Journal of Biomedical Science (2005) 12:701–710
DOI 10.1007/s11373-005-9011-4

Structural analysis of reverse transcriptase mutations at codon 215 explains the predominance of T215Y over T215F in HIV-1 variants selected under antiretroviral therapy

Nouara Yahi[1,*], Jacques Fantini[1], Mireille Henry[2], Christian Tourrès[2] & Catherine Tamalet[2]

[1]*Laboratoire de Biochimie et Physicochimie des Membranes Biologiques, Faculté des Sciences et Techniques St-Jérôme, Université Paul Cézanne, 13013, Marseille, France;* [2]*Laboratoire de Virologie, UF SIDA, Hôpital de la Timone, 13005, Marseille, France*

Received 23 March 2005; accepted 19 July 2005

Key words: aromatic, antiretroviral, CH–π interaction, fitness, HIV-1, π–π interaction, mutation, resistance, therapy

Summary

Mutations at codon 215 of HIV-1 reverse transcriptase (RT) confer resistance to nucleoside analogs through RT-catalyzed ATP-dependent phosphorolysis. We showed that mutation T215Y is predominant over T215F (respectively 38.8 vs. 7.04% of 7312 sequences from a cohort of patients receiving antiretroviral therapy in France). Ambiguous mixtures at codon 215 (e.g. TNYS and TFSI) were resolved by cloning and sequencing representative clinical samples. Mutation T215F was preferentially associated with K70R (>71%), D67N (>73%) and K219Q/E/N (>76%), whereas T215Y was associated with M41L (>84%) and L210W (>58%). A similar distribution was observed with RT sequences stored in the Stanford HIV Drug Resistance Database. The structural background of these two distinct mutational patterns was investigated by molecular modeling of ATP-mutant RT complexes, on the basis of known ATP–protein interactions. We found that the aromatic side chain of tyrosine (Y) – but not phenylalanine (F) – optimally stacked with the adenine ring of ATP. Mutation L210W further stabilized this aromatic π–π stacking interaction, increasing the affinity of the T215Y/L210W double mutant for ATP. Overall, this study provides a biochemical basis accounting for the evolutionary pathway of T215 mutations in HIV-1 RT, leading to the preferential selection of T215Y vs. T215F.

Introduction

Mutations at codon 215 of HIV-1 RT play a central role in HIV-1 resistance to thymidine analogs such as zidovudine (AZT) and stavudine (d4T) [1]. HIV-1 reverse transcriptase (RT) amino acid position 215 is monomorphic for threonine (Thr, T) in the wild-type virus population. Drug resistant variants containing aromatic amino acids at position 215 arise under drug selective pressure [2]. The main mutated amino acids are tyrosine (Tyr, Y) or phenylalanine (Phe, F) [3]. As two nucleotide changes (ACC to TAC or TTC) are required to substitute T by Y or F, intermediate genotypes with single nucleotide mutations can also be detected during the course of antiretroviral therapy [3]. Interestingly, these transient mutations do not confer by themselves resistance to thymidine analogs but represent a step towards the

*To whom correspondence should be addressed. Fax +33-491-288236; E-mail: nouara.yahi@univ.u-3mrs.fr

emergence of T215Y/F [4, 5]. Accordingly, it is widely admitted that AZT resistance requires the introduction of an aromatic residue at position 215 [6, 7]. In contrast with mutation M184V, which decreases the affinity of the mutated RT for the nucleoside analog lamivudine (3TC) through a steric hindrance mechanism [8], thymidine analog mutations (TAMs), including T215Y/F, do not prevent drug binding to RT [9]. Instead, TAMs confer increased rates of phosphorolytic removal of chain-terminating AZT from the 3′ terminus of the primer [10]. Thus, the AZT-resistant mutant RT exhibits an increased rate of RT-catalyzed ATP-dependent phosphorolysis [11]. In this respect, TAMs act essentially by increasing the affinity of RT for ATP, so that the excision reaction can occur at physiological ATP concentrations. In particular, the presence of an aromatic side chain at position 215 is likely to allow a stacking interaction with the aromatic cycles of adenine through π–π interactions [11, 12]. TAMs generally associated with T215Y/F (i.e. M41L, D67N, K70R, L210W, and K219Q) may also contribute to increase the affinity of RT for ATP through a fine tuning of the three-dimensional structure of the enzyme [12]. However, TAMs cannot be stochastically combined: if some of them are very frequently associated (e.g. M41L and T215Y), some appear to be mutually exclusive (e.g. K70R and L210W) [3, 13]. Correspondingly, two specific sets of TAMs combinations have been characterized: (i) 41L + 210W + 215Y, and (ii) 67N + 70R + 215F + 219Q/E/N [14]. These profiles reveal that 215Y can be associated with 210W, but not with 70R, whereas 215F can be associated with 70R, and not with 210W [15]. Considering that Y and F only differ by a hydroxyl (–OH) group in their side chains (–CH_2-phenyl for F, –CH_2-phenol for Y), such a high level of selectivity in these mutational patterns may be somewhat surprising. The aim of the present study was to analyze the mutational patterns associated with 215F and 215Y in two independent large HIV-1 RT of sequence databases: our own database, which has been initiated in 1997 in Marseille and counts more than 7000 RT sequences collected during the clinical follow-up of HIV-infected patients under treatment [3, 16, 17], and the Stanford Drug Resistance Database developed on a world-wide basis by Robert W. Shafer [18]. This allowed us to provide a precise assessment of the percentage of associations between T215Y and T215F with other TAMs. Moreover, a significant percentage of sequence data could not be directly interpreted because of ambiguities at position 215 due to multiple codon possibilities. To clarify this point, representative samples were cloned and sequenced. This strategy confirmed the predominance of T, Y and F at codon 215, leading to a clear-cut simplification of sequence electropherogram interpretations in routine clinical practice. Finally, the respective impact of T215F and T215Y mutations on ATP positioning in the mutated RT has been studied by modeling approaches. The mutational background in which these mutations occurred has also been studied by molecular modeling, giving rise to a biochemical explanation for the emergence of specific mutational profiles.

Materials and methods

Sequence analysis

Plasma samples from HIV-1 infected patients under antiretroviral therapy were obtained from routine clinical practice in specialized AIDS hospitals, from January 1997 to May 2004 [16]. Viral RNA extraction, polymerase chain reaction (PCR) amplification in the RT gene, and sequencing were carried out as described previously [17]. A 1316 bp fragment encompassing the RT coding region was sequenced in both the 5′ and 3′ directions using the ABI Prism Big Dye Terminator cycle sequencing Ready Reaction kit version 2.0 (Applied Biosystems). Overall, 7312 RT sequences were available for analysis and the data were stored in a specifically designed database [16, 17]. Specific queries were made in this database and in the Stanford HIV Drug Resistance Database ([18], http://hivdb.stanford.edu/). Sequence differences were compared with an HIV-1 standard laboratory strain (HXB2) and were analyzed at key positions associated with drug resistance as well as other polymorphic sites. TAMs, also referred to as nucleoside excision mutations (NEMs), were defined as M41L, D67N, K70R, L210W, T215Y or F, and K219Q, E or N [19].

Clonal sequencing

Amplified RT fragments were inserted into PGEM-easy plasmid (Promega) and cloned in competent *Escherichia coli*. The recombinant plasmids were purified using Wizard plus SV minipreps DNA purification system (Promega). The inserted pol gene fragments were amplified using inner primers. Bidirectional overlapping dideoxynucleoside sequencing reactions were performed and products were resolved electrophoretically on an ABI 3100 sequencer (Applied Biosystems). Approximately 20 clones per isolate were sequenced.

Molecular modeling

Modeling of RT mutations into the p66 subunit of the ternary RT-template primer-dTTP complex [20] was carried out with the Swiss Model program [21]. The A chain of the RT (PDB entry 1RTD:A) was chosen as template for the generation of the model. In our models, the orientation of aromatic 215 mutant residues was obtained from crystal structures of mutant RT (PDB Entry 1RT3 for Phe215 and 1LW0 for Tyr215) [11, 22]. For Trp210 and Arg70 side chains, minor adjustments were obtained by selecting adequate rotamers in the PDB library and/or using the torsion command of the software. This allowed us to obtain these residues in the configuration proposed by Boyer et al. [11]. The Hyperchem 7 package was used for studying the geometry of the ATP site of mutated HIV-1 RT (1RTD:A) and docking in ATP. After calculating the minimal energy of each ATP–RT complex with the Polak–Ribiere algorithm, the Hyperchem files were saved in a PDB compatible format, visualized and further analyzed with the Deep View program ([23], http://www.expasy.org/spdbv/).

Results

Of the 7312 RT sequences stored in our database, 2837 (38.8 %) were unambiguously 215Y, and 515 (7.04%) 215F. The percentages of TAMs associated with either 215Y or 215F are shown in Table 1. The data were obtained from our database (Marseille database) or from the Stanford HIV Drug Resistance Database (Stanford database). The results were remarkably convergent. In both cases, 215Y was often accompanied by 41L and 210W mutations, but not with 70R. In contrast, 215F was often associated with 67N, 70R and 219Q/E/N.

As a matter of fact, careful examination of sequence electrophoregrams at position 215 revealed a significant proportion of ambiguous base calling, indicating possible nucleotide mixtures. Representative types of ambiguities at codon 215 are shown in Figure 1 (left panel). In all cases, the very low background in these electrophoregrams is consistent with the detection of HIV-1 variants within a mixture. Despite the multiple theoretical codon possibilities listed in the central panel of Figure 1, two main profiles corresponding to either TNYS or TFSI can be interpreted. In our database, these profiles were observed in 152 and 47 sequences, respectively for TNYS and TFSI. To discriminate between these 4 possibilities for each profile, 2 representative samples of each profile were cloned and sequenced. For TFSI, 21 and 23 clones were generated from 2

Table 1. Associations of T215Y or T215F with other TAMs.

	% Of sequences with RT gene mutation				
T215 mutation	41L	67N	70R	210W	219Q/E/N
Marseille database					
T215Y	90	48	10	70	27
T215F	54	84	71	11	79
Stanford database					
T215Y	84	40	13	58	14
T215F	48	73	76	8	76

Associations of mutations were analyzed in HIV-1 RT sequences obtained in Marseille patients since 1997 and stored in our own database (7312 sequences) or in RT sequences available through the Stanford HIV resistance database.

	ACC	Thr (100% wild-type)
	TAC	Tyr (100% mutant)
	TTC	Phe (100% mutant)
	ACC, TTC TCC, ATC	Thr, Phe Ser, Ile (TFSI profile)
	ACT, TTT TCT, ATT	Thr, Phe Ser, Ile (TFSI profile)
	ACC, ACT, TTC, TTT ATC, ATT, TCC, TCT	Thr, Thr, Phe, Phe Ile, Ile, Ser, Ser (TFSI profile)
	ACC, AAC TAC, TCC	Thr, Asn Tyr, Ser (TNYS profile)
	ACC, ACT, AAC, AAT TAC, TAT, TCC, TCT	Thr, Thr, Asn, Asn Tyr, Tyr, Ser, Ser (TNYS profile)
	ACT, AAT TAT, TCT	Thr, Asn Tyr, Ser (TNYS profile)

Figure 1. Representative electrophoregrams of codon 215 of HIV-1 RT sequences. The theoretical codons are listed in the middle column, and the corresponding amino acids in the right column.

clinical isolates, and the results were: (i) for the first sample, 17 clones with 215F, 5 with 215T, and 1 with 215L; for the second sample, 12 clones with 215T, and 9 clones with 215F. For the TNYS profile, 23 and 25 clones were generated from 2 isolates, and the results were : (i) 19 clones with

215T, and 4 with 215Y; (ii) 9 clones with 215T, 14 clones with 215Y, and 2 with a deleted 215 codon. Overall, these data demonstrated the high prevalence of 215T, 215Y and 215F in the ambiguous mixtures, excluding other hypothetical variants such as S, I, and N. Moreover, detection of mutants that were not predicted from the mixtures (L and the deletion) was anecdotic (3 cases over 92 clones, i.e. 3.26%). Overall, these data reinforce the conclusion that Y and F are the most frequent mutations at codon 215. Thus, the modeling approach used thereafter has been conducted in order to shed some light on the biochemical logic of the T215Y vs. T215F preference that clearly emerged from the data in Table 1.

In Figure 2a and 2b, the p66 subunit of the RT complexed with its substrates (template-primer + dTTP) has been modeled with a Tyr or a Phe residue instead of Thr at position 215. It is clear by looking at both the front view (Figure 2a) and the rear view (Figure 2b) of the enzyme that the orientation of the aromatic side chain is different for Phe and Tyr. It is important to state that these models have been reproduced with the original 215 atom coordinates, i.e. without further refinement in the direction of the side chains. This indicates that Phe and Tyr mutations of codon 215 have not the same impact on the three-dimensional structure of HIV-1 RT. The positions of the 210W and 70R mutants are also indicated in the model. The side chain of the 210W mutant has been finely tuned in order to project the aromatic indole ring towards the aromatic side chain at position 215. Although all these mutations cluster around the dNTP binding pocket, their main effect is to increase the affinity of ATP for RT by defining a new binding site that is required for the ATP-dependent excision reaction.

Now, let us consider that the first TAM to appear is K70R. As shown in Figure 2c and 2e, the delocalized guanidinium charge of the arginine side chain may act as a donor H group, forming a H-bond with the oxygen atom of the ribofuranose cycle. This kind of interaction may not be optimal, since it does not allow any stabilization for the aromatic ring of the adenine moiety of ATP. However, if mutation T215F occurs after the K70R mutation, ATP will be wedged between the side chains of Phe215 and Arg70, and the interaction will be stabilized by an adenine-Phe stacking. In this context, mutation L210W does not appear, probably because it would stack on Phe 215 and displace the complex in the direction of Arg70 (i.e. on the right in Figure 2e), which would result in a steric clash with the guanidinium cation of Arg70. On the other hand, if mutation 215Y appears first, ATP will bind to the Tyr side chain through aromatic stacking with the adenine ring, but also with the ribofuranose cycle (Figure 2d and 2f). Indeed, both cycles are in the range of 4–5Å from Tyr, a distance compatible with both π–π (aromatic–aromatic) and CH–π (sugar–aromatic) stacking interactions [24]. In this context, K70R is unlikely to appear, because ATP, trapped by Tyr 215 (i.e. on the left in Figure 2f), would not have any atom available for an interaction with the arginine side chain. Thus, there is no particular pressure for selecting K70R mutants from the Y215 population. In contrast, the interaction between Tyr215 and adenine needs a reinforcement provided by the 210W mutation, which may: (i) contribute to stabilizing the aromatic interaction between ATP and Tyr by allowing the formation of a complex network of π–π stacking interactions (Trp + Tyr stronger than Tyr alone); (ii) lock the aromatic side chain of Tyr so that it cannot move outside the ATP binding site, for instance by forming a H bond with Asp 218 (data not shown). Finally, the two distinct patterns of TAMs (41L + 67N + 210W + 215Y + 219Q, or 41L + 67N + 70R + 215F + 219Q) have been modeled in the background of HIV-1 RT (Figure 2c–f). These models were obtained by injecting the sequence of two clinical RT isolates harboring the adequate pattern of mutations into the Swiss Model program, using 1RTD:A as template. The side chains Y215 and F215 were positioned according to the crystal structure of the corresponding RT mutants [15, 22]. Significant differences in the ATP binding site can be visualized between both patterns, in agreement with the concept that there is biochemical logic in the development of specific mutational pathways involving positions 70, 210 and 215 of HIV-1 RT. In particular, if one assumes that the π–π stacking interaction between the aromatic side chain at position 215 is the key determinant of the ATP–RT interaction, the nature of this side chain (Tyr or Phe) may therefore affect the affinity of mutant RT for ATP. This may in turn influence the selective pressure that applies on residues 67 and 219 of HIV-1 RT.

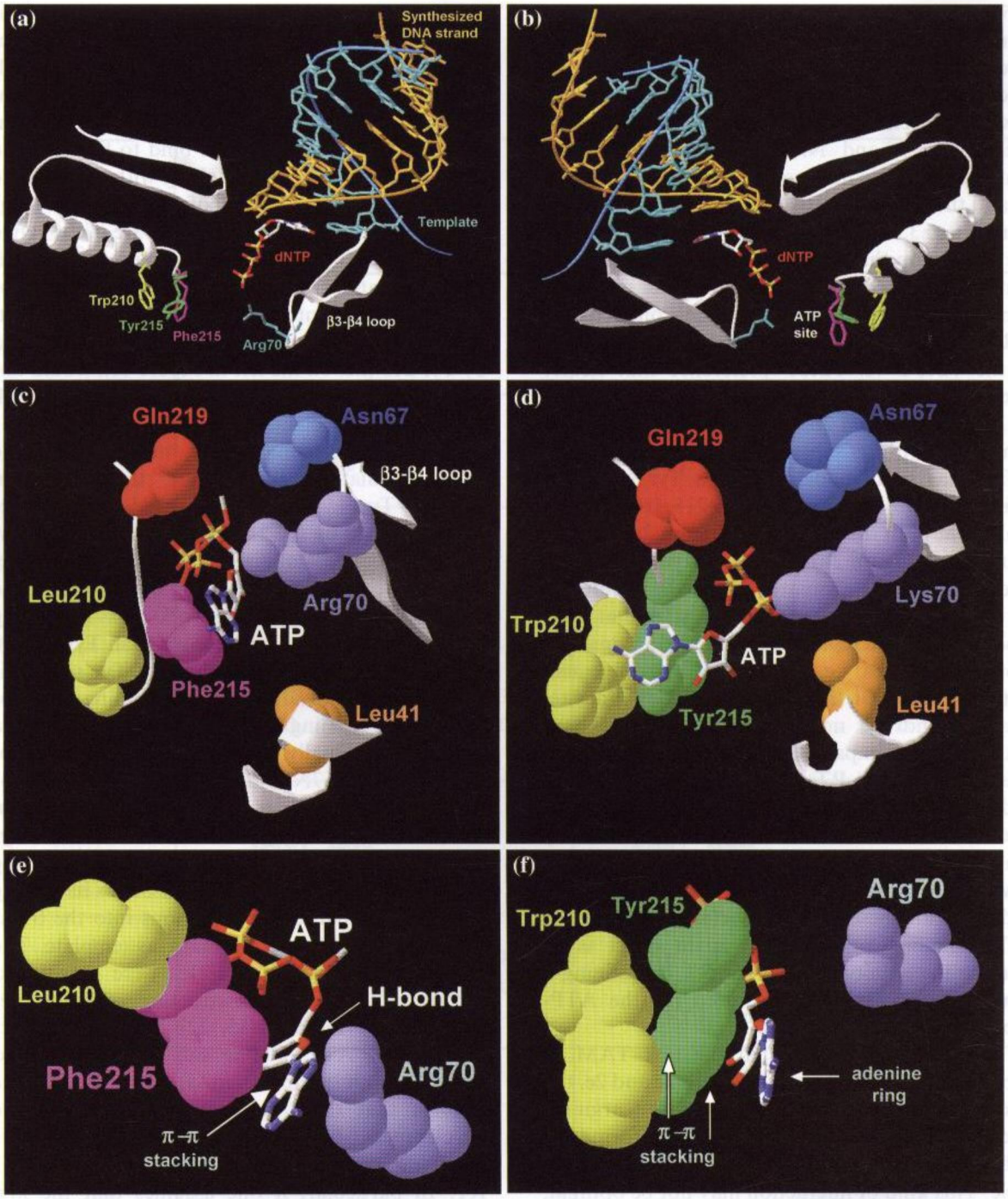

Figure 2. Models showing the HIV-1 RT with mutations T215F and T215Y. The ternary complex RT-template/primer-dTTP is shown in panel (a) (front view) and panel (b) (rear view) with mutations T215Y and 215F superimposed. The position of the $\beta3-\beta4$ loop in the finger subdomain, mutation K70R, and the ATP site of mutant RT are also described. The model was obtained with the wild-type HIV-1 RT ternary complex (PDB entry 1RTD) used as template. Panels (c) and (d): Overall structure of the ATP site determined by two distinct pathways of mutations involving either T215F or T215Y. The model was obtained from the sequence of two representative clinical isolates bearing either the T215F or T215Y pattern. Panel (e): potential interaction between ATP and mutant T215F/K70R shown in panel (c), based on molecular mechanics simulations. Panel (f): potential interaction between ATP and mutant T215Y/L210W shown in panel (d), based on molecular mechanics simulations.

Discussion

With the introduction of HIV-1 RT sequencing in clinical practice, it has been possible to constitute large databases of genotypic data and associated resistance patterns. Our own database, created in January 1997, allows the retrieval of sequences that meet specific criteria such as the occurrence and frequency of a particular mutation, the nature and frequency of the amino acid substitution at a given codon, and/or the rate of association of two resistance mutations [3, 16, 17]. As soon as in 1999, a cross-sectional study of 287 HIV-infected patients of this database demonstrated that T215Y/F was the most common point mutation (53%), consistent with the widespread use of AZT (about 40% of patients) in the studied population [3]. In the present study, conducted five years later, the prevalence of this mutation remains very high, as 46.8% of sequences displayed either F (7.04%) or Y (38.8%) instead of wildtype T at this position. This is indicative of an unexpected level of resistance despite the widespread use of suppressive anti-HIV combination therapies [16]. Indeed, recent studies in both Europe and the United States showed that more than 75% of HIV-1 infected patients have virus strains that are resistant to at least one NRTI [1, 16]. The relative proportion between the mutations at codon 215 was 72% for T215Y and 18% for T215F in the Marseille database (7312 genotypes), and 78% and 22%, respectively in the Stanford database (2284 genotypes). This distribution is in full agreement with the study of Chappey et al. [14] who observed a prevalence of 72% of T215Y and 21% of T215F in a population of 9645 genotypes.

In clinical practice, it is of primary importance to detect transitional mutations as well as genotypic mixtures because at the time of their detection, the virus still displays a sensitive phenotype [3, 19, 25]. This is the case for transitional 215 mutations, which are found in viruses fully susceptible to NRTI in phenotypic assays [14]. However, these mutants have higher ability than wild-type viruses to select 215Y/F under selective drug pressure [19]. The prevalence of such transitional 215 mutants in our database was estimated to 3.98% (291 sequences). A key feature of the sequencing methodology is that it allows the detection of the simultaneous presence of wild-type and mutant nucleotides at any position of the sequence. In the case of mutation 215, the situation is particularly complex because the ambiguities may concern one, two or even the three bases of the codon (Table 1). This property was already reported in the first cross-sectional study of HIV-1 RT mutational patterns, but at this time all theoretical codon possibilities were considered, corresponding to two main hypothetical amino acid mixtures TNYS and TFSI [3]. Since then, it was occasionally observed during the follow-up of genotyping testing that TNYS evolved towards Y, whereas TFSI evolved towards F. This interpretation was logical, as TNYS for instance can be resolved into wild-type T, transitional N and S, and finally mutant Y. A decisive outcome of the present study is that transitional mutants are almost not detected after cloning representative samples. Instead, only T, F or Y was found in the sequenced clones, indicating that in these mixtures, transitional mutants are not abundant enough (if ever present) to be detected, at least within the scale of ca. 20 clones per sample. This important result will simplify the interpretation of such electrophoregrams which were in the past considered as ambiguous data. Now, it can be reasonably considered that TNYS profile means T/Y, and TFSI means T/F, so that mutant viruses are likely to be already selected.

To understand why T215Y is more frequent in mutant viruses than T215F, it is necessary to determine in which genetic background these mutations appeared. In the present study, two different and specific patterns associated with a 215 mutation were observed. These patterns were: (i) 215Y with 41L and 210W, and (ii) 215F with 67N, 70R, and 219Q/E/N. These data are consistent with two recent studies [14, 15]. The mutual exclusion of mutations K70R and L210W has also been reported in the cross-sectional study published by our laboratory in 1999 [3] and has been further documented in a specific study [13]. Since K70R generally precedes mutation 215 during the course of NRTI therapy [26], these data suggest that K70R preselects T215F instead of T215Y. Conversely, in absence of K70R, this is T215Y rather than T215F that is selected first, followed by L210W. This sequence of mutations explains why K70R and L210W are very rarely observed together in the same viral genome. Overall, these data suggested that the combination of mutations T215Y/L210W confers a higher level of resistance

to thymidine analogs than K70R/T215F. Interestingly, none of these combinations appeared to affect the binding of AZT to the dNTP binding site of HIV-1 RT [6]. In particular, molecular modeling approaches showed that the azido group linked to the 3′C of ribose has room in both wild-type and mutated RTs. Thus, the concept that TAMs (including mutation 215) acted by decreasing the affinity of mutant RT for AZT had to be abandoned. Instead, TAMs may confer increased rates of phosphorolytic removal of chain-terminating AZT from the 3′ terminus of the primer [27]. Originally, two potential phosphorolytic substrates have been considered: PPi and ATP [10]. However, the involvement of PPi in the excision reaction is not very likely, at least for three reasons: (i) the hydrolysis of PPi, which occurs immediately after the incorporation of the last nucleotide into the growing DNA chain, is the driving force of the DNA synthesis reaction. Indeed, this reaction is thermodynamically unfavored, and only the rapid hydrolysis of PPi by cellular pyrophosporylases can prevent the reverse reaction, i.e. DNA hydrolysis. In other words, the cellular PPi concentration during reverse transcription is not consistent with a role for PPi in the excision of AZT-MP, (ii) the excision reaction driven by PPi would regenerate AZT-TP, (iii) TAMs have no detectable effect on PPi binding to mutated RT [11]. In contrast, the candidature of ATP is more serious. The excision product is distinct from AZT-TTP. The cellular concentrations of ATP, even in infected cells, are consistent with an ATP-dependent phosphorolysis. Finally, TAMs may increase the affinity of RT for the ATP substrate by delineating an ATP binding site distinct from the dNTP binding pocket [11].

Now it remains to explain why ATP prefers T215Y/L210W rather than K70R/T215F. Hopefully, the way proteins recognize ATP has been the subject of in-depth structural studies [28–30]. ATP consists of three building units: adenine, which has a largely hydrophobic protein interface, phosphate, which interacts primarily with hydrophilic residues, and ribose, which is intermediate (Figure 2c and 2d). Consequently, the main problem to solve in order to create a binding pocket for ATP is to provide a complementary hydrophobic surface for adenine. Tyrosine and phenylalanine residues are particularly suited for this function, because their aromatic side chains can share π electrons with the adenine rings, forming a geometrically optimized π–π stacking interaction. However, the respective propensities of Tyr and Phe to interact with adenine are not similar: according to the calculations of Moodie et al. [30], based on the study of a non-homologous dataset of 18 high-resolution protein/nucleotide crystal structures, the propensity Π was 0.225 for Tyr, and 0.118 for Phe. Since Π represents the frequency of a given amino acid found in the ATP binding site of these proteins, one could conclude that Tyr is significantly more efficient than Phe for interacting with adenine. There are several biochemical reasons for this preference. For instance, the –OH group of Tyr can form a H-bond with ribose (generally with the –OH linked to the C′2 atom), whereas Phe cannot. Correspondingly, the propensity of Phe to interact with ribose is close to 0, compared to 0.18 for Tyr (see Table 8 in ref. [30]). Such a double interaction (π–π stacking + H-bond) can be observed in carbamate kinase (PDB entry 1E19), between Tyr244 and ADP [31]. However, the orientation of ATP in modeled RT mutants is not compatible with this kind of interaction, suggesting that other parameters may be involved (discussed below). On the other hand, phosphate predominantly hydrogen bonds with the main-chain NH groups and the side chains of Arg or Lys. Yet the preference is clearly for Arg rather than Lys, as shown by the calculation of their propensity to interact with phosphate [30]. These structural data, together with the molecular modeling study shown in Figure 2, strongly suggest that the ATP binding surface formed by T215Y/L210W is particularly functional for two reasons (Figure 2d): (i) the aromatic side chain of Trp (210W) strengthens the π–π stacking interaction of Tyr (215Y) with the adenine ring. This is important because Tyr has a slightly weaker stacking capability than Phe, due to polarization of the π electronic system toward the phenolic –OH [32]. This is not incompatible with the recorded preference of Tyr vs. Phe in ATP recognition, because a slightly weaker interaction may allow a greater flexibility of the side chain, which may be important to find an optimal fit, (ii) The ribofuranose cycle stacks onto the phenol ring, so that not only π–π but also CH–π stacking interactions are involved in ATP binding, forming a complex network of stabilizing bonds. The interaction of ATP with the K70R/T215F mutant

is less sophisticated (Figure 2e): ribose forms a H-bond with the Arg side chain, and the adenine ring interacts with Phe215 through a suboptimal, because nonparallel, π–π tacking. Each combination of mutations determines a specific geometry for the ATP binding site, as shown in Figure 2c and 2d. Both orientations are compatible with the ATP-dependent phosphorolysis reaction, although the Tyr215 mutant may be more fit than T215F [15]. It should be emphasized that the marked difference in the orientation of the aromatic side chains of Tyr215 and Phe 215 is based on previous structural studies of mutant RT [33], and not from molecular modeling simulations. Therefore, it can be reasonably hypothesized that a suboptimal stacking interaction between adenine and Phe215, compared to an optimal one with Tyr215/Trp210, is sufficient *per se* to explain the Tyr vs. Phe preference of 215 mutants. Moreover, the data of Table 1 do support the view that there is a steric constraint in the ATP site so that only T215F can occur after K70R, and L210W after T215Y. Further adjustments of the binding site may then explain the preference of M41L for T215Y, and D67N + K219Q/E/N for T215F (Table 1). In the wild-type RT complex, there is an electrostatic interaction between side chains D67 and K219 [20]. This bridge may impair the accessibility of the γ-phosphate of ATP to the excision site, so that either D67 or K219 has to be mutated to facilitate the ATP-dependent phosphorolysis process [11].

In conclusion, this study based on the analysis of a large number of RT sequences from two independent databases confirmed the existence of two main mutational patterns according to the nature of the aromatic substitution at codon 215. The cloning of ambiguous samples due to multiple mixtures at codon 215 has allowed to solve the theoretical formulas TNYS and TFSI into T/Y and T/F, respectively, so that for the clinician such samples should now be considered as mutated T215F or T215Y. Molecular modeling approaches in the background of known protein–ATP interactions have shed some light on the biochemical logic explaining the preferential selection of a combination of TAMs with T215Y/L210W rather than K70R/T215F. This study may provide a rational basis for the analysis of sequence data with a mutation at codon 215 and a biochemical reassessment of our understanding of the resistance mechanisms associated with this mutation.

References

1. Susman E., Many HIV patients carry mutated drug-resistant strains. Lancet 359: 49, 2002.
2. Larder B.A., Kellam P. and Kemp S.D., Zidovudine resistance predicted by direct detection of mutations in DNA from HIV-infected lymphocytes. AIDS 5: 137–144, 1991.
3. Yahi N., Tamalet C., Tourrès C., Tivoli N., Ariasi F., Volot F., Gastaut J.A., Gallais H., Moreau J. and Fantini J., Mutation patterns of the reverse transcriptase and protease genes in human immunodeficiency virus type 1-infected patients undergoing combination therapy: survey of 787 sequences. J. Clin. Microbiol. 37: 4099–4106, 1999.
4. Violin M., Cozzi-Lepri A., Velleca R., Vincenti A., D'Elia S., Cjiodo F., Ghinelli F., Bertoli A., d'Arminio Monforte A., Perno C.F., Moroni M. and Balotta C., Risk of failure in patients with 215 HIV-1 revertants starting their first thymidine analog-containing highly active antiretroviral therapy. AIDS 18: 227–235, 2004.
5. Yerly S., Rakik A., De Loes S.K., Hirschel B., Descamps D., Brun-Vézinet F. and Perrin L., Switch to unusual amino acids at codon 215 the human immunodeficiency virus type 1 reverse transcriptase gene in seroconvertors infected with zidovudine-resistant variants. J. Virol. 72: 3520–3523, 1998.
6. Lacey S.F. and Larder B., Mutagenic studies of codons 74 and 215 of the human immunodeficiency virus type 1 reverse transcriptase which are significant in nucleoside analog resistance. J. Virol. 68: 3421–3424, 1994.
7. Larder B., Interactions between drug resistance mutations in human immunodeficiency virus type 1 reverse transcriptase. J. Gen. Virol. 75: 951–957, 1994.
8. Gao H.Q., Boyer P.L., Sarafianos S.G., Arnold E. and Hughes S.H., The role of steric hindrance in 3TC resistance of human immunodeficiency virus type-1 reverse transcriptase. J. Mol. Biol. 300: 403–418, 2000.
9. Lacey S.F., Reardon J.E., Furfine E.S., Kunkel T.A., Bebeneck K., Eckert K.A., Kemp S.D. and Larder B.A., Biochemical studies on the reverse transcriptase and RNase H activities from human immunodeficiency virus strains resistant to 3′-azido-3′deoxythymidine. J. Biol. Chem. 267: 15789–15794, 1992.
10. Sluis-Cremer N., Arion D. and Parniak M.A., Molecular mechanisms of HIV-1 resistance to nucleoside reverse transcriptase inhibitors (NRTIs). Cell. Mol. Life Sci. 57: 1408–1422, 2000.
11. Boyer P.L., Sarafianos S.G., Arnold E. and Hugues S.H., Selective excision of AZTMP by drug-resistant human immunodeficiency virus reverse transcriptase. J. Virol. 75: 4832–4842, 2001.
12. Meyer P.R., Matsuura S.E., Tolun A.A., Pfeifer I., So A.G., Mellors J.W. and Scott W.A., Effects of specific zidovudine resistance mutations and substrate structure on nucleotide-dependent primer unblocking by human immunodeficiency virus type 1 reverse transcriptase. Antimicrob. Agents Chemother. 46: 1540–1545, 2002.
13. Yahi N., Tamalet C., Tourrès C., Tivoli N. and Fantini J., Mutation L210W of HIV-1 reverse transcriptase in patients

receiving combination therapy. Incidence, association with other mutations and effects of the structure of mutated reverse transcriptase. J. Biomed. Sci. 7: 507–513, 2000.
14. Chappey C., Wrin T., Deeks S. and Petropoulos C.J., Evolution of amino acid 215 in HIV-1 reverse transcriptase in response to intermittent drug selection. XII International HIV Drug Resistance Workshop, Los Cabos, Mexico, Abstract 32, June 10–14, 2003.
15. Hu Z.X., Reid P., Lu J. and Kuritzkes D.R., Fitness of T215Y vs T215F mutants in HIV-1 RT: comparison of specific thymidine analogue-resistance mutation patterns. Antiviral Ther. 59: S68, 2004.
16. Tamalet C., Fantini J., Tourrès C. and Yahi N., Resistance of HIV-1 to multiple antiretroviral drugs in France: a 6-year survey (1997–2002) based on an analysis of over 7000 genotypes. AIDS 17: 2383–2388, 2003.
17. Yahi N., Fantini J., Tourrès C., Tivoli N., Koch N. and Tamalet C., Use of drug resistance sequence data for the systematic detection of non-B human immunodeficiency virus type 1 (HIV-1) subtypes: how to create a sentinel site for monitoring the genetic diversity of HIV-1 at a country scale. J. Infect. Dis. 183: 1311–1317, 2001.
18. Shafer R.W., Dupnik K.M., Winters M.A. and Eshleman S.H., A guide to HIV-1 reverse transcriptase and protease sequencing for drug resistance studies. Human Retroviruses and AIDS, Theoretical Biology and Biophysics. Los Alamos National Laboratories, 2001.
19. Shafer R.W., Genotypic testing for human immunodeficiency virus type 1 drug resistance. Clin. Microbiol. Rev. 15: 247–277, 2002.
20. Huang H., Chopra R., Verdine G.L. and Harrison S.C., Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase. Implications for drug resistance. Science 282: 1669–1675, 1998.
21. Schwede T., Kopp J., Guex N. and Peitsch M.C., SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31: 3381–3385, 2003.
22. Ren J., Esnouf R.M., Hopkins A.L., Jones E.Y., Kirby I., Keeling J., Ross C.K., Larder B.A., Stuart D.I. and Stammers D.K., 3′Azido-3′deoxythymidine drug resistance mutations in HIV-1 reverse transcriptase can induce long range conformational changes. Proc. Natl. Acad. Sci. USA 95: 9518–9523, 1998.
23. Guex N. and Peitsch M.C., SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18: 2714–2723, 1997.
24. Nishio M., Umezawa Y., Hirota M. and Takeuchi Y., The CH/π interaction: significance in molecular recognition. Tetrahedron 51: 8665–8701, 1995.
25. Tamalet C., Yahi N., Tourrès C., Colson P., Quinson A.M., Poizot-Martin I., Dhiver C. and Fantini J., Multidrug resistance genotypes (insertions in the beta3-beta4 finger subdomain and MDR mutations) of HIV-1 reverse transcriptase from extensively treated patients: incidence and association with other resistance mutations. Virology 270: 310–316, 2000.
26. Boucher C.A., Sullivan E.O., Mulder J.W., Ramautarsing C., Kellam P., Darby G., Lange J.M., Goudsmit J. and Larder B.A., Ordered appearance of zidovudine resistance mutations furing treatment of 18 human immunodeficiency virus-positive subjects. J. Infect. Dis. 165: 105–110, 1992.
27. Arion D., Kaushik N., McCormick S., Borkow G. and Parniak M.A., Phenotypic mechanism of HIV-1 resistance to 3′-azido-3′-deoxythymidine (AZT): increased polymerization processivity and enhanced sensitivity to pyrophosphate of the mutant viral reverse transcriptase. Biochemistry 37: 15908–15917, 1998.
28. Babor M., Sobolev V. and Edelman M., Conserved positions for ribose recognition: importance of water bridging interactions among ATP, ADP and FAD-protein complexes. J. Mol. Biol. 323: 523–532, 2002.
29. Mao L., Wang Y., Liu Y. and Hu X., Molecular determinants for ATP-binding in proteins: a data mining and quantum chemical analysis. J. Mol. Biol. 336: 787–807, 2004.
30. Moodie S.L., Mitchell J.B.O. and Thornton J.M., Protein recognition of adenylate: an example of a fuzzy recognition template. J. Mol. Biol. 263: 486–500, 1996.
31. Ramon-Maiques S., Marina A., Uriarte M., Fita I. and Rubio V., The 1.5 Å resolution crystal structure of the carbamate kinase-like carbamoyl phosphate synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus*, bound to ADP, confirms that this thermostable enzyme is a carbamate kinase, and provides insight into substrate binding and stability in carbamate kinases. J. Mol. Biol. 299: 463–476, 2000.
32. Kaushik N., Singh K., Alluru K.I. and Modak M.J., Tyrosine 222, a member of the YXDD motif of MuLV RT, is catalytically essential and is a major component of the fidelity center. Biochemistry 38: 2617–2627, 1999.
33. Chamberlain P.P., Ren J., Nichols C.E., Douglas L., Lennerstrand J., Larder B.A., Stuart D.L. and Stammers D.K., Structures of zidovudine- or lamivudine-resistant human immunodeficiency virus type 1 reverse transcriptases containing mutations at codons 41, 184, and 215. J. Virol. 76: 10015–10019, 2002.