A peer-reviewed version of this preprint was published in PeerJ on 2 December 2019.

View the peer-reviewed version (peerj.com/articles/pchem-6), which is the preferred citable publication unless you specifically need to cite this preprint.

Spectrometric and computational studies of the binding of HIV-1 integrase inhibitors to viral DNA extremities

Lea El Khoury 1, Krystel El Hage 2, Jean-Philip Piquemal 3, Serge Fermandjian 4, Richard Maroun 5, Nohad Gresh Corresp., 6, Zeina Hobaika 5

1 Pharmaceutical Sciences, University of California, Irvine, Irvine, California, United States
2 Laboratoire Structure-Activite des Biomolecules, Université d’Evry Val d’Essonne, Evry, Essonne, France
3 Laboratoire de Chimie Théorique, Université Panthéon-Sorbonne (Paris I), Paris, France
4 UFR Biomedicale, Université René Descartes (Paris V), Paris, France
5 Faculté des Sciences, Université Saint Joseph, Beirut, Lebanon
6 Laboratoire de Chimie Théorique, Sorbonne Universite, Paris, France

Corresponding Author: Nohad Gresh
Email address: gresh@lct.jussieu.fr

Three Integrase (IN) strand transfer inhibitors are in intensive clinical use, raltegravir, elvitegravir and dolutegravir. However, the onset of IN resistance mutations limits their therapeutic efficiency. As put forth earlier, the drug affinity for the intasome could be improved by targeting preferentially the retroviral nucleobases, which are little, if at all, mutation-prone. We report experimental results of anisotropy fluorescence titrations of viral DNA by these three drugs. These show that the ranking of their inhibitory activities of the intasome corresponds to that of their free energies of binding, D Gs, to retroviral DNA, and that such a ranking is only governed by the binding enthalpies, D H, the entropy undergoing marginal variations. This ranking can therefore be directly correlated to that of model Quantum Chemistry (QC) calculations of intermolecular interaction energies of the sole halobenzene ring with the highly conserved retroviral nucleobases G4 and C14, using Density Functional Theory. This DE(QC) ranking is in turn reproduced by the corresponding DE tot values computed with a polarizable molecular mechanics/dynamics procedure, SIBFA (Sum of Interactions Between Fragments Ab initio computed). Such validations should enable polarizable molecular dynamics simulations on more potent inhibitors in their complexes with the complete intasome. Such derivatives should principally encompass modified halobenzene rings.
Spectrometric and computational studies of the binding of HIV-1 integrase inhibitors to viral DNA extremities

Léa El Khoury \( ^{a, b, \Phi} \), Krystel El Hage \( ^{b, c, \Psi} \), Jean-Philip Piquemal \( ^{a, d, e} \), Serge Fermandjian \( ^{c} \), Richard G. Maroun \( ^{b} \), Nohad Gresh \( ^{a*} \), and Zeina Hobaika \( ^{b*} \)

\( ^{a} \) Laboratoire de Chimie Théorique, Sorbonne Université, UMR7616 CNRS, Paris 75005, France

\( ^{b} \) UR EGP, Centre d’Analyses et de Recherche, Faculté des Sciences, Université Saint-Joseph de Beyrouth, B.P. 11-514 Riad El Solh, Beirut 1107 2050, Lebanon

\( ^{c} \) Chemistry and Biology Nucleo(s)tides and Immunology for Therapy (CBNIT), UMR 8601 CNRS, UFR Biomedicale, Paris 75006, France

\( ^{d} \) Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX 78712, United States

\( ^{e} \) Institut Universitaire de France, Paris Cedex 05, 75231, France

\( ^{\Phi} \) Present address: Department of Pharmaceutical Sciences, University of California, Irvine, California 92697, United States

\( ^{\Psi} \) Present address: Laboratoire Structure-Activité des Biomolécules Normales et Pathologiques, Université Evry-Val d’Essonne, Evry, France, INSERM, U829, Evry, France.

*Corresponding authors Nohad Gresh (SU), gresh@lct.jussieu.fr; Zeina Hobaika (USJ): zeina.hobaika@usj.edu.lb
Abstract. Three Integrase (IN) strand transfer inhibitors are in intensive clinical use, raltegravir, elvitegravir and dolutegravir. However, the onset of IN resistance mutations limits their therapeutic efficiency. As put forth earlier, the drug affinity for the intasome could be improved by targeting preferentially the retroviral nucleobases, which are little, if at all, mutation-prone. We report experimental results of anisotropy fluorescence titrations of viral DNA by these three drugs. These show that the ranking of their inhibitory activities of the intasome corresponds to that of their free energies of binding, $\Delta Gs$, to retroviral DNA, and that such a ranking is only governed by the binding enthalpies, $\Delta H$, the entropy undergoing marginal variations. This ranking can therefore be directly correlated to that of model Quantum Chemistry (QC) calculations of intermolecular interaction energies of the sole halobenzene ring with the highly conserved retroviral nucleobases G4 and C14, using Density Functional Theory. This $\Delta E(QC)$ ranking is in turn reproduced by the corresponding $\Delta E_{tot}$ values computed with a polarizable molecular mechanics/dynamics procedure, SIBFA (Sum of Interactions Between Fragments Ab initio computed). Such validations should enable polarizable molecular dynamics simulations on more potent inhibitors in their complexes with the complete intasome. Such derivatives should principally encompass modified halobenzene rings.

I. Introduction.

HIV-1 Integrase (IN) is a key element in viral replication. In addition to its essential role in viral DNA (vDNA) integration into host genomic DNA, IN appears involved, directly or indirectly, in reverse transcription [1], nuclear import [2-3] and HIV-1 particle maturation [4-5]. Since in addition IN has no counterpart in human cells, it could represent a privileged target for the design of potent antiretroviral drugs.

The integration step is carried by a multimer of INs assembled on vDNA ends, referred to as the intasome. In a first step, denoted as 3'-processing, a 3'GT dinucleotide is removed from each end of the Long Terminal Repeats (LTRs) of vDNA. This occurs in the cytoplasm within a multi-component pre-integration complex (PIC) which gathers the vDNA and several viral and cellular proteins. DNA strand transfer occurs in a second step, after the PIC is chaperoned into the nucleus and results in integration of vDNA as a provirus into the host genome. This requires cutting of two phosphodiester bonds five base pairs apart on opposite strands of the host DNA and is done by free 3'-OH groups that were liberated following LTR processing [6-8].
IN strand transfer inhibitors (INSTIs) proved to be much more effective than processing inhibitors and enabled the development of a successful class of antiretroviral drugs [9]. Three inhibitors inspired by the original diketo acids (DKAs) have been successively approved and are commonly used in HIV-1 treatment, namely raltegravir (RAL; MK-0518), elvitegravir (EVG; GS-9137) and dolutegravir (DTG; S/GSK1349572) [10-13]. The latter is a second generation INSTI aimed at maintaining an effective efficacy to IN variants resistant to RAL and EVG [14-15].

As was the case for protease and reverse transcriptase inhibitors [16-17], IN inhibitors can generate several resistance mutations. These were recently reported following RAL, EVG and DTG treatment [18-21]. They affect, not only IN residues in direct interactions with INSTIs but also "outer-shell" residues indirectly bound [14, 22-23]. Among these are double mutations such as Q148R/H coupled to G140S/A that produce important synergetic effects on the efficacy of RAL, EVG and even DTG [24-25].

EVG, RAL and DTG (Figure 1) selectively bind at the interface of IN and the viral DNA ends, within the intasome and have in common two distinct structural motives: a) a large centralized pharmacophore contributing its keto oxygen and a coplanar neighboring oxygen to coordinate both two IN catalytic Mg(II), structural water molecules, and, either directly or through water, IN residues; and b) a halobenzyl group targeting the highly conserved 5'CpA 3'/5'TpG 3' step on the viral DNA ends [9]. However, the surface and oxygen arrangement of the central pharmacophore differ among the three INSTIs, as well as the nature and position of the halogenation of their terminal aromatic ring: a single F is attached in para to RAL's halobenzene ring, whereas EVG has an F in ortho and a Cl in meta, and DTG has two F atoms in ortho and para.

Therefore, within the intasome, the binding of all three drugs targets both viral DNA and the viral protein. In addition to the well-known established interactions with the catalytic and non-catalytic site of IN, following the 3' processing reaction, the X-ray structures of IN-DNA-inhibitor complexes (Figure 1B) show that the halogenated benzene ring stacks over the C base (C14) upstream, while the C-X bond points toward the center of the G base (G4) downstream of the second strand [26-28].

On another note, it was reported that, compared with RAL and EVG, DTG displays a more potent in vitro anti-HIV activity and a distinct resistance profile [14-15, 29-30]. Furthermore, we have ourselves reported that an increase in the drug-vDNA complex stability correlates with an increase in drug activity and a decrease in viral resistance [14, 27-28], highlighting the important contribution of the vDNA end recognition for the binding affinity of INSTIs to the intasome [31-
In this connection, we have shown a narrow correlation between the strongest DTG-DNA affinity and DTG’s highest barrier to resistance [32].

Maximizing shape complementarity at the IN-vDNA-inhibitor surface could serve as a guiding principle for the development of new INSTIs [9, 26, 31, 35]. An important feature of second generation INSTIs consists into their increased contacts not only with IN active and non-active site, but also with processed viral DNA, right before the strand transfer step. This is supported by the fact that the sequence of the nucleic bases at the ends of the LTR represents stringent requirements concerning retroviral integration and that there is no evidence of mutations in the LTRs that could lead to resistance to INSTIs [35-38].

Could attempts to design novel INSTI’s with enhanced affinities focus on the viral DNA, which would render the new inhibitors much less sensitive to mutations occurring on the IN protein? As a continuation of our previous studies, we experimentally measure the ΔG values for the binding of RAL, EVG and DTG to viral DNA. This is done by fluorescence anisotropy experiments at three different temperatures. We also analyze, by both ab initio quantum chemistry (QC) and polarizable molecular mechanics/dynamics, the intermolecular interaction energies of their halobenzyl rings with G4 and C14. The individual energy contributions of ΔE(QC) are also compared to their SIBFA counterparts. Such analyses and validations of inhibitor interactions within the core of vDNA binding site constitute a necessary step toward long-duration PMD of drug-intasome complexes.

An outstanding feature of the CX ring in halobenzenes, discovered on the basis of quantum chemistry [39, 40] is the existence of a zone of electron depletion along the extension of the bond with a magnitude increasing along the series F> Cl > Br > I. This ‘sigma-hole’ goes along with a zone of electronic build-up on a cone around the halogen. It has been earlier shown that atom-centered point charges used in ‘classical’ force-fields cannot account for the impact of the sigma-hole on the Coulomb electrostatic contribution E_C: this could be only partly remedied upon resorting to an additional fictitious atom prolonging the CX bond with a partial charge and a distance to the X bearer that have to be fit on the basis of QC calculations [41-43] On the other hand, anisotropic potentials such as SIBFA, with distributed atomic multipoles up to quadrupoles, were shown to closely account for the impact of the sigma-hole along the F, Cl, and Br series on the magnitude of E_C both along and around the CX- bond without extra calibration effort [44]. Along these lines, recent work showed that another possibility, resorting to a distributed charge model to reproduce the local quadrupole around the CX could also enable
to account for the impact of the sigma-hole on $E_C$ [45, 46]. An additional incentive to validate SIBFA in the present work is the perspective of its applications to the entirety of the inhibitor-INT complex, in which the presence of two Mg(II) cations close to one another and of structured water molecules render it preferable to resort to polarizable potentials than to ‘classical’, non-polarizable ones.

At this stage we do not intend to compute true $\Delta G$s to compare the binding affinities of the three inhibitors to vDNA and a fortiori to INT. These would only be meaningful at the outcome of long-duration Molecular Dynamics, and such an outcome could be very sensitive to the accuracy of the intermolecular potential. As a first step toward this evaluation, we focus here on the sole halobenzene-G4/C14 interactions expressed in terms of actual enthalpies of binding and evaluate if two points are satisfied: a) whether the ranking of $\Delta E(QC)$ intermolecular interaction energies parallels that of the experimental binding enthalpies; b) whether in turn the corresponding ranking of $\Delta E(SIBFA)$ values parallels the QC one.

II. Materials and Methods

1. DNA sample and inhibitors. The oligonucleotide LTR32 (Figure 2) was purchased from Eurogentec (Belgium). It was designed to adopt a folded double-stranded hairpin structure even under the low concentrations ($10^{-9}$ to $10^{-5}$ M) used in fluorescence anisotropy experiments. RAL, EVG and DTG were purchased from AdooQ and Medchemexpress, respectively and their structures are represented in Figure 1A.

LTR32 is a linear oligonucleotide sequence designed to adopt a double strand hairpin structure in solution upon folding around a loop created by a purposely added thymine triplet (TTT in green) with the sensitive fluorescein reporter (F, in orange) grafted to its central T. The latter allows fluorescence studies in solution at low concentrations. The stem that reproduces the 3' processed LTR end comprises a 17-nucleotide strand and a 15-nucleotide strand corresponding to the unreactive strand and the reactive strand, respectively. Their pairing leaves an unpaired dinucleotide 5' AC 3' at the 5' end on the unreactive strand. In each strand the nucleotide numbering goes from the 5' to the 3' extremity. The highly conserved doublet of base pairs, here numbered C14-G4 and A15-T3, is colored in red.
2. Fluorescence measurements. Thermodynamic parameters of ligand-processed DNA complexes were identified using fluorescence anisotropy [47-48] on a Jobin-Yvon Fluoromax II instrument. RAL was purchased from AdooQ and EVG and DTG were purchased from Medchemexpress. The LTR32 oligonucleotide, reproducing the processed viral DNA, was purchased from Eurogentec (Belgium). It contains a thymine loop bearing the fluorescein reporter for fluorescence studies. During titrations, labeled DNA (LTR32) was dissolved in phosphate buffer (10mM, pH 6, I=0.1) and placed in thermally jacketed quartz cells (1cm) at 5°C, 15°C and 25°C; increasing concentrations of the inhibitors (RAL, EVG or DTG) were then added. The excitation was recorded at 488 nm and the emission at 516 nm. The equilibrium dissociation constants (K_d) were determined with GraphPAD Prism 5 applying the non-linear regression (curve fit)-Least square procedure. This analysis led to the calculation of binding free energies using the following equation: ΔG = -RT ln (1/K_d).

3. PDB entries. In the used computational approaches, all complexes were extracted from the X-ray structures of the PFV intasome (IN-viral DNA-Mg^{2+}) in complex with EVG [26] (PDB code: 3L2U), DTG [28] (PDB code: 3S3M) and RAL [27] (PDB code: 3OYA).

4. Ab-initio QC Computations. The systems were first energy-minimized at the correlated level using the dispersion – corrected B97-D functional by Grimme et al. [49] and the cc-pVTZ basis set [50, 51] with the Gaussian 9 (G09) software [52]. During the minimization C4 and G14 were kept fixed and only the halogenated ring was allowed to move.

41. Energy decomposition analysis. Two sets of calculations were performed at the Hartree-Fock and DFT-d levels in order to obtain all five contributions to the interaction energy.

4.1.1. Hartree-Fock (HF) calculations. The energy decomposition analysis (EDA) was done using the reduced variational space (RVS) procedure of Stevens and Fink [53]. It separates the total interaction energy into four contributions: the first-order (E_1) Coulomb (E_C) and short-range exchange− repulsion (E_{exch}) and the second-order (E_2) polarization (E_{pol}) and charge transfer (E_{ct}). The basis set superposition error (BSSE) [54, 55] is evaluated within the virtual orbital space. EDA was done with the GAMESS software [56] at the HF/cc-pVTZ level of theory. This basis set was shown to closely reproduce the results from the more extended aug-cc-pVTZ
basis set in several test calculations bearing on inter- as well as intramolecular interactions [57].

4.1.2. Correlated Calculations. In order to obtain the dispersion contribution, the intermolecular interaction energies $\Delta E$ were computed at the correlated level using the dispersion-corrected B97-D functional with the G09 software and the cc-pVTZ basis set. The values of $\Delta E$ were also corrected for BSSE. For a given complex, the dispersion contribution ($E_{\text{disp}}$) is evaluated as the difference between the BSSE-corrected B97-D intermolecular interaction energies and the HF ones.

5. SIBFA Computations. The intermolecular interaction energy ($\Delta E_{\text{tot}}$) is computed as the sum of five contributions: electrostatic multipolar ($E_{\text{MTP}}$), short-range repulsion ($E_{\text{rep}}$), polarization ($E_{\text{pol}}$), charge transfer ($E_{\text{CT}}$), and dispersion ($E_{\text{disp}}$) [58]. $E_{\text{MTP}}$ is computed with distributed multipoles derived from the QC molecular orbitals of the individual ligands [59, 60, 61], augmented with penetration [62]. The anisotropic polarizabilities are distributed on the centroids of the localized orbitals [63]. $E_{\text{rep}}$ and $E_{\text{CT}}$ are computed using representations of the molecular orbitals on the chemical bonds and the lone-pairs. $E_{\text{disp}}$ has an expansion into $1/R^6$, $1/R^8$, and $1/R^{10}$ along with an exchange-dispersion component [64]. The parameters for F and Cl were reported in [45].

III. Results and discussion.

We have previously reported that the strong INSTIs bind tightly to the ends of the 3’-processed vDNA, at the LTR ends [32]; thus, justifying their function as a blocker of the strand transfer step [26]. Table 1 reports the Kd values for the binding of RAL, EVG and DTG to the LTR32 oligonucleotide, as determined by fluorescence anisotropy titrations in a phosphate buffer at 5°, 15° and 25°C. The fluorescence anisotropy titration curves of LTR32 for increasing concentrations of drugs are reported in Figures S.II., S.III. and S.IV. Those of LTR32 by DTG (S.II.A.) and EVG (S.III.A.) at 5°C already reported in El Khoury et al., 2017 [32], are added to this study for completeness. The DTG > EVG > RAL Kd ranking of inhibitor-vDNA binding affinities is the same, regarding these three inhibitors, as the one reported for their binding to the complete intasome [65]. Including the desolvation energies of the halobenzene rings is not expected to alter this ranking. We have computed their continuum solvation energies $\Delta G_{\text{solv}}$ with a Polarizable Continuum Model (PCM) [66]. These could represent an upper bound to the ‘real’ solvation energies, on account of a lesser exposure to solvation of the ring when it is integrated.
in the entire drug. The present values amount to -5.1, -6.0, and -3.9 kcal/mol for RAL, EVG and DTG, respectively. The relative differences are much smaller than the $\Delta E_{\text{tot}}$ ones, and their inclusion would actually favor DTG.

We have further unraveled the enthalpy and entropy components of the free energies of LTR32-INSTI complexation. The $\Delta G$, $\Delta H$ and $\Delta S$ values are reported in Tables II.a-c.

These results show the affinity ranking of the three inhibitors for end vDNA to be dominated by the enthalpy component. This finding is fully consistent with the microcalorimetry study reported by Chaires et al., which covered 26 DNA ligands [67] and led to the conclusion that formation of DNA-intercalator complexes is enthalpy-driven, while that of DNA-groove binder complexes is entropy-driven. RAL-LTR32, EVG-LTR32 and DTG-LTR32 interactions are characterized by a mean $\Delta H / \Delta G$ ratio in the 1.3-1.4 range. This is within the 0.83-1.97 $\Delta H / \Delta G$ ratio range, considered as a signature of an enthalpy-dominated interaction.

The above experimental results are an incentive for SIBFA polarizable molecular dynamics simulations of complexes of various halogenated drugs with retroviral DNAs, which should benefit from the massively parallel Tinker-HP software, co-developed in one of our Laboratories [68]. We deemed it necessary, however, to perform a prior validation in addressing the question: to which extent would the binding of the drug halobenzyl rings to G4 and C14 be accountable for the DTG > EVG > RAL ranking, and how well could the outcome from high-level QC computations be accounted for by the SIBFA polarizable molecular mechanics procedure?

The considered complexes have small sizes and, for the present purposes, energy-minimizations bore on the sole halobenzyl ring. An evaluation of the SIBFA accuracy is nevertheless mandatory, as there would be little hope that inconsistencies between the SIBFA and QC results at this early stage could be obliterated or restored by subsequent large-scale MD simulations on the entire drug-IN-vDNA complex. It also is in line with our previous analyses on the binding of a series of mono- and poly-halogenated rings to G4/C14 and the sensitivity of $\Delta E$ and its individual contributions to diverse chemical substitutions [69, 70]. Table III lists the calculated values of $\Delta E(\text{QC/B97-D})$ and $\Delta E_{\text{tot}}(\text{SIBFA})$. The results are summarized in Table III, which lists the values of $\Delta E(\text{QC/B97-D})$ and $\Delta E_{\text{tot}}(\text{SIBFA})$. Supp. Info SI lists the individual
contributions of QC and SIBFA $\Delta E$ values and compares their trends. It is first observed that regarding the halobenzene rings, $\Delta E(B97-D)$ for the inhibitor-G4-C14 complexes has the same DTG > EVG > RAL ranking as the $\Delta H$ values for the DKA-vDNA complexes. This attests to the key role of the halobenzene ring as a modulator of inhibitor affinity for vDNA. It is also noted that the QC results can themselves be closely accounted for by $\Delta E_{\text{tot}}(\text{SIBFA})$, regarding both the magnitudes of $\Delta E$ and the DKA ranking. Close agreements between SIBFA and QC values for the G4/C14 complexes with several halobenzene derivatives were previously reported by us [68]. Supp. Info SI shows that such agreements carry out regarding the individual energy contributions and their trends.

IV. Conclusions and Perspectives.

This study focused on three INSTIs successively used in anti-HIV-1 therapy, raltegravir (RAL), elvitegravir (EVG), and dolutegravir (DTG). We carried out measurements of their free energies of binding, $\Delta G$, to the vDNA end in solution, and unraveled their enthalpy and entropy components in solution. We found that the $\Delta G$ ranking DTG > EVG > RAL parallels that inferred for the intasome [65]. The $\Delta G$ ranking is also paralleled by the $\Delta H$ one, a signature for intercalation-driven binding. It is also the same as the one computed by high-level QC for the binding of the halobenzene ring to the sole G4 and C14 dimer, as well as by the SIBFA polarizable molecular mechanics procedure. The consistency between SIBFA and QC was previously supported by the complexes of G4/C14 with a diversity of substituted halobenzenes [69].

As put forth in [69], it is possible to leverage the ‘Janus-like’ properties of the CX bond (X=F, Cl, Br), electron-deficient along the bond and electron-rich in a cone around it, to target respectively and simultaneously electron-rich and electron-deficient sites of the nucleotide bases. Polarizable molecular mechanics is responsive to the electronic changes brought about by substitutions as these impact the magnitude of both QC-derived distributed multipoles and polarizabilities used to compute the $E_{\text{MTP}}$ and $E_{\text{pol}}$ contributions. These should enable to fine-tune and further evolve the affinity of halobenzenes for targeted HIV-1 DNA bases.
Along these lines, several novel compounds were recently designed and endowed with more favorable $\Delta E(QC)$ and $\Delta E(SIBFA)$ values than the DTG ring. They will be reported in a forthcoming paper (El Darazi et al., manuscript in preparation).

In addition to handling the halobenzene interactions [69, 70], a further asset of SIBFA and related polarizable potentials [71] is the reliable handling of poly-ligated complexes of divalent cations [72] and interactions involving ‘discrete’ structural waters [73, 74]. Such structural motives are also encountered in the intasome-INSTI complexes.

Grounded on these validations, we plan to undertake long-duration polarizable MD simulations on a diversity of INSTI complexes with intasome, resorting to the massively parallel computer code Tinker-HP. These will enable to quantify the extent to which the halobenzene-G4/C14 interactions are modulated by the conformational flexibilities of each partner within the complex, by the electrostatic potentials and fields exerted by the neighboring INT residues and viral DNA bases, and possibly as well by the two neighboring divalent Mg(II) cations and by the structural waters.

**Figure captions.**

**Figure 1.** FDA approved Integrase strand transfer inhibitors INSTIs (from left to right): raltegravir, elvitegravir and dolutegravir. (A) 2D structures of the inhibitors. The red dashed circle indicates the halobenzyl moiety. (B) 3D structure of each inhibitor in complex with the Integrase (IN) and the viral DNA (vDNA). (C) Close-up on the interactions involving the halobenzene, cytosine 16 and guanine 14.

**Figure 2:** LTR32 is a linear oligonucleotide sequence designed to adopt a double strand hairpin structure in solution upon folding around a loop created by a purposely added thymine triplet (TTT in green) with the sensitive fluorescein reporter (F, in orange) grafted to its central T. The latter allows fluorescence studies in solution at low concentrations. The stem that reproduces the 3’ processed LTR end comprises a 17-nucleotide strand and a 15-nucleotide strand corresponding to the unreactive strand and the reactive strand, respectively. Their pairing leaves an unpaired dinucleotide 5’ AC 3’ at the 5’ end on the unreactive strand. In each strand
the nucleotide numbering goes from the 5’ to the 3’ extremity. The highly conserved doublet of base pairs, here numbered C14-G4 and A15-T3, is colored in red.

Acknowledgements.

We wish to thank the Grand Equipement National de Calcul Intensif (GENCI): Institut du Développement et des Ressources en Informatique Scientifique (IDRIS), Centre Informatique de l’Enseignement Supérieur (CINES), France, project No. x2009-075009, and the Centre de Ressources Informatiques de Haute Normandie (CRIHAN, Rouen, France), project 1998053. We also thank the Research Council of Saint-Joseph University of Beirut, Lebanon (Project FS71), the Lebanese National Council for Scientific Research, CNRS-L (Project FS80), as well as the French Institute-Lebanon and the French-Lebanese Program CEDRE (Project 35327UJ) for their continuous financial support.

IV. References.


inhibitors disrupt core maturation and induce a reverse transcription block in target

A.N.; and Steven, A.C. Distribution and redistribution of HIV-1 nucleocapsid protein in
immature, mature, and integrase-inhibited virions: a role for integrase in


[9] Zhao, X.Z.; Smith, S.J.; Maskell, D.P.; Métifiot, M.; Pye, V.E.; Fesen, K.; Marchand, C.;
Pommier, Y.; Cherepanov, P.; Hughes, S.H.; and Burke Jr, T.R. Structure-guided
7315-7332.

Gonzalez, C.J.; Chen, J.; Harvey, C.M.; Isaacs, R.D. Safety and efficacy of the HIV-1
integrate inhibitor raltegravir (MK-0518) in treatment-experienced patients with
multidrug-resistant virus: a phase II randomised controlled trial. *Lancet*. 2007,
369(9569), 1261-1269.


1627-1637.


[15] Wainberg, M.A.; Han, Y.S. Will drug resistance against dolutegravir in initial therapy ever

Hellinger, J.; Shirvani, V.; Zolopa, A.; Shafer, R.W. HIV-1 protease mutations and
4261.

sites and resistance mutations in the three-dimensional structure of HIV-1 reverse


of raltegravir and elvitegravir on HIV-1 integrase catalytic reactions and on a series of

[20] Hazuda, D.J.; Feloek, P.; Witmer, M.; Wolfe, A.; Stillmock, K.; Grobler, J.A.; Espeseth,
A.; Gabryeliski, L.; Schleif, W.; Blau, C.; Miller, M.D.; Inhibitors of strand transfer that
prevent integration and inhibit HIV-1 replication in cells. *Science*. 2000, 287(5453), 646-
650.


[38] Zargarian, L.; Benleumi, M.S.; Renisio, J.G.; Merad, H.; Maroun, R.G.; Weber, F.;
Mauffret, O.; Porumb, H.; Troalen, F.; and Fermandjian, S. Strategy to discriminate
between high and low affinity bindings of human immunodeficiency virus, type 1
integrase to viral DNA. *J. Biol. Chem.* 2003, 278(22), pp.19966-19973.


[40] Politzer, P.; Murray, J. S.; Clark, T., Halogen bonding: an electrostatically-driven highly


anisotropic molecular mechanics/dynamics potential account for sigma hole effects in

[45] Devereux, M.; Raghunathan, S.; Fedorov, D. G.; Meuwly, M.. A Novel, Computationally
Efficient Multipolar Model Employing Distributed Charges for Molecular Dynamics

Theory Comput.* 2016, 12, 3008-3019.

[47] Heyduk, T.; Lee, J.C. Application of fluorescence energy transfer and polarization to
monitor Escherichia coli cAMP receptor protein and lac promoter interaction.


[49] Grimme, S. Semiempirical GGA-type density functional constructed with a long-range

[50] Dunning Jr, T.H. Gaussian basis sets for use in correlated molecular calculations. I. The

[51] Feller, D., 1996. The role of databases in support of computational chemistry

[52] Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman,
J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; Nakatsuji, H. Gaussian

[53] Stevens, W.J.; Fink, W.H. Frozen fragment reduced variational space analysis of
139(1), 15-22.

[54] Boys, S.F.; Bernardi, F.D. The calculation of small molecular interactions by the
differences of separate total energies. Some procedures with reduced errors. *Mol.

11024-11031.

Koseki, S.; Matsunaga, N.; Nguyen, K.A.; Su, S.; Windus, T.L. General atomic and


El Khoury, L.; Naseem-Khan, S.; Kwapien, K.; Hobaiaka, Z.; Maroun, R.G.; Piquemal, J.P.; Gresh, N., Importance of explicit smeared lone-pairs in anisotropic polarizable...


INT Inhibitors
Figure 1. FDA approved Integrase strand transfer inhibitors INSTIs: raltegravir, elvitegravir and dolutegravir. (A) 2D structures of the inhibitors. The red dashed circle indicates the halobenzyl moiety. (B) 3D structure of each inhibitor in complex with the IN active site and the vDNA. (C) Close-up on the interactions involving the halobenzene, cytosine 14 and guanine 4 near the processed 3′-end of the viral DNA.
**Figure 2:** LTR32 is a linear oligonucleotide sequence designed to adopt a double strand hairpin structure in solution upon folding around a loop created by a purposely added thymine triplet (TTT in green) with the sensitive fluorescein reporter (F, in orange) grafted to its central T. The latter allows fluorescence studies in solution at low concentrations. The stem that reproduces the 3' processed LTR end comprises a 17-nucleotide strand and a 15-nucleotide strand corresponding to the unreactive strand and the reactive strand, respectively. Their pairing leaves an unpaired dinucleotide 5' AC 3' at the 5' end on the unreactive strand. In each strand the nucleotide numbering goes from the 5' to the 3' extremity. The highly conserved doublet of base pairs, here numbered C14-G4 and A15-T3, is colored in red.

| 5' A C T G C T A G A G A T T T T C C T | T-F |
| 3' HO-A C G A T C T C T A A A A G G T | 15 14 1 |
**Table 1** (on next page)

Dissociation contents of LTR32-INSTI interactions
<table>
<thead>
<tr>
<th>Temperature</th>
<th>RAL (Molar)</th>
<th>EVG (Molar)</th>
<th>DTG (Molar)</th>
</tr>
</thead>
<tbody>
<tr>
<td>278</td>
<td>5.90E-09</td>
<td>9.50E-11</td>
<td>1.63E-12</td>
</tr>
<tr>
<td>288</td>
<td>7.69E-09</td>
<td>1.44E-10</td>
<td>2.74E-12</td>
</tr>
<tr>
<td>298</td>
<td>1.26E-08</td>
<td>2.65E-10</td>
<td>6.07E-12</td>
</tr>
</tbody>
</table>

Kd (Molar)

Temperature (Kelvin)
Table 2 (on next page)

Thermodynamic contributions of LTR32-INSTI interactions
Table II.a. Thermodynamic contributions of LTR32-raltegravir interaction

<table>
<thead>
<tr>
<th>Temperature</th>
<th>ΔG</th>
<th>ΔH</th>
<th>TΔS</th>
</tr>
</thead>
<tbody>
<tr>
<td>278</td>
<td>-10.07</td>
<td>-14.16</td>
<td>4.09</td>
</tr>
<tr>
<td>288</td>
<td>-10.29</td>
<td>-14.53</td>
<td>4.24</td>
</tr>
<tr>
<td>298</td>
<td>-10.37</td>
<td>-14.75</td>
<td>4.39</td>
</tr>
</tbody>
</table>

Values (kcal/mol) of ΔG, ΔH and TΔS

Temperature (Kelvin)

Table II.b. Thermodynamic contributions of LTR32-eltreviravir interaction

<table>
<thead>
<tr>
<th>Temperature</th>
<th>ΔG</th>
<th>ΔH</th>
<th>TΔS</th>
</tr>
</thead>
<tbody>
<tr>
<td>278</td>
<td>-12.27</td>
<td>-16.36</td>
<td>4.09</td>
</tr>
<tr>
<td>288</td>
<td>-12.48</td>
<td>-16.71</td>
<td>4.23</td>
</tr>
<tr>
<td>298</td>
<td>-12.56</td>
<td>-16.94</td>
<td>4.38</td>
</tr>
</tbody>
</table>

Values (kcal/mol) of ΔG, ΔH and TΔS

Temperature (Kelvin)

Table II.c. Thermodynamic contributions of LTR32-dolutegravir interaction

<table>
<thead>
<tr>
<th>Temperature</th>
<th>ΔG</th>
<th>ΔH</th>
<th>TΔS</th>
</tr>
</thead>
<tbody>
<tr>
<td>278</td>
<td>-14.43</td>
<td>-18.50</td>
<td>4.07</td>
</tr>
<tr>
<td>288</td>
<td>-14.66</td>
<td>-18.88</td>
<td>4.22</td>
</tr>
<tr>
<td>298</td>
<td>-14.72</td>
<td>-19.08</td>
<td>4.37</td>
</tr>
</tbody>
</table>

Values (kcal/mol) of ΔG, ΔH and TΔS

Temperature (Kelvin)
Table 3 (on next page)

QC and SIBFA computations of the INSTI(s)=G4/C14 interactions
Table III. QC and SIBFA computations of the INSTI(s)-C14/G4 interactions.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>RAL SIBFA</th>
<th>RAL QC</th>
<th>EVG SIBFA</th>
<th>EVG QC</th>
<th>DTG SIBFA</th>
<th>DTG QC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔE tot</td>
<td>-31.5</td>
<td>-30.1</td>
<td>-37.1</td>
<td>-38.3</td>
<td>-39.0</td>
<td>-40.8</td>
</tr>
</tbody>
</table>