Computational development of rubromycin-based lead compounds for HIV-1 reverse transcriptase inhibition

The binding of several rubromycin-based ligands to HIV1-reverse transcriptase was analyzed using molecular docking and molecular dynamics simulations. MM-PBSA analysis and examination of the trajectories allowed the identification of several promising compounds with predicted high affinity towards reverse transcriptase mutants which have proven resistant to current drugs. Important insights on the complex interplay of factors determining the ability of ligands to selectively target each mutant have been obtained.

PeerJ Reviewing Manuscript

Carlos E. P. Bernardo and Pedro J. Silva*

REQUIMTE/Faculdade de Ciências da Saúde, Universidade Fernando Pessoa, Rua Carlos da Maia, 296, 4200-150 Porto,Portugal

PeerJ Reviewing Manuscript

PeerJ Reviewing Manuscript

Introduction

HIV reverse transcriptases are multifunctional enzymes which use the virus single-stranded RNA genome as template to build a double-stranded DNA which may later be incorporated into the host's genome. They are composed of two subunits: p66 acts both as a DNA polymerase and as a RNAase which cleaves RNA/DNA hybrid molecules and p51 (whose sequence is equal to that of p66, but lacks the last 124 aminoacids) plays mostly a structural role. Due to its crucial role in the virus life cycle, HIV reverse transcriptase (RT) has been the target of several successful drug-developing efforts. These drugs may be grouped in several classes based on their mechanism of action (thoroughly reviewed in(Jochmans, 2008; Sarafianos et al., 2009; Singh et al., 2010)): nucleoside analogue RT inhibitor (NRTI), like azidothymidine(Mitsuya et al., 1985) (the first successful drug against HIV) act as a alternative substrates and block the synthesis of the viral DNA due to their lack of a free 3' OH- group; nucleotide-competing RT inhibitors (NcRTI) like INDOPY-1(Jochmans et al., 2006) bind the active site in an as-yet-undisclosed manner; and non-nucleoside RT inhibitors (NNRTI) in contrast bind to the enzyme in a hydrophobic pocket 10 Å away from the active site(Kohlstaedt et al., 1992; Ding et al., 1998) and prevent the enzyme from attaining a catalytically competent conformation. Since reverse transcriptases lack a proofreading ability, very high rates of mutation are observed and mutants resistant to one or more drugs frequently arise. To decrease the probability of selection of drug-resistant strains, a combination therapy including drugs with different targets and modes of action is most often used in clinical practice. Still, newer drugs must be continually developed to fight resistant strains.

Rubromycins are a small class of compounds containing naphtoquinone and 8-hydroxyisocoumarin moieties(Brasholz et al., 2007). In 1990, β- and γ-rubromycin were shown to inhibit HIV-1 reverse transcriptase(Goldman et al., 1990), although at levels that were also

toxic to human T lymphocytes. γ-rubromycin was later shown to be an inhibitor of human telomerase(Ueno et al., 2000), fueling interest in its use as an anti-cancer agent. The development of less toxic variants of these lead compounds has long been prevented due to the difficulty of their laboratory synthesis, but several synthetic routes to these interesting molecules have recently become available (Akai et al., 2007; Rathwell et al., 2009; Wu, Mercado, & Pettus, 2011; Wilsdorf & Reissig, 2014), enabling the evaluation of many simpler derivatives as candidates for the inhibition of telomerase (Yuen et al., 2013). As far as we could ascertain, no derivatives of γ-rubromycin with substitution patterns as complex as those observed in the natural molecule have yet been synthesized. As we envisage that such derivatives might afford higher selectivity towards selected reverse-transcriptase mutants or more favorable pharmacokinetic properties, we decided to evaluated several not-yet-synthesized γ-rubromycin derivatives using computational docking and molecular dynamics simulations of the most promising candidates. The results are compared to those of the commercially-available, 2nd-generation NNRTI drug rilpivirine.

Computational methods

All computations were performed in YASARA(Krieger et al., 2004) using the crystal structure of the rilpivirine-inhibited HIV1 reverse transcriptase published by Das *et al.* (PDB: 2ZD1)(Das et al., 2008). A double-mutant structure, (p66)K103N/(p66)Y181C and a quadruple mutant (p51p66)M184I/(p51p66)E138K, were also generated to evaluate the robustness of the ligand binding to reverse transcriptase variants with increased resistance to NNRTIs: K103N is known to strongly reduce susceptibility to efavirenz and nevirapine(Bacheler et al., 2001; Rhee et al., 2004; Eshleman et al., 2006; Zhang et al., 2007; Melikian et al., 2014) and E138K has a similar effect towards rilpivirine, which is increased by M184I(Kulkarni et al., 2012); Y181C reduces

PeerJ Reviewing Manuscript

PeerJ Reviewing Manuscript

suceptibility to efavirenz, etravirine and rilpivirine(Reuman et al., 2010; Tambuyzer et al., 2010; Rimsky et al., 2012). γ-Rubromycin-based ligands (Figure 1 and Supporting Information) were docked to the wild-type structure with AutoDock 4.2.3(Morris et al., 2009) using default docking parameters and point charges assigned according to the AMBER03 force field(Duan et al., 2003). The highest scoring ligands and poses were selected for molecular dynamics simulations. Initial structures for molecular dynamics simulations of mutant proteins were generated from the corresponding ligand-bound wild-type structures through mutation of the corresponding aminoacids. All simulations were run with the AMBER03 forcefield(Duan et al., 2003), using a multiple time step of 1.25 fs for intramolecular and 2.5 fs for intermolecular forces. Simulations were performed in cells 5 Å larger than the solute along each axis (final cell dimensions 127.3 × 102.6 × 78.8 Å), and counter-ions (88 Cl^- and 77 Na^+) were added to a final concentration of 0.9 % NaCl. In total, the simulation contained approximately 106,500 atoms. A 7.86 Å cutoff was taken for Lennard-Jones forces and the direct space portion of the electrostatic forces, which were calculated using the Particle Mesh Ewald method(Essmann et al., 1995) with a grid spacing <1 Å, 4th order B-splines and a tolerance of 10^{-4} for the direct space sum. Simulated annealing minimizations started at 298 K, velocities were scaled down with 0.9 every ten steps for a total time of 5 ps. After annealing, simulations were run at 298 K. Temperature was adjusted using a Berendsen thermostat(Berendsen et al., 1984) based on the time-averaged temperature, i.e., to minimize the impact of temperature control, velocities were rescaled only about every 100 simulation steps, whenever the average of the last 100 measured temperatures converged. Substrate parameterization was performed with the AM1BCC protocol(Jakalian et al., 2000; Jakalian, Jack, & Bayly, 2002). All simulations were run for 30 ns. Differences in ligand binding energies between wild-type and mutant proteins were evaluated using the MM-PBSA methodology(Srinivasan et al., 1998). Although MM-PBSA is unable to afford accurate absolute binding energies and the high standard deviation of MM-PBSA energies limits its ability to

discriminate between ligands with similar binding-affinities(Weis et al., 2006) to a protein, its application to the analysis of the affinity of a single molecule to a series of protein mutants affords high quality results(Moreira, Fernandes, & Ramos, 2007; Martins et al., 2013), presumably due to better cancellation of errors (as the effect of a point-mutation on a large protein is, in relative terms, much smaller than that of a substitution in a small molecule). For each snapshot (taken at 0.25 ns intervals from the last 15 ns of the simulation) we computed the molecular mechanics energy of the protein-ligand complex, the electrostatic contribution to solvation energy (using the Adaptive Poisson-Boltzmann Solver (Baker et al., 2001)) and nonelectrostatic contributions to solvation (with a surface-area-dependent term(Wang et al., 2001)). These computations were repeated for each snapshot for the ligand-free protein and the protein-free ligand, to obtain an estimate of the average binding energy of each ligand. Normal mode analysis computations were performed using the Webnm@ server at http://apps.cbu.uib.no/webnma/home (Hollup, Salensminde, & Reuter, 2005).

III. Results

Computational docking allows the fast screening of a large number of candidate ligands, which may afterwards be analyzed through more demanding computational techniques in the search for suitable leads for further development and experimental characterization. Out initial screen analyzed the docking performance of γ-rubromycin derivatives with/without truncated rings, substitution of the oxygen atoms appended to the spirocyclic ring and different substitution patterns around the rings. Twenty-six of the tested γ-rubromycin-based ligands bind preferentially to an exposed pocket in subunit p51 formed by the Glu89-His96 loop and the Pro157-Leu187 helix-turn-sheet motif. This pocket lies very far away from the nucleic acid binding surface

PeerJ Reviewing Manuscript

(Figure 2), which completely prevents this binding mode from competitively inhibiting the reaction mechanism.

This distant binding pocket might still affect the catalytic activity of the enzyme by triggering a conformational change from the active "open" conformation (Ding et al., 1998) to an inactive conformation. Since such transitions are usually too slow to be observed with molecular dynamics simulations, we analyzed the available vibrational modes of HIV-1 reverse transcriptase using the efficient algorithm and simplified force field described by Hinsen(Hinsen, 1998). In this method, the protein is simulated as a coarse-grained series of springs connecting every Cα with all other Cα atoms with exponentially-decaying force-constants. Despite its conceptual simplicity, the computed vibrational modes and vibration frequencies have been shown to correlate very well with those observed in explicit molecular dynamics simulations. Furthermore, important conformational changes can most frequently be explained by the first few non-trivial vibrational modes, which enables its use in the location of allosteric transitions (Tama & Sanejouand, 2001; Zheng & Brooks, 2005; Zheng, Brooks, & Thirumalai, 2006, 2007; Rodgers et al., 2013; Sanejouand, 2013). Application of this method to the catalytically active "open" conformation of HIV-1 reverse transcriptase (PDB: 2HMI) (Ding et al., 1998) shows that inclusion of a coarse-grained representation of γ-rubromycin in the proposed binding site does not affect the protein flexibility: indeed, hardly any changes in vibrational modes are observed, as confirmed by the very high correlation coefficients between the normal modes of ligand-bound and empty reverse transcriptase, which always exceeding 0.9977. Figure 3 shows the contributions of each aminoacid to the first six non-translational, non-rotational modes obtained by this method, and clearly highlights the negligible contribution of the aminoacids lining this proposed binding pocket to the overall flexibility of the enzyme.

PeerJ Reviewing Manuscript

PeerJ Reviewing Manuscript

Several γ-rubromycin-based ligands (**11**, **12**, **18**, **21**, **30**, **31**, **33 - 46**) may bind the previously defined NNRTI-binding pocket with affinities exceeding those of this distant, inactive, binding pocket. The most promising leads (Table 1) generally had (like the NNRTI drug rilpivirine) a nitrile group appended to the ligand. The behavior of these molecules in the reverse transcriptase binding pocket of wild-type and mutant reverse transcriptase was then evaluated through 30 ns-long molecular dynamics simulations and compared to that of rilpivirine. The worst-scoring ligands towards the NNRTI binding pocket were those where any of the rings had been removed, as well as the ones where the oxygen at the R_6 position was substituted by nitrogen or carbon. Surprisingly, substitution of the =CH- at the R_4 position by an isoelectronic =N- (ligand **32**) also led to a dramatic loss of binding affinity. Binding affinities of each ligand to wild-type and mutant HIV-1 RT s were computed with the MM-PBSA methodology using the last 15 ns of each molecular dynamics simulation (Table 2). This method, while not accurate enough to produce reliable absolute binding free energies, has been shown to provide good estimates of binding affinity trends provided that either the ligands or the protein targets under comparison are very similar (Massova & Kollman, 2000). The computed data for rilpivirine agree with the experimentally observed sensitivity of its binding to E138K / M184I variants, and to the relative insensitivity of its effect on the presence/absence of K103N or Y181C mutation, which supports the applicability of the MM-PBSA approach to this system. Ligands **13**, **27, 36** and **45** are computed to bind significantly stronger to the rilpivirine-resistant E138K/M184I HIV1-RT variant than to the wild-type protein, and may therefore be suitable lead compounds for further pharmaceutical developments against rilpivirine-resistant strains. Further insight to the determinants of binding affinity was obtained through close inspection of each simulation.

As observed in the crystal structure(Das et al., 2008), rilpivirine remains bound to RT throughout the simulation through a large number of hydrophobic contacts and two very stable hydrogen bonds with the backbone of Lys101, whether in the wild-type or any of the tested mutants. Its

PeerJ Reviewing Manuscript

high hydrophobicity strongly favor it to adopt a very buried conformation and low solvent-accessible area throughout the simulation. The high stability of the hydrogen bonds does not change in the mutated variants, but the total number of close hydrophobic contacts between rilpivirine and the protein does become smaller in the E138K/M184I mutant, which is consistent with the experimentally observed lower affinity of this drug towards it (Singh et al., 2012), and the computed MM-PBSA binding energy.

γ-rubromycin is a much larger and less flexible ligand than rilpivirine: as it binds to the NNRTI binding patch, the methoxy-bearing end of γ-rubromycin remains in contact with the solvent through its hydrophilic surface (Figure 4), whereas the oxygen atoms in its naphtoquinone moiety establish stable hydrogen bonds with Lys101 and Lys103. In the K103N/Y181C mutant, γ-rubromycin becomes less exposed to the solvent, since the shorter sidechain of Asn103 (compared to the wild-type Lys 103) forces the naphtoquinone moiety of the ligand to penetrate deeper into the crevice in order to establish a stabilizing hydrogen bond with Asn103. The buried conformation of γ-rubromycin removes the methoxy group from its favored solvent-exposed environment leading to a binding mode which is computed by MM-PBSA to be markedly less favored than observed in the wild-type protein, but which remains stable due to the difficulty in breaking the large number of favorable hydrogen bonds to Asn103 and Lys101. γ-rubromycin binding to the E138K/M184I is very similar to the wild-type protein: hydrogen bonds between the ligand and Lys101 and Lys103 are also present (though ca. 0.4 Å longer), and subtle cavity rearrangements due to the loss of the ionic bridge between Lys101 and (p51)Glu138 (which is mutated to a Lys) lead to the possibility of intermitent H-bonded interactions between the carbonyl of Ile180 (or the sidechain of (p51)Thr139) and the naphtoquinone moiety.

The binding of ligand **13** to wild type RT differs more from that of γ-rubromycin than would be expected from the very small difference in their structures (the single substitution of a methoxy

group in γ-rubromycin by an ethyl): since the ethyl group is less hydrophilic than a methoxy, it initially tends to establish a hydrophobic interaction with the sidechain of Val179, instead of protruding (like the methoxy group) in the direction of the solvent, leading to a binding mode where the stabilizing hydrogen-bonds between the ligand and the protein are due to Glu138 instead of Lys101. In contrast to what is observed in the binding of γ-rubromycin to the K103N/Y181C, the replacement of the Lys-based H-bonds does not lead to an unfavorable buried conformation of the ligand because, as the simulation progresses, the interaction with Glu138 causes subtle changes in the local environment which becomes more exposed to the solvent than originally: indeed, there is in average one more water molecule near ligand **13** than near γ-rubromycin, leading to a smaller desolvation penalty when **13** binds to the protein. Binding of **13** to the mutants is strongly favored over binding to the wild-type due to the formation of hydrogen bonding to the backbone of Ile180 (especially in E138K/M184I) and especially by the changes in the electrostatic component of ligand solvation caused by the presence of two intra-molecular H-bonds in **13** when bound to the mutant proteins.

The sp^3 hybridization in the acetyl-bearing carbon of the isocoumarin-moiety in **27** introduces a deviation from full planarity in that region of the ligand, which facilitates its interactions with the Trp229 and Ty188 aminoacids on that end of the NNRTI-binding cavity. Ligand **27** is found to bind much more favorably to the quadruple mutant E138K/M184I (with a very large number of very short and stable hydrogen bonds with Lys101, Lys103, Lys138 and Thr139) than to wild-type or K103N/Y181C, where the only stable hydrogen bonds available are those with Glu138. The electrostatic component of the solvation energy of **27** follows the opposite trend as the protein is changed from WT to the mutants, but the smaller variation of this factor simply dampens the magnitude of the change in binding affinities brought about by the variation in protein-ligand interactions.

PeerJ Reviewing Manuscript

Ligand **36** bears a fluorine atom in place of the methoxy group carried by γ-rubromycin. Like ligand **27**, **36** has higher affinity to the E138K/M184I mutant than to either the wild-type and, especially, the K103N/Y181C mutant. The minute size of the fluorine substituent allows Lys101 and Glu138 (which lie on opposite sides of the crevice where the ligands bind) to approach each other and form a strong ionic bridge which pushes the ligand further inside the cavity. This ionic bridge cannot form in the E138K/M184I mutant, leading to a binding mode where the ligand is slightly more exposed and strongly binds to Lys101, Lys103 and Lys 138. In the K103N/Y181C, the interactions between ligand and protein are weaker due to the strong deviations from 180º in the possible H-bonding partners in the binding cavity.

Ligands **37** and **38** bear a chlorine and a cyanide (respectively) in place of the fluorine present on **36**. The intermediate size of these substituents (relative to the fluorine in **36** and the methoxy in γ-rubromycin) leads to an intermediate degree of penetration in the binding cavity, between those of **36** and in γ-rubromycin. As observed in most cases, Lys101 is responsible for the most stable interaction between protein and ligand. No single contribution is, however, determinant in the observed trend of binding affinity of **37** to the proteins, as the correlation of total binding energies to either electrostatic components of solvation or to protein-ligand interaction is insignificant: the overall effect is rather the result of subtle interplay of the electrostatic component of solvation and the protein-ligand interaction. Solvation effects, in contrast are determinant in the binding trends observed for ligand **38**, as the higher affinity to the M184I/E138K mutant is correlated to its much smaller desolvation penalty, which is due to the considerable exposure of its nitrile group to solvent when the entrance to the binding channel is not blocked by the Glu138-Lys101 ionic bond (Fig. 5).

Ligand **45** bears, like ligands **38** and **46**, a nitrile group in the position occupied by a methoxy in γ-rubromycin. It differs from **38** by the replacement of the acetyl substituent of the isocoumarin by a hydroxymethyl and by switching the orientation of the lactone group in isocoumarin from

PeerJ Reviewing Manuscript

-O-C=O to O=C-O. The replacement of acetyl from hydroxymethyl makes the isocoumarin end of **45** significantly smaller and less hydrophilic, leading that end of the molecule towards the inside of the crevice and the nitrile-bearing naphtoquinone portion of **45** to protrude from the other end of the cavity into the solvent. The only H-bonds between ligand and protein now involve the backbone atoms of Lys101 and Glu138. These H-bonds weaken considerably in both mutants, but this destabilizing effect is overtaken by sizable stabilizing effects due to favorable solvation, leading to overall better binding to K103N/Y181C and (especially) E138K/M184I.

Other than the lack of H-bond donating ability in its isocoumarin moiety (due to the replacement of its hydroxyl by a carbonyl), ligand **46** is identical to ligand **38**. Unlike ligand **38**, its ability to bind the K103N/Y181C mutant is not inferior to its affinity to the wild-type protein: the presence of a carbonyl instead of a hydroxyl allows it to accept a hydrogen bond from Tyr183, which is able to rotate into position in the mutant due to the smaller size occupied by Cys181 (compared to the original tyrosine present in the wild type).

We also performed molecular dynamics simulation of ligands **32** and **24** bound to the NNRTI-binding pocket, as these ligands are ranked by AutoDock among the weakest in affinity towards this binding site. These simulations are expected to shed light into structural factors that disfavor binding to HIV-1 reverse transcriptase.

Ligand **24** (like the weakly-binding ligands **22**, **23** and **26**) lacks the spirocyclic scaffold characteristic of rubromycins. The longer head-to-tail distance due to the opening of the central acetal leads to a binding mode where the naphtoquinone moiety is buried inside the enzyme in a position similar to that taken by isocoumarin in the other simulations and the isocoumarin end becomes much more exposed to the solvent than observed for naphtoquinone in the binding modes of the other ligands. This portion of the molecule now remains continuously exposed to the solvent, significantly decreasing the number of stable hydrophobic ligand-protein interactions. (Figure 6, upper panel).

PeerJ Reviewing Manuscript

Ligand **32** differs from γ-rubromycin only in the replacement of the =CH group at the R_4 position by a nitrogen atom. This small isoelectronic change does not, however, lead to very dramatic differences in dynamic behavior, as evidenced by the high similarity of the resulting simulation to that of the original γ-rubromycin: Lys103 still maintains a strong and stable interaction to the naphtoquinone moiety, and Lys101 now interacts with the appended methoxy group in R_{11}, rather than the –OH group in R_8 (Figure 6). The weaker binding of **32** predicted by the docking analysis seems to be due mostly to the cost of burying the polar =N- group away from the solvent, rather than on any large differences of ligand-protein interactions.

Influence of ligand binding on protein dynamics

In the absence of NNRTI, reverse transcriptase may adopt either a compact structure(Hsiou et al., 1996) or an open structure (Ding et al., 1998) which allows the binding of a RNA template and the polymerization of DNA (Figure 7). The binding of an NNRTI acts as a "wedge"(Ivetac & McCammon, 2009) that further separates the catalytic triad (Asp110, Asp185 and Asp186) from Met230, which is believed to be part of the primer-recognition region. The same "wedge" effect is observed for all tested ligands bound to the NNRTI binding pocket (Figure 7 and supporting Information). The only instance where the diagnostic Lys154-Asn255 and Glu67-Gln242 distances stably assume values as short as those observed in the catalytically active conformation is observed in the simulation of the K103N/Y181C mutant in the presence of ligand **46.**

Conclusions

Our computational study confirms that γ-rubromycin -based ligands are able to bind to HIV1-reverse transcriptase at the previously defined NNRTI-binding site, and allowed the identification

PeerJ Reviewing Manuscript

of ligands that are predicted to bind very strongly to RT mutants which have shown high resistance towards other NNRTI compounds. The best compounds (**13**, **27**, **36** and **45**) achieve selective binding to the highly resistant mutant E138K/M184I through very subtle variations on the degree of exposure to solvent, and on the number and strength of hydrogen bonds and hydrophobic interactions with the protein. These molecules should therefore become good candidates in the quest for suitable γ-rubromycin-based drugs.

References

Akai, S., Kakiguchi, K., Nakamura, Y., Kuriwaki, I., Dohi, T., Harada, S., Kubo, O., Morita, N., & Kita, Y. 2007. *Total synthesis of (+/-)-gamma-rubromycin on the basis of two aromatic Pummerer-type reactions*. Angewandte Chemie (International ed. in English) 46(39):7458–61.

Bacheler, L., Jeffrey, S., Hanna, G., D'Aquila, R., Wallace, L., Logue, K., Cordova, B., Hertogs, K., Larder, B., Buckery, R., et al. 2001. *Genotypic correlates of phenotypic resistance to efavirenz in virus isolates from patients failing nonnucleoside reverse transcriptase inhibitor therapy.* Journal of virology 75(11):4999–5008.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. 2001. *Electrostatics of nanosystems: application to microtubules and the ribosome.* Proceedings of the National Academy of Sciences of the United States of America 98(18):10037–41.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. 1984. *Molecular dynamics with coupling to an external bath*. The Journal of Chemical Physics 81(8):3684.

Brasholz, M., Sörgel, S., Azap, C., & Reißig, H.-U. 2007. *Rubromycins: Structurally Intriguing, Biologically Valuable, Synthetically Challenging Antitumour Antibiotics*. European Journal of Organic Chemistry 2007(23):3801–3814.

Das, K., Bauman, J. D., Clark, A. D., Frenkel, Y. V, Lewi, P. J., Shatkin, A. J., Hughes, S. H., & Arnold, E. 2008. *High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: strategic flexibility explains potency against resistance mutations.* Proceedings of the National Academy of Sciences of the United States of America 105(5):1466–71.

Ding, J., Das, K., Hsiou, Y., Sarafianos, S. G., Clark, A. D., Jacobo-Molina, A., Tantillo, C., Hughes, S. H., & Arnold, E. 1998. *Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 A resolution.* Journal of molecular biology 284(4):1095–111.

Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., et al. 2003. *A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations.* Journal of computational chemistry 24(16):1999–2012.

Eshleman, S. H., Jones, D., Galovich, J., Paxinos, E. E., Petropoulos, C. J., Jackson, J. B., & Parkin, N. 2006. *Phenotypic drug resistance patterns in subtype A HIV-1 clones with nonnucleoside reverse transcriptase resistance mutations.* AIDS research and human retroviruses 22(3):289–93.

PeerJ Reviewing Manuscript

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. 1995. *A smooth particle mesh Ewald method*. The Journal of Chemical Physics 103(19):8577.

Goldman, M., Salituro, G., Bowen, J., Williamson, J., Zink, D., Schleif, W., & Emini, E. 1990. *Inhibition of human immunodeficiency virus-1 reverse transcriptase activity by rubromycins: competitive interaction at the template.primer site*. Mol. Pharmacol. 38(1):20–25.

Hinsen, K. 1998. *Analysis of domain motions by approximate normal mode calculations.* Proteins 33(3):417–29.

Hollup, S. M., Salensminde, G., & Reuter, N. 2005. *WEBnm@: a web application for normal mode analyses of proteins.* BMC bioinformatics 6(1):52.

Hsiou, Y., Ding, J., Das, K., Clark, A. D., Hughes, S. H., & Arnold, E. 1996. *Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms.* Structure (London, England : 1993) 4(7):853–60.

Ivetac, A., & McCammon, J. A. 2009. *Elucidating the inhibition mechanism of HIV-1 non-nucleoside reverse transcriptase inhibitors through multicopy molecular dynamics simulations.* Journal of molecular biology 388(3):644–58.

Jakalian, A., Bush, B. L., Jack, D. B., & Bayly, C. I. 2000. *Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method*. Journal of Computational Chemistry 21(2):132–146.

Jakalian, A., Jack, D. B., & Bayly, C. I. 2002. *Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation*. Journal of Computational Chemistry 23(16):1623–1641.

Jochmans, D. 2008. *Novel HIV-1 reverse transcriptase inhibitors.* Virus research 134(1-2):171–85.

Jochmans, D., Deval, J., Kesteleyn, B., Van Marck, H., Bettens, E., De Baere, I., Dehertogh, P., Ivens, T., Van Ginderen, M., Van Schoubroeck, B., et al. 2006. *Indolopyridones inhibit human immunodeficiency virus reverse transcriptase with a novel mechanism of action.* Journal of virology 80(24):12283–92.

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., & Steitz, T. A. 1992. *Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor.* Science (New York, N.Y.) 256(5065):1783–90.

Krieger, E., Darden, T., Nabuurs, S. B., Finkelstein, A., & Vriend, G. 2004. *Making optimal use of empirical energy functions: Force-field parameterization in crystal space*. Proteins-Structure Function and Bioinformatics 57:678–683.

Kulkarni, R., Babaoglu, K., Lansdon, E. B., Rimsky, L., Van Eygen, V., Picchio, G., Svarovskaia, E., Miller, M. D., & White, K. L. 2012. *The HIV-1 reverse transcriptase M184I mutation enhances the E138K-associated resistance to rilpivirine and decreases viral fitness.* Journal of acquired immune deficiency syndromes (1999) 59(1):47–54.

Martins, S. A., Perez, M. A. S., Moreira, I. S., Sousa, S. F., Ramos, M. J., & Fernandes, P. A. 2013. *Computational Alanine Scanning Mutagenesis: MM-PBSA vs TI*. Journal of Chemical Theory and Computation 9(3):1311–1319.

Massova, I., & Kollman, P. A. 2000. *Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding*. Perspectives in Drug Discovery and Design 18(1):113–135.

Melikian, G. L., Rhee, S.-Y., Varghese, V., Porter, D., White, K., Taylor, J., Towner, W., Troia, P., Burack, J., Dejesus, E., et al. 2014. *Non-nucleoside reverse transcriptase inhibitor (NNRTI)*

PeerJ Reviewing Manuscript

cross-resistance: implications for preclinical evaluation of novel NNRTIs and clinical genotypic resistance testing. The Journal of antimicrobial chemotherapy 69(1):12–20.

Mitsuya, H., Weinhold, K. J., Furman, P. A., St Clair, M. H., Lehrman, S. N., Gallo, R. C., Bolognesi, D., Barry, D. W., & Broder, S. 1985. *3'-Azido-3'-deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro.* Proceedings of the National Academy of Sciences of the United States of America 82(20):7096–7100.

Moreira, I. S., Fernandes, P. A., & Ramos, M. J. 2007. *Computational alanine scanning mutagenesis--an improved methodological approach.* Journal of computational chemistry 28(3):644–54.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. 2009. *AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.* Journal of computational chemistry 30(16):2785–91.

Rathwell, D. C. K., Yang, S.-H., Tsang, K. Y., & Brimble, M. A. 2009. *An efficient formal synthesis of the human telomerase inhibitor (+/-)-gamma-rubromycin.* Angewandte Chemie (International ed. in English) 48(43):7996–8000.

Reuman, E. C., Rhee, S.-Y., Holmes, S. P., & Shafer, R. W. 2010. *Constrained patterns of covariation and clustering of HIV-1 non-nucleoside reverse transcriptase inhibitor resistance mutations.* The Journal of antimicrobial chemotherapy 65(7):1477–85.

Rhee, S.-Y., Liu, T., Ravela, J., Gonzales, M. J., & Shafer, R. W. 2004. *Distribution of human immunodeficiency virus type 1 protease and reverse transcriptase mutation patterns in 4,183 persons undergoing genotypic resistance testing.* Antimicrobial agents and chemotherapy 48(8):3122–6.

Rimsky, L., Vingerhoets, J., Van Eygen, V., Eron, J., Clotet, B., Hoogstoel, A., Boven, K., & Picchio, G. 2012. *Genotypic and phenotypic characterization of HIV-1 isolates obtained from patients on rilpivirine therapy experiencing virologic failure in the phase 3 ECHO and THRIVE studies: 48-week analysis.* Journal of acquired immune deficiency syndromes (1999) 59(1):39–46.

Rodgers, T. L., Townsend, P. D., Burnell, D., Jones, M. L., Richards, S. A., McLeish, T. C. B., Pohl, E., Wilson, M. R., & Cann, M. J. 2013. *Modulation of global low-frequency motions underlies allosteric regulation: demonstration in CRP/FNR family transcription factors.* PLoS biology 11(9):e1001651.

Sanejouand, Y. 2013. *Elastic network models: theoretical and empirical foundations.* Methods in molecular biology (Clifton, N.J.) 924:601–16.

Sarafianos, S. G., Marchand, B., Das, K., Himmel, D. M., Parniak, M. a, Hughes, S. H., & Arnold, E. 2009. *Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition.* Journal of molecular biology 385(3):693–713.

Singh, K., Marchand, B., Kirby, K. a, Michailidis, E., & Sarafianos, S. G. 2010. *Structural Aspects of Drug Resistance and Inhibition of HIV-1 Reverse Transcriptase.* Viruses 2(2):606–638.

Singh, K., Marchand, B., Rai, D. K., Sharma, B., Michailidis, E., Ryan, E. M., Matzek, K. B., Leslie, M. D., Hagedorn, A. N., Li, Z., et al. 2012. *Biochemical mechanism of HIV-1 resistance to rilpivirine.* The Journal of biological chemistry 287(45):38110–38123.

Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A., & Case, D. A. 1998. *Continuum Solvent Studies of the Stability of DNA , RNA , and Phosphoramidate - DNA Helices* 7863(98).

Tama, F., & Sanejouand, Y. H. 2001. *Conformational change of proteins arising from normal mode calculations.* Protein engineering 14(1):1–6.

PeerJ Reviewing Manuscript

Tambuyzer, L., Vingerhoets, J., Azijn, H., Daems, B., Nijs, S., de Béthune, M.-P., & Picchio, G. 2010. *Characterization of genotypic and phenotypic changes in HIV-1-infected patients with virologic failure on an etravirine-containing regimen in the DUET-1 and DUET-2 clinical studies.* AIDS research and human retroviruses 26(11):1197–205.

Ueno, T., Takahashi, H., Oda, M., Mizunuma, M., Yokoyama, A., Goto, Y., Mizushina, Y., Sakaguchi, K., & Hayashi, H. 2000. *Inhibition of Human Telomerase by Rubromycins: Implication of Spiroketal System of the Compounds as an Active Moiety*. Biochemistry 39(20):5995–6002.

Wang, J., Morin, P., Wang, W., & Kollman, P. A. 2001. *Use of MM-PBSA in Reproducing the Binding Free Energies to HIV-1 RT of TIBO Derivatives and Predicting the Binding Mode to HIV-1 RT of Efavirenz by Docking and MM-PBSA*. Journal of the American Chemical Society 123(22):5221–5230.

Weis, A., Katebzadeh, K., Söderhjelm, P., Nilsson, I., & Ryde, U. 2006. *Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field.* Journal of medicinal chemistry 49(22):6596–606.

Wilsdorf, M., & Reissig, H.-U. 2014. *A Convergent Total Synthesis of the Telomerase Inhibitor (±)-γ-Rubromycin.* Angewandte Chemie (International ed. in English) 53(17):4332–6.

Wu, K.-L., Mercado, E. V, & Pettus, T. R. R. 2011. *A convergent total synthesis of (±)-γ-rubromycin.* Journal of the American Chemical Society 133(16):6114–7.

Yuen, T.-Y., Ng, Y.-P., Ip, F. C. F., Chen, J. L.-Y., Atkinson, D. J., Sperry, J., Ip, N. Y., & Brimble, M. a. 2013. *Telomerase Inhibition Studies of Novel Spiroketal-Containing Rubromycin Derivatives*. Australian Journal of Chemistry 530–533.

Zhang, Z., Xu, W., Koh, Y.-H., Shim, J. H., Girardet, J.-L., Yeh, L.-T., Hamatake, R. K., & Hong, Z. 2007. *A novel nonnucleoside analogue that inhibits human immunodeficiency virus type 1 isolates resistant to current nonnucleoside reverse transcriptase inhibitors.* Antimicrobial agents and chemotherapy 51(2):429–37.

Zheng, W., & Brooks, B. 2005. *Identification of dynamical correlations within the myosin motor domain by the normal mode analysis of an elastic network model.* Journal of molecular biology 346(3):745–59.

Zheng, W., Brooks, B. R., & Thirumalai, D. 2006. *Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations.* Proceedings of the National Academy of Sciences of the United States of America 103(20):7664–9.

Zheng, W., Brooks, B. R., & Thirumalai, D. 2007. *Allosteric transitions in the chaperonin GroEL are captured by a dominant normal mode that is most robust to sequence variations.* Biophysical journal 93(7):2289–99.

PeerJ Reviewing Manuscript

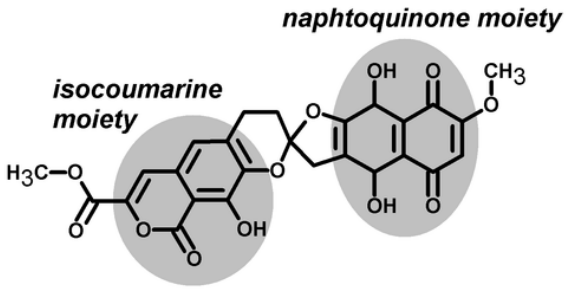
Figure 1

Figure 1

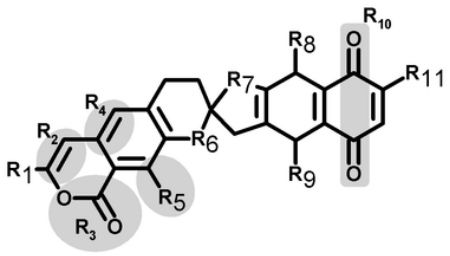
Structures of the tested γ-rubromycin-based ligands. Substitution patterns in molecules 1-14; 17; 20-21 and 26-46 are shown as Supporting Information.

PeerJ Reviewing Manuscript

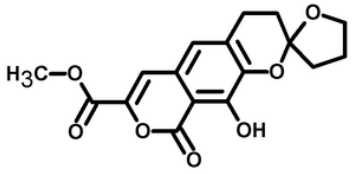
PeerJ Reviewing Manuscript



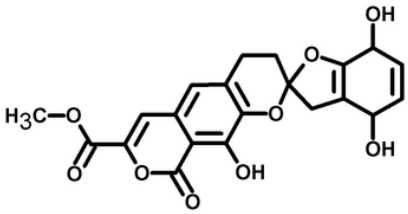
γ−rubromycin



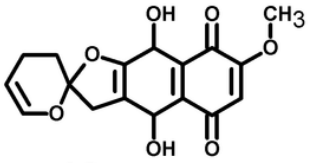
1-14 ; 17 ; 20-21 ; 26-46



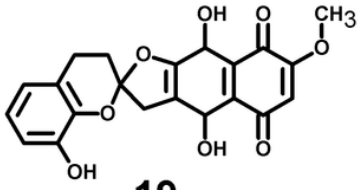
15



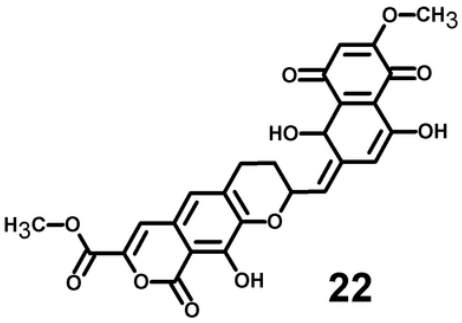
16



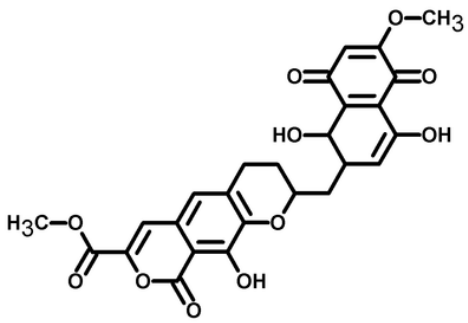
18



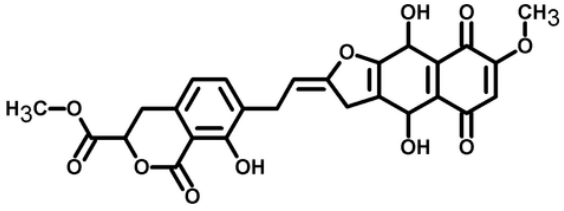
19



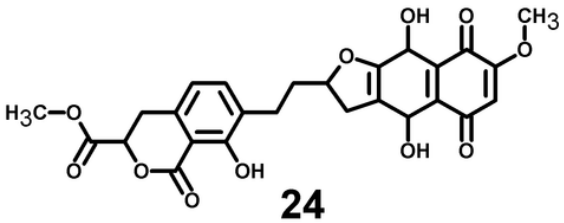
22



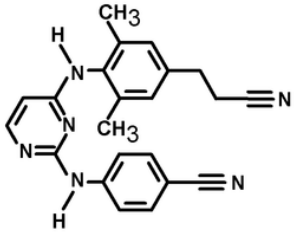
25



23



24

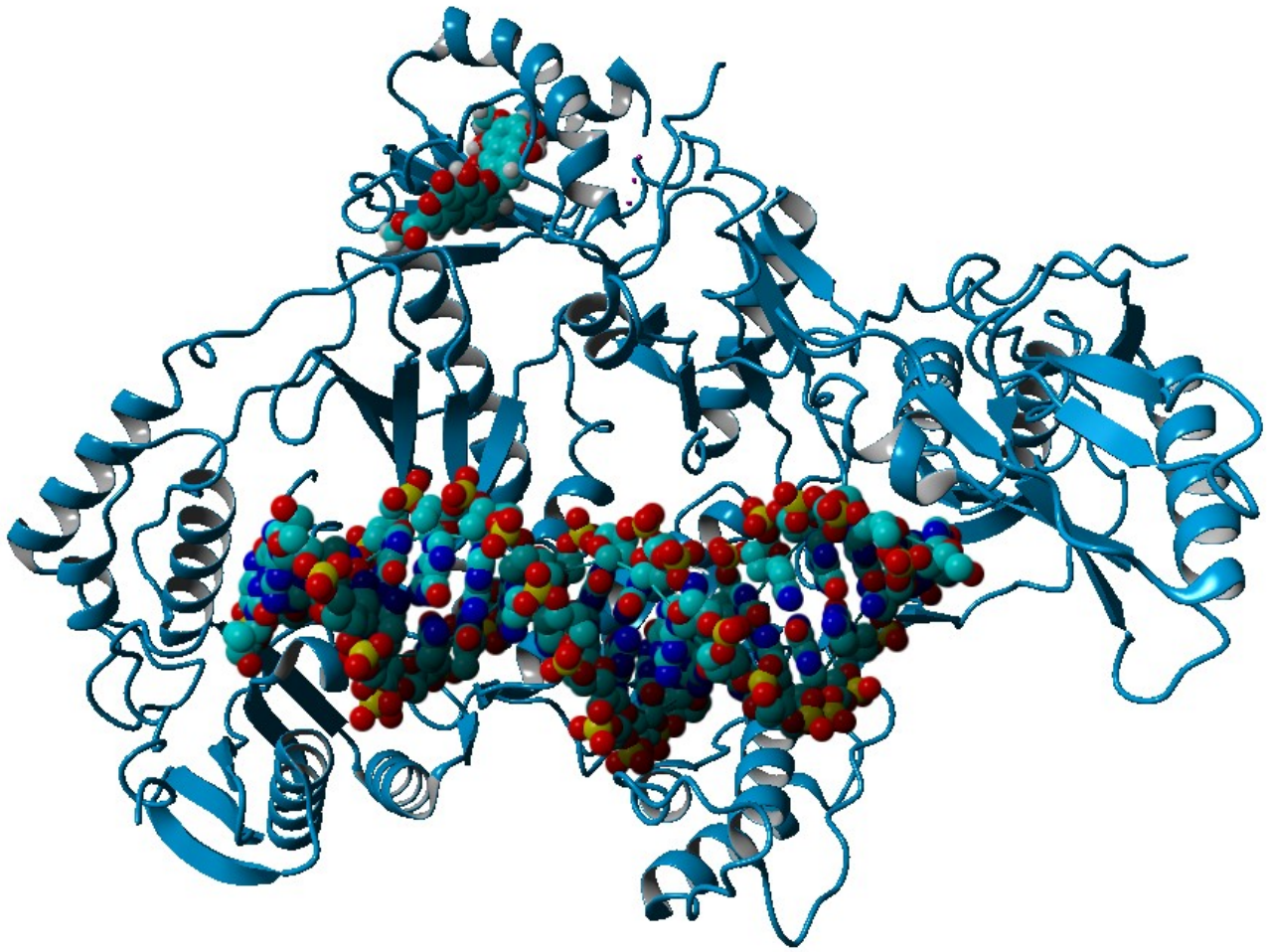


rilpivirine

Figure 2

Figure 2

Preferential binding mode of ligand **28** to HIV reverse transcriptase, as computed with AutoDock, with superposed DNA molecule taken from the DNA-bound HIV structure (PDB: 2HMI) (Ding et al., 1998).



PeerJ Reviewing Manuscript

Table 1(on next page)

Table 1

Substitution patterns and AutoDock-computed binding energies of the best-scoring ligands to the previously characterized NNRTI binding pocket. Only differences from the parent compound (γ-rubromycin) are shown. The binding energy of the drug rilpivirine, computed with the same methods, amounts to -13.25 kcal.mol^{-1}. Data for all ligands is available as Supporting Information.

PeerJ Reviewing Manuscript

Ligand:	**γ-rubromycin**	**46**	**36**	**27**	**45**	**13**	**38**	**37**
R_1=	-COOCH3				-CH_2OH			
R_2=	-C=C-H			(*S*) HC-CH_2				
R_3=	-C=O-O				-O-C=O			
R_4=	=C=C-H	-C-CH_2-						
R_5=	-C-OH	-C=O						
R_6=	-O-							
R_7=	-O-							
R_8=	-OH							
R_9=	-OH							
R_{10}=	-C=O -C=O							
R_{11}=	-O-CH_3	-CN	-F		-CN	-CH_2-CH_3	-CN	-Cl
Binding energy	-12.95	-13.71	-13.72	-13.82	-13.82	-13.91	-14.25	-14.29

PeerJ Reviewing Manuscript

Table 2(on next page)

Table 2

Binding affinity (average ± standard error of the mean) of the best-scoring ligands to reverse-transcriptase mutants, relative to the binding affinity of each ligand to the wild-type enzyme. Values in kcal.mol^{-1}. Negative values show stronger binding than observed to the wild-type protein.

PeerJ Reviewing Manuscript

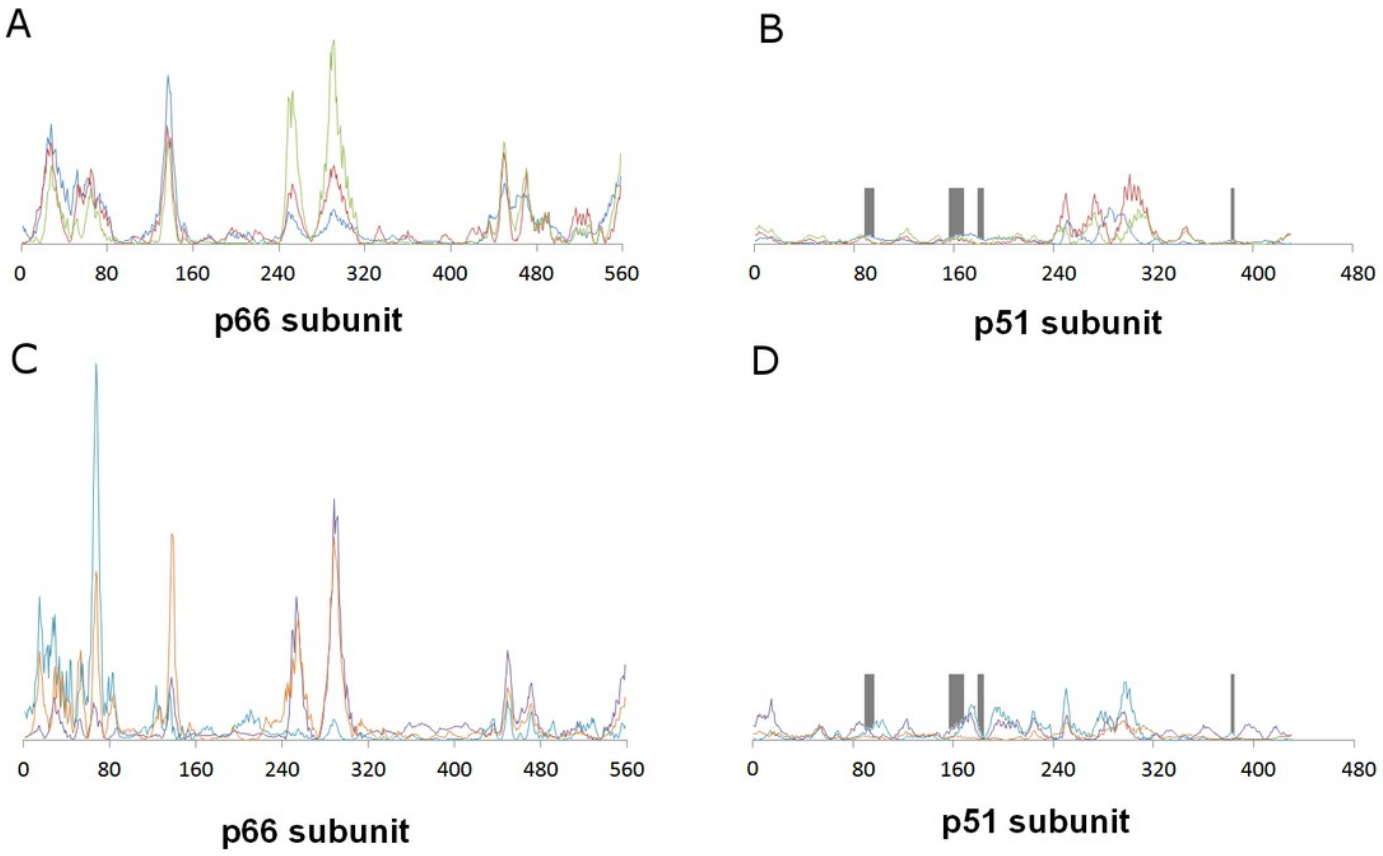
	K103N / Y181C	E138K / M184I
Rilpivirine	1.6±0.9	3.6±0.8
γ-rubromycin	9.8±1.1	0.3±1.1
13	-6.7±1.4	-16.8±1.4
27	7.7±1.4	-13.0±1.2
36	10.3±1.0	-7.1±0.9
37	-4.0±1.2	-5.2±1.4
38	4.6±1.0	-4.1±0.9
45	-3.6±1.0	-7.0±1.2
46	-1.9±1.2	-3.8±1.0

PeerJ Reviewing Manuscript

Figure 3

Figure 3

Relative contribution of each aminoacid displacement to the first six non-trivial normal modes of HIV-1 reverse transcriptase. A and B: modes 7 (blue), 8 (red) and 9 (green). C and D: modes 10 (violet), 11 (light blue) and 12 (orange). The regions lining the proposed binding pocket are highlighted in dark gray.

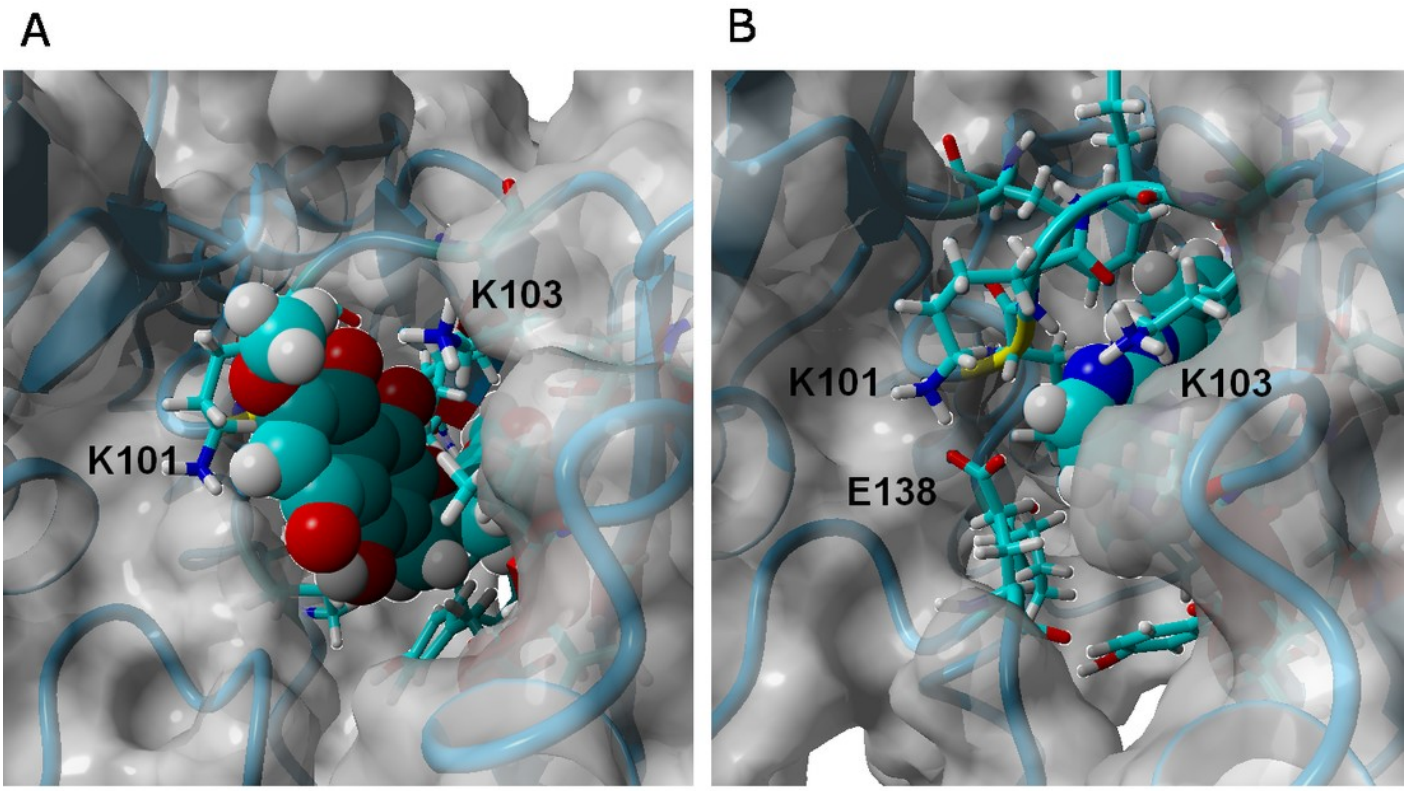


PeerJ Reviewing Manuscript

Figure 4

Figure 4

ɣ-rubromycin (A) and rilpivirine (B) bound to wild-type HIV-1 reverse transcriptase. Snapshots were taken from random points in the last 15ns of molecular dynamics simulations.



PeerJ Reviewing Manuscript

Figure 5

Figure 5

Ligand 38 bound to wild-type (A and B), K103N/Y181C (C and D) and E138K/M184I (E and F). Snapshots were taken from random points in the last 15ns of molecular dynamics simulations. As far as possible, the bonds depicted in the graphs have been highlighted with the same color in the corresponding image (the exceptions are the C181•••OH hydrogen bond shown as the red line in D, and the two hydrogen bonds between K101 backbone NH and nearby oxygen atoms in the inhibitor shown as red and blue lines in F)

PeerJ Reviewing Manuscript

PeerJ Reviewing Manuscript

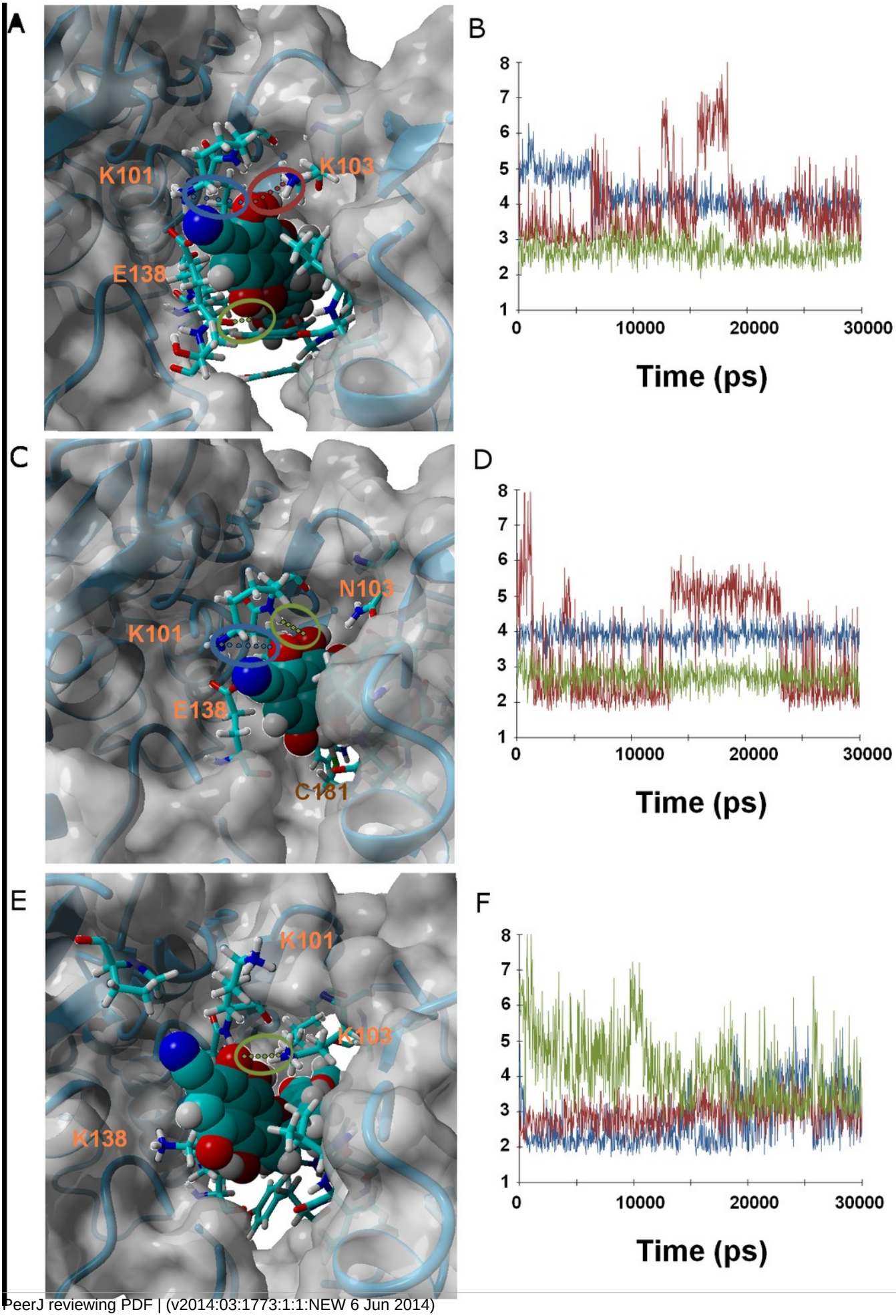


Figure 6

Figure 6

Ligands **24** (A and B), **32** (C and D) and γ-rubromycin (E and F) bound to wild-type reverse transcriptase. Snapshots were taken from random points in the last 15 ns of molecular dynamics simulations. As far as possible, the bonds depicted in the graphs have been highlighted with the same color in the corresponding image (the exceptions are the I180 carbonyl•••OH hydrogen bond shown as the blue line in the B, and the hydrogen bond between K101 carbonyl and methoxy oxygen in γ-rubromycin shown as a blue line in F)

PeerJ Reviewing Manuscript

PeerJ Reviewing Manuscript

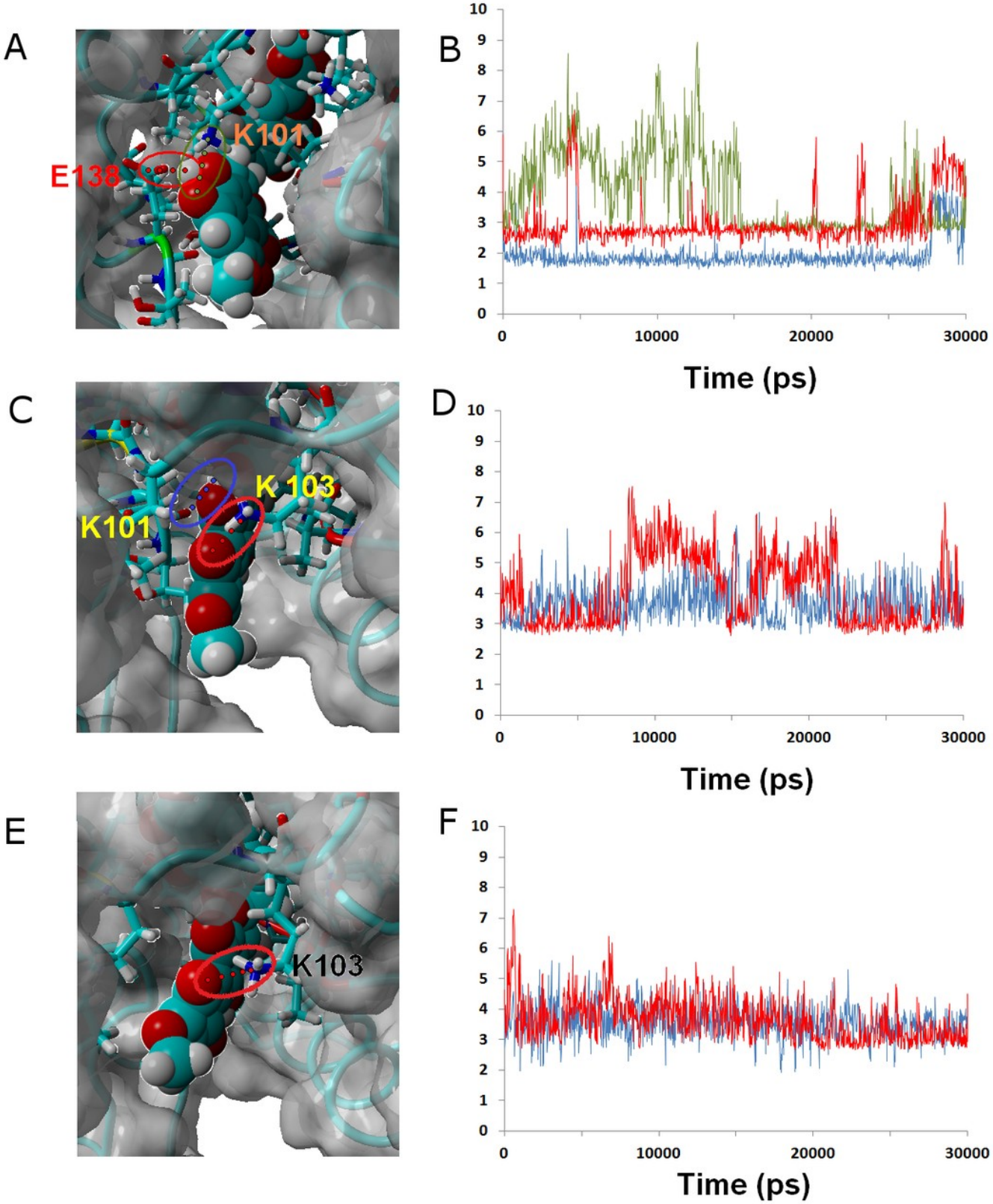


Figure 7

Figure 7

Ligand 38 bound to wild-type (A and B), K103N/Y181C (C and D) and E138K/M184I (E and F). Snapshots were taken from random points in the last 15ns of molecular dynamics simulations. As far as possible, the bonds depicted in the graphs have been highlighted with the same color in the corresponding image (the exceptions are the C181•••OH hydrogen bond shown as the red line in D, and the two hydrogen bonds between K101 backbone NH and nearby oxygen atoms in the inhibitor shown as red and blue lines in F)

PeerJ Reviewing Manuscript

PeerJ Reviewing Manuscript