Published in final edited form as:
Biochemistry. 2009 August 18; 48(32): 7646. doi:10.1021/bi900790x.

Solution Structural Dynamics of HIV-1 Reverse Transcriptase Heterodimer

James M. Seckler[‡], **Kathryn J. Howard**[§], **Mary D. Barkley**[*,‡,§], and **Patrick L. Wintrode**[*,‡]

[‡]Department of Physiology & Biophysics, School of Medicine, Case Western Reserve University, Cleveland, 10900 Euclid Avenue, Ohio 44106

[§]Department of Chemistry, Case Western Reserve University, Cleveland, 10900 Euclid Avenue, Ohio 44106

Abstract

Crystal structures and simulations suggest that conformational changes are critical for the function of HIV-1 reverse transcriptase. The enzyme is an asymmetric heterodimer of two subunits, p66 and p51. The two subunits have the same N-terminal sequence, with the p51 subunit lacking the C-terminal RNase H domain. We used hydrogen exchange mass spectrometry to probe the structural dynamics of RT. H/D exchange revealed that the fingers and palm subdomains of both subunits form the stable core of the heterodimer. In the crystal structure, the tertiary fold of the p51 subunit is more compact than that of the polymerase domain of the p66 subunit, yet both subunits show similar flexibility. The p66 subunit contains both the polymerase and RNase H catalytic sites. H/D exchange indicated that the RNase H domain of p66 is highly flexible. The β-sheet β12–β13–β14 lies at the base of the thumb subdomain of p66, and contains highly conserved residues involved in template/primer binding and NNRTI binding. Using the unique ability of hydrogen exchange mass spectrometry to resolve slowly interconverting species, we found that β-sheet β12–β13–β14 undergoes slow cooperative unfolding with a $t_{1/2}$ of 6.6 s. The H/D exchange results are discussed in relation to existing structural, simulation, and sequence information.

Reverse transcriptase performs the first step in replication of HIV.[1] RT copies the single-stranded viral RNA genome into a double-stranded proviral DNA prior to insertion by integrase into a chromosome of the infected cell (1). RT has RNA- and DNA-dependent DNA polymerase activities and RNase H activity. The mature enzyme is a heterodimer of p66 and p51 subunits; p51 has the same N-terminal polymerase domain as p66, but lacks the C-terminal

[*]Correspondence: MDB, mdb4@case.edu; PLW, patrick.wintrode@case.edu

[1]Abbreviations:

ANM	anisotropic network model
EDTA	ethylenediaminetetraacetic acid
GNM	Gaussian network model
HIV-1	human immunodeficiency virus type 1
HXMS	hydrogen exchange mass spectrometry
NNRTI	nonnucleoside reverse transcriptase inhibitor
NRTI	nucleoside reverse transcriptase inhibitor
RT	reverse transcriptase
Tris	tris(hydroxymethyl)aminomethane

RNase H domain. The enzyme has an asymmetric structure (2). The p66 subunit contains the enzyme active sites (3), whereas the p51 subunit appears to have a structural function. The p66 polymerase domain has a right-handed conformation like other polymerases with fingers (residues 1–84, 120–150), palm (residues 85–119, 151–243), and thumb (residues 244–322) subdomains plus a connection subdomain (residues 323–427). The fingers, palm, and thumb subdomains of p66 form the template/primer binding cleft with the polymerase active site residues (D110, D185, and D186) in the palm subdomain (4). Although the four subdomains of the polymerase domain have similar folds in p66 and p51, their relative orientations differ in the two subunits (Figure 1).

Numerous crystal structures are available for wild-type and mutant HIV-1 RTs in the absence and presence of various substrates and inhibitors. The conformation of p51 subunit is essentially the same in all the structures, whereas the p66 polymerase domain adopts both open and closed positions of the fingers and thumb subdomains, suggesting that RT is quite flexible. The crystallographic *B*-factors of two unliganded wild-type RT structures (same space group) identify the same mobile regions in p51 subunit, but somewhat different mobile regions in p66 subunit (5,6). In addition to structural evidence, computational studies have explored the flexibility of RT and its complexes and proposed possible functional roles for protein dynamics in translocation of template/primer substrate (7-10), inhibition by NRTIs (11) and NNRTIs (7,9,10,12-14), and drug resistance (13). However, experimental studies of RT conformation and dynamics in solution are few. Both open and closed conformations of the p66 polymerase domain were observed by site-directed spin labeling (15), and the solution structure and backbone dynamics of the isolated RNase H domain were determined by NMR (16-18).

This paper uses hydrogen exchange mass spectrometry to examine the solution conformation and dynamics of HIV-1 RT. HXMS studies provide a medium resolution snapshot of the exchange of amide protons on the peptide backbone (19). The H/D exchange was monitored as a function of time after dilution of RT heterodimer into deuterated buffer. The rate of exchange in each peptide depends on the extent of solvent accessibility, which in turn reports surface versus buried residues and lability of hydrogen-bonded secondary structures. Exchange rates were measured separately for p66 and p51 subunits of the heterodimer, mapped onto the amino acid sequences, and compared to crystallographic and computational results. The implications of the protein dynamics for RT function are discussed.

EXPERIMENTAL PROCEDURES

Protein Preparation

Biochemical reagents and chemicals were obtained from Roche Applied Science (Indianapolis, IN) and Sigma Chemicals (St. Louis, MO) unless otherwise specified. RT buffer D is 0.05 M Tris (RNase, DNase-free, pH 7.2), 25 mM NaCl, 1 mM EDTA, and 10% (v/v) glycerol (molecular biology grade redistilled). PBS is 0.1 M phosphate buffer (pH 7.2) and 0.15 M NaCl (Pierce, Rockford, Il). D_2O was purchased from Cambridge Isotope Laboratories (Andover, MA). HIV-1 RT subunits with N-terminal hexahistidine extensions were expressed separately in *E. coli* M15 strains carrying pDM1.1 plasmids p6H RT for p66 and p6H RT51 for p51. Cells were cultured at 37 °C to an optical density of 0.7 at 600 nm, induced with 200 μg/mL of isopropyl-β-D-1-thiogalactoside, grown for 3 h, and harvested. RT proteins were purified as described (20) and modified (21).

The N-termini of p66 and p51 were labeled with biotin using the EC-Link-NHS-PEO_4-Biotinylation Kit (Pierce, Rockford, IL). Protein samples were dialyzed overnight against 2 × 1 L of PBS, incubated overnight on a rocker with 10-fold molar excess of biotin, and then dialyzed against RT buffer D containing 50% glycerol at 4 °C. Labeling was confirmed to be 100% by the HABA avidin assay provided with the kit, SDS-PAGE mobility shift, and mass

spectrometry. The p66/p51 heterodimer with one biotin-labeled subunit was prepared by incubating equimolar mixtures of labeled and unlabeled subunits in RT buffer D containing 50% glycerol for at least 5 days. Activity of biotin-labeled p66/p51 was confirmed using EnzCheck Reverse Transcriptase Assay Kit (Invitrogen Corp., Carlsbad, CA). Total protein concentrations of p66/p51 solutions are 20–30 μM (monomer units). Based on the dimerization constants (21), the p66/p51 solutions contain >84–87% heterodimer.[2]

Peptide Mapping by HPLC–Tandem Mass Spectrometry

Peptide mapping experiments were carried out as described previously (22) with the following modification. Five μg (0.08 nmol) of p66 in 100 μL of RT buffer D was mixed with 500 μL of 100 mM NaH_2PO_4 (pH 2.4) and was digested with 5 μL of 1 mg/mL porcine pepsin in H_2O. Sequencing by tandem mass spectrometry was carried out using an LCQ-DECA quadrupole ion trap mass spectrometer (ThermoElectron). Additional peptide mapping experiments were conducted on an LTQ-FT ion cyclotron resonance mass spectrometer (ThermoElectron) in order to confirm peptide identification by exact mass.

H/D Exchange of p66/p51 Heterodimer

RT heterodimer with either the p66 or p51 subunit containing an N-terminal biotin tag (11.2 μg of labeled p66 or 9.0 μg of labeled p51) in RT buffer D-H_2O containing 50% glycerol was diluted 10-fold into RT buffer D-D_2O (pD 7.2) containing 5% glycerol and incubated for different times from 5 s to 2 h at 25 °C. Exchange was quenched by diluting the protein sample 5-fold with 100 mM NaH_2PO_4 (pH 2.4) at 4 °C. RT is reversibly denatured by acidic pH (23), and the biotin-labeled subunit was removed using a strepavidin column or beads. For column separations, the protein sample was loaded onto a 1 mL HiTrap Streptavidin HP column (GE Healthcare Bio-Sciences, Piscataway, NJ) and the column was washed with 1 mL of 100 mM NaH_2PO_4 (pH 2.4). The biotin-labeled subunit was bound to the column; the unlabeled subunit was collected in the flow-through. The column was then washed with 20 mL of 100 mM NaH_2PO_4 (pH 2.4) to remove any residual free subunit prior to reuse. SDS gel electrophoresis showed a single band corresponding to the unlabeled subunit. For bead separations, the protein sample was mixed with 20 μL of Ultralink Immobilized NeutraAvidin Protein beads (Pierce, Rockford, Il) and vortexed for 30 s. The mixture was then centrifuged for 30 s and the supernatant was collected. SDS gel electrophoresis showed that ~85% of the labeled subunit was removed (Figure S1). The amount of residual unlabeled subunit estimated form the mass spectra is ~10%. This level of background signal did not compromise our ability to quantify deuterium uptake; signal from the residual labeled subunit shifts the centroid mass by ≤ 0.1 Da.

Isotope Analysis by HPLC-ESI Mass Spectrometry

The deuterium-labeled protein was digested on ice with 5 μL of 1 mg/mL porcine pepsin in H_2O for 5 min and analyzed by HPLC-MS as described elsewhere (24). Deuterium levels for each peptide were corrected for back exchange using the equation

$$D = \frac{m - \mathrm{m0\%}}{m - \mathrm{m100\%}} \times \mathrm{N} \quad (1)$$

where D is the number of amide hydrogens exchanged with deuterium, m is the centroid mass of the peptide at a given time point, $m0\%$ is the mass of the undeuterated peptide, $m100\%$ is the mass of the fully deuterated peptide, and N is the number of amide hydrogens in the peptide.

[2]Estimated using K_d values determined in RT buffer D at 5 °C. Preliminary data indicate that increasing glycerol concentration favors dimerization.

D is then plotted versus time and the resulting curve is fitted to a triple exponential function using nonlinear least squares (OriginLab).

$$D = N - N_{fast}e^{-k_{fast}t} - N_{medium}e^{-k_{medium}t} - N_{slow}e^{-k_{slow}t} \quad (2)$$

where N_{fast}, N_{medium}, N_{slow} are the number of fast, medium, and slow exchanging amide hydrogens in the peptide, and k_{fast}, k_{medium}, k_{slow} are the rate constants of the observed exchange.

RESULTS

Peptide Mapping and Identification

A total of 50 peptic fragments were identified by tandem mass spectrometry giving a total sequence coverage of 75% for the p66 subunit and 77% for the p51 subunit. The longest of these peptic fragments has 30 exchangeable amide hydrogen atoms and the shortest has five, with an average among all of the peptides of 15 exchangeable amide hydrogens; 27 fragments were analyzed in HXMS experiments. These are well-distributed throughout the entire protein sequence (Figure 2). The remaining 23 peptic fragments provide largely redundant information; different peptides with substantial sequence overlap show similar exchange behavior. The analyzed peptic fragments cover 63% of the amino acids in the dimer interface of the p66 subunit and 87% of the interface residues of the p51 subunit (21). The sequence coverage of residues involved in template/primer binding is 85% of the total DNA/DNA and RNA/DNA contacts made by both subunits, including the catalytic triad in the polymerase active site, four of the six residues which make contact with dTTP substrate, and three of the four conserved residues in the RNase H active site (25,26). The sequence coverage of the hydrophobic NNRTI binding pocket is 83% (27).

Kinetics of H/D Exchange

Figure 3A shows raw mass spectra of the peptide from residues 257–282 of RT incubated for different times in deuterated buffer. The shift in mass as deuteriums replace hydrogens along the amide backbone is easily seen. These spectra were used to make a plot of deuterium uptake *vs* time. The solid line shows a good fit of these data to eq 2 (Figure 3B). Rate constants obtained from curve fitting to the analyzed peptides for p66 and p51 are given in Tables S1 and S2. The rate constants of fast exchanging hydrogens give exchange half-lives on the order of seconds, which is comparable to the values expected for completely unprotected hydrogens (28). Thus the fast exchanging hydrogens are most likely exposed to solvent and not involved in hydrogen bonding. The rate constants for medium exchanging hydrogens give half-lives on the order of 1–10 min, and these hydrogens are likely in structured regions that undergo substantial conformational fluctuations. Slow hydrogens exchange with half-lives on the order of 1–100 h and are therefore located in regions of highly stable structure in solution. It was not possible to fit all peptide exchange curves to eq. 2 with high confidence. In some cases, this was because the deuterium level showed a rapid rise at early labeling times followed by a very abrupt leveling off. In other cases, the slow rate constant was indeterminate due to extremely stable secondary structure. However, in both of these cases, the number of slow exchanging hydrogens could be determined from the number of hydrogens that remained unexchanged at the longest incubation times. Rate constants, of course, could not be determined for these, only the fact that they exchange considerably more slowly than the fast or medium hydrogens.

Percent exchange at various incubation times is mapped onto the sequence of the p66 subunit of RT in Figure 4A. The fingers and palm subdomains have stable secondary structure, as evidenced by the slow rates of H/D exchange. A notable exception is the peptide spanning

residues 210–231, which shows little protection from exchange. The catalytic triad and dNTP binding site are both wholly contained within the fingers and palm subdomains (25). The catalytic triad is encompassed by two small peptides, one of which (183–187) remains almost completely protected from exchange for the duration of the time course, while the other (110–115) is reasonably stable, but does show exchange. Three of the six residues in the dNTP binding site are part of the short peptide (110–115) that contains a single residue from the catalytic triad. The fourth dNTP binding residue is contained in a third peptide (147–160) that also shows a low rate of exchange (Figure 3C, lower curve). This indicates that both the polymerase active site and the dNTP binding site are structured and stable.

The NNRTI binding pocket exists only in structures of RT—NNRTI complexes. It includes two β-sheets that are also present in structures in the absence of drug. The first of these sheets, β6–β10–β9, is covered by four different peptides (88–109, 110–115, 183–187, 187–192). All of these peptides show low rates of exchange, suggesting that this β-sheet is quite rigid. In contrast, the second sheet, β12–β13–β14, is covered by two peptides (210–231, 232–246), which both show extremely high rates of exchange. As discussed below, the peptide covering β13–β14 (232–246) undergoes slow cooperative unfolding. All peptides within the thumb subdomain show intermediate rates of exchange, suggesting that the entire subdomain is structured but flexible. Rates of H/D exchange in the connection subdomain are faster on average than in the fingers, palm, and thumb subdomains (except residues 210 to 246), but are consistent with stable secondary structure. Of the five peptides covering the RNase H domain, only one (492–500) shows a high degree of protection against exchange. This indicates that the RNase H domain is highly flexible. Three of the RNase H catalytic residues (E478, D496, and D549) are covered by three different peptides (469–479, 492–500, 534–560). Two of these three peptides (469–479, 534–560) show high rates of exchange. Some active site residues in both the polymerase and RNase H domains coordinate magnesium ions. Note that the buffers used in the experiments reported here do not contain magnesium.

The percent exchange at various incubation times for the p51 subunit are given in Figure 4B. Although the tertiary structures of p51 and the p66 polymerase domain differ, the secondary structures are largely identical and this is reflected in similar H/D exchange rates, albeit with some local differences. The degree of exchange at short labeling times (5–10 s) reflects primarily solvent exposure and hydrogen bonding, and the levels of exchange at 5 s are nearly identical in p66 and p51, consistent with the fact that they possess similar secondary structure. Local differences are seen at longer labeling times, indicating that local flexibility differs between the two subunits. The long stretch of residues from 210 to 246 in the palm and a portion of the thumb are somewhat more flexible in p51, while residues 340–370 in the connection are slightly more stable. Interestingly, p51 does not show slower exchange overall than p66, indicating that despite being more compact, p51 is not significantly more rigid than p66.

The dimer interface includes residues from both subunits of the heterodimer. All peptides in the p66 polymerase domain that contain a significant portion of the dimer interface show very slow exchange at the earliest time points, indicating that they are shielded from solvent. This implies that the dimer interface observed in crystals is stable and present in solution. The loop (426–439) that links the connection subdomain to the RNase H domain shows a much greater degree of initial protection than expected for a solvent-exposed loop, consistent with being part of the dimer interface. Because the loop has no secondary structure, the initial protection from exchange is due completely to residues in the dimer interface. However, this peptide exchanges rapidly, indicating that the dimer interface provides little long-term protection from solvent. Four peptides in the RNase H domain also cover residues in the dimer interface. One of these (492–500) shows almost complete protection and another shows some initial protection (501–517), while the other two peptides (518–533, 534–560) show very little initial protection, suggesting that this region of the dimer interface is more labile.

The peptides in p51 that contain portions of the dimer interface likewise show low levels of initial exchange, with two notable exceptions (283–300, 417–425). Interface residues in both of these peptides make contacts with the RNase H domain of p66; p51 residues 288–290 contact p66 residues 435–439, while p51 residues 417–425 form contacts with p66 residues 503–504 and 532–537. The relative lack of protection in these regions even at short labeling times further supports the conclusion that the dimer interface formed between the RNase H domain and p51 is highly dynamic. Of residues contained in the two unprotected peptides (417–425, 426–440) at the C-terminus of p51, only residues 415–421 appear to be essential for heterodimer formation, although C-terminal truncation of p51 after residue 421 or 426 affects enzyme activity (29).

Figure 5 presents the percent exchange for two time points (5 and 900 s) mapped onto the crystal structure of unliganded HIV-1 RT (5). At 5 s the whole protein shows a great degree of protection from exchange (Figure 5A). By 900 s the fingers and palm subdomains of both subunits form a stable core of the protein that is largely protected from deuterium exchange, while the rest of the protein appears to be much more flexible allowing its amide hydrogens to exchange with solvent (Figure 5B).

Comparison to Crystal Structure

Coloring the structure by percent exchange as in Figure 5 has the advantage of presenting the data without making assumptions about the solution structure of the protein. Because of limited resolution, a given peptide may contain secondary structural elements as well as coil regions, skewing the results so that a region with stable secondary structure may appear to be unstructured. This problem is avoided by comparing the number of slow exchanging amide hydrogens from Tables S1 and S2 with the number of hydrogen-bonded amide hydrogens predicted from the crystallographic model of RT using the program WhatIf (5,30). The ratio of slow exchanging to hydrogen-bonded amide hydrogens serves as a measure of protein flexibility. Table 1 lists the flexibility ratios for each of the analyzed peptides. In cases where the ratio is greater than 1, the additional protection can be explained by solvent shielding. These ratios were used to color the structure of RT (Figure 6). It appears that the stable core formed by the fingers and palm subdomains of both subunits is quite rigid, while the regions adjacent to this core (thumb subdomain of p66 and connection subdomains of both subunits) are more flexible but still stable. The regions distant from the core (RNase H domain of p66 and portions of connection and thumb subdomains of p51) are unstable. Exceptions to this are the peptides 210–231, 232–246, and 283–300 (Figure 3C, upper curve) in the p66 subunit and peptide 232–246 in the p51 subunit. All of these regions are much more flexible than would be expected if the secondary structures shown in the crystal structure were stable. It is worth noting that a significant difference in secondary structure between the p66 and p51 subunits is that residues corresponding to β-strands 12 and 14 in p66 do not form β-strands in p51. Instead they are either missing electron density (strand 12) or adopt a coiled conformation (strand 14).

Slow cooperative unfolding at the base of the p66 thumb domain

A particular strength of HXMS is its ability to monitor cooperative conformational changes by directly resolving distinct structural populations in solution. If a protein or protein region exists in two distinct, slowly interconverting conformations, the resulting spectrum of the deuterium labeled sample will exhibit a two-peak pattern or “double isotopic envelope”, with each peak representing one of the conformations. In the case of protein unfolding, a lower and higher m/z envelope will correspond to the folded and unfolded forms, respectively. If the protein or region undergoes cooperative unfolding, the lower m/z peak will disappear during continuous incubation in D_2O, while the higher m/z peak will grow, and the rate of the disappearance of the lower peak will provide a measure of the rate of cooperative unfolding. This will be observed even if the unfolding is reversible, since deuterium uptake is effectively

irreversible in a large excess of D_2O. Under physiological conditions and in the absence of denaturant, such slow cooperative unfolding is rarely observed because the timescale of unfolding/refolding is generally much faster than the intrinsic rate of H/D exchange. However, it has been observed in some cases, most notably in certain SH3 domains (31).

We have observed this pattern of slow cooperative unfolding in residues 232–246 of p66, which is located at the base of the thumb subdomain and corresponds to β-strands 13 and 14 and a loop. Figure 7A shows the spectra of this peptide at 5, 10, 20, and 30 s of incubation in D_2O. The 5–30 s spectra show a double isotopic envelope as well as the disappearance of the lower m/z peak over time. Quantifying the relative amounts of folded and unfolded forms from the areas under Gaussian peaks indicates that the folded form disappears with a $t_{1/2} = 6.6$ s (Figure 7B). Because the isotopic envelopes are not well resolved, the fitting to two Gaussians is uncertain. Therefore, we also employed peak width analysis (32). Depending on the height at which peak widths were determined, a $t_{1/2} \sim 15–20$ s for unfolding is obtained. This is approximately 2-fold larger than the value determined for Gaussian fitting.

DISCUSSION

Comparing the distribution of slowly exchanging amide hydrogens with the distribution of hydrogen bonds indicates that the majority of secondary structural elements seen in the crystal structure of RT are relatively stable in solution (Figure 6). A major exception is the RNase H domain and the p51 thumb. Both of these regions exchange far more quickly than the crystal structure of RT or the solution structure of the isolated RNase H domain suggest. This high degree of conformational flexibility may allow the RNase H domain to reorient itself to accommodate different template/primer substrates and binding orientations (33). It is unlikely that absence of magnesium in the H/D exchange experiments accounts for this flexibility, as the dynamics appeared similar under a variety of solution conditions in NMR relaxation studies of the isolated RNase H domain (18). The isolated domain is well-structured on the NMR time scale, except for 24 residues at the C-terminus (17). The absence of resonances for most of the C-terminal helix αE', which is present in the crystal structure of the isolated domain (34), is indicative of a slow conformational exchange process. The HXMS time scale (s to h) is much slower than the NMR time scale (ms to ps). The fast exchanging amide hydrogens in peptide 534–560 of p66 observed by H/D exchange are consistent with an unstable C-terminus of the RNase H domain. In addition, the p66 thumb, while stable at short labeling times, shows substantial flexibility at longer times (Figure 5). This flexibility may allow for structural adaptation of thumb subdomain residues to binding sites on template/primer during polymerization.

Much of our current knowledge regarding the structural dynamics of RT comes from simulations. Bahar *et al.* employed a coarse-grained approach called GNM to examine the global dynamics of RT (7). They found that the lowest frequency modes were dominated by motions of the fingers and thumb subdomains and RNase H domain of p66, whereas p51 was essentially rigid. While whole domain motions will not necessarily be reflected in H/D exchange rates, it is nonetheless instructive to compare our results with the results of a recent ANM analysis (35). The two approaches are in agreement that the palm and connection subdomains are relatively rigid and that the RNase H domain and p66 thumb domain are flexible (Figure 6C). In particular, the p51 thumb—RNase H interface shows little protection from exchange, indicating weak contacts between the two domains. The weakness of these contacts might facilitate the large scale RNase H domain motions that are apparent in the lowest frequency normal modes.

The β-sheet β12–β13–β14, lies at the base of the p66 thumb subdomain. It contains the highly conserved WMG loop that forms part of the primer grip, as well as residues F227, W229, and

L234 that form part of the NNRTI binding pocket. The sheet is characterized as a hinge-bending center in the GNM analysis of RT—nevirapine complex (7). The slow cooperative unfolding in peptide 232–246 indicates that this β-sheet is not stable in solution, but is disrupted in a concerted manner, with an unfolding $t_{1/2}$ of 6.6 s. The p66 thumb subdomain assumes somewhat different orientations in different crystal structures. In the template/primer-bound structures it adopts an "open" conformation (4,25,26), in contrast to the "closed" conformation seen in the unliganded structures (5,6). Of the RT—template/primer structures currently available, three show disruption of either β-strand 12 or 14 (25,26,36-40). Based on these observations, we suggest that the instability seen in β-strands 12, 13 and 14 may facilitate thumb subdomain motions.

Both analysis of viral genomes isolated from patients and directed mutagenesis studies have provided extensive information on the degree to which different regions of RT tolerate sequence variation. Ceccherini-Silberstein *et al.* (41) examined patterns of sequence variation in the first 320 residues of RT isolated from 1,704 HIV-positive individuals and identified several contiguous stretches of high sequence conservation. Our H/D exchange experiments provide substantial sequence coverage for six of the nine regions. Of these six regions (residues 91–97, 107–117, 124–135, 147–157, 181–94 and 216–244), five are located in regions of rigid/stable structure, as evidenced by their relatively slow rates of H/D exchange. This is consistent with the general observation that proteins are less tolerant to mutations in regions that are tightly packed and/or involved in strong intramolecular interactions. The exception is the region defined by residues 216–244, which shows very rapid H/D exchange and contains β-sheet β12–β13–β14 mentioned above. Ceccherini-Silberstein *et al*. (41) analyzed RT sequences from both drug-naïve and drug-treated patients, and compared patterns of sequence conservation in the two populations. The stretch of residues 216–244 is one of three regions that are not conserved in drug-treated patients despite their high conservation in naïve patients. Two peptides, which together include residues 209–246, comprise the largest contiguous stretch of fast exchanging residues in the polymerase domain of RT, and also contain a number of mutations associated with drug resistance K219, P225, F227, and K238.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to Valerie A. Braz for discussion of HXMS data.

This work was supported by NIH grants GM071267 (MDB) and Case/UHC Center for AIDS Research AI36219 (PLW).

REFERENCES

(1). Coffin, JM.; Hughes, SH.; Varmus, HE. Retroviruses. Cold Spring Harbor Laboratory Press; Plainview, New York: 1997.

(2). Wang J, Smerdon SJ, Jäger J, Kohlstaedt LA, Rice PA, Friedman JM, Steitz TA. Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer. Proc. Natl. Acad. Sci. U.S.A 1994;91:7242–7246. [PubMed: 7518928]

(3). Le Grice SSFJ, Naas T, Wohlgensinger B, Schatz O. Subunit-selective mutagenesis indicates minimal polymerase activity in heterodimer-associated p51 HIV-1 reverse transcriptase. EMBO J 1991;10:3905–3911. [PubMed: 1718745]

(4). Jacobo-Molina A, Ding J, Nanni RG, Clark AD Jr. Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P, Hizi A, Hughes SH, Arnold E. Crystal structure of human immunodeficiency virus

type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. Proc. Natl. Acad. Sci. U. S. A 1993;90:6320–6324. [PubMed: 7687065]

(5). Hsiou Y, Ding J, Das K, Clark AD Jr. Hughes SH, Arnold E. Structure of unliganded HIV-1 reverse transcriptase at 2.7 Å resolution: implications of conformational changes for polymerization and inhibition mechanisms. Structure 1996;4:853–860. [PubMed: 8805568]

(6). Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison SC. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. Proc. Natl. Acad. Sci. U. S. A 1995;92:1222–1226. [PubMed: 7532306]

(7). Bahar I, Erman B, Jernigan RL, Atilgan AR, Covell DG. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. J. Mol. Biol 1999;285:1023–1037. [PubMed: 9887265]

(8). Madrid M, Jacobo-Molina A, Ding J, Arnold E. Major subdomain rearrangement in HIV-1 reverse transcriptase simulated by molecular dynamics. Proteins 1999;35:332–337. [PubMed: 10328268]

(9). Madrid M, Lukin JA, Madura JD, Ding J, Arnold E. Molecular dynamics of HIV-1 reverse transcriptase indicates increased flexibility upon DNA binding. Proteins 2001;45:176–182. [PubMed: 11599020]

(10). Shen L, Shen J, Luo X, Cheng F, Xu Y, Chen K, Arnold E, Ding J, Jiang H. Steered molecular dynamics simulation on the binding of NNRTI to HIV-1 RT. Biophys. J 2003;84:3547–3563. [PubMed: 12770866]

(11). Carvalho ATP, Fernandes PA, Ramos MJ. The excision mechanism in reverse transcriptase: pyrophosphate leaving and fingers opening are uncoupled events with the analogues AZT and d4T. J. Phys. Chem. B 2007;111:12032–12039. [PubMed: 17887788]

(12). Rodriguez-Barrios F, Perez C, Lobaton E, Velazquez S, Chamorro C, San-Felix A, Perez-Perez MJ, Camarasa MJ, Pelemans H, Balzarini J, Gago F. Identification of a putative binding site for [2′,5′-bis-O-(tert-butyldimethylsilyl)-beta-D-ribofuranosyl]-3′-spiro-5”-(4”-amino-1”,2”-oxathiole-2”, 2”-dioxide)thymine (TSAO) derivatives at the p51-p66 interface of HIV-1 reverse transcriptase. J. Med. Chem 2001;44:1853–1865. [PubMed: 11384232]

(13). Rodríguez-Barrios F, Balzarini J, Gago F. The molecular basis of resilience to the effect of the Lys103Asn mutation in non-nucleoside HIV-1 reverse transcriptase inhibitors studied by targeted molecular dynamics simulations. J. Am. Chem. Soc 2005;127:7570–7578. [PubMed: 15898808]

(14). Zhou Z, Madrid M, Evanseck JD, Madura JD. Effect of a bound non-nucleoside RT inhibitor on the dynamics of wild-type and mutant HIV-1 reverse transcriptase. J. Am. Chem. Soc 2005;127:17253–17260. [PubMed: 16332074]

(15). Kensch O, Restle T, Wohrl BM, Goody RS, Steinhoff HJ. Temperature-dependent equilibrium between the open and closed conformation of the p66 subunit of HIV-1 reverse transcriptase revealed by site-directed spin labelling. J. Mol. Biol 2000;301:1029–1039. [PubMed: 10966802]

(16). Powers R, Clore GM, Stahl SJ, Wingfield PT, Gronenborn A. Analysis of the backbone dynamics of the ribonuclease H domain of the human immunodeficiency virus reverse transcriptase using ^{15}N relaxation measurements. Biochemistry 1992;31:9150–9157. [PubMed: 1382587]

(17). Pari K, Mueller GA, DeRose EF, Kirby TW, London RE. Solution structure of the RNase H domain of the HIV-1 reverse transcriptase in the presence of magnesium. Biochemistry 2003;42:639–650. [PubMed: 12534276]

(18). Mueller GA, Pari K, DeRose EF, Kirby TW, London RE. Backbone dynamics of the RNase H domain of HIV-1 reverse transcriptase. Biochemistry 2004;43:9332–9342. [PubMed: 15260476]

(19). Wales TE, Engen JR. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass. Spectrom. Rev 2006;25:158–170. [PubMed: 16208684]

(20). Le Grice SFJ, Cameron CE, Benkovic SJ. Purification and characterization of human immunodeficiency virus type 1 reverse transcriptase. Methods Enzymol 1995;262:130–144. [PubMed: 8594344]

(21). Venezia CF, Howard KJ, Ignatov ME, Holladay LA, Barkley MD. Effects of efavirenz binding on the subunit equilibria of HIV-1 reverse transcriptase. Biochemistry 2006;45:2779–2789. [PubMed: 16503633]

(22). Tsutsui Y, Liu L, Gershenson A, Wintrode PL. The conformational dynamics of a metastable serpin studied by hydrogen exchange and mass spectrometry. Biochemistry 2006;45:6561–6569. [PubMed: 16716066]

(23). Divita G, Rittinger K, Restle T, Immendorfer U, Goody RS. Conformational stability of dimeric HIV-1 and HIV-2 reverse transcriptases. Biochemistry 1995;34:16337–16346. [PubMed: 8845359]

(24). Zhang Z, Smith DL. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. Protein. Sci 1993;2:522–531. [PubMed: 8390883]

(25). Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science 1998;282:1669–1675. [PubMed: 9831551]

(26). Sarafianos SG, Das K, Tantillo C, Clark AD Jr. Ding J, Whitcomb JM, Boyer PL, Hughes SH, Arnold E. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. EMBO J 2001;20:1449–1461. [PubMed: 11250910]

(27). Das K, Lewi PJ, Hughes SH, Arnold E. Crystallography and the design of anti-AIDS drugs: conformational flexibility and positional adaptability are important in the design of non-nucleoside HIV-1 reverse transcriptase inhibitors. Prog. Biophys. Mol. Biol 2005;88:209–231. [PubMed: 15572156]

(28). Englander SW, Kallenbach NR. Hydrogen exchange and structural dynamics of proteins and nucleic acids. Q. Rev. Biophys 1983;16:521–655. [PubMed: 6204354]

(29). Jacques PS, Wohrl BM, Howard KJ, Le Grice SFJ. Modulation of HIV-1 reverse transcriptase function in “selectively deleted” p66/p51 heterodimers. J. Biol. Chem 1994;269:1388–1393. [PubMed: 7507107]

(30). Hooft RW, Sander C, Vriend G. Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. Proteins 1996;26:363–376. [PubMed: 8990493]

(31). Engen JR, Smithgall TE, Gmeiner WH, Smith DL. Identification and localization of slow, natural, cooperative unfolding in the hematopoietic cell kinase SH3 domain by amide hydrogen exchange and mass spectrometry. Biochemistry 1997;36:14384–14391. [PubMed: 9398156]

(32). Weis DD, Wales TE, Engen JR, Hotchko M, Ten Eyck LF. Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. J. Am. Soc. Mass Spectrom 2006;17:1498–1509. [PubMed: 16875839]

(33). Abbondanzieri EA, Bokinsky G, Rausch JW, Zhang JX, Le Grice SFJ, Zhuang X. Dynamic binding orientations direct activity of HIV reverse transcriptase. Nature 2008;453:184–189. [PubMed: 18464735]

(34). Davies JF II, Hostomska Z, Hostomsky Z, Jordan SR, Matthews DA. Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase. Science 1991;252:88–95. [PubMed: 1707186]

(35). Bahar, I. Anisotropic Network Model web server. 2009. http://ignmtest.ccbb.pitt.edu/cgi-bin/anm/anm1.cgi

(36). Jaeger J, Restle T, Steitz TA. The structure of HIV-1 reverse transcriptase complexed with an RNA pseudoknot inhibitor. EMBO J 1998;17:4535–4542. [PubMed: 9687519]

(37). Sarafianos SG, Das K, Clark AD Jr. Ding J, Boyer PL, Hughes SH, Arnold E. Lamivudine (3TC) resistance in HIV-1 reverse transcriptase involves steric hindrance with β-branched amino acids. Proc. Natl. Acad. Sci. U. S. A 1999;96:10027–10032. [PubMed: 10468556]

(38). Sarafianos SG, Clark AD Jr. Das K, Tuske S, Birktoft JJ, Ilankumaran P, Ramesha AR, Sayer JM, Jerina DM, Boyer PL, Hughes SH, Arnold E. Structures of HIV-1 reverse transcriptase with pre- and post-translocation AZTMP-terminated DNA. EMBO J 2002;21:6614–6624. [PubMed: 12456667]

(39). Peletskaya EN, Kogon AA, Tuske S, Arnold E, Hughes SH. Nonnucleoside inhibitor binding affects the interactions of the fingers subdomain of human immunodeficiency virus type 1 reverse transcriptase with DNA. J. Virol 2004;78:3387–3397. [PubMed: 15016861]

(40). Tuske S, Sarafianos SG, Clark AD Jr. Ding J, Naeger LK, White KL, Miller MD, Gibbs CS, Boyer PL, Clark P, Wang G, Gaffney BL, Jones RA, Jerina DM, Hughes SH, Arnold E. Structures of HIV-1 RT-DNA complexes before and after incorporation of the anti-AIDS drug tenofovir. Nat. Struct. Mol. Biol 2004;11:469–474. [PubMed: 15107837]

(41). Ceccherini-Silberstein F, Gago F, Santoro M, Gori C, Svicher V, Rodriguez-Barrios F, d'Arrigo R, Ciccozzi M, Bertoli A, d'Arminio Monforte A, Balzarini J, Antinori A, Perno CF. High sequence conservation of human immunodeficiency virus type 1 reverse transcriptase under drug pressure despite the continuous appearance of mutations. J. Virol 2005;79:10718–10729. [PubMed: 16051864]

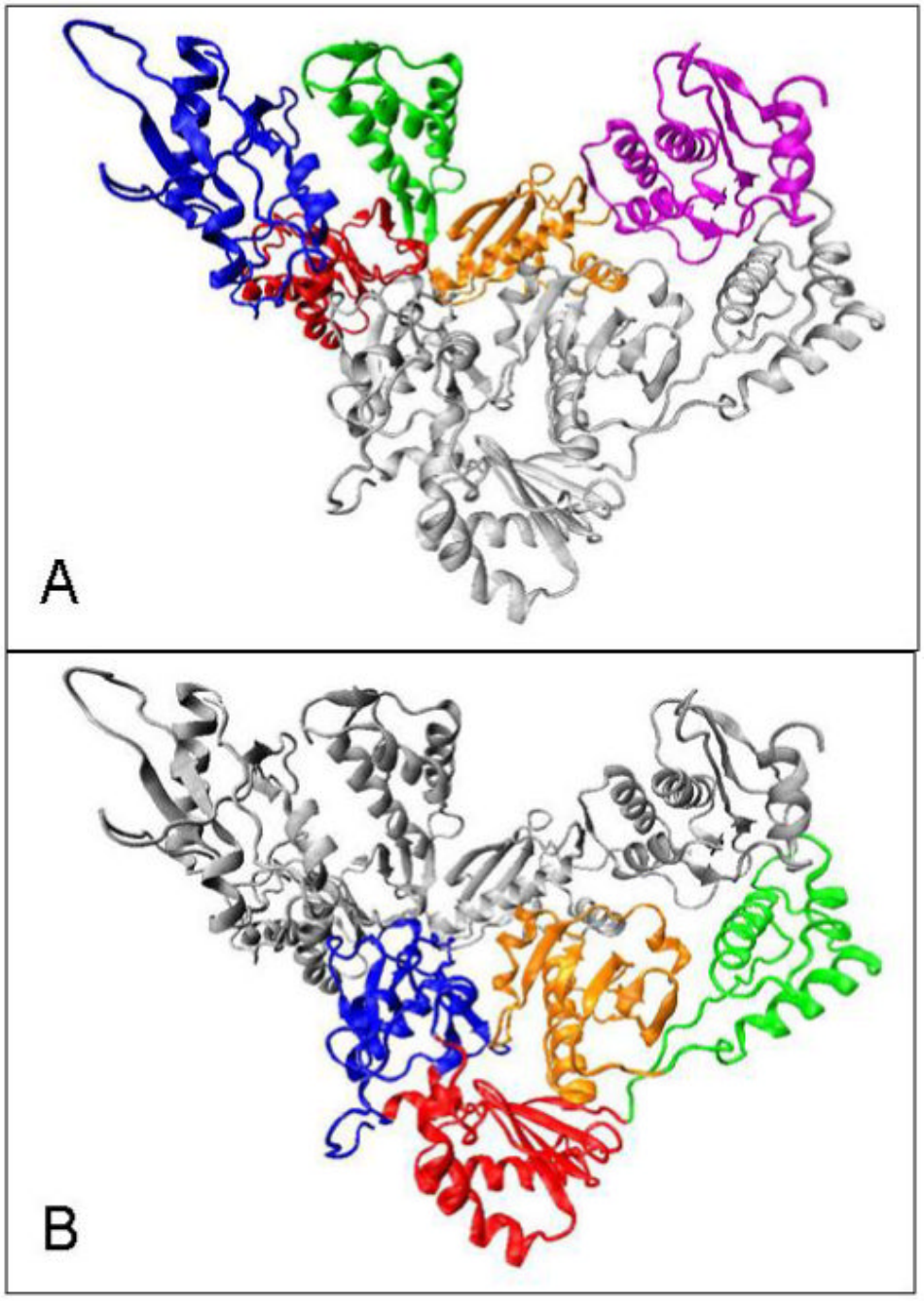


Figure 1.
Structure of unliganded HIV-1 RT. (A) p66 subunit (colored), p51 subunit (gray) and (B) p66 subunit (gray), p51 subunit (colored). Four subdomains of the polymerase domain: fingers (blue), palm (red), thumb (green), connection (orange); RNase H domain (magenta). Protein Databank ID 1DLO (5).

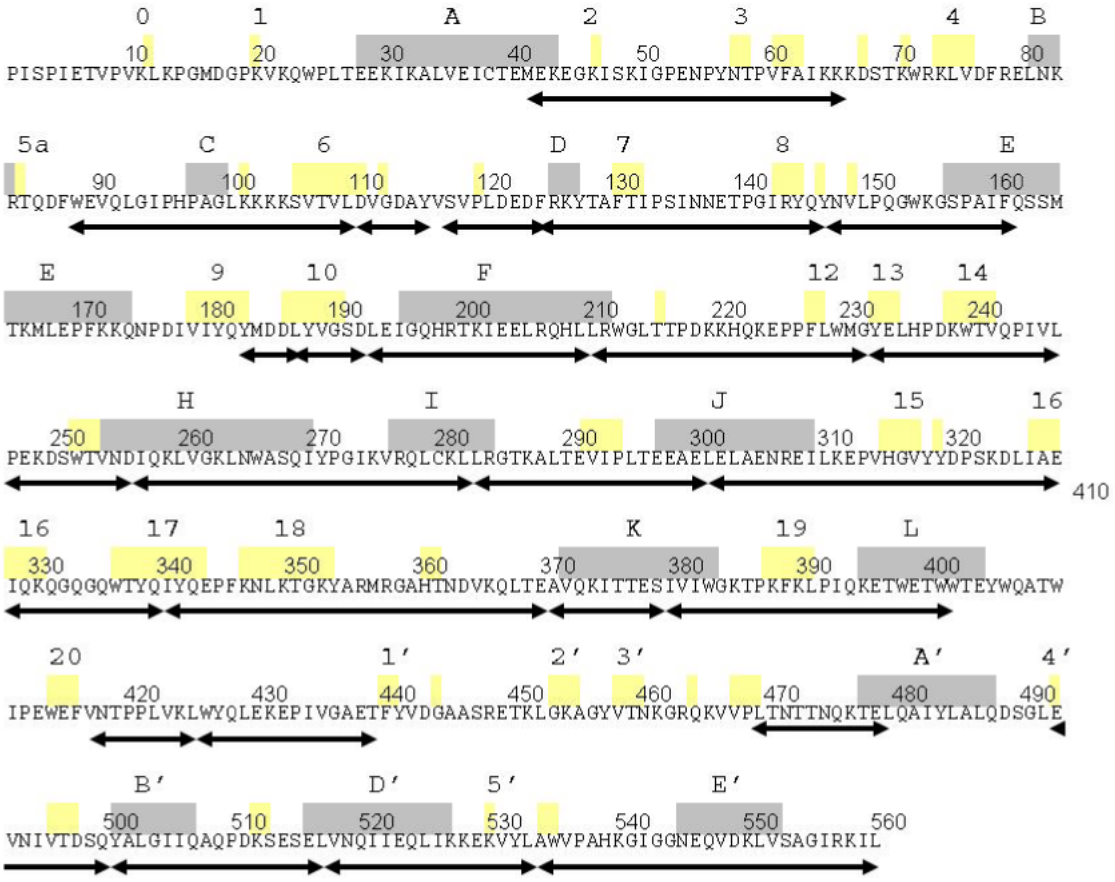


Figure 2.

Peptides used for analysis of H/D exchange data. Peptides are indicated by double-headed arrows under the sequence of HIV-1 p66. Structural elements from 1DLO: α-helices (gray bars), letters, β-sheets (yellow bars), numbers.

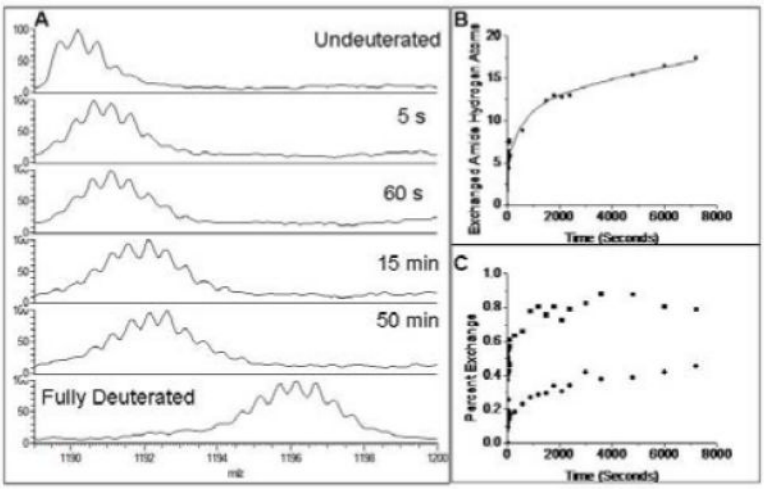
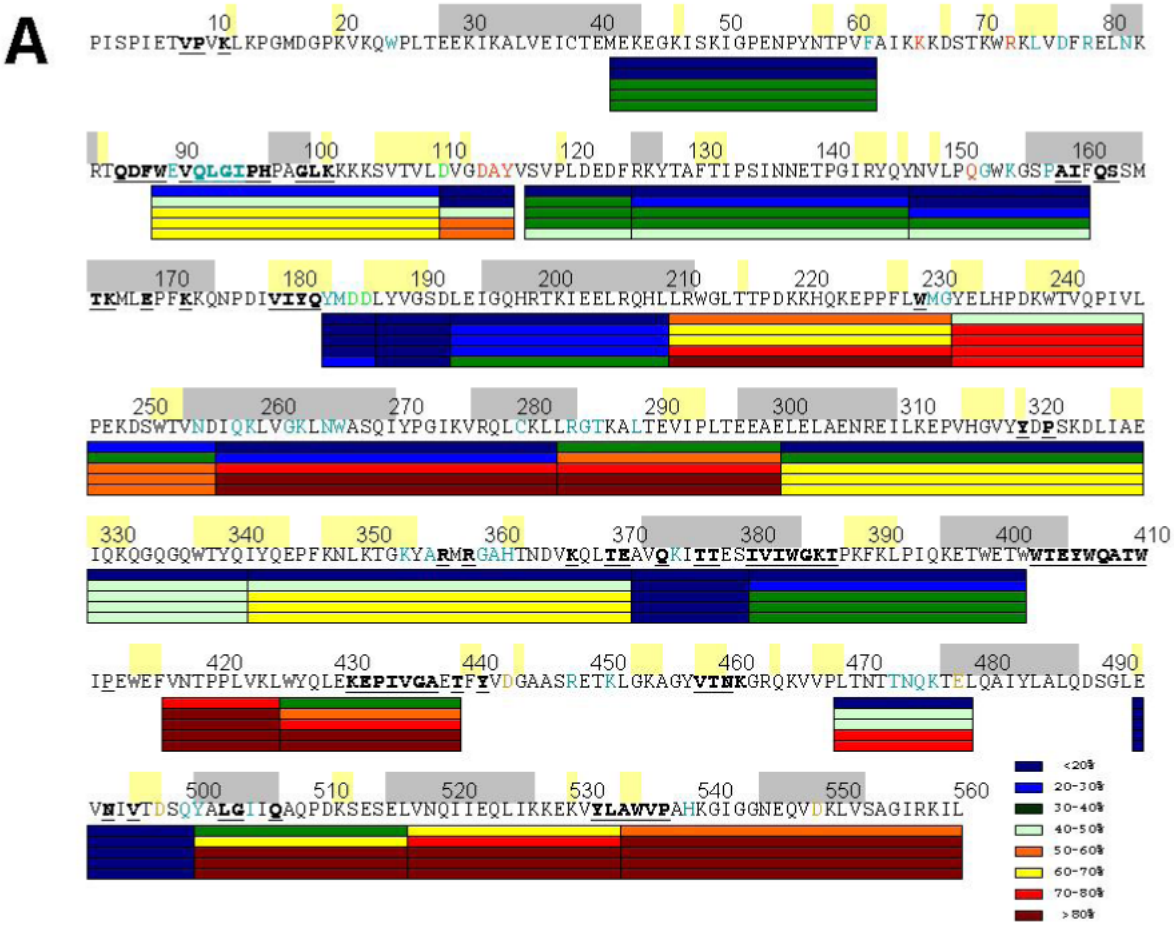


Figure 3.

HXMS data and curve fitting of H/D exchange data for p66 subunit. (A) Raw spectra showing a peak for peptide 41–61 at different incubation times in RT buffer D-D_2O. Spectra of undeuterated and fully deuterated reference samples also shown. (B) Number of amide hydrogens exchanged with deuteriums vs time. Solid line is fit of the data to eq 2. (C) Percent exchange vs time of two peptides. Peptide 147–160 (lower curve) is in a buried region of the fingers/palm subdomains in the crystal structure of HIV-1 RT (1DLO). Peptide 283–300 (upper curve) is in a surface-exposed region of the thumb subdomain.

A



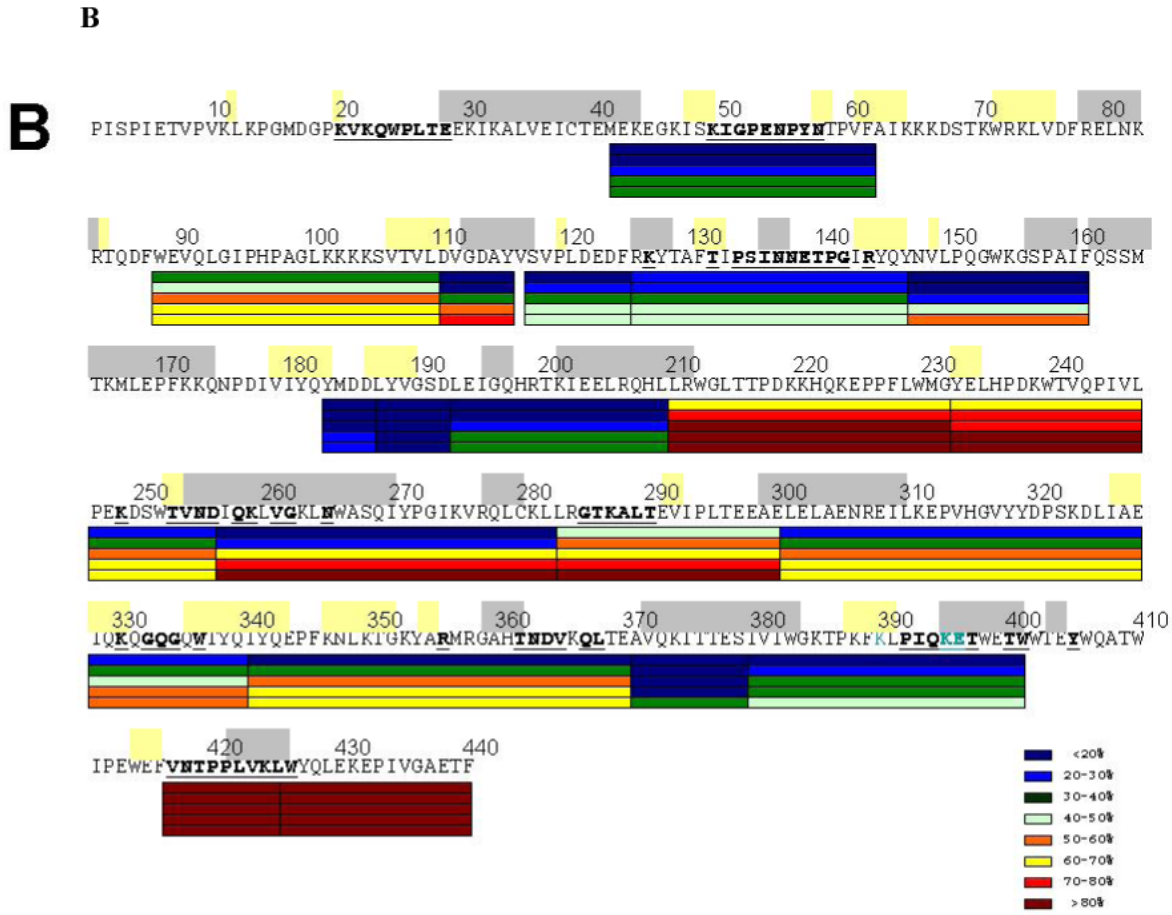


Figure 4.

Percent exchange of peptides from (A) p66 subunit and (B) p51 subunit of HIV-1 RT. Colored bars below sequence from top to bottom give exchange at 5, 60, 600, 3000, and 7200 s. Gray and yellow bars above sequence correspond to α-helices and β-sheets. Contact residues: dimer interface (underlined bold face), template/primer substrate binding site (blue), dTTP substrate binding site (red), polymerase active site (green), RNase H active site (brown).

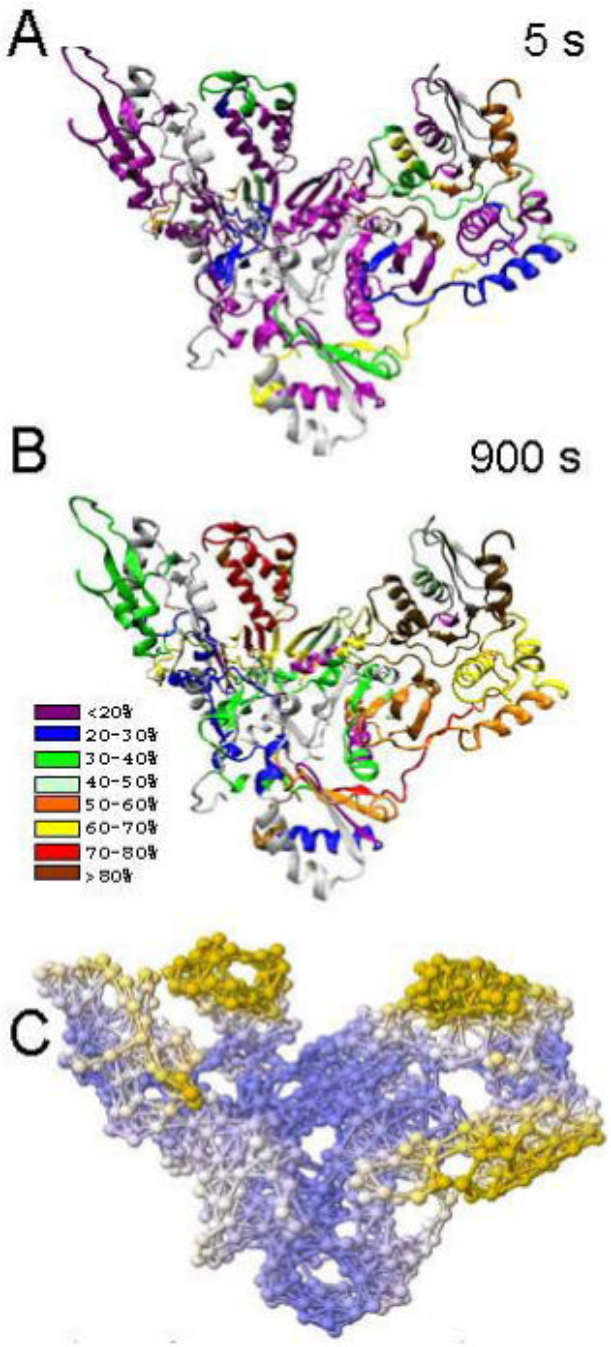


Figure 5.

Percent exchange for both subunits mapped onto the structure of HIV-1 RT. (A) 5 s and (B) 900 s incubation in RT buffer D-D_2O. Lower percentages correspond to a greater degree of protection from exchange. (C) Mobility in the second mode as calculated by ANM (35). Gold regions represent highly mobile regions; blue regions represent rigid regions.

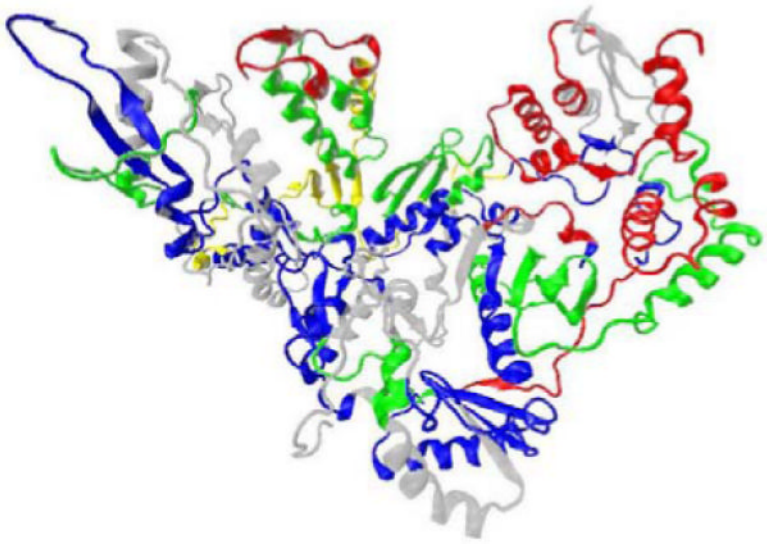


Figure 6.
Stability of hydrogen bonds in RT. Structure of HIV-1 RT colored by the ratio of slowly exchanged to hydrogen-bonded amide hydrogens: missing coverage (gray), very flexible regions (red), flexible regions (yellow), rigid regions (green), very rigid regions (blue).

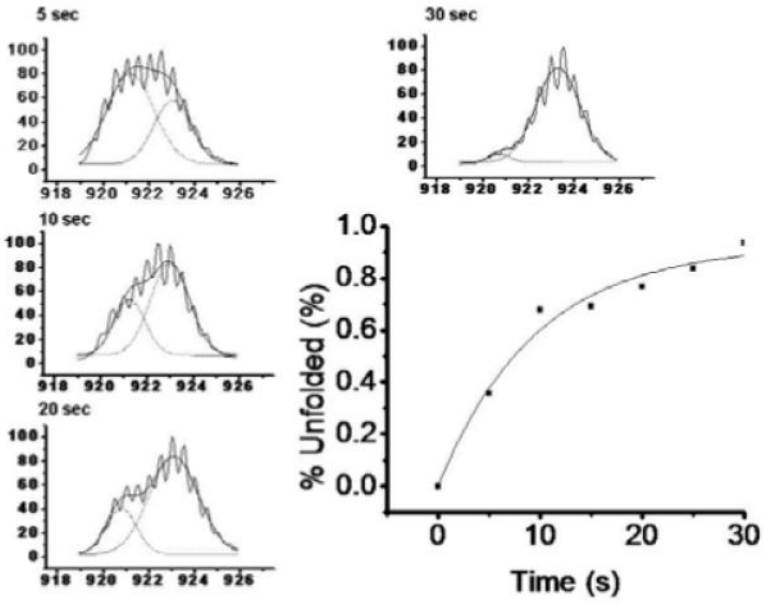


Figure 7.

(A) Mass spectra of peptide 232–246 at different incubation times in RT buffer D-D_2O. The high and low m/z peaks are fit to Gaussians (dotted lines). (B) Fraction unfolded as a function of time. Solid line is fit to a single exponential.

Table 1

Comparison of the number of slow exchanging amide hydrogens[a] with the number of hydrogen bonds[b]

Sequence	residue	p66 subunit			p51 subunit		
	number	n-slow	n-whatif	ratio[c]	n-slow	n-whatif	ratio[c]
MEKEGKISKIGPENPYNTPVF	41–61	11	14	0.79	11	13	0.85
WEVQLGIPHPAGLKKKKSVTVL	88–109	7	10	0.70	8	10	0.80
DVGDAY	110–115	2	3	0.67	3	5	0.60
SVPLDEDF	117–124	4	4	1.00	3	3	1.00
FRKYTAFTIPSINNETPGIRYQY	124–146	13	12	1.08	13	16	0.81
NVLPQGWKGSPAIF	147–160	7	7	1.00	5	8	0.63
YMDDL	183–187	4	2	2.00	3	2	1.50
LYVGSD	187–192	5	4	1.25	5	4	1.25
LEIGQHRTKIEELRQHL	193–209	11	13	0.85	12	11	1.09
LRWGLTTPDKKHQKEPPFLWMG	210–231	5	11	0.45	4	5	0.80
YELHPDKWTVQPIVL	232–246	3	7	0.43	0	3	0.00
PEKDSWTVND	247–256	3	4	0.75	4	5	0.80
IQKLVGKLNWASQIYPGIKVRQLCKL	257–282	13	20	0.65	7	18	0.39
LRGTKALTEVIPLTEEAE	283–300	3	9	0.33	5	8	0.63
LELAENREILKEPVHGVYYDPSKDLIAE	301–328	10	20	0.50	11	17	0.65
IQKQGQGQWTYQ	329–340	5	8	0.63	6	9	0.67
IYQEPFKNLKTGKYARMRGAHTNDVKQLTE	341–370	11	14	0.79	14	19	0.74
AVQKITTES	371–379	8	8	1.00	8	8	1.00
IVIWGKTPKFKLPIQKETWETW	380–401	12	15	0.80	13	16	0.81
VNTPPLVKL	417–425	1	2	0.50	0	3	0.00
WYQLEKEPIVGAET	426–439	2	2	1.00	-	-	-
WYQLEKEPIVGAETF	426–440	-	-	-	2	1	2.00
LTNTTNQKTEL	469–479	2	6	0.33	-	-	-
EVNIVTDSQ	492–500	8	6	1.33	-	-	-
YALGIIQAQPDKSESEL	501–517	3	12	0.25	-	-	-
VNQIIEQLIKKEKVYL	518–533	0	13	0.00	-	-	-

Sequence	residue	p66 subunit			p51 subunit		
	number	n-slow	n-whatif	ratio[c]	n-slow	n-whatif	ratio[c]
AWVPAHKGIGGNEQVDKLVSAGIRKIL	534–560	4	11	0.36	-	-	-

[a] Slow exchanging amide hydrogens in Tables S1 and S1.

[b] Number of hydrogen bonds determined by WhatIf using the RT structure 1DLO

[c] Ratio of slow exchanging hydrogens to the number of hydrogen bonds.