Docking Analysis and Resistance Evaluation of Clinically Relevant Mutations Associated with the HIV-1 Non-nucleoside Reverse Transcriptase Inhibitors Nevirapine, Efavirenz and Etravirine

Stefano Alcaro,[a] Claudia Alteri,[b] Anna Artese,*[a] Francesca Ceccherini-Silberstein,[b] Giosuè Costa,[a] Francesco Ortuso,[a] Ada Bertoli,[b] Federica Forbici,[c] Maria Mercedes Santoro,[b] Lucia Parrotta,[a] Philippe Flandre,[d] Bernard Masquelier,[e] Diane Descamps,[f] Vincent Calvez,[d] Anne-Genevieve Marcelin,[d] Carlo Federico Perno,[b] Tobias Sing,[g] and Valentina Svicher[b]

An integrated computational and statistical approach was used to determine the association of non-nucleoside reverse transcriptase inhibitors (NNRTIs) nevirapine, efavirenz and etravirine with resistance mutations that cause therapeutic failure and their impact on NNRTI resistance. Mutations detected for nevirapine virological failure with a prevalence greater than 10% in the used patient set were: K103N, Y181C, G190A, and K101E. A support vector regression model, based on matched genotypic/phenotypic data ($n = 850$), showed that among 6365 analyzed mutations, K103N, Y181C and G190A have the first, third, and sixth greatest significance for nevirapine resistance, respectively. The most common indicator of treatment failure for efavirenz was K103N mutation present in 56.7% of the patients where the drug failed, followed by V108I, L100I, and G190A. For efavirenz resistance, K103N, G190, and L100I have the first, fourth, and eighth greatest significance, respectively, as determined in support vector regression model. No positive interactions were observed among nevirapine resistance mutations, while a more complex situation was observed with treatment failure of efavirenz and etravirine, characterized by the accumulation of multiple mutations. Docking simulations and free energy analysis based on docking scores of mutated human immunodeficiency virus (HIV) RT complexes were used to evaluate the influence of selected mutations on drug recognition. Results from support vector regression were confirmed by docking analysis. In particular, for nevirapine and efavirenz, a single mutation K103N was associated with the most unfavorable energetic profile compared to the wild-type sequence. This is in line with recent clinical data reporting that diarylpyrimidine etravirine, a very potent third generation drug effective against a wide range of drug-resistant HIV-1 variants, shows increased affinity towards K103N/S mutants due to its high conformational flexibility.

Introduction

Acquired immunodeficiency syndrome (AIDS) is caused by human immunodeficiency virus (HIV). More than 33 million people are currently infected with HIV, which is a retrovirus. HIV reverse transcriptase (RT) is a crucial enzyme responsible for the synthesis of double-stranded DNA from the single-stranded RNA genome. RT is an asymmetric heterodimer composed of two subunits, p66 and p51. Due to its essential role in HIV replication, RT is a major target in antiretroviral therapy.[1] More than half of the currently approved drugs for the treatment of HIV-1 infection are RT inhibitors. They can be classified into two main groups: nucleoside RT inhibitors (NRTIs) and non-nucleoside RT inhibitors (NNRTIs).[1–4] NRTIs compete with the natural deoxynucleoside triphosphate (dNTP) substrate for RT binding, and thus inhibit HIV replication by terminating chain elongation. NNRTIs bind to an allosteric site in HIV-1 RT and indirectly interfere with dNTP incorporation, probably by altering the conformation of residues located in the active site and inducing structural distortion of RT. However, drug-resistant mutants have emerged because of pharmacological selection pressure and limit the effectiveness of these

[a] *Prof. S. Alcaro, Dr. A. Artese, Dr. G. Costa, Dr. F. Ortuso, Dr. L. Parrotta*
Dipartimento di Scienze Farmacobiologiche
Università degli Studi "Magna Græcia" di Catanzaro
Complesso Ninì Barbieri, 88021 Roccelletta di Borgia (CZ) (Italy)
E-mail: artese@unicz.it

[b] *Dr. C. Alteri, Dr. F. Ceccherini-Silberstein, A. Bertoli, M. M. Santoro,*
Prof. C. F. Perno, Dr. V. Svicher
Dipartimento di Medicina Sperimentale e Biochimica
Università "Tor Vergata", Via Montpellier 1, 00133, Roma (Italy)

[c] *Dr. F. Forbici*
INMI "L. Spallanzani", Via Portuense 292, 00149 (Italy)

[d] *Dr. P. Flandre, Prof. V. Calvez, Prof. A.-G. Marcelin*
Department of Virology, AP-HP, Pitié-Salpêtrière Hospital, INSERM U943,
Université Pierre et Marie Curie
56 Bd Vincent Auriol, 75625 Paris (France)

[e] *Dr. B. Masquelier*
Laboratoire de Virologie CHU de Bordeaux and UMR5234 Université Bordeaux Segalen, 146 rue Léo-Saignat, 33076 Bordeaux (France)

[f] *Prof. D. Descamps*
Department of Virology, AP-HP, Groupe Hospitalier Bichat-Claude Bernard, Université Denis Diderot-Paris 7, Paris (France)

[g] *Dr. T. Sing*
Max–Planck Institute for Informatics
Building E1.4, Stuhlsatzenhausweg 85, 66123 Saarbrücken (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cmdc.201100362.

highly potent inhibitors. In vitro experiments determined that resistance to most NNRTIs is caused by multiple mutations in or near the NNRTI binding pocket of HIV-1 RT, including K103, Y181, and Y188.[5]

Several structural studies have demonstrated that NNRTI binding affects the conformation of the catalytic carboxylates able to bind metal cofactors (Mg^{2+} cations) causing a shift in the position of the primer grip[6,7] and a decrease of thumb mobility in the p66 subunit.[8–10] These inhibitors are stabilized in the binding pocket by 1) stacking interactions between the aromatic rings of the drug and side chains of Y181, Y188, W229 and Y318; 2) electrostatic forces, particularly significant for K101, K103 and the p51 residue, E138; 3) van der Waals interactions with L100, V106, V179, Y181, G190, W229, L234 and Y318; and 4) hydrogen bonds between the NNRTI and the RT main chain.

The NNRTI binding pocket is an elastic pocket situated between the β6–β10–β9 and β12–β13–β14 sheets in the palm subdomain of the p66 subunit, approximately 10 Å away from the catalytic site. Its conformation depends on the size, specific chemical structure, and binding mode of the drug. Structural studies showed that nevirapine (NVP) and other first-generation NNRTIs adopt a "two-ring" binding mode[8,11] and are characterized by a low genetic barrier to the development of resistance. NVP binding is stabilized by RT amino acids Y181 and Y188 through stacking interactions between their aromatic side chains and the pyridine groups of the inhibitor. Not surprisingly, mutations at position 181 (e.g., Y181C or Y181I) or position 188 (e.g., Y188C, Y188I, Y188L or Y188H) are responsible for conferring high-level resistance to the inhibitor.

Second-generation inhibitors such as efavirenz (EFV) were designed to reduce the contribution of stacking interactions in their stabilization of the NNRTI binding pocket.[12–14] In particular, upon binding, EFV makes direct or water-mediated hydrogen bonds with the protein backbone of K101 and K103.[13] The most common mutation found in patients failing to respond to EFV is K103N, but there are other frequently observed substitutions, such as V108I, P225H or L100I, K101E, K101Q, Y188H, Y188L, G190S, G190A and G190E.[15] Molecular modeling studies using the crystal structure of the unbound HIV-1 K103N mutant RT have shown that a hydrogen bond is formed between the hydroxy group of Y188 and the N103 NH group. Such an interaction is responsible for the stabilization of the apo-RT conformation and for an increase in the energy barrier for NNRTI binding.[16]

In order to avoid cross-reactivity with NVP and EFV and to obtain a better resistance profile with an increased genetic barrier, third-generation NNRTIs have been developed. Etravirine (ETR), formerly TMC-125, is a diarylpyrimidine (DAPY) characterized by a high torsional flexibility.[17] Structural studies have shown the ability of ETR, as well as other DAPY analogues, to assume distinct conformational variants associated with improved potency against wild-type HIV-1 RT and a wide range of drug-resistant mutants.[18] ETR was found to be effective against viruses containing combinations of common drug resistance mutations, such as K101E/K103N or K103N/Y181C.[17] In ETR-resistant isolates, several substitutions were observed, such as L100I, V179I/F, Y181C, G190E, M230L and Y318F. Clinical studies have demonstrated the accumulation of at least two drug resistance mutations associated with high-level resistance to ETR. Structural analyses have indicated that in the presence of single mutations, ETR reduces its alternative binding configurations restricting the potential for accommodating additional RT mutations. These observations support the finding that ETR resistance requires multiple mutations.[18,19]

With the aim to optimize DAPY compounds, rilpivirine (RPV) was designed with the focus of establishing interactions with the conserved amino acids in the NNRTI binding pocket, in particular W229.[20] In vitro, RPV shows a resistance profile and genetic barrier to the development of resistance comparable to those of ETR.[21] High-resolution crystal structures of RPV in complex with L100I/K103N HIV-1 RT show the inhibitor reaching deeper into the NNRTI binding pocket toward amino acid W229 compared with ETR.[22] Both DAPY compounds exhibit a similar flexibility in adapting to the altered binding site caused by resistance mutations.[23]

Lersivirine (LSV) is a next-generation NNRTI belonging to the pyrazole family characterized by an in vitro resistance profile comparable to that of ETR and RPV.[24] LSV has shown good antiviral activity and tolerability in a proof of concept monotherapy study with once- and twice-daily administration to treatment-naïve patients for 7 days.[25] LSV has a novel mode of binding to RT resulting in largely non-overlapping resistance with existing agents from the class. It is selective against a range of human targets, has broad-spectrum activity against primary and lab-derived NNRTI-sensitive or -resistant strains.[26] Furthermore, it is being currently being studied in a phase IIB dose finding study, with once- and twice-daily administration.[25]

Though several studies have contributed to our current knowledge of drug-related RT variants,[27] the factors contributing to NNRTI resistance are numerous and have not yet been completely elucidated. Therefore, rationalizing the molecular recognition of some of the currently approved NNRTIs in the presence of known RT mutations associated with resistance is a critical step for the development of superior inhibitors. Here, we analyzed the clinical occurrence of RT mutations associated with drug resistance to three approved NNRTIs using computational analysis of HIV-1 RT crystallographic models. Such an analysis, based on the comparison of structural data to drug resistance profiles, has allowed us to understand the influence of some drug resistance-associated mutations in the recognition of NVP, EFV and ETR by RT.

Results and Discussion

In this study, we focused our attention to RT mutations associated with resistance to the NNRTIs currently in clinical use. Initially, we determined the prevalence and patterns of RT mutations detected in a large group of patients ($n = 1265$) failing for the first time a treatment regime contain an NNRTI. The prevalence of these resistance-associated mutations in subsets of patients failing for the first time a treatment course containing NVP ($n = 568$), EFV ($n = 647$), or ETR ($n = 50$) is reported in

Table 1. Prevalence of NNRTI resistance mutations in patients failing treatment regimes with NVP ($n = 568$), EFV ($n = 647$) or ETR ($n = 50$) for the first time.

NVP[b]			EFV[b]			ETR[b]		
Mutation[a]	n	[%]	Mutation[a]	n	[%]	Mutation[a]	n	[%]
K103N	213	(37.5)	K103N	362	(56.0)	Y181C	17	(34.0)
Y181C	195	(34.3)	V108I	92	(14.2)	G190A	10	(20.0)
G190A	141	(24.8)	L100I	80	(12.4)	K103N	9	(18.0)
K101E	57	(10.0)	G190A	77	(11.9)	L100I	6	(12.0)
V106A	45	(7.9)	P225H	64	(9.9)	V108I	6	(12.0)
V108I	42	(7.4)	K238T	42	(6.5)	K101E	5	(10.0)
K238T	22	(3.9)	Y181C	42	(6.5)	K101I	4	(8.0)
Y188L	20	(3.5)	Y188L	37	(5.7)	G190S	3	(6.0)
G190S	16	(2.8)	G190S	33	(5.1)	K103S	2	(4.0)
Y181I	15	(2.6)	K101P	26	(4.0)	V179F	2	(4.0)
M230L	7	(1.2)	K103S	13	(2.0)	Y181I	2	(4.0)
K103S	6	(1.1)	G190E	11	(1.7)	Y181V	2	(4.0)
Y188C	4	(0.7)	V106A	10	(1.5)	Y188L	2	(4.0)
Y188H	3	(0.5)	V106M	9	(1.4)	M230L	2	(4.0)
K101P	3	(0.5)	M230L	9	(1.4)	V106A	1	(2.0)
Y181V	2	(0.4)	Y188H	8	(1.2)	V106M	1	(2.0)
L100I	2	(0.4)	Y188C	2	(0.3)	E138K	1	(2.0)
V106M	2	(0.4)	E138K	2	(0.3)	–	–	–
E138K	2	(0.4)	V179F	1	(0.2)	–	–	–
–	–	–	Y181V	1	(0.2)	–	–	–

[a] Mutation data taken from References [29] and [30]; only mutations detected in our dataset are shown.
[b] The number of patients in a given mutation subset is give as an absolute value (n) and as a proportion (%) of the total number of patients failing the treatment regime. Nevirapine (NVP), total $n = 568$, efavirenz (EFV), total $n = 647$; etravirine (ETR), total $n = 50$.

Table 1. Mutations detected in more than 10 % of the patient subset for NVP were: K103N (37.5 %), Y181C (34.3 %), G190A (24.8 %), and K101E (10.0 %). Similarly, the most common RT mutation found in patients failing a treatment regime that included EFV was K103N, present in 56.7 % of the patient subset, followed by V108I (14.2 %), L100I (12.4 %), and G190A (11.9 %). Conversely, Y181C was present only in 6.5 % of patients in the EFV subset. Of the 50 patients failing a treatment regimen containing ETR, NNRTI resistance mutations were detected in 39 of them, with Y181C being the most common mutation followed by G190A (Table 1). In contrast, we observed only a low prevalence of Y181I and Y181V; this is consistent with a previous study showing that these two mutations, although conferring high-level resistance to ETR, rarely occur in patients where ETR failed (prevalence < 1 %).[28]

Interaction among NNRTI resistance mutations

In order to provide a comprehensive definition of mutational patterns underlying NNRTI in vivo virological failure, defined as two consecutive determinations of plasma HIV-RNA < 50 copies mL^{-1}, we also evaluated the association of NNRTI resistance mutations in subsets of patients failing treatment regimes containing NVP ($n = 568$), EFV ($n = 647$) and ETR ($n = 50$). In particular, we calculated the phi coefficient and its statistical significance for each pair of mutations (Table 2). A positive and statistically significant correlation between mutations at two specific positions ($0 < $ phi $ < 1$, $P < 0.05$) indicates that under drug pressure, these two positions mutate in a correlated manner in order to confer an advantage in term of drug resist-

ance or viral fitness. This indicates that the co-occurrence of mutations does not take place randomly.

No positive correlations among NNRTI resistance mutations were observed in patients where a NVP-containing regimen failed (only exceptions are K101E and G190A that strongly correlated with each other; phi = 0.35, $P = 1.1$ E−10, Table 2). This result is fully consistent with a recent study showing the presence of very weak interactions among NVP resistance mutations.[31] Conversely, we observed strong negative correlations for Y181C with Y188L (phi = −0.18, $P = 6.5$ E−4), and for K103N with either K101E (phi = 0.23, $P = 1.4$ E−5), G190A (phi = 0.25, $P = 4.1$ E−6), or G190S (phi = 0.15, $P = 1.6$ E−2). These results indicate that the virological failure of NVP, a first-generation NNRTI, is generally mediated by a single mutation confirming the low HIV genetic barrier to this drug.

A more complex situation was observed for EFV, a second-generation NNRTI, characterized by a higher genetic barrier to resistance compared with NVP. Indeed, we observed strong positive correlations for K103N with either L100I (phi = 0.22, $P = 2.8$ E−6) or P225H (phi = 0.15, $P = 1.3$ E−2), for Y181C with G190A (phi = 0.24, $P = 1.3$ E−4), and for K101E with either G190A (phi = 0.21, $P = 8.8$ E−4) or G190S (phi = 0.31, $P = 5.3$ E−6). The correlation between K101E and G190S is consistent with a recent study showing that the presence of K101E + G190S correlated with increased EFV resistance when compared with either of the single mutants and improved viral fitness under EFV pressure.[32] The negative correlation of K103N with K101E and mutations at RT position 190, observed in patients failing treatment with NVP, was also confirmed in EFV virological failure (Table 2).

Since pair-wise analysis suggested the existence of specific pathways of mutations underlying EFV resistance, we performed an average linkage hierarchical agglomerative cluster analysis to investigate this hypothesis in more detail. The topology of the dendrogram, shown in Figure 1, suggests the existence of at least two distinct clusters involving K103N along with L100I and P225H (bootstrap = 0.98) and Y181C along with G190A/S and K101E (bootstrap = 0.96). These results support our hypothesis that the virological failure of EFV is mediated by the accumulation of multiple mutations. This data are consistent with a previous study showing that the point mutation K103N was most often the first NNRTI resistance mutation observed in patients for whom EFV combination therapy failed, followed by the eventual observation of double- and even

Table 2. Significantly correlated pairs of NNRTI mutations in patients failing a treatment regime containing NVP ($n = 568$), EFV ($n = 647$), or ETR ($n = 50$) for the first time.
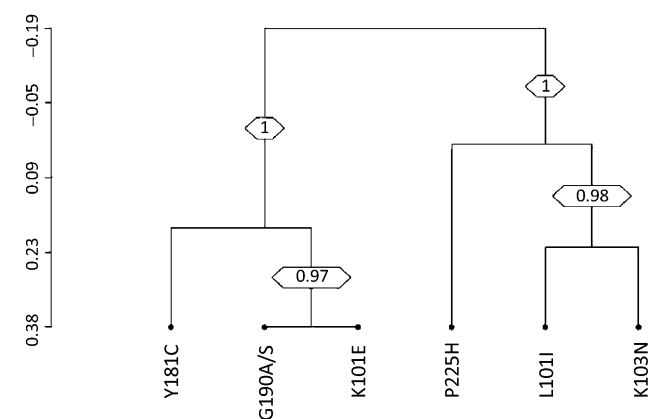
Mutation[a]	Frequency[b]		Correlated mutation[a]	Frequency[b]		Co-variation frequency[b]		Phi[c]	P value[d]
	n	[%]		*n*	[%]	*n*	[%]		
Nevirapine (NVP)									
K101E	57	10.0	G190A	141	24.8	42	7.4	0.35	1.11 E−10
K103N	213	37.5	G190A	141	24.8	39	6.8	−0.25	4.08 E−06
			G190S	16	2.8	1	1.7	−0.15	1.62 E−02
			K101E	57	10.0	9	1.6	−0.23	1.43 E−05
Y181C	195	34.3	Y188L	20	3.5	0	0.0	−0.18	6.46 E−04
Efavirenz (EFV)									
K101E	44	6.8	G190A	77	11.9	18	2.8	0.21	8.76 E−04
			G190S	33	5.1	14	2.7	0.31	5.35 E−06
G190A	77	11.9	Y181C	42	6.5	19	2.9	0.24	1.31 E−04
K103N	362	52.0	P225H	64	9.9	60	9.2	0.15	1.29 E−02
			L100I	80	12.4	79	12.2	0.22	2.77 E−06
			G190A	77	11.9	47	7.3	−0.20	7.63 E−04
			G190E	11	1.7	1	0.1	−0.26	5.14 E−05
			G190S	33	5.1	7	1.1	−0.40	5.65 E−12
			K101E	44	6.8	12	1.8	−0.42	3.60 E−13
			Y188L	37	5.7	18	2.8	−0.23	2.80 E−04
L100I	80	12.4	K103N	362	52.0	79	12.2	0.22	2.77 E−06
			V108I	92	14.2	5	0.8	−0.16	4.46 E−03
Etravirine (ETR)									
G190S	3	6.0	K101E	5	10.0	2	66.7	0.48	2.00 E−02
			Y181C	17	34.0	3	100.0	0.35	4.00 E−02
L100I	6	12.0	K103N	9	18.0	4	66.7	0.58	1.00 E−03

[a] Mutation data taken from References [29] and [30]. [b] The frequency of patients failing a treatment regime containing NVP ($n = 568$), EFV ($n = 647$), or ETR ($n = 50$) for the first time. [c] Positive and negative correlations with phi > 0.15 and phi $< −0.15$, respectively, are shown. [d] All *P* values for co-variation were significant at a false discovery rate of 0.01 following correction for multiple-hypothesis testing (Benjamini–Hochberg method).

tients on a treatment regime that includes ETR is still limited. However, new patterns of NNRTI resistance mutations were observed. In particular, the mutation G190S strongly correlated with both K101E and Y181C (phi = 0.48 and 0.35, $P = 0.02$ and 0.04, respectively). In addition, ETR resistance mutation L100I and the NVP/EFV resistance mutation K103N strongly correlated each other (phi = 0.58, $P = 0.001$). In all, the four patients carrying both the NNRTI resistance mutations, K103N was already present at the baseline, suggesting that this mutation could help the selection of L100I. This is the first study analyzing the patterns of mutations associated with ETR resistance. Our results confirm the requirement of multiple resistance mutations to overcome the pressure imposed by ETR, a recently approved, third-generation NNRTI.

Interaction among NNRTI and NRTI resistance mutations

To assess the co-variation of NNRTI- and NRTI-selected resistance mutations, we performed an analysis on RT sequences from 697 patients failing a treatment regimen containing an NNRTI. We identified 16 positive and 5 negative statistically significant correlations between NNRTI and NRTI mutations. The strongest positive correlations involved the NNRTI mutation L100I and the NRTI mutations L74V, K65R and T69I (phi > 0.12, $P \leq 0.001$). Another notable positive correlation exists between P225H and M184V (phi = 0.11, $P \leq 0.001$). The EFV resistance mutation P225H was also strongly negatively correlated with the thymidine analogue mutations M41L, L210W and T215Y (phi $\leq −0.13$, $P < 0.0001$). The full results of this analysis are available in Table S1 in the Supporting Information.

Impact of mutations on resistance to nevirapine and efavirenz

To investigate the contribution of mutations to viral resistance to NVP and EFV, we analyzed an independent data set of 850 matched genotype–phenotype pairs for each NNRTI using a feature ranking system based on support vector regression (SVR). This multivariate procedure allows the quantification of the impact of a mutation on resistance to a specific drug relative to other mutations. Among the 6365 mutations analyzed, our model showed that the major NNRTI resistance muta-

Figure 1. Dendrogram obtained from average linkage hierarchical agglomerative clustering showing clusters of NNRTI resistance mutations in 647 patients where EFV treatment failed. The length of branches reflects distances between mutations in the original distance matrix. Bootstrap values indicating the significance of clusters are reported in the boxes.

triple-mutant viruses with higher levels of drug resistance than the K103N single-mutant virus.[15]

Co-variation analysis was also performed on the data obtained from the ETR patient subset ($n = 50$). Due to the recent introduction of this drug into clinical use, the number of pa-

tions[30] were among the 25 mutations determined to have the greatest weight on resistance to NVP and/or EFV. This result indicates that our model might accurately capture the current knowledge regarding resistance to NNRTIs. The ranking of mutations by SVR provided evidence for the different contributions of a specific mutation to NVP and EFV resistance (Table 3).

1) *Contribution to NVP resistance.* Our model identified the top eight mutations with a Z-score[33] greater than 0.3 directly conferring resistance to NVP (Table 3). In particular, the strongest contribution was observed for K103N (Z-score = 1.1882), followed by Y188L (Z-score = 1.0052), and Y181C (Z-score = 0.7571). Both of the NNRTI resistance mutations at position 190

Table 3. Z-score[33] resistance data calculated for the reverse transcriptase (RT) complexes in the presence of nevirapine (NVP) and efavirenz (EFV).

NVP		EFV	
Model	Z-score	Model	Z-score
WT	0.0000	WT	0.0000
K103N	1.1882	K103N	1.1247
Y188L	1.0052	Y188L	0.9028
Y181C	0.7571	G190S	0.6776
G190S	0.6131	G190A	0.6319
V106A	0.5624	K103S	0.5452
G190A	0.5440	K101E	0.3832
K103S	0.5283	K101P	0.3457
K101P	0.3748	L100I	0.3399
–	–	Y181C	0.3171
–	–	V179D	0.3006

(G190S and G190A) were included in the top eight mutations contributing to NVP resistance ranking in the 4th (G190S: Z-score = 0.5440) and 6th position (G190A: Z-score = 0.6131). Finally, we included the I135T polymorphism, even thought the associated Z-score was less than 0.3, as feedback for the prediction quality (Z-score = 0.2351).

2) *Contribution to EFV resistance.* Our model identified the top ten mutations with a Z-score greater than 0.3 (Table 3). Similar to NVP, the mutation that makes the greatest contribution to EFV resistance was K103N (Z-score = 1.1247) and Y188C (Z-score = 0.9028), followed by G190S (Z-score = 0.6776) and G190A (Z-score = 0.6319). Y181C, showing the third most significant contribution to NVP resistance in the SVR model, was ranked in the 9th position in the context of EFV treatment failure. Finally, K101E and L100I showed a Z-score of greater than 0.3 only in the context of viral resistance to EFV and not NVP. Again, to check the prediction quality, we included the I135T mutation in our analysis (Z-score = 0.1390).

RT–NNRTIs complexes structural analysis

Z-score resistance data provided us with the rationale to evaluate the influence of the common mutations among those selected with respect to RT molecular recognition of NVP and EFV. Our computational approach was subsequently applied to ETR complexes, analyzing the same substitutions identified.

ETR correlating Z-score values were not available due to the low statistical significance of therapeutic data, but recent literature observations supported our findings.[28]

In order to analyze NVP–RT, EFV–RT and ETR–RT complexes, crystal structures deposited in the Protein Data Bank (PDB)[34] with the codes 1VRT[11] (resolution 2.20 Å), 1IKW[35] (resolution 3.00 Å) and 1SV5[18] (resolution 2.90 Å) were selected. Structural studies comparing a series of NNRTIs with widely variant inhibitory potencies against wild-type RT showed significant conformational differences in the NNRTI binding pocket.[36] In particular, Y181 and Y188 are known to contribute strongly to the stabilization of the drug–enzyme complex through hydrophobic interactions so that a single mutation at these residues gives a drastic reduction in inhibitor binding affinity due to a loss of aromatic ring stacking between the first-generation inhibitors such as NVP and the two tyrosine residues of the NNRTI binding site.[14]

By contrast, amino acids K101 and K103, located at the opening of the NNRTI binding pocket, play crucial roles in hydrophilic interactions with the inhibitor. Therefore, substitutions K103N and K101E cause strong resistance to multiple NNRTIs, including the first-generation drug NVP.[37,38] Analyzing the crystallographic structures of the wild-type and K103N mutant RTs in complex with EFV,[35] the K103N mutation was found to have minimal influence on the bound conformation of this inhibitor. However, several studies have showed that in the absence of an NNRTI, a hydrogen bond forms between Y188 and N103 closing the entrance to the binding pocket, thereby efficiently reducing its access to multiple NNRTIs[13,35,39] and stabilizing the unbound enzyme state.[39,40] A mutation less commonly observed at position 103 included K103S, which was correlated to a reduction in viral susceptibility to first- and second-generation NNRTIs. On the other hand, mutations at this position cause no significant resistance to ETR, since this inhibitor was designed to interact with this residue; in particular, ETR displays enhanced binding to the K103N mutant RT.[41,16]

Furthermore, residue L100 helps define the putative solvent-accessible entrance to the NNRTI binding pocket at the RT heterodimer interface. In particular, the increase in resistance caused by mutation at position 100 from leucine to isoleucine (L100I) is related to steric hindrance. Increased steric bulk causes changes in the conformation of the inhibitor within the binding site and/or in the geometry of the binding pocket itself.[42] As reported in the literature, resistance mutation L100I was associated with a decrease in side chain length and distortion of the binding pocket region with a complete loss of protein–drug interactions.[43] Such an effect is particularly evident for second- and third-generation NNRTIs.[15,28,44]

Another mutation that deserves attention is K101P, shown to directly contribute to ETR resistance. In particular, it has been observed that this mutation is the 3rd most significant mutation involved in resistance to this NNRTI after Y181I and Y181V.[28]

For all NNRTIs, the analyzed mutant models were built from crystallographic data by replacing the appropriate residues; all structures were subsequently energy optimized. The minimiza-

tion procedure was validated by superimposing the optimized mutant RT structures on the crystal structures of related mutant RTs, and the obtained root mean square (RMS) values were always less than 1.5 Å (Tables S2 and S3 in the Supporting Information).

With the aim to investigate the influence of the analyzed mutations on the conformational properties of the enzyme, we submitted the above-mentioned complexes to AutoDock Vina[45] docking simulations. To test the accuracy and reliability of the docking procedure, we firstly performed AutoDock Vina redocking calculations starting from the original [NVP·RT], [EFV·RT], and [ETR·RT] crystal structures in order to geometrically reproduce the experimental data. In such an analysis, we also included [RPV·RT] and [LSV·RT] PDB co-crystallographic structures crystal structures even though they were not considered in our study. For all complexes, the program was able to reproduce the experimental geometries well with RMS values always below 1 Å (for full details, see Table S4 and Figures S1–S10 in the Supporting Information). Moreover, we carried out also cross-docking simulations and found that all NNRTIs were well accommodated in the binding pocket of both wild-type and mutant RT (Figure S11–S15 in the Supporting Information).

Consequently, we applied the same procedure for RT recognition of NVP, EFV and ETR, and the obtained configurations were evaluated in terms of their interaction free energy (Table 4). The analysis revealed that in NVP–RT and EFV–RT complexes, K103N and Y188L mutated enzymes were associated with reduced inhibitor binding affinities compared with wild-type enzyme, thus indicating decreased stability according to their high resistance index (Table 3).

Analyzing ETR–RT complexes, the bound DAPY drug was found to be less stabilized within the NNRTI binding pocket of the K101P mutant enzyme compared with the wild type, while in presence of K103N and K103S substitutions, the binding configuration was found to be more favorable, in agreement with literature data.[18, 19] With respect to RT in complex with EFV and ETR, the L100I mutation caused a decrease in drug affinity compared with wild type (based on SVR results). This is consistent with SVR results for EFV (Table 3) and with recent finding showing that L100I is among the four mutations with the greatest contribution to viral resistance to ETR.[28] Such a

finding was not noticed for NVP; again, this is consistent with SVR results and the extremely low prevalence of L100I in patients failing treatment with NVP compared with those failing treatment regimes containing EFV or ETR (0.4% versus 12.4% and 12%, respectively).

Moreover, in order to evaluate the only positive association of NNRTI resistance mutations, we included in our thermodynamics analysis the double mutant K101E+G190A for NVP and EFV recognition. In both cases, we observed a reduced binding affinity with respect to the wild type (data not shown), in accordance with the experimental profiles.

To examine the obtained theoretical complexes for the analyzed NNRTIs, the interactions established between the studied inhibitors and the enzyme active site were evaluated for either wild type or high-resistance-associated mutants (Tables S5–S7 in the Supporting Information). For the best wild-type configuration of the NVP–RT complex, several non-bonding contacts with different RT residues, such as L100, K101, K103, V106, Y181, Y188, G190, F227, W229 and Y318, were observed (Figure 2A). In the best configuration with the K103N mutant enzyme, NVP was found to make several van der Waals interactions with the same NNRTI binding pocket residues. However, accommodation of NVP in the wild-type enzyme was well stabilized through an interaction with G190, located in the binding pocket β-sheet and found to be essential for interactions with this drug.[46] Such a contribution is missing in the K103N mutant complex where a reduced interaction network between the dipyridodiazepinone inhibitor and the RT amino acids V106, Y181 and L234 was observed when compared with the wild-type sequence. Furthermore, in the complex of NVP and the Y188L mutant RT, the drug was deficient of some crucial hydrophobic interactions with V106 and Y181 (Figure S16 in the Supporting Information). These geometric considerations were consistent with the calculated binding affinity profiles.

In the wild-type EFV–RT global minimum complex, a hydrogen bond and several van der Waals interactions were detected (Figure 2B). In particular, the benzoxazinone inhibitor established a hydrogen bond with the backbone carbonyl oxygen of K101, found to be crucial for ligand binding. Recently, a computational method was used to study the interactions between EFV and the binding pocket of wild-type HIV-1 RT.[47] The results showed a net attractive interaction between EFV and the surrounding residues, with K101 predicted to have a stronger interaction with the drug than the other residues. Such an interaction is crucial for the stability of the inhibitor in the binding site and strengthens the affinity of EFV compared with other NNRTIs. It was suggested that weaker binding of EFV to the K103N mutant can be attributed to a slightly weakened attractive interaction between the drug and the K101 residue.[48]

Examining the interactions of the most stable EFV–RT complex in the presence of the K103N mutation, the drug lacks some hydrophobic contacts with the binding site residues, specifically with V106 and P236. Nevertheless, the hydrogen bond with K101 remained the strongest interaction. Moreover, it was observed that the N103 mutated residue in K103N/

Table 4. Evaluation of the AutoDock Vina[45] interaction energy calculated for all studied reverse transcriptase (RT) complexes in the presence of the three inhibitors nevirapine (NVP), efavirenz (EFV), and etravirine (ETR).

Model	ΔG_{bind} [kcal mol^{-1}]		
	NVP	EFV	ETR
WT	−10.3	−11.9	−9.6
L100I	−10.7	−10.4	−9.1
K101P	−10.4	−11.8	−8.5
K103N	−9.6	−9.8	−10.3
K103S	−10.2	−11.7	−10.2
I135T	−10.3	−11.8	−9.5
Y181C	−10.0	−11.5	−9.4
Y188L	−9.5	−10.9	−9.7
G190A	−9.8	−11.6	−9.6
G190S	−9.7	−11.5	−9.7
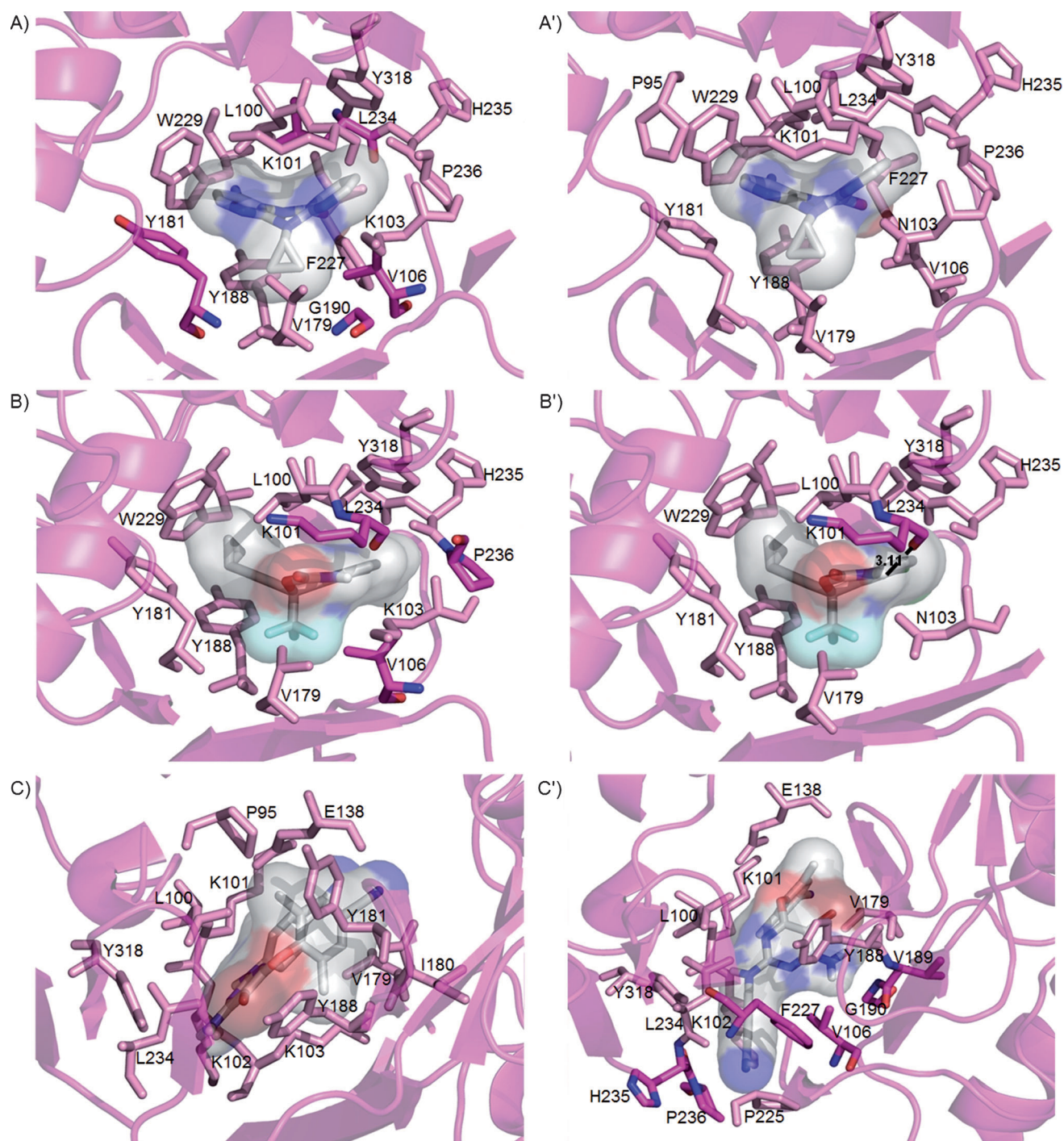
Figure 2. Three-dimensional representation of NNRTI interactions in the RT binding pocket. The panels show the best poses for NVP, EFV and ETR docked to wild-type RT (A, B, C) and K103N mutant RT (A', B', C'). The drug is shown as grey carbon sticks with its surface area. The RT NNRTI binding pocket residues involved in interactions are shown in pink, while amino acids able to establish additional contacts or hydrogen bonds are given in magenta. Panel B reports the distance (Å) between the atoms involved in the hydrogen bond (-----).

Y181C HIV-1 RT altered the hydrogen-bond distance between the carbonyl group of the EFV benzoxine-2-one ring and the backbone amino hydrogen (NH) of K101.[48] As reported in Figure 2B, the measured distance of this interaction also showed the single K103N substitution related to such a perturbation ($d_{WT} = 3.05$ Å; $d_{K103N} = 3.11$ Å). The increased distance was associated with the observation that N103 creates repulsive inter-

actions with EFV when compared to the interaction between K103 and the drug.[48] Analyzing the EFV–RT Y188L mutant complex, a decreased number of good contacts between the inhibitor and W229 were noticed (Figure S17 in the Supporting Information); L234 and Y318 residues were missing in this configuration. In the EFV–RT L100I mutant global minimum complex, the drug was less stabilized due to a reduced interaction

network with Y188 and W229 and due to the loss of a crucial hydrophobic contact with Y318. These structural findings and the obtained thermodynamics results can justify the poor drug stabilization within the NNRTI binding pocket and its high-level resistance profile in the presence of the L100I, K103N and Y188L mutations.

In the most probable wild-type ETR–RT configuration, the DAPY was well stabilized in the enzyme active site through a large number of non-bonding contacts (Figure 2C). Specifically, ETR was found to interact with E138 of the p51 subunit; such an interaction was missing for the first- and second-generation NNRTIs. E138 was the only residue of the p51 subunit interacting either directly or indirectly with the bound NNRTI. Such a residue is part of the β7–β8 loop in the highly conserved amino acid sequence of HIV-1 RT. Mutagenesis analyses highlighted the pivotal structural and functional roles of E138 in NNRTI binding pocket stabilization in that its mutation significantly reduced viral infectivity.[49]

Interestingly, analyzing the most stable configuration in the presence of the K103N substitution, additional non-bonding interactions were observed with amino acids V106, V189, G190, F227, H235 and P236, compared with the wild-type complex. Moreover the best poses for both the wild-type and K103N enzymes in complex with ETR showed relevant conformational differences. In particular, the wild-type complex assumed a more compact configuration, with its aromatic rings implicated in intramolecular stacking interactions. By contrast, in the K103N mutant complex, due to the high flexibility of the mutated enzyme, ETR appeared conformationally more opened, thus allowing its aromatic moieties to be involved in additional contacts with F227 and P236.

Examining the ETR–RT configuration in the presence of the K101P substitution, the drug lacked crucial interactions with K102, K103, L234, H235, P236 and Y318, decreasing its binding affinity. The K101P destabilizing effect on the drug–enzyme complex was associated with a lower number of van der Waals contacts compared with the wild-type sequence, even if ETR was able to establish three hydrogen bonds with E138 and I180 (Figure S18 in the Supporting Information). In the L100I ETR–RT global minimum complex, a decreased number of favorable interactions between the inhibitor and K101, E138, I180 and Y181 was observed. In particular, the reduced E138 van der Waals network contributed to the destabilization of the DAPY inhibitor in the NNRTI binding pocket. Furthermore, K102, L234 and Y318 residues were missing in this configuration with a subsequent unfavorable binding affinity profile.

Conclusions

Non-nucleoside reverse transcriptase inhibitor (NNRTI) resistance mutations have been shown to either directly hinder drug binding to the pocket by altering the size, shape, and polarity of different parts of the NNRTI binding site or to indirectly prevent access to the pocket. However, some of the most clinically relevant substitution mutations have also been found to influence reverse transcriptase functions other than DNA polymerization. Here, we applied a computational protocol to

combinations of mutated and wild-type RT sequences and three approved NNRTIs with the aim to combine structural and thermodynamic data with relevant clinical results. Resistance, expressed as a Z-score index, was related to interaction free energies in the fully optimized ensembles. Several structural features resulted useful in rationalizing most of the clinical observations related to resistance development. Such an approach can be adopted to obtain additional information for the rational design of novel NNRTIs active against resistant RT mutants. The encouraging results obtained in our preliminary study have motivated us to extend our analysis to include a larger set of RT–NNRTI complexes and eventually to apply this computational strategy to other enzymatic targets subjected to resistance-inducing mutations. The results of these ongoing studies will be reported in due course.

Experimental Section

Clinical methods

For this study, conducted in different clinical departments in central Italy, we analyzed 1273 pol sequences from 697 HIV-1 subtype B patients failing for the first time an antiretroviral regimen containing at least one NNRTI: specifically, EFV ($n = 647$), NVP ($n = 568$) and ETR ($n = 50$). Data for all patients were stored in a specifically designed anonymous database that included genotypic, demographic, immunologic, virologic and therapeutic parameters.

HIV sequencing: HIV genotype analysis was performed on plasma samples by means of a commercially available kit (ViroSeq HIV-1 Genotyping System, AB Pharmaceuticals, LLC, Chesterfield, MO, USA). Briefly, RNA was extracted, retrotranscribed by murine leukemia virus (MULV) RT, and amplified with AmpliTaq Gold polymerase using two different sequence-specific primers for 40 cycles. Pol-amplified products (containing the entire protease and the first 320 amino acids of the RT open reading frame) were full-length sequenced in sense and antisense orientations using seven different overlapping sequence-specific primers by an automated sequencer (ABI 3100). Sequences having a mixture of wild-type and mutant residues at single positions were considered to have the mutation(s) at that position. The isolates were subtyped by comparing them to reference sequences of known subtype.[30]

Statistical analysis

Mutation prevalence: We calculated the prevalence of mutations known to give resistance to the NNRTI in isolates from patients failing their last treatment regime containing a given NNRTI: namely, EFV ($n = 647$), NVP ($n = 568$) and ETR ($n = 50$).

Mutation co-variation: We exhaustively analyzed patterns of pairwise interactions among mutations associated with NNRTI treatment. Specifically, for each pair of mutations and corresponding wild-type residues, a Fisher's exact test was performed to assess whether co-occurrence of the mutated residues differed significantly from what would be expected under an independence assumption. The Benjamini–Hochberg method was used to correct for multiple testing at a false discovery rate of 0.01. Samples having a mixture of two or more mutations at a given pair of positions were ignored in calculating the co-variation due to the impossibility of identifying whether these mutations are indeed locat-

ed in the same viral genome. We performed this analysis in the EFV (*n* = 647) and NVP (*n* = 568) patient subsets.

Cluster analysis: In order to identify and summarize higher order interactions of mutations, we transformed the pair-wise phi correlation coefficients into dissimilarity values. Based on these pair-wise dissimilarity values, a dendrogram was computed by hierarchical clustering. Finally, the stability of the resulting dendrogram was assessed from 100 bootstrap replicates. The details of this explorative data analysis procedure have been described in detail elsewhere.[50–52] Briefly, the transformation from phi coefficients to dissimilarity values was done by mapping phi = 1 (maximal positive association) to dissimilarity = 0, and phi = −1 (maximal negative association) to dissimilarity = 1, with linear interpolation between the two. The dissimilarity of mutations at the same position was left undefined; as such, pairs never co-occur in a single sequence (except from mixtures) and would lead to artifacts in the resulting dendrogram. The resulting partial dissimilarity matrix was then used as input for average linkage hierarchical agglomerative clustering, and undefined dissimilarity values were ignored in computing average dissimilarities between clusters. In order to assess the stability of the resulting dendrogram, confidence values for all subtrees in the dendrogram were computed by 100 replicates of the clustering procedure on sequence sets bootstrapped from the two sets of sequences. For instance, a bootstrap value of 1 for an edge in the dendrogram means that the set of mutations in the induced subtree occurs as a subtree in all dendrograms from the different bootstrap replicates. Thus, higher bootstrap values indicate that the association of mutations into a group is not due to sampling bias. Finally, we also examined the extent to which the NNRTI-selected mutations co-varied with mutations at 18 NRTI resistance positions.[30] Sequences were analyzed using the same methods described for assessing co-variation within the NNRTI mutation dataset.

Computational methods

Association with phenotypic NNRTI susceptibility: We measured the impact of EFV and NEV resistance mutations in multivariate computational models for predicting phenotypic resistance from genotype. In contrast to the univariate setting, multivariable analyses allow the assessment of the impact of mutations relative to other mutations. Specifically, we analyzed the support vector regression (SVR) models used in the geno2pheno web-based prediction service,[53, 54] by exploiting the bilinearity of the kernel used in geno2pheno, according to the method described in the literature.[51, 55] These models are based on approximately 850 matched genotype–phenotype pairs derived from another recombinant assay.[56]

Structural analysis and docking simulations: For all studied NNRTIs, the analyzed mutant models were built from crystallographic data by replacing residues. Each complex was then placed in a cubic cell with size adjusted to maintain a minimum distance of 10 Å to the cell boundary and soaked with a pre-equilibrated box of water using the System Builder module of the Desmond package.[57] All overlapping solvent molecules were removed and an appropriate number of counter ions were added to maintain charge neutrality. In order to optimize the geometries, all complexes were energy minimized using OPLS 2005 as the force field,[58] a steepest descent algorithm with a maximum of 200 steps and a convergence threshold of 1.0 Kcal mol^{-1} Å^{-1}.

In order to evaluate the reliability of the minimization procedure, we superimposed our generated optimized models on the related experimentally determined crystallographic mutant structures. In particular, for both NVP and EFV in complex with mutated RT, we considered K103N and Y181C mutant enzyme crystal structures with the PDB[34] ID codes 1FKP[59] and 1JLB[14] for NVP, and 1IKV[35] and 1JKH[14] for EFV. Due to the poor availability of ETR–RT crystal structures, we only took into account the K103N mutated complexes, with PDB codes 1V5[18] and 3MED.[60] The obtained RMS values, calculated on the enzyme Cα atoms, are reported in Table S2 in the Supporting Information. Moreover, the RMS values calculated in the NNRTI binding pocket considering all residues within 15 Å from the inhibitor and including the side chains also, are showed in Table S3 in the Supporting Information.

AutoDock Vina[45] redocking simulations were performed with the aim to validate the docking application. For each analyzed inhibitor, we selected two PDB co-crystal structures considering both wild-type and mutated sequences, and we superimposed the best poses, relative to the reference crystal structures, generated by docking. In these calculations, we selected wild-type and mutated structures of RT co-crystallized with NVP, EFV, ETR, RPV and LSV, with PDB codes 1VRT,[11] 2HNY,[61] 1IKW,[35] 1JKH,[14] 3MEC,[60] 1V5,[18] 3MEE,[60] 3MEG,[60] 2WON[26] and 2WOM,[26] respectively. The RMS values, calculated considering the analyzed ligands C sp^2/sp^3 united atoms, are indicated in Table S4 in the Supporting Information.

Moreover, cross-docking calculations were carried out. In cross-docking simulations, the ligand is docked to a receptor conformational state co-crystallized with a different ligand. For the five pairs of ligand–receptor complexes, we docked each of the ligands to the receptor structure co-crystallized with the other NNRTI (Figures S11–S15 in the Supporting Information). In both redocking and cross-docking calculations, the graphic interface AutoDockTools (ADT) was used to manipulate the models and to perform input/output procedures. After removing the experimentally complexed ligands, the targets were converted into the "Protein Data Bank-like file format with charges and atom types" (PDBQT) format by the ADT program. The procedure was repeated for both the receptor and the ligands using Gasteiger PEOE partial charges.[62] This method applied a Lamarckian model of genetics generating 100 allowed configurations per ligand. The search space was defined as a cubic box of 8000 Å^3 centered on the p66 L100 side chain. For our simulations, the exhaustiveness setting was increased to 128. After the validation of minimization and docking procedures starting from the energy optimized geometry, all the analyzed complexes were submitted to AutoDock Vina[45] docking simulations using the same protocol mentioned above. The first step of the computational work was to select appropriate structures of the enzymes from the PDB.[34] For our analysis, PDB entries 1VRT,[11] 1IKW[35] and 1V5[18] were considered for RT co-crystallized with NVP, EFV and ETR, respectively. The LigPlot program[63] was used to identify the most relevant interactions between the NNRTIs and both RT subunits. All three-dimensional figures were prepared using PyMOL graphics and modeling package version 0.99.[64] All calculations were performed on a Linux cluster with 16 dual-Xeon processor nodes.

Acknowledgements

[1] M. A. Parniak, N. Sluis-Cremer, *Adv. Pharmacol.* **2000**, *49*, 67–109.

[2] E. De Clercq, *J. Clin. Virol.* **2004**, *30*, 115–133.

[3] J. Balzarini, *Curr. Top. Med. Chem.* **2004**, *4*, 921–944.

[4] E. De Clercq, *Med. Res. Rev.* **2009**, *29*, 611–645.

[5] a) F. Clavel, A. J. Hance, *N. Engl. J. Med.* **2004**, *350*, 1023–1035; b) L. Menéndez-Arias, *Antiviral Res.* **2010**, *85*, 210–231.

[6] J. Ding, S. H. Hughes, E. Arnold, *Biopolymers* **1997**, *44*, 125–138.

[7] K. Das, S. G. Sarafianos, A. D. Clark Jr., P. L. Boyer, S. H. Hughes, E. Arnold, *J. Mol. Biol.* **2007**, *365*, 77–89.

[8] L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice, T. A. Steitz, *Science* **1992**, *256*, 1783–1790.

[9] G. Tachedjian, S. P. Goff, *Curr. Opin. Investig. Drugs* **2003**, *4*, 966–973.

[10] C. Tantillo, J. Ding, A. Jacobo-Molina, R. G. Nanni, P. L. Boyer, S. H. Huang, R. Pauwels, K. Andries, P. A. Janssen, E. Arnold, *J. Mol. Biol.* **1994**, *243*, 369–387.

[11] J. Ren, R. Esnouf, E. Garman, D. Somers, C. Ross, I. Kirby, J. Keeling, G. Darby, Y. Jones, D. Stuart, D. Stammers, *Nat. Struct. Biol.* **1995**, *2*, 293–302.

[12] S. D. Young, S. F. Britcher, L. O. Tran, L. S. Payne, W. C. Lumma, T. A. Lyle, J. R. Huff, P. S. Anderson, D. B. Olsen, S. S. Carroll, D. J. Pettibone, J. A. O'Brien, R. G. Ball, S. K. Balani, J. H. Lin, I.-W. Chen, W. A. Schleif, V. V. Sardana, W. J. Long, V. W. Byrnes, E. A. Emini, *Antimicrob. Agents Chemother.* **1995**, *39*, 2602–2605.

[13] J. Ren, J. Milton, K. L. Weaver, S. A. Short, D. I. Stuart, D. K. Stammers, *Structure* **2000**, *8*, 1089–1094.

[14] J. Ren, C. Nichols, L. Bird, P. Chamberlain, K. Weaver, S. Short, D. I. Stuart, D. K. Stammers, *J. Mol. Biol.* **2001**, *312*, 795–805.

[15] L. T. Bacheler, E. D. Anton, P. Kudish, D. Baker, J. Bunville, K. Krakowski, L. Bolling, M. Aujay, X. V. Wang, D. Ellis, M. F. Becker, A. L. Lasut, H. J. George, D. R. Spalding, G. Hollis, K. Abremski, *Antimicrob. Agents Chemother.* **2000**, *44*, 2475–2484.

[16] J. Ren, D. K. Stammers, *Virus Res.* **2008**, *134*, 157–170.

.

[17] K. Andries, H. Azijn, T. Thielemans, D. Ludovici, M. Kukla, J. Heeres, P. Janssen, B. De Corte, J. Vingerhoets, R. Pauwels, M. P. de Béthune, *Antimicrob. Agents Chemother.* **2004**, *48*, 4680–4686.

[18] K. Das, A. D. Clark, Jr., P. J. Lewi, J. Heeres, M. R. De Jonge, L. M. Koymans, H. M. Vinkers, F. Daeyaert, D. W. Ludovici, M. J. Kukla, B. De Corte, R. W. Kavash, C. Y. Ho, H. Ye, M. A. Lichtenstein, K. Andries, R. Pauwels, M. P. De Béthune, P. L. Boyer, P. Clark, S. H. Hughes, P. A. Janssen, E. Arnold, *J. Med. Chem.* **2004**, *47*, 2550–2560.

[19] J. Vingerhoets, H. Azijn, E. Fransen, I. De Baere, L. Smeulders, D. Jochmans, K. Andries, R. Pauwels, M. P. de Béthune, *J. Virol.* **2005**, *79*, 12773–12782.

[20] J. Guillemont, E. Pasquier, P. Palandjian, D. Vernier, S. Gaurrand, P. J. Lewi, J. Heeres, M. R. de Jonge, L. M. H. Koymans, F. F. D. Daeyaert, M. H. Vinkers, E. Arnold, K. Das, R. Pauwels, K. Andries, M.-P. de Béthune, E. Bettens, K. Hertogs, P. Wigerinck, P. Timmerman, P. A. J. Janssen, *J. Med. Chem.* **2005**, *48*, 2072–2079.

[21] L. T. Rimsky, H. Azijn, I. Tirry, J. Vingerhoets, R. Mersch, G. Kraus, M.-P. de Béthune, G. Picchio, *Antiviral Ther.* **2009**, *14 (Suppl. 1)*, A141.

[22] K. Das, J. D. Bauman, A. D. Clark, Jr., Y. V. Frenkel, P. J. Lewi, A. J. Shatkin, S. H. Hughes, E. Arnold, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1466–1471.

[23] M.-P. de Béthune, *Antiviral Res.* **2010**, *85*, 75–90.

[24] a) J. Mori, R. Corbau, D. Lewis, S. Ellery, H. Mayer, M. Perros, M. Westby, *15th Conference on Retroviruses and Opportunistic Infections*, (Boston, USA), **2008**, abstract P728; b) A. S. Thornberry, J. Mori, M. Perros, M. Westby, C. Craig, *7th European HIV Drug Resistance Workshop*, (Stockholm, Sweden), **2009**, abstract 63.

[25] G. Fätkenheuer, S. Staszewski, A. Plettenburg, F. Hackman, G. Layton, L. McFadyen, T. Davis, J. Jenkins, *4th International AIDS Society conference on HIV Pathogenesis, Treatment and Prevention*, (Sydney, Australia), **2007**, abstract WESS202.

[26] R. Corbau, J. Mori, C. Phillips, L. Fishburn, A. Martin, C. Mowbray, W. Panton, C. Smith-Burchnell, A. Thornberry, H. Ringrose, T. Knöchel, S. Irving, M. Westby, A. Wood, M. Perros, *Antimicrob Agents Chemother.* **2010**, *54*, 4451–4463.

[27] S. Alcaro, C. Alteri, A. Artese, F. Ceccherini-Silberstein, G. Costa, F. Ortuso, L. Parrotta, C. F. Perno, V. Svicher, *Drug Resist. Updates* **2011**, *14*, 141–149.

[28] J. Vingerhoets, L. Tambuyzer, H. Azijn, A. Hoogstoel, S. Nijs, M. Peeters, M. P. de Béthune, G. De Smedt, B. Woodfall, G. Picchio, *AIDS* **2010**, *24*, 503–514.

[29] *Update of the drug resistance mutations in HIV-1: December 2009*, V. A. Johnson, F. Brun-Vezinet, B. Clotet, H. F. Gunthard, D. R. Kuritzkes, D. Pillay, J. M. Schapiro, D. D. Richman, *Top HIV Med.* **2009**, *17*, 138–145.

[30] Stanford University, HIV Drug Resistance Database, last update: June 10, 2008; website: http://hivdb.Stanford.edu (last accessed: September 12, 2011); S.-Y Rhee, M. J. Gonzales, R. Kantor, B. J. Betts, J. Ravela, R. W. Shafer, *Nucleic Acids Res.* **2003**, *31*, 298–303.

[31] J. Zhang, T. Hou, W. Wang, J. S. Liu, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1321–1326.

[32] J. Wang, H. Liang, L. Bacheler, H. Wu, K. Deriziotis, L. M. Demeter, C. Dykes, *Virology* **2010**, *402*, 228–237.

[33] N. Beerenwinkel, T. Sing, T. Lengauer, J. Rahnenführer, K. Roomp, I. Savenkov, R. Fischer, D. Hoffmann, J. Selbig, K. Korn, H. Walter, T. Berg, P. Braun, G. Fätkenheuer, M. Oette, J. Rockstroh, B. Kupfer, R. Kaiser, M. Däumer, *Bioinformatics* **2005**, *21*, 3943–3950.

[34] The Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB); http://www.rcsb.org/pdb.

[35] J. Lindberg, S. Sigurdsson, S. Lowgren, H. O. Andersson, C. Sahlberg, R. Noreen, K. Fridborg, H. Zhang, T. Unge, *Eur. J. Biochem.* **2002**, *269*, 1670–1677.

[36] A. L. Hopkins, J. Ren, R. M. Esnouf, B. E. Willcox, E. Y. Jones, C. Ross, T. Miyasaka, R. T. Walker, H. Tanaka, D. K. Stammers, D. I. Stuart, *J. Med. Chem.* **1996**, *39*, 1589–1600.

[37] R. T. D'Aquila, J. M. Schapiro, F. Brun-Vézinet, B. Clotet, B. Conway, L. M. Demeter, R. M. Grant, V. A. Johnson, D. R. Kuritzkes, C. Loveday, R. W. Shafer, D. D. Richman, *Top HIV Med.* **2002**, *10*, 21–25.

[38] J. Ren, C. E. Nichols, P. P. Chamberlain, K. L. Weaver, S. A. Short, J. H. Chan, J. P. Kleim, D. K. Stammers, *J. Med. Chem.* **2007**, *50*, 2301–2309.

[39] Y. Hsiou, J. Ding, K. Das, A. D. Clark, Jr., P. L. Boyer, P. Lewi, P. A. Janssen, J. P. Kleim, M. Rösner, S. H. Hughes, E. Arnold, *J. Mol. Biol.* **2001**, *309*, 437–445.

[40] F. Rodríguez-Barrios, F. Gago, *J. Am. Chem. Soc.* **2004**, *126*, 15386–15387.

[41] P. R. Harrigan, T. Mo, B. Wynhoven, J. Hirsch, Z. Brumme, P. McKenna, T. Pattery, J. Vingerhoets, L. T. Bacheler, *AIDS* **2005**, *19*, 549–554.

[42] D. G. Prajapati, R. Ramajayam, M. R. Yadav, R. Giridhar, *Bioorg. Med. Chem.* **2009**, *17*, 5744–5762.

[43] N. Sluis-Cremer, N. A. Temiz, I. Bahar, *Curr. HIV Res.* **2004**, *2*, 323–332.

[44] M. T. Lai, M. Lu, P. J. Felock, R. C. Hrin, Y. J. Wang, Y. Yan, S. Munshi, G. B. McGaughey, R. M. Tynebor, T. J. Tucker, T. M. Williams, J. A. Grobler, D. J. Hazuda, P. M. McKenna, M. D. Miller, *Antimicrob Agents Chemother.* **2010**, *54*, 4812–4824.

[45] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455–461.

[46] S. Paolucci, F. Baldanti, G. Campanini, R. Cancio, A. Belfiore, G. Maga, G. Gerna, *Antiviral Res.* **2007**, *76*, 99–103.

[47] P. Nunrium, M. Kuno, S. Saen-oon, S. Hannongbua, *Chem. Phys. Lett.* **2005**, *405*, 198–202.

[48] P. Srivab, S. Hannongbua, *ChemMedChem* **2008**, *3*, 803–811.

[49] P. K. Pandey, N. Kaushik, K. Singh, B. Sharma, A. K. Upadhyay, S. Kumar, D. Harris, V. N. Pandey, *BMC Biochemistry* **2002**, *3*, 18.

[50] V. Svicher, T. Sing, M. M. Santoro, F. Forbici, F. Rodríguez-Barrios, A. Bertoli, N. Beerenwinkel, M. C. Bellocchi, F. Gago, A. d'Arminio Monforte, A. Antinori, T. Lengauer, F. Ceccherini-Silberstein, C. F. Perno, *J. Virol.* **2006**, *80*, 7186–7198.

[51] F. Ceccherini-Silberstein, V. Svicher, T. Sing, A. Artese, M. M. Santoro, C. Gori, A. Bertoli, S. Alcaro, A. d'Arminio Monforte, A. Antinori, T. Lengauer, C. F. Perno, *J. Virol.* **2007**, *81*, 11507–11519.

[52] V. Svicher, S. Aquaro, R. D'Arrigo, A. Artese, S. Dimonte, S. Alcaro, G. Di Perri, S. Lo Caputo, R. Bellagamba, M. Zaccarelli, U. Visco-Comandini, A.

Antinori, P. Narciso, F. Ceccherini-Silberstein, C. F. Perno, *J. Infect. Dis.* **2008**, *197*, 1408–1418.

[53] N. Beerenwinkel, M. Daumer, M. Oette, K. Korn, D. Hoffmann, R. Kaiser, T. Lengauer, J. Selbig, H. Walter, *Nucleic Acids Res.* **2003**, *31*, 3850–3855.

[54] Max–Planck Institut Informatik, Genafor—Society for Sustained Research (Bonn, Germany); Geno2pheno [resistance] 3.2, www.geno2pheno.org (last accessed: September 12, 2011).

[55] T. Sing, V. Svicher, N. Beerenwinkel, F. Ceccherini-Silberstein, M. Daumer, R. Kaiser, H. Walter, K. Korn, D. Hoffmann, M. Oette in *Knowledge Discovery in Databases: PKDD 2005*, (Eds.: A. Jorge, L. Torgo, P. Brazdil, R. Camacho, J. Gama), Springer, New York, **2005**, p. 285–296..

[56] H. Walter, B. Schmidt, K. Korn, A. M. Vandamme, T. Harrer, K. Uberla, *J. Clin. Virol.* **1999**, *13*, 71–80.

[57] a) K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan, D. E. Shaw in *SC '06 Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, Association for Computing Machinery, New York, **2006**; b) Desmond Molecular Dynamics System, version 2.2, D. E. Shaw Research, New York, NY, **2009**. Maestro–Desmond Interoperability Tools, version 2.2, Schrödinger, New York, NY, **2009**; K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I.

Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan, D. E. Shaw, *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*, (Tampa, USA), November 11–17, 2006.

[58] a) W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236; b) G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, *J. Phys. Chem. B* **2001**, *105*, 6474–6487.

[59] J. Ren, J. Milton, K. L. Weaver, S. A. Short, D. I. Stuart, D. K. Stammers, *Structure Fold. Des.* **2000**, *8*, 1089–1094.

[60] E. B. Lansdon, K. M. Brendza, M. Hung, R. Wang, S. Mukund, D. Jin, G. Birkus, N. Kutty, X. Liu, *J. Med. Chem.* **2010**, *53*, 4295–4299.

[61] J. Ren, C. E. Nichols, A. Stamp, P. P. Chamberlain, R. Ferris, K. L. Weaver, S. A. Short, D. K. Stammers, *Febs J.* **2006**, *273*, 3850–3860.

[62] J. Gasteiger, M. Marsili, *Tetrahedron* **1980**, *36*, 3219–3228.

[63] A. C. Wallace, R. A. Laskowski, J. Thornton, *Protein Eng.* **1995**, *8*, 127–134.

[64] W. L. DeLano, The PyMOL Molecular Graphics System, (Version 1.2), Schrödinger, LLC.; http://www.pymol.org.