Ab initio molecular dynamics studies on HIV-1 reverse transcriptase triphosphate binding site: Implications for nucleoside–analog drug resistance

FRANK ALBER1 AND PAOLO CARLONI1,2

1International School for Advanced Studies (SISSA) and Istituto Nazionale di Fisica della Materia (INFM), Via Beirut 2-4, 34014, Trieste, Italy
2International Centre for Genetic Engineering and Biotechnology, AREA Science Park, Padriciano 99, 34012, Trieste, Italy

(Received June 6, 2000; Final Revision October 5, 2000; Accepted October 11, 2000)

Abstract

Quantum-chemical methods are used to shed light on the functional role of residues involved in the resistance of HIV-1 reverse transcriptase against nucleoside-analog drugs. Ab initio molecular dynamics simulations are carried out for models representing the adduct between the triphosphate substrate and the nucleoside binding site. The triphosphate is considered either deprotonated or protonated at the γ-position. Although the protonated form already experiences large rearrangements in the ps time scale, the fully deprotonated state exhibits a previously unrecognized low-barrier hydrogen bond between Lys65 and γ-phosphate. Absence of this interaction in Lys65→Arg HIV-1 RT might play a prominent role in the resistance of this mutant for nucleoside analogs (Gu Z et al., 1994b, Antimicrob Agents Chemother 38:275–281; Zhang D et al., 1994, Antimicrob Agents Chemother 38:282–287). Water molecules present in the active site, not detected in the X-ray structure, form a complex H-bond network. Among these waters, one may be crucial for substrate recognition as it bridges Gln151 and Arg72 with the β-phosphate. Absence of this stabilizing interaction in Gln151→Met HIV-1 RT mutant may be a key factor for the known drug resistance of this mutant toward dideoxy-type drugs and AZT (Shirasaka T et al., 1995, Proc Natl Acad Sci USA 92:2398–2402; Iversen AK et al., 1996, J Virol 70:1086–1090).

Keywords: ab initio molecular dynamics; Car-Parrinello; HIV reverse transcriptase; low-barrier hydrogen-bond; mutations; resistance

Reverse transcriptase from the human immunodeficiency virus type 1 (HIV-1 RT) is a major target for anti-AIDS therapy (Larder et al., 1987; Frankel & Young, 1998). This enzyme catalyzes the transcription of the single-stranded RNA viral genome into double-stranded DNA, which subsequently becomes integrated into host cell chromosomes (Frankel & Young, 1998; Turner & Summers, 1999). HIV-1 RT is a complex machinery, acting both as RNA/DNA dependent DNA-polymerase and as a ribonuclease H (Frankel & Young, 1998; Huang et al., 1998; Turner & Summers, 1999).

The large substrate diversity of HIV-1 RT (Tantillo et al., 1994), in comparison to the cellular isoenzymes (Cheng et al., 1987), represents the basis of the successful therapeutic use of nucleotide-analog drugs (NRTIs) (Fig. 1). The triphosphorylated form of the NRTIs are believed to inhibit HIV-1 RT activity by both competitive inhibition as well as chain termination (Quan et al., 1998). In the latter case, the viral enzyme is able to incorporate into the growing DNA strand the triphosphorylated form of several analogs whose 3′-OH group is missing or replaced with different functions, such as the azide group in AZT (Fig. 1). The viral DNA elongation is then blocked thereby inhibiting replication of the virus particles (Taylor et al., 1990; Goody et al., 1991; Herdevijn, 1992; Schinazi, 1993).

Drug benefit is severely limited by the capability of the virus to develop mutations, which ultimately lead to drug resistance (Larder, 1993; Richman, 1993; De Clercq, 1994; Sarafianos et al., 1999). The spectrum of alterations is rather broad, as evinced by genetic and biochemical studies performed either in the laboratory or in clinical trials (Boyer et al., 1994; Tantillo et al., 1994; Larder & Stammers, 1999). Single mutations effective against drug action are usually accompanied sequentially by three to four additional mutations so that several highly resistant mutation patterns are observed. Resistant strains against NRTIs include mutations at amino acids: (1) Gln151→Met, Ala62→Val, Val75→Ile,
Phe77\rightarrow Leu, Phe116\rightarrow Tyr (AZT, ddI, and ddC multi-drug resistance; Shirasaka et al., 1995; Iversen et al., 1996); (2) Met41\rightarrow Leu, Asp67\rightarrow Asn, Lys70\rightarrow Arg, Thr215\rightarrow Tyr/Phe, Lys219\rightarrow Gln (AZT resistant; Larder & Kemp, 1989; Kellam et al., 1992; Carroll et al., 1994; Krebs et al., 1997); (3) Lys65\rightarrow Arg (ddl, ddC, and 3TC cross-resistant; Gu et al., 1994b; Zhang et al., 1994; Arion et al., 1996); (4) Met184\rightarrow Val/Leu (3TC resistant; Boucher et al., 1993; Gao et al., 1993; Tisdale et al., 1993).

Drug suppression is governed by diverse factors, which range from changes of DNA processivity (resistance against ddI, ddC: Arion et al., 1996; and 3TC: Boyer & Hughes, 1995), and of drug/substrate molecular recognition (resistance against ddl, AZT: Martin et al., 1993; and ddl, ddC: Gu et al., 1994a) to removal of chain-terminating residues from blocked primers (resistance against AZT: Meyer et al., 1999).

The recent determination of the crystal structure of a ternary catalytic complex of HIV-1 RT with a substrate (dTTP) and the DNA-primer and template (Huang et al., 1998) (Fig. 2A) has provided the structural basis of drug resistance. Several mutations causing resistance to NRTIs turn out to be located close to the nucleoside binding site (Fig. 2B), indicating that this type of resistance is related to a modification of the molecular recognition between phosphate and/or nucleobase-sugar moieties (Zhang et al., 1994) and the enzyme.

First principles quantum mechanics offers powerful techniques to study biochemical problems. Detailed quantum mechanical treatments of molecular and electronic structure are emerging as tools for the prediction of molecular geometry, reaction paths of enzymatic transformations, electrostatic effects, and interpretations of spectroscopic probes of biomolecular environments at very accurate levels. Ab initio calculations are also increasingly being applied to systems of pharmaceutical relevance, helping to shed light on intricate and subtle aspects of substrate- and drug-enzyme intermolecular interaction (Cheney et al., 1988; Beveridge & Heywood, 1993; Perakyla & Pakkanen, 1995, 1996; Beveridge, 1996; Alber et al., 1998; De Santis & Carloni, 1999; Piana & Carloni, 2000).

![Fig. 1. Chemical formulas of selected nucleotide analogs binding to HIV-1 RT currently in clinical use for anti-AIDS therapy.](image1)

![Fig. 2. A: dTTP–DNA–primer/template–HIV-1 RT p66 ternary complex (Huang et al., 1998). dTTP and residues important for triphosphate binding are represented as balls and sticks. DNA strands as thin lines. The p66 unit is 560 residues long, with its DNA polymerase domain (blue, right) and Rnase H functional domain (red, left), respectively. B: In the dTTP binding site pI H-bonds to Lys65, Arg72 side chains, and Asp113, Ala114 backbone amides. Figures were made with MOLSCRIPT (Kraulis, 1991).](image2)
Here, we use ab initio quantum chemical methods to study the nature of the intermolecular interactions between HIV-1 RT and nucleosidic phosphate, which plays a major role in drug resistance (Zhang et al., 1994) for the AZT and dideoxy-type of drugs (d4l, ddC). Our aim is to gain insights into the functional role of key amino acids essential for nucleoside binding, as well as to provide information on the stability and protonation state of the triphosphate moiety.

Our calculations are carried out within the Car–Parrinello ab initio molecular dynamics scheme (Car & Parrinello, 1985). This approach appears favorable in several respects. Being based on the density functional theory with gradient corrections, it allows a rather accurate description of intermolecular forces (such as H-bonding) and of charge distribution in relatively large systems, as required here. Furthermore, it fully incorporates temperature effects, which are crucial for biological systems (Brooks et al., 1988). Finally, it has already shown to reproduce accurately structural and dynamical properties of phosphate moieties (Alber et al., 1999a, 1999b) and of enzyme active sites (Alber et al., 1998; De Santis & Carloni, 1999; Piana & Carloni, 2000).

**Computational procedure**

**Structural models**

Our starting model was the catalytic HIV-1 RT complex, whose crystal structure has been solved at 3.2 Å resolution (Huang et al., 1998) (Protein Data Bank (PDB) entry code: 1RTD) (Fig. 2A), immersed in water.

The initial positions of water molecules were constructed as follows: the ternary complex of HIV-1 RT p66 with substrate (dTTP) and the DNA-primer and template (Huang et al., 1998) was immersed in a 10 Å thick shell of waters. Sodium counterions were added to ensure electroneutrality. Water molecules and protein polar hydrogens were equilibrated by performing 100 ps of classical MD simulation using the GROMOS96 suite of programs (Van Gunsteren et al., 1996) at 300 K. (All the heavy atoms of the protein/DNA complex were kept constrained at their X-ray positions.) The united atom GROMOS-A force field was used. Constant temperature simulation was achieved by coupling the system with a Berendsen’s thermal bath (Berendsen et al., 1984) with a relaxation time of 0.2 ps.

Water molecules interacting to the biomolecule exhibited residence times of few tens ps, which is a much longer timescale than that of our Car–Parrinello calculations. All water molecules, which were within a distance of 5 Å to the triphosphate at the last MD snapshot, were included in our quantum chemical model. We may expect that adding second-shell water molecules to the system will not change qualitatively our results.

At the present stage, it is impossible to perform ab initio simulations on the entire protein. Thus, our investigation was limited to a portion of the triphosphate (tP)-binding site (Fig. 2B), namely to complexes I and II (Fig. 3A). This choice appears to capture all the crucial interactions with the triphosphate moiety (the focus of the present work), yet it is still manageable from a computational point of view. The effect of the environment is then estimated by including the protein electrostatic field in the calculations.

Complexes I and II were made up of several ingredients (Fig. 3A).

**The sugar-triphosphate moiety**

The protonation state of tP is not known. In aqueous solution, the primary phosphate ionization state (one negative charge per phosphate unit) prevails up to a pH of around 7, where the last proton is removed from the terminal phosphate group (Saenger, 1983). Thus, both the fully deprotonated (I) and the mono-protonated (II) forms could exist in the protein at physiological pH. The proton in II was located on the OC3(tP) atom of the γ-phosphate

---

Fig. 3. A: Model complexes I, II used in the quantum-mechanical calculations. They include: (1) dTTP sugar–triphosphate moiety; (2) the two Mg$^{2+}$ ions; (3) Asp110, Asp185 side chains (modeled as formiate anions), Lys65 (modeled as methylammonium cation) and Arg72 (modeled as guanidinium cation) side chains; (4) Val111, Gly112, Asp113 protein backbones; (5) Ala 114; (6) selected water molecules; (7) primer 3'-hydroxy group modeled as methanol. For the sake of clarity, H atoms (except those of Asp110 and Asp185) are not shown. B: Atom labeling of tP binding site. Mg$^{2+}$ ion–donor atom coordination bonds are represented in dashed lines. For the sake of clarity, only H-bonds formed by tP and waters WAT1, WAT3, and WAT7 are shown (dotted lines). It is fully deprotonated; II is protonated on the OC3 atom.
group (Fig. 3B) as (1) the latter has higher pK_a (pK_a = 7.5) than the α- and β-moieties (Saenger, 1983); (2) protonation of γ-phosphate-OC1 and γ-phosphate-OC2 atoms partially disrupts H-bond patterns and Mg(II) coordination, respectively.

Two Mg(II) coordination polyhedra

These are composed of water molecules and the essential Asp110 and Asp185 residues (Kaushik et al., 1996), modeled as formate anions (Fig. 3A) and the primer-hydroxy group. The latter, which was not present in the X-ray structure, was added according to standard bond lengths and angles. [Asp186, the third member of the “Asp triade” in the active site (Kaushik et al., 1996) was not included, as it neither binds to the Mg^{2+} ions nor to the substrate (Larder et al., 1987; Kaushik et al., 1996).]

Groups interacting with the triphosphate anion

Arg72 and Lys65 side chains (modeled as guanidinium and methylammonium cations, respectively); the entire Ala114 residue; Val111 carbonyl group and Gly112, Asp113, Tyr115 backbone units.

Selected (16) water molecules

These are present in the active site channel from the classical simulation and located close to the tP (see above).

Charges of I and II were 0 and +1, respectively.

Calculations

Quantum-chemical calculations were performed within the framework of the density functional theory (DFT). Gradient corrected exchange and correlation functionals with the Becke (1988) and Lee-Yang-Parr (Lee et al., 1988) parameterizations were used. The basis set consisted of a plane-wave expansion up to an energy cutoff of 70 Ry. Only the Γ-point of the Brillouin zone was included. Ionic cores and valence electrons interactions were described by Troullier–Martins type (Troullier & Martins, 1991) pseudopotentials. Isolated system conditions were used as in Barnett and Landman (1993).

Car–Parrinello DFT-based MD simulations (Car & Parrinello, 1985) of complexes I and II were performed for 1.5 ps at 298 K. A time step of 0.121 fs and a fictitious electron mass of 1,000 a.u. were used. Constant temperature simulations were achieved by coupling the systems with a Nosé thermostat (Nosé, 1984) at 300 K with a frequency of 500 cm^{-1}.

The protein–DNA external potential Φ(r) at point r is evaluated according to:

$$
\Phi_{prot}(r) = \sum_{i} \frac{q_i}{|r_i - r|}
$$

where q_i represents the RESP (Cornell et al., 1993) atomic point charges of atom i with coordinates r_i, following the procedure of our previous work (De Santis & Carloni, 1999; Piana & Carloni, 2000). Charges on protein residues that are included in our quantum model have been set to zero. The protein/DNA-electrostatic potential is included in the electron/nucleus attraction integral of the Kohn–Sham Hamiltonian. The calculation with the external potential is carried out for a single-point calculation of a specific conformation of I. This conformation was obtained by geometry optimization in vacuo.

All calculations, performed with the parallel version of the Car–Parrinello (Car & Parrinello, 1985) code CPMD V3.0h (Hutter et al., 1995), required 500 h on the 64-node Cray T3E supercomputer.

Calculated properties

The electronic structure was characterized in terms of centers of maximally localized Wannier orbitals (WOc) (Silvestrelli et al., 1998; Silvestrelli & Parrinello, 1999a; Mazzini & Vanderbilt, 1997), recently implemented in the Car–Parrinello scheme (Silvestrelli & Parrinello, 1999a). The relative position of these centers have been shown to be related to change in Pauling electronegativity of two atoms forming a bond and thus provide an estimate for the bond polarity (Alber et al., 1999b). Bond polarity indices (BI) were calculated following a procedure of Alber et al. (1999b).

The WOCs were also used to calculate molecular electric dipole moments. Following the procedure of Silvestrelli and Parrinello (1999b), we assume that the electronic charge distribution is concentrated in point charges located at the WOCs. This allows separation of the charge distribution in contributions from distinct elements. Changes of dipole moments can be a measure of polarization effects due to the molecular interactions in our complexes.

Velocity autocorrelation functions (VAC) were calculated for the atoms of selected bonds (OH bonds in the water molecules, the PO and OH bond of the γ-phosphate and the N-H bonds in K65) as in Alber et al. (1999a) and Allen and Toldesley (1986), where \( t_{max} = 0.6 \) ps. The corresponding power spectrum was obtained as Fourier transform of the VAC. The frequency interval of the transformation was chosen as 10 cm^{-1}. The VAC was calculated over the trajectory of the last 1.2 ps. The dependence of frequency sampling was investigated by calculating the VAC in the (0.3–0.9) and (0.9–1.5) ps time ranges. The maximum difference among all calculated frequencies turned out to be as small as 20 cm^{-1}.

H-bond frequency (HbF) is calculated as the number of H-bonds for each atom normalized to the total number of trajectory steps; that is, HbF = 1 implies that a H-bond is always maintained during the dynamics. Hydrogen bonds are assumed if d(hydrogen – H-bond acceptor) ≤ 2.5 Å and 120 ≤ θ(H-bond donor – hydrogen H-bond acceptor) ≥ 180.

Results

In this section, we describe the structural and dynamical properties of complexes representing the triphosphate binding site in HIV-1 RT in different protonation states, the fully deprotonated (I) and
the monoprotonated (II) forms, in the absence or in the presence of the protein electrostatic field.

Fully deprotonated form (complex I)

Triphosphate (tP) conformational stability and H-bond network at the active site

The relative orientation of phosphate to the ribose is conveniently described in terms of the angle $\xi$ between atoms OC3'(ribose), PA(tP), and PB(tP), and distance $D$ between atoms OC3'(ribose) and PC(tP) (see Fig. 4A). After 0.5 ps $\xi$ and $D$ fluctuates around average values (Fig. 4B; Table 1). Most intermolecular enzyme-tP contacts are maintained during the dynamics: Table 2 shows that tP keeps its H-bonds with Arg72, Lys65, Asp113, and Ala114 backbone.

Our simulation also allows to establish the interactions of the ribose-3'-hydroxy group. Table 1 indicates that a stable intramolecular hydrogen bond to the $\beta$-phosphate unit is formed during the entire length of the MD simulation (Table 2). The rotational fluctuations around the 3'-OH in the ribose 3'-hydroxy group allows for H-bonding either to OB1 or OB2.

tP also interacts extensively with water molecules, forming and breaking several H-bonds during the dynamics (Table 2). Table 3 compares selected properties of the strongly bound water molecules (namely WAT1, WAT3, and WAT7, with time averaged H-acceptor distances less than 2 Å) with those obtained from simulations of water in the bulk (Silvestrelli & Parrinello, 1999b) and in the gas phase (Silvestrelli & Parrinello, 1999b). The MD averaged O—H bonds of these waters are significantly longer than those in vacuo and slightly longer than those in bulk, presumably as a result of stronger H-bond interactions (Jeffrey & Saenger, 1991). Consistently, these ordered waters also cause a decrease of the O—H stretching frequency with respect to water in bulk (Table 3).

A good indication of polarization effects on the water molecules, and thus a qualitative estimate on its electrostatic binding contribution, is given by the size of the molecular dipole moment of the waters. Therefore, we have calculated the molecular dipole moments of all waters derived by the maximally localized Wannier functions (see Computational procedure section) (Silvestrelli & Parrinello, 1999b). For our analysis we have used the geometry optimized structure of a MD snapshot.

The averaged dipole moment of all water molecules in our model turns out to be smaller than for water in bulk ($\mu_{\text{water}}$ = 2.6 Debye, in comparison to a recent investigations on water in bulk $\mu_{\text{bulk}}$ = 3.0 Debye). This small value arises from contributions of some of the waters relatively far from the biomolecule, whose H-bond valences are partially unsaturated ($\mu_{\text{water}}$ = 2.5 Debye). In contrast, waters coordinated to Mg$^{2+}$ and/or binding to phosphate show relatively large polarization ($\mu_{\text{water}}$ = 3.0 Debye). Waters that contribute significantly to tP stabilization (namely WAT1, WAT3, and WAT7) are more polarized in comparison to the averaged water dipole moment. In particular WAT3 reveals a dipole moment as large as $\mu_{\text{WAT3}}$ = 3.2 Debye, due to the strong electric field of Arg72 side chain and $\beta$-phosphate charged groups. As a result, WAT3 is very well ordered, and it fluctuates around an average position. Due to its electrostatic and well-ordered structure, we would expect significant contribution to tP recognition.

However, in the crystal structure, no water molecules are resolved at the present resolution. Inspection of the X-ray structure (Huang et al., 1998) suggests that a water at the favorable WAT3 position would be accommodated in a small cavity formed by tP and the side chains of Arg72, Lys65, Gln151, and Phe116.
Presumably the Gln151 residue is positioned to further hydrogen bond with this water molecule (Fig. 5).

**Lys65-tP interactions**

The Lys65-γ-phosphate salt bridge is well maintained during the dynamics (Tables 1, 2). The salt bridge is characterized by very short Lys-tP distances (Table 1); interestingly, proton hopping between OC1(tP) and ϵ(Lys65) occurs several times in the ps timescale (Fig. 6A). Each proton transfer event is correlated to local minima in the (OC1(tP)-Ne(Lys65) distance.

This type of interaction, characterized by fast proton hopping between H-bond donor and acceptor, is described as “low-barrier hydrogen bond” (LBHB). Theoretical work (Piana & Carloni, 2000) and suggestions from experimental studies (Rodriguez et al., 1993) have pointed to the existence of a LBHB in the HIV-1 protease cleavage site, for the biological function of serine proteases (Lin et al., 1998; De Santis & Carloni, 1999) and in enzyme model substrates (Schiotto et al., 1998).
LBHBs are characterized by decreased donor⋯H stretching frequencies. Our calculations provide a value of 2,400 cm\(^{-1}\) for both \(\nu_{\text{OH}}\) and \(\nu_{\text{NH}}\) stretching vibrational modes, in good accord with experimental data for a variety of LBHBs (Perrin & Nielson, 1997).

To provide a picture of the LBHB chemical bonding, we calculate the electron localization function (ELF) (Becke & Edgcombe, 1990; Silvi & Savin, 1994), which has been shown to be a useful tool to study the nature of the chemical bond in LBHBs (De Santis & Carloni, 1999; Schiott et al., 1998). Figure 6C–E plots the projection of ELF on the plane of \(\text{Ne}(\text{Lys65})\cdot\cdot\cdot\text{H}_2\cdot\cdot\cdot\text{OC1(tP)}\) atoms before (Fig. 6C), during (Fig. 6D), and after (Fig. 6E).

### Table 3. Comparison of selected properties of WAT1, WAT3, WAT7 in model I with isolated and liquid water

<table>
<thead>
<tr>
<th></th>
<th>Ab initio MD (WAT1, WAT3, WAT7)</th>
<th>Optimized structure, isolated water (Silvestrelli &amp; Parrinello, 1999b)</th>
<th>Ab initio MD liquid water (Silvestrelli &amp; Parrinello, 1999b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D(O-H)</td>
<td>1.00(0)</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>(\nu_{\text{OH}})</td>
<td>2,700–3,000</td>
<td>3,300(^{a})</td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\)Distances are reported in Å; vibrational frequencies in cm\(^{-1}\).

\(^{b}\)Estimated from the \(\text{D}_2\text{O}\) vibrational spectrum (Silvestrelli et al., 1997).

---

**Fig. 6.** tP-Lys65 interactions. \(\text{NeH}_2(\text{Lys65})\cdot\cdot\cdot\text{OC1(tP)}\) distances as a function of time in (A) I and (B) II. (C–E) Electron localization function (ELF) projected on the OC1⋯H⋯Ne plane (C) before, (D) during, and (E) after the proton transfer in the LBHB (complex I). The ELF ranges from blue (0.20) to red (0.98). For the sake of clarity only the Lys65 and γ-phosphate atoms are shown. (F) Three-centered molecular orbital in the transition state involving the proton shared by OC1(tP) and Ne(Lys65). Isodensity contour level corresponds to \(-5.5\) a.u. (yellow) and 5.5 a.u (green).
proton transfer. Red areas indicate spatial regions of strong electron localization (ELF close to 1), in which the Pauli exclusion principle has little influence on the electron distribution. In the initial state (Fig. 6C), high values of ELF on atom Ne(Lys65) indicate the electron lone pair, the red area between OC1(tP) and Hα specify the covalent chemical bond. In the “transition state” (Fig. 6D), chemical bonding changes dramatically. The significant increase of ELF between atoms Hα and Ne(Lys65) denotes the formation of a new covalent bond, while ELF is still significant (and, hence, the bonding highly covalent) along the OC1(tP)···Hα bond. Calculation of the electronic structure in terms of Kohn–Sham one-electron levels provides a consistent and an independent evidence of a Ne(Lys65)···Hα···OC1(tP) covalent interaction. A covalent σ-type molecular orbital stabilizes the transition state (Fig. 6F), confirming the ideas of Gilli and Cleland (Cleland, 1992; Schiott et al., 1998).

Environment effects
Our MD simulations do not take into account the presence of the outer protein electrostatic field. To estimate its effect, we have analyzed the electronic structure of a geometry optimized MD snapshot in terms of maximally localized Wannier centers (WOC) (see Computational procedure section). We compare the geometrical variations of the WOC along a chemical bond as a result of the presence of the protein/DNA electrostatic potential. This analysis has been shown to be a precise indicative of environmental effects on the electronic structure and can be related to changes in bond polarity (Alber et al., 1999b). We focus here on changes of electronic structure of the LBHB and of WAT3.

Figure 7 compares the positions of the WOCs with and without the inclusion of the protein electrostatic field. We observe only a minimal shift of the WOC along the Ne(Lys65)–Hα bond that corresponds to only a small variation in bond polarity due to the effect of the electric field (Fig. 7). We calculate the difference in bond polarity to be as small as ΔB_NeHα(N) = 0.002, which corresponds to only a very small change in Pauling electronegativity (∼0.02 units) (Alber et al., 1999b). In comparison, the difference in BI within all three Lys65–NH bonds is as much as ΔB_NH(N) = 0.114. Thus, inclusion of the protein electric field appears not to affect the electronic properties in the Ne(Lys65)···Hα···OC1(tP) interactions.

We note also only minor changes in electronic structure of WAT3, and consequently, the size of the molecular dipole moment is not affected by the inclusion of the protein electrostatic field.

Magnesium coordination
After 0.3 ps, both magnesium(II) ions are stable octahedral coordinated. Whereas the complete coordination shell of Mg(2) is already known from the X-ray structure, the exact position and coordination of Mg(1) are unknown, as the primer 3’-OH terminus, which is expected to be coordinated to Mg(1) is included in the model and it is not present in the X-ray structure (Huang et al., 1998).

Therefore, significant structural arrangements in the Mg(1) position take place, as demonstrated by the large difference of the time-averaged position in comparison to X-ray data (Table 1).

The Mg(1)–OA2(tP) coordination distance is reduced during our simulation up to a value comparable to the other Mg(1)-oxygen distances [average 2.2(0.1) Å] in the same complex.

The coordination shell includes OD1B(Asp110), OD2A(Asp185), OA2(tP), two water molecules (WAT2, WAT9), and the primer 3’-OH terminus (Fig. 8A). WAT2 hydrogen bonds also to the α-phosphate and could play a role in phosphate binding (Fig. 8A). During the dynamics Mg(2) is coordinated to OA2(tP), OB2(tP), OC2(tP), OV(Val111), OD2B(Asp185), and OD1B(Asp110) (Fig. 8B), as already shown in the X-ray structure.

Vibrational frequencies can probe the binding of metal ions (Wang et al., 1998). The non-ester P–O(tP) stretching frequencies are calculated to decrease by ≈200 cm⁻¹ upon binding to Mg²⁺, in agreement with experimental data (Wang et al., 1998).

Monoprotonated form (complex II)
Protonation of the γ-phosphate in model II has a dramatic effect on the stability of the complex, already in the very short time scale investigated.

H-bond pattern and conformational flexibility
The interactions between the protein and tP are highly affected by the protonation of the γ-phosphate. H-bonds are partially or totally lost between the β-phosphate and the enzyme backbone Ala114 and Asp113 and residue Arg72 as well as between the β-phosphate and the C3’-OH group (Tables 1, 2).

As a consequence, the D and ξ geometric parameters are significantly larger than those of model I, and the tP conformation differs largely from the X-ray structural template (Fig. 4B; Table 1).

Solvent structuring is similar to I. Several waters bind to tP, among which only WAT3 is particularly well ordered in both complexes, fluctuating around an average position during the dynamics. This is due presumably to strong interactions with Arg72 and tP charged groups.

Lys65–tP interactions
No proton exchange takes place between Lys65 and the γ-phosphate during the timescale investigated (Fig. 6B); instead, Lys65 turns out to interact preferentially with the β-phosphate

\[ γ-phosphate \]
\[ Lys65 \]

Fig. 7. Centers of maximally localized Wannier orbitals (WOC) calculated for a geometry optimized MD snapshot of complex I (see Computational procedure section). Positions of WOCs are depicted as yellow spheres in vacuo and green spheres with the inclusion of the protein electric field.
H-bond acceptors OBC(tP) and OB1(tP) (Tables 1, 2), which causes large structural rearrangements.

Overall, the structural differences to the X-ray structure are significantly larger than for I. The distance between the Arg72 and Lys65 side chain is rather large compared with that of the X-ray structure and the MD-averaged distance in I. This is presumably a consequence of the unfavorable electrostatic Lys65–Arg72 interactions.

**Magnesium coordination**

The Mg(1) averaged position is rather similar to that in complex I (Table 1). However, after 0.7 ps of simulation, we observe some rearrangements in its coordination sphere, which results in the loss of a water ligand (WAT9) (Fig. 8C). This rearrangement is initiated by structural variations of the Asp185 group that compensates for the octahedral coordination of Mg(1). Presumably, this is a consequence of the larger conformational fluctuations in tP in model II, which affects also the positions of the charged Asp residues. The coordination polyhedra for Mg(2) is stable and similar to complex I (Fig. 8D), whereas its averaged position deviates significantly from the X-ray position (Table 1) as a result of the large structural arrangement in tP conformation.

In conclusion, our calculations suggest that the triphosphate–enzyme pocket complex in its X-ray structural conformation is much more flexible if the γ-phosphate is protonated, and that intermolecular interactions provide less stabilization of the adduct in terms of H-bonding.

**Discussion**

Our quantum-mechanical calculations, carried out for the fully deprotonated form (complex I) and the monoprotonated form (in the γ position) (complex II), provide information on chemical bonding and intermolecular interactions of the substrate–triphasate moiety at the HIV-1 RT nucleoside binding site.

Within the limitations of the relatively small model (such as the limited representation of hydration) of the short timescale explored, we find a high sensitivity of tP conformational properties on changes in the chemical environment of the γ-phosphate.

The fully deprotonated form of tP turns out to be the more stable and is in good agreement to the HIV-1 RT X-ray structure. The conformational stability of the triphosphate (tP) anion could be achieved not only—as expected—by strong electrostatic interactions with the positively charged enzyme pocket, but also by means of previously unrecognized fast proton exchange with the highly conserved Lys65 residue (Barber et al., 1990). Proton hopping between the OC1(tP) and NE(Lys65) atoms occurs in the sub-ps timescale and is characterized by a very short heteroatomic...
distance (Table 1). Analysis of chemical bonding, in terms of electron localization functions (Fig. 6C–E; Becke & Edgecombe, 1990) and Kohn–Sham one-electron levels (Fig. 6F), indicates that the LBHB is covalent (and thus highly directional) in nature. Similar findings have recently been reported for other enzyme models (Schiott et al., 1998; De Santis & Carloni, 1999), confirming the results obtained with an atoms in molecule analysis (Schiott et al., 1998). However, we have to mention that a quantitative description of the proton motion in the LBHB would require a quantum treatment of the proton (Marx et al., 1999), an explicit treatment of the protein and thus may be an important factor in stabilizing the bioactive conformation.

Inclusion of the external protein electrostatic field appears not to affect the electronic properties in the Lys65–phosphate interactions. As the description of the motion of the protons (as well as of all the other atoms of the system) are classical, the quantum nature of proton dynamics is expected to modify the picture quantitatively but not qualitatively.

The energetic contribution of the LBHB in binding is difficult to assess. The increased covalent bonding contribution between Lys65 and the γ-phosphate is accompanied by partial charge neutralization, and thus could be easily counterbalanced by the expected loss of electrostatic stabilization in the binding pocket. However, a large degree of covalency is expected to be strongly directional and thus may be an important factor in stabilizing the bioactive conformation.

A crucial point regards the reliability of our computational approach in describing the LBHB. To the best of our knowledge, these are the first calculations pointing to a LBHB between a positively charged group and thus may be an important factor in stabilizing the bioactive conformation.

Thus, the directionality to the covalent Lys65–tP interaction (Table 1) could play a prominent role in stabilizing the bioactive phosphate conformation.

Protonation of γ-phosphate largely affects the conformational properties of tP and Lys65 (complex II in Table 1), as a consequence of the loss of the covalent LBHB and of the decrease of electrostatic interactions between tP and the protein active site positively charged groups. The deprotonated form turns out to be the most stable in the enzymatic pocket of HIV-1 RT and, therefore, our calculations point to the presence of this form in the active site of HIV-1 RT. Notice that although the fraction of the unprotonated form at physiological pH in aqueous solution is only ~30% (Saenger, 1983), in the protein the pKₐ of phosphate might increase because of the strong interactions with a variety of positively charged groups (Mg²⁺, Lys65, Arg72).

Water molecules, not detected in the X-ray structure, form a complex H-bond network at the active site. A well-orderd water (WAT3 in Fig. 5), highly polarized by the environment, appears to play a role in recognition of tP, by bridging the latter with Arg72 and presumably Gln151. Mutation of the latter with nonpolar residues could affect the WAT3 H-bond network and may be a factor in the well-known resistance of mutation Gln151→Met toward dideoxy type drugs (ddI, ddC, and 3TC) and AZT (Shirasaka et al., 1995).

In conclusion, our quantum chemical calculations reveal a variety of functionally important characteristics of tP/HIV-1 RT interactions, which cannot be discerned by visual inspection of the molecular structure. Several ingredients, such as the polarization forces, the treatment of bond-forming/breaking processes in the LBHB and temperature effects might play a critical role for substrate binding and possibly for drug-resistance mechanisms.

Acknowledgments
This work was funded by the ICARUS project of the European Commission, and the calculations were performed at CINECA in Bologna, Italy. We thank Sergio Pantano for running the GROMOS96 program.

References


