viral resistance to human immunodeficiency virus type 1-specific pyridinone reverse transcriptase inhibitors. *J. Virol.* **65**, 4887–4892.
14. Emini, E.A., Graham, D.J., Gotlib, L., Condra, J.H., Byrnes, V.W. & Schleif, W.A. (1993). HIV and multidrug resistance. *Nature* **364**, 679.
15. 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.
16. Jacobo-Molina, A., *et al.*, & Arnold, E. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. USA* **90**, 6320–6324.
17. Smerdon, S.J., *et al.*, & Steitz, T.A. (1994). Structure of the binding site for non-nucleoside inhibitors of the reverse transcriptase of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **91**, 3911–3915.
18. Nanni, R.G., Ding, J., Jacobo-Molina, A., Hughes, S.H. & Arnold, E. (1993). Review of HIV-1 reverse transcriptase three-dimensional structure: implications for drug design. *Perspect. Drug Discov. Design* **1**, 129–150.
19. Jäger, J., Smerdon, S., Wang, J., Boisvert, D.C. & Steitz, T.A. (1994). Comparison of three different crystal forms shows HIV-1 reverse transcriptase displays an internal swivel motion. *Structure* **2**, 869–876.
20. Rodgers, D.W., *et al.*, & Harrison, S.C. (1995). The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **92**, 1222–1226.
21. Unge, T., *et al.*, & Strandberg, B. (1994). 2.2 Å resolution structure of the amino-terminal half of HIV-1 reverse transcriptase (fingers and palm subdomains). *Structure* **2**, 953–961.
22. Brünger, A.T. (1993). *X-PLOR: A System for X-ray Crystallography and NMR. Version 3.1.* Yale University, New Haven and London.
23. 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 Crystallogr. A* **47**, 110–119.
24. Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149.
25. Collaborative Computational Project, No.4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763.
26. Ding, J., Jacobo-Molina, A., Tantillo, C., Lu, X., Nanni, R.G. & Arnold, E. (1994). Buried surface analysis of HIV-1 reverse transcriptase p66/p51 heterodimer and its interaction with dsDNA template/primer. *J. Mol. Recogn.* **7**, 157–161.
27. Wang, J., *et al.*, & Steitz, T.A. (1994). Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer. *Proc. Natl. Acad. Sci. USA* **91**, 7242–7246.
28. Burley, S.K. & Petsko, G.A. (1985). Aromatic–aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**, 23–28.
29. Hunter, C.A., Singh, J. & Thornton, J.M. (1991). π–π interactions: the geometry and energetics of phenylalanine–phenylalanine interactions in proteins. *J. Mol. Biol.* **218**, 837–846.
30. Byrnes, V.W., *et al.*, & Emini, E.A. (1993). Comprehensive mutant enzyme and viral variant assessment of human immunodeficiency virus type 1 reverse transcriptase resistance to non-nucleoside inhibitors. *Antimicrob. Agents Chemother.* **37**, 1576–1579.
31. Sardana, V.V., *et al.*, & Condra, J.H. (1992). Functional analysis of HIV-1 reverse transcriptase amino acids involved in resistance to multiple non-nucleoside inhibitors. *J. Biol. Chem.* **267**, 17526–17530.
32. Boyer, P.L., Ding, J., Arnold, E. & Hughes, S.H. (1994). Drug resistance of human immunodeficiency virus type 1 reverse transcriptase: subunit specificity of mutations that confer resistance to non-nucleoside inhibitors. *Antimicrob. Agents Chemother.* **38**, 1909–1914.
33. Balzarini, J., Kleim, J.P., Riess, G., Camarasa, M.J., De Clercq, E. & Karlsson, A. (1994). Sensitivity of (138Glu→Lys) mutated human immunodeficiency virus type 1 (HIV-1) reverse transcriptase (RT) to HIV-1-specific RT inhibitors. *Biochem. Biophys. Res. Commun.* **201**, 1305–1312.
34. Debyser, Z., *et al.*, & De Clercq, E. (1991). An antiviral target on reverse transcriptase of human immunodeficiency virus type 1 revealed by tetrahydroimidazo-[4,5,1-*jk*][1,4]benzodiazepinone-2(1*H*)-one and -thione derivatives. *Proc. Natl. Acad. Sci. USA* **88**, 1451–1455.

35. Wu, J.C., *et al.*, & Grob, P.M. (1991). A novel dipyridodiazepinone inhibitor of HIV-1 reverse transcriptase acts through a nonsubstrate binding site. *Biochemistry* **30**, 2022–2026.
36. Shih, C.-K., Rose, J., Hansen, G.L., Wu, J.C., Bacolla, A. & Griffin, J.A. (1991). Chimeric human immunodeficiency virus type 1/type 2 reverse transcriptases display reversed sensitivity to non-nucleoside analog inhibitors. *Proc. Natl. Acad. Sci. USA* **88**, 9878–9882.
37. Condra, J.H., *et al.*, & Sardana, V.V. (1992). Identification of the human immunodeficiency virus reverse transcriptase residues that contribute to the activity of diverse non-nucleoside inhibitors. *Antimicrob. Agents Chemother.* **36**, 1441–1446.
38. Hizi, A., *et al.*, & McMahon, J.B. (1993). Specific inhibition of the reverse transcriptase of human immunodeficiency virus type 1 and the chimeric enzymes of human immunodeficiency virus type 1 and type 2 by non-nucleoside inhibitors. *Antimicrob. Agents Chemother.* **37**, 1037–1042.
39. Bacolla, A., *et al.*, & Griffin, J.A. (1993). Amino acid substitutions in HIV-1 reverse transcriptase with corresponding residues from HIV-2. *J. Biol. Chem.* **268**, 16571–16577.
40. Boyer, P.L., Currens, M.J., McMahon, J.B., Boyd, M.R. & Hughes, S.H. (1993). Analysis of non-nucleoside drug-resistant variants of human immunodeficiency virus type 1 reverse transcriptase. *J. Virol.* **67**, 2412–2420.
41. Peeters, M., Piot, P. & Groen, G. (1991). Variability among HIV and SIV strains of African origin. *AIDS* **5** (suppl. 1), S29–S36.
42. Hirsch, V.M., Olmsted, R.A., Murphey-Corb, M., Purcell, R.H. & Johnson, P.R. (1989). An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* **339**, 389–392.
43. Huet, T., Cheynier, R., Meyerthans, A., Roelants, G. & Wain-Hobson, S. (1990). Genetic organization of a chimpanzee lentivirus related to HIV-1. *Nature* **345**, 356–359.
44. Gojobori, T., *et al.*, & Yokoyama, S. (1990). Evolutionary origin of human and simian immunodeficiency viruses. *Proc. Natl. Acad. Sci. USA* **87**, 4108–4111.
45. Debyser, Z., *et al.*, & Desmyter, J. (1992). Differential inhibitory effects of TIBO derivatives on different strains of simian immunodeficiency virus. *J. Gen. Virol.* **73**, 1799–1804.
46. Gopalakrishnan, V. & Benkovic, S. (1994). Effect of a thiobenzimidazolone derivative of DNA strand transfer catalyzed by HIV-1 reverse transcriptase. *J. Biol. Chem.* **269**, 4110–4115.
47. Davies, J.F., Hostomska, Z., Hostomsky, Z., Jordan, S.R. & Matthews, D.A. (1991). Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. *Science* **252**, 88–95.
48. Chattopadhyay, D., *et al.*, & Einspahr, H.M. (1993). Crystallographic analysis of an active HIV-1 ribonuclease H domain shows structural features that distinguish it from the inactive form. *Acta Crystallogr. D* **49**, 423–427.
49. Jacques, P.S., Wohrl, B.M., Ottmann, M., Darlix, J.L. & Le Grice, S.F. (1994). Mutating the 'primer grip' of p66 HIV-1 reverse transcriptase implicates tryptophan-229 in template-primer utilization. *J. Biol. Chem.* **269**, 26472–26478.
50. Arnold, E., *et al.*, & Hughes, S.H. (1992). Structure of HIV-1 reverse transcriptase/DNA complex at 7 Å resolution showing active site locations. *Nature.* **357**, 85–89.
51. Clark, A.D., Jr, Jacobo-Molina, A., Clark, P., Hughes, S.H. & Arnold, E. (1995). Crystallization of HIV-1 reverse transcriptase in the presence and absence of nucleic acid substrates, inhibitors, and an antibody Fab fragment. *Methods Enzymol.*, in press.
52. Matthews, B.W. & Czerwinski, E.W. (1975). Local scaling: a method to reduce systematic errors in isomorphous replacement and anomalous scattering measurements. *Acta Crystallogr. A* **31**, 480–487.
53. Brünger, A.T. (1990). Extension of molecular replacement: a new strategy based on Patterson correlation refinement. *Acta Crystallogr. A* **46**, 46–57.
54. Kleywegt, G.J. & Jones, T.A. (1994). Halloween, masks and bones. In *From First Map to Final Model.* (Bailey, S., Hubbard, R. and Waller, D., eds), pp. 59–66, SERC Daresbury Laboratory, Warrington, UK.
55. Brünger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.
56. Luzzati, V. (1952). Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Crystallogr.* **5**, 802–810.
57. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
58. Schleif, W.A., *et al.*, & Byrnes, V.W. (1992). Development and analysis of human immunodeficiency virus type 1 resistant to HIV-1 specific pyridinone reverse transcriptase inhibitors. *J. Cell. Biochem.* Suppl. **16E**, 87.
59. Vasudevachari, M.B., *et al.*, & Salzman, N.P. (1992). Prevention of the spread of HIV-1 infection with non-nucleoside reverse transcriptase inhibitors. *Virology* **190**, 269–277.
60. Balzarini, J., *et al.*, & De Clercq, E. (1993). HIV-1-specific reverse transcriptase inhibitors show differential activity against mutant strains containing different amino acid substitutions in the reverse transcriptase. *Virology* **192**, 246–253.
61. Richman, D., *et al.*, & Griffin, J. (1991). Human immunodeficiency virus type 1 mutants resistant to non-nucleoside inhibitors of reverse transcriptase arise in tissue culture. *Proc. Natl. Acad. Sci. USA* **88**, 11241–11245.
62. Kleim, J.-P., *et al.*, & Paessens, A. (1993). Activity of a novel quinoxaline derivative against human immunodeficiency virus type 1 reverse transcriptase and viral replication. *Antimicrob. Agents Chemother.* **37**, 1659–1664.
63. Dueweke, T.J., *et al.*, & Tarpley, W.G. (1993). A mutation in reverse transcriptase of bis(heteroaryl)piperazine-resistant human immunodeficiency virus type 1 that confers increased sensitivity to other non-nucleoside inhibitors. *Proc. Natl. Acad. Sci. USA* **90**, 4713–4717.
64. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.
65. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. (1985). Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**, 9–17.
66. Clavel, F., Guyader, M., Guetard, D., Salle, M., Montagnier, L. & Alizon, M. (1986). Molecular cloning and polymorphism of the human immune deficiency virus type 2. *Nature* **324**, 691–695.
67. Franchini, G., *et al.*, & Reitz, M.J., Jr. (1987). Sequence of simian immunodeficiency virus and its relationship to the human immunodeficiency viruses. *Nature* **328**, 539–543.
68. Baier, M., Garber, C., Muller, C., Cichutek, K. & Kurth, R. (1990). Complete nucleotide sequence of a simian immunodeficiency virus from African green monkeys: a novel type of intragroup divergence. *Virology* **176**, 216–221.
69. Tsujimoto, H., *et al.*, & Hayami, M. (1989). Sequence of a novel simian immunodeficiency virus from a wild-caught African mandrill. *Nature* **341**, 539–541.

Received: **30 Jan 1995**; revisions requested: **23 Feb 1995**: revisions received: **1 Mar 1995**. Accepted: **2 Mar 1995**.