JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

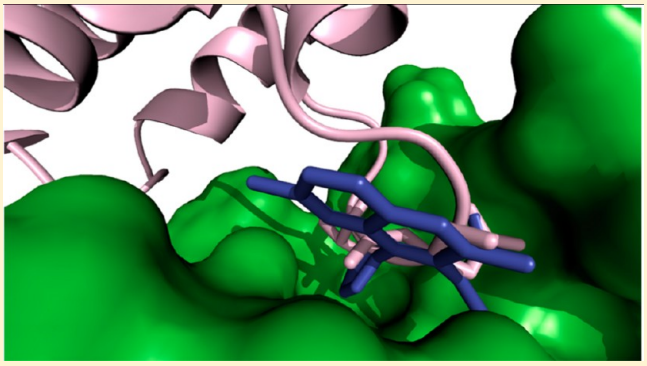
pubs.acs.org/jcim

Insight into the Fundamental Interactions between LEDGF Binding Site Inhibitors and Integrase Combining Docking and Molecular Dynamics Simulations

Laura De Luca,* Francesca Morreale,* and Alba Chimirri

Dipartimento di Scienze del Farmaco e Prodotti per la Salute, Università di Messina, Viale Annunziata, I-98168 Messina, Italy

ABSTRACT: In recent years, HIV-1 integrase (IN) has emerged as an attractive target for novel anti-AIDS agents. In particular, nonactive-site-binding IN inhibitors would display synergy with current strand-transfer-specific IN inhibitors and other antiretroviral drugs in clinical use. An effective allosteric inhibitory approach would be the disruption of protein−protein interaction (PPI) between IN and cellular cofactors, such as LEDGF/p75. To date, several small molecules have been reported to be inhibitors of the PPI between IN and LEDGF/p75. In this study, we investigated the most relevant interactions between five selected PPI inhibitors and IN comparing them to the naturally occurring IN-LEDGF/p75 complex. We calculated the binding free energies by using the method of molecular mechanics-generalized Born surface area (MM-GBSA). Total energy was decomposed on per residue contribution, and hydrogen bond occupancies were monitored throughout the simulations. Considering all these results we obtained a good correlation with experimental activity and useful insights for the development of new inhibitors.



INTRODUCTION

Acquired immunodeficiency syndrome (AIDS) is the most challenging pandemic of the 21st century. Although HAART has brought about a substantial decrease in the death rate, its efficacy has been limited by the emergence of drug-resistant viral strains and drug toxicity, so a refining of the current therapies and the developing of new therapeutic paradigms are still warranted.[1]

HIV-1 integrase (IN), the viral enzyme that catalyzes the integration of proviral cDNA into the host cell genome, has emerged as an attractive target for novel anti-AIDS agents.[2] Raltegravir (Merck Laboratories) is the first IN inhibitor approved for clinical use.[3] Other drugs in this class, such as elvitegravir (Gilead Sciences) and dolutegravir (GlaxoSmithKline), are in an advanced stage of development.[4,5]

The inevitability of drug resistance prompts the need to discover altogether new classes of IN inhibitors that target regions distant from the enzyme active site which are effective at disrupting IN function in vivo. A possible strategy could be the clinical use of nonactive-site-binding IN inhibitors that would also display synergy with current strand-transfer-specific IN inhibitors and other antiretroviral agents in clinical use.

IN is composed of three functional domains: the N terminal region, the catalytic core, and the C terminal region. The IN enzyme is a highly amenable target for allosteric drug development, as numerous allosteric inhibitory approaches are possible. One of them would be the disruption of interactions between IN and host cellular cofactors from which the integration process in vivo depends. Disruption of the protein−protein interactions (PPI) between IN and cellular cofactors represents a largely unexplored approach for the design and development of allosteric HIV/AIDS antiretroviral agents.[6] The association between IN and the cellular cofactor LEDGF/p75 is currently the most promising IN interaction for the design of PPI inhibitors (PPIIs).[6−8] LEDGF is a pro-survival protein that is closely associated with condensed chromatin in the nucleus acting as a chromatin-tethering factor to facilitate the interaction between IN and nuclear chromatin.[9,10]

Cherepanov et al. reported the crystal structure (PDB code: 2B4J) of the dimeric catalytic core domain of HIV-1 IN (IN_{CCD}) complexed to the IN binding domain of LEDGF ($LEDGF_{IBD}$). A number of peptides and small molecules have already been reported as PPIIs against the IN−LEDGF/p75 interaction. For two of them, the crystal structure of the complex between IN_{CCD} and the ligand was reported (PDB codes: 3LPT, 3LPU) confirming that the compounds occupy the same space in the LEDGF/p75-binding pocket of integrase; thus, the binding is mutually exclusive.

To further explore the crucial interactions between these inhibitors and HIV-1 IN, we decided to perform docking studies and molecular dynamics (MD) simulations on some complexes. Applying MD simulations to the final complexes of a docking study can have a dual use; they can refine the final structures and also be used to predict accurate binding free

Received: July 31, 2012
Published: November 29, 2012

dx.doi.org/10.1021/ci300361e | J. Chem. Inf. Model. 2012, 52, 3245−3254

energies. Furthermore, the time-dependent evolution of the system during the simulation provides a dynamic picture of the complex.[11]

This article reports the results of our study investigating the fundamental interactions between five selected inhibitors and the HIV-1 IN enzyme comparing them to the naturally occurring IN-LEDGF/p75 complex. Binding free energies were calculated using the molecular mechanics-generalized Born surface area (MM-GBSA) method.[12] Total energy was decomposed on per residue contribution, and hydrogen bond occupancies were monitored throughout the simulations, with the aim to correlate the obtained results with experimental activity and gain useful insights for the development of new inhibitors.

MATERIALS AND METHODS

Docking Studies. The crystal structure of the dimeric catalytic core domain of HIV-1 IN complexed with the LEDGF/p75 IBD was retrieved from the RCSB Protein Data Bank (entry code 2B4J)[13] and used for our docking simulations. The LEDGF structure was removed, and hydrogen atoms were added to the IN protein in Discovery Studio 2.5.5.[14] Docking studies were performed using the genetic optimization for ligand docking (GOLD) software package version 4.1.1 from the Cambridge Crystallographic Data Centre (CCDC).[15] In order to evaluate the performance of our docking calculations, we tried to reproduce the crystallographic position of compounds **1** and **2** (PDB codes 3LPT and 3LPU, respectively)[16] using the five different scoring functions implemented in GOLD. GoldScore proved to be the most effective scoring function in reproducing the experimentally determined binding modes (RMSD values 1.01 Å for compound **1** and 1.07 Å for compound **2**). The standard default settings were used in all calculations. For each of the 100 independent genetic algorithm runs, a default maximum of 100 000 genetic operations was performed using the default operator weights and a population size of 100 chromosomes. Default cutoff values of 2.5 Å for hydrogen bonds and 4.0 Å for van der Waals interactions were employed. Automatic bond settings were used, allowing the torsion angles of all acyclic, rotatable bonds in the ligand to vary except for amide bonds. Results differing by less than 0.75 Å in ligand-all-atom RMSD were clustered together. A 20.0 Å radius active site was drawn on the original position of the LEDGF/p75$_{IBD}$ dipeptide Ile365−Asp366, and automated cavity detection was used. Two rotatable bonds of the diketo acid moiety of the ligands were kept fixed during docking calculation. The two molecules containing the diketo acid moiety were simulated in the monoionized form.[17]

Molecular Dynamics Simulations. The starting model for simulations of IN-LEDGF/p75 was prepared from the X-ray structure 2B4J of the dimeric IN$_{CCD}$ (chains A and B), complexed with the LEDGF$_{IBD}$ (chains C and D).[13] The water molecules and chain D were discarded. The missing residues in chains A and B of the IN-LEDGF/p75 complex were reconstructed by superimposing chain C of the HIV-1 IN 1BL3[18] structure and energy-minimized using the Maestro[19] protein preparation wizard to release possible steric clashes and overlaps of side chains, until an RMSD of 0.30 Å was reached. This IN-LEDGF complex containing chains A, B, and C was used for MD simulations. For the simulation of IN-PPIIs complexes, chain C was discarded, and the crystallographic position for compounds **1** and **2** and docking poses for compound **3**, CHIBA-3053, and CHI-1043 were used.

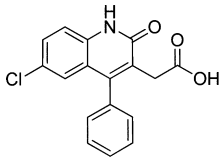
MD simulations were run using the PMEMD module of AMBER 11[20] and *parm99.dat* and *frcmod.ff03* parameter files.[21] General Amber force field (GAFF)[22] parameters were assigned to the ligands, while partial charges were calculated using the AM1-BCC method as implemented in the Antechamber suite of AMBER 11. An appropriate number of counterions were added to neutralize the system with the XLEAP module of AMBER 11. The system was immersed in a parallelepiped box of TIP3P water molecules, extending at least 10 Å from the protein. The geometry of the system was minimized in two steps. First, the water molecules were refined through 1000 steps of steepest descent followed by conjugate gradient until an RMSD of 0.10 Å was reached, keeping the protein fixed with a constraint of 500 kcal/mol. Second, the whole system was optimized with 1000 steps of steepest descent followed by conjugate gradient until an RMSD of 0.10 Å was reached, applying a constraint of 10 kcal/mol on the α carbons. MD trajectories were run using the minimized structure as a starting input. Constant volume simulations were performed for 500 ps, during which time the temperature was raised from 0 to 300 K using the Langevin dynamics method. Then, 1500 ps of constant-pressure MD simulations were performed at 300 K in three steps of 500 ps each. During the three periods of this second stage, the α carbons were blocked with harmonic force constants of 10, 5, and 1 kcal/mol·Å. Finally, a 8 ns MD simulation without restraint was run at a constant temperature of 300 K and a constant pressure of 1 atm. During the simulations, the particle mesh Ewald method was employed to calculate the long-range electrostatic interactions, while the SHAKE method was applied to constrain all covalent bonds involving hydrogen atoms. A 10 Å cutoff value was used for the nonbonded interactions, and a time step of 2 fs was used for the simulations. A similar procedure has been successfully applied to study analogue systems.[23,24]

Analysis of MD Trajectories and Free Energy Calculation. Snapshots of the complexes during the simulation and the average structures, with solvent and counterions stripped away, were obtained with the Ptraj module of the AMBER 11 suite.[20] Ptraj was also employed to obtain time-dependence of the RMSD of the α carbons and of the ligand; furthermore the hydrogen bonds were detected when the acceptor−donor atom distance was lower than 3.5 Å and the acceptor−H−donor angle was more than 120°. The MM-GBSA method,[12] implemented in the AMBER program, has been used for free energy calculation and to investigate the energetic contributions of each residue to ligand binding. In the MM-GBSA calculations, the free energy of the ligand binding is estimated by taking into account the solvation energies of the interacting molecules, in addition to molecular mechanics (MM) energies. The contribution of polar solvation energy is calculated with the generalized Born (GB) implicit solvent model, whereas the nonpolar part of the solvation energy is dependent on the solvent accessible surface area (SA).[12] For MM-GBSA analysis, snapshots at 40 ps intervals were extracted from the last 4 ns of the MD trajectory, and the binding free energies were averaged over the ensemble of conformers produced (100 snapshots for each trajectory).

RESULTS AND DISCUSSION

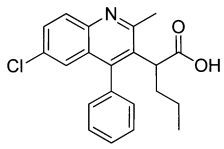
A total of six complexes were considered for this study: first, the complex between the two proteins (IN$_{CCD}$ and LEDGF$_{IBD}$)

Chart 1. Chemical Structures and Biological Activities[a]



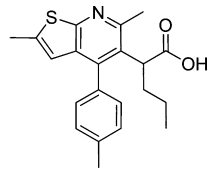
1

IC_{50} = 12.2 μM [a]



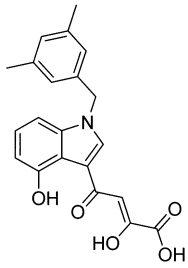
2

IC_{50} = 1.37 μM [a]



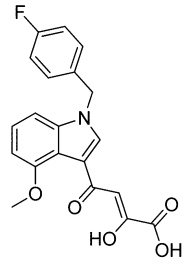
3

IC_{50} = 0.58 μM [a]



CHIBA-3053

IC_{50} = 3.5 μM [a]



CHI-1043

IC_{50} = 36.16 μM [a]

[a]Concentration required to inhibit *in vitro* IN−LEDGF/p75 interaction by 50% when tested in AlphaScreen assay.[25]

whose interaction has to be blocked; second, five inhibitors belonging to two different classes that bind to IN, preventing the interaction with LEDGF/p75. All the selected inhibitors (Chart 1) show micromolar and submicromolar activities when tested in AlphaScreen Assay for PPI inhibition,[25] together with a good profile when tested in cells.

The first class of inhibitors (**1**−**3**) includes a series of 2-(quinolin-3-yl)acetic acid derivatives named, by the authors, "LEDGINs," which are small molecules that target the IN−LEDGF/p75 interaction and thereby inhibit HIV replication.[16] The best results were obtained for derivative **3** (Chart 1). The crystal structures for **1** and **2** confirmed that they are bound in a cleft between the two monomers of the IN core dimer and occupy the same space in the LEDGF/p75-binding pocket.

The second class of inhibitors (CHIBA-3053 and CHI-1043) is comprised of indole derivatives that have been designed on the basis of computational studies considering the crucial role of several hot spot residues of the $LEDGF_{IBD}$, such as Ile365, Asp366, and Phe406.[26−28] The most active compound was 4-[1-(3,5-dimethylbenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid, named CHIBA-3053. Another representative of this class of diketoacid inhibitors, 4-[1-(4-fluorobenzyl)-4-methoxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic, named CHI-1043,[29] has shown a better tolerability in cells even if associated with a lower inhibition of IN−LEDGF interaction.[30] These two molecules were used in our study and compared to LEDGINs and LEDGF protein.

Static Structural Analysis of Binding Mode. The binding interface in the X-ray crystal structure of the IN−LEDGF/p75 complex has been widely described in the literature.[9,13] Specifically, the side chain of LEDGF residue Ile365 projects into a hydrophobic pocket formed by residues Leu102, Ala128, Ala129, and Trp132 of IN chain B and Thr174 and Met178 of chain A. LEDGF residues Phe406 and Val408 contact and occlude the solvent from the exposed Trp131 in IN chain B. Moreover, LEDGF Asp366 forms a bidentate hydrogen bond with the main chain amide of Glu170 and His171 of IN chain A. Furthermore, the backbone amide of LEDGF residue Ile365 establishes a hydrogen bond with the backbone carbonyl group of Gln168 in the IN chain A (Figure 1A).[13]

Overlaying the crystal structures of derivatives **1** and **2** bound to IN_{CCD} with the IN_{CCD}−$LEDGF_{IBD}$ cocrystal structure confirms that the compounds occupy the same space in the LEDGF/p75-binding pocket. In particular, the main chain nitrogens of residues Glu170 and His171 and the side-chain OH of Thr174 of integrase form hydrogen bonds with the carboxyl moiety of **1** and **2** (mimicking Asp366 of LEDGF/p75). Furthermore, the apolar side chain of Ala128 of integrase packs directly against compound **2** (clarifying the resistance of HIV-1 mutant Ala128Thr to these quinoline derivatives). The phenyl group of **1** and **2** occupies the same hydrophobic area of the Ile365 side chain. Moreover, the binding of the compounds does not lead to any significant conformation changes of the side chains in the pocket. Only minor side chain changes can be observed for the Gln95 and the Gln168.[16]

Looking at the crystal structures of the two LEDGINs (**1** and **2**), we decided to test our previously reported docking protocol together with the other scoring functions implemented in GOLD software.[15] We carried this out in order to find the most appropriate method to predict the binding mode of ligands whose crystal structure in complex with IN_{CCD} was not known. GoldScore proved to be the most effective scoring function in reproducing the experimentally determined binding modes and was thus chosen to predict the binding mode of compound **3** as well as of indole derivatives CHIBA-3053 and CHI-1043. After docking we analyzed clusters and poses, rescoring them with the MM-GBSA method implemented in AMBER. Usually, with respect to free energy calculations, scoring functions implemented within docking programs are not sufficiently accurate to identify the most stable conformation of a given ligand with the highest binding affinity among a set of

dx.doi.org/10.1021/ci300361e | J. Chem. Inf. Model. 2012, 52, 3245−3254

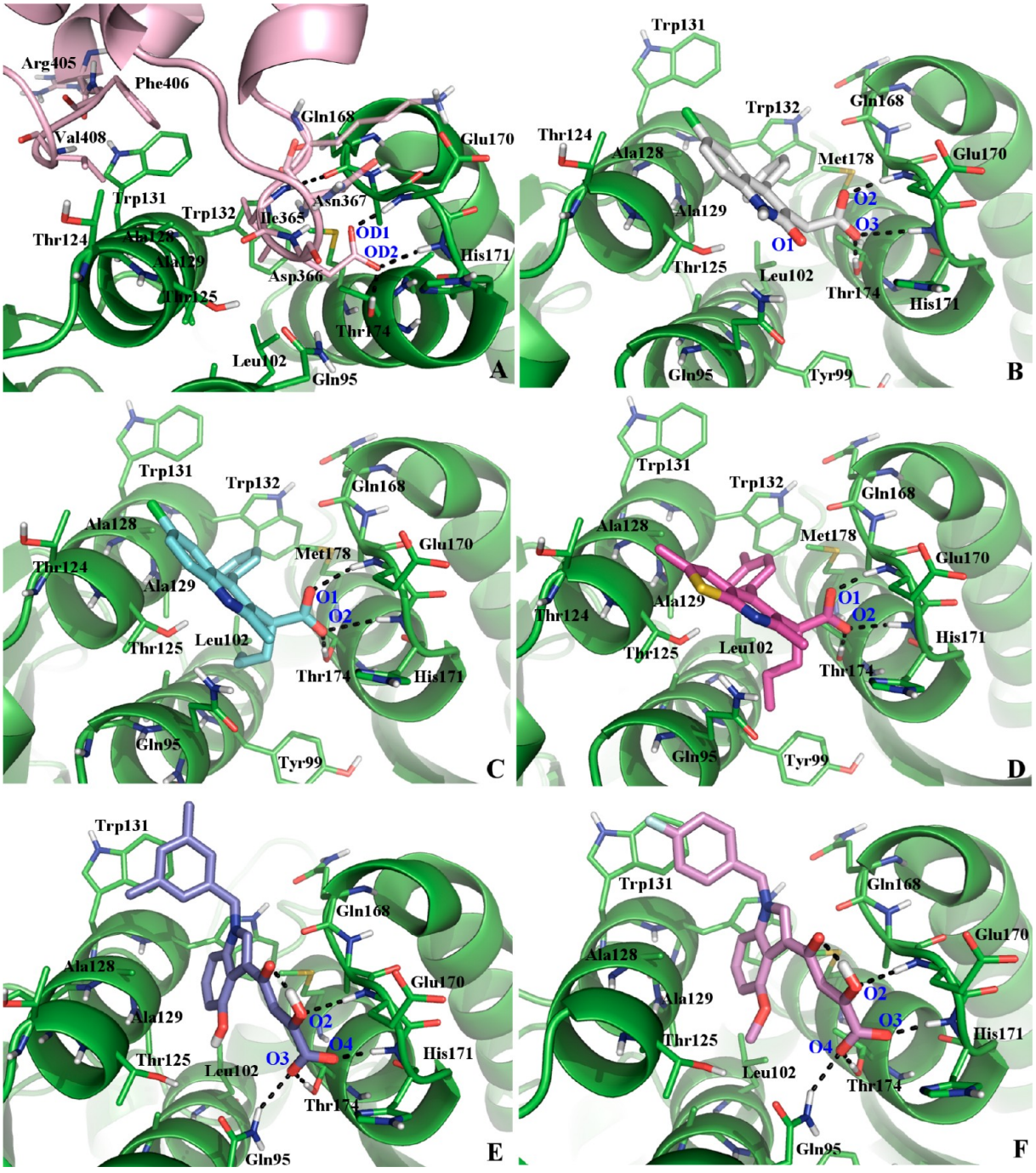


Figure 1. Static structural analysis of IN−LEDGF/p75 interface (A) and binding mode of compounds **1** (B), **2** (C), **3** (D), CHIBA-3053 (E), and CHI-1043 (F). Blue labels indicate oxygen atom numbers used in Tables 1−3. Black dots represent hydrogen bonds. This figure was produced with Pymol.[33]

compounds. In fact, in our case, docking calculations were not able to explain the better or worse activity of compounds, thus suggesting the need for more accurate molecular modeling studies. The best pose obtained is shown in Figure 1D−F, which was then chosen for MD simulations.

Compound **3** showed a binding mode that was almost identical to analogue **2**. CHIBA-3053 and CHI-1043 show similar binding modes (Figure 1). The diketo acid moiety forms hydrogen bonds with the main chain nitrogens of residues Glu170 and His171 and the side chain of residues Thr174 and Gln95. The fused benzene ring of the indole nucleus projects into the IN hydrophobic pocket, and the N-benzyl substituent presents hydrophobic contacts for the crucial Trp131 residue of chain B that were more consistent with 3,5-dimethyl substitution. Moreover, the 4-methoxy group of CHI-1043 could create an accessory hydrophobic contact with the IN hydrophobic pocket.

MD Simulations. The main objective of this study was to highlight the important interactions between the selected IN−LEDGF interaction inhibitors and to correlate them, and other computational results, with experimental activity. Molecular dynamics was chosen as a tool to monitor the time-dependent evolution of the described systems during the simulation. As a reference, we first performed MD simulation on the complex

dx.doi.org/10.1021/ci300361e | J. Chem. Inf. Model. 2012, 52, 3245−3254

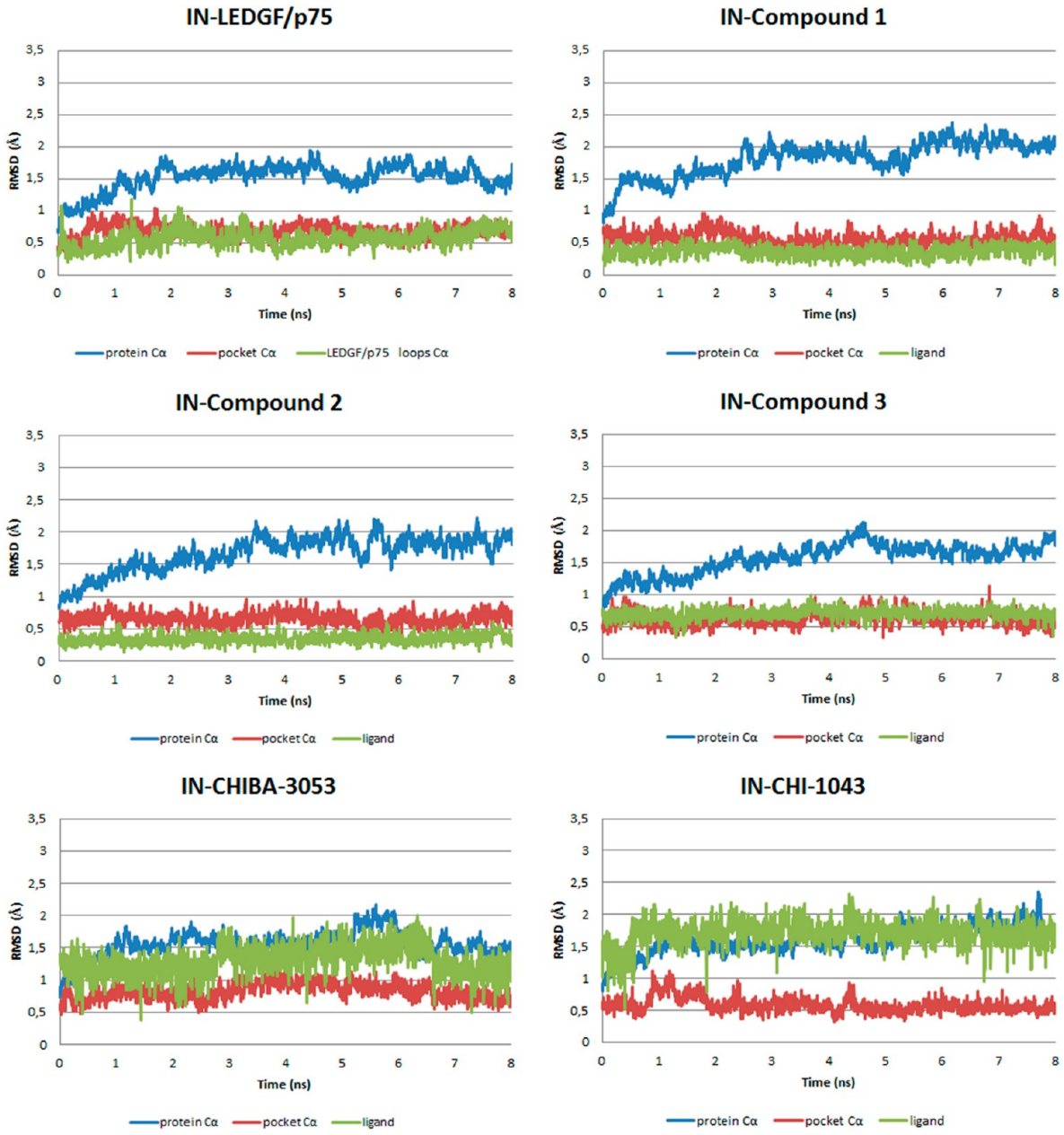


Figure 2. RMSDs of Cα atoms of the protein, Cα of binding pocket, and ligands as a function of the simulation time. For the IN−LEDGF complex, Cα atoms of the protein, Cα of binding pocket, and Cα of LEDGF loops binding IN were considered.

between the $LEDGF_{IBD}$ and IN_{CCD} dimer because understanding interfacial contributions of single amino acids and stable hydrogen bonds between the two proteins could be useful for the explanation of the different experimental activities of reported inhibitors. We then considered the crystallized complexes of **1** and **2** and the resulting docking complexes for **3**, CHIBA-3053, and CHI-1043.

System Stability. The stability of the six systems was determined by RMSD analysis of the MD trajectories obtained from simulations with respect to the initial structure considering the ligand, the Cα of the whole complex, and Cα of the binding pocket residues (residues within a 6 Å sphere from the bound inhibitor). The results are plotted in Figure 2.

As can be seen in the plots, the RMSDs of CHIBA-3053 and CHI-1043 are relatively higher than those of the LEDGINs derivatives. The IN−CHIBA-3053 complex shows a higher RMSD value for the ligand among 3 and 7 ns, that is consistent with a temporary variation in hydrogen bonds. After 4 ns, the RMSD of Cα tends to converge, indicating that the systems are stable and equilibrated. Thus, the MD trajectories from the last 4 ns of simulation of all systems were taken for MM-GBSA analysis.[12]

Hydrogen-Bonding Interactions. With the aim of highlighting stable and unstable hydrogen bonding of the starting structure, H-bond interactions were calculated considering all steps of the MD simulation and are shown in Tables 1−3, when the occupancies were more than 20% in the investigated time period; distances between involved atoms are also indicated.

In the IN-LEDGF/p75 interface there were three very stable hydrogen bonds whose occupancies were more than 90%, they are (1) the hydrogen bond between the backbone NH of

dx.doi.org/10.1021/ci300361e | J. Chem. Inf. Model. 2012, 52, 3245−3254

Table 1. Hydrogen Bond Analysis from the Results of MD Simulation for IN-LEDGF/p75 Complex

IN complex with	donor	acceptor	occupancy (%)[a]	distance (Å)[b]
LEDGF/p75	Glu170(A) NH	Asp366(C) OD1	99.76	2.826 (0.13)
	Thr174(A) OH	Asp366(C) OD2	98.36	2.763 (0.16)
	Ile365(C) NH	Gln168(C) O	96.36	3.033 (0.18)
	His171(A) NH	Asp366(C) OD2	65.96	3.207 (0.18)
	Gln95(B) NH_2	Asp366(C) O	55.08	2.924 (0.16)
	Glu170(A) NH	Asp366(C) OD2	46.44	3.246 (0.18)
	Thr125(B) OH	Ile365(C) O	42.04	2.748 (0.16)
	Asn367(C) NH_2	Glu170(A) OE2	35.64	2.959 (0.18)
	Asn367(C) NH_2	Glu170(A) OE1	30.56	2.956 (0.18)
	Trp131(B) NH	Arg405(C) O	30.16	3.046 (0.21)

[a]The listed donor and acceptor pairs satisfy the criteria for the hydrogen bond over 20.0% occupancy during the entire simulation. [b]The average distance with standard error (SE = standard deviation/$N^{1/2}$) in parentheses between the hydrogen-acceptor atom and hydrogen-donor atom in the investigated time period.

Table 2. Hydrogen Bond Analysis from the Results of MD Simulations for IN in Complex with Compounds 1−3

IN complex with	donor	acceptor	occupancy (%)[a]	distance (Å)[b]
compound **1**	Thr174(A) OH	ligand O3	98.76	2.700 (0.14)
	Glu170(A) NH	ligand O2	95.48	2.868 (0.14)
	Glu170(A) NH	ligand O3	61.52	3.159 (0.19)
	Gln95(B) NH_2	ligand O1	51.44	2.939 (0.17)
	His171(A) NH	ligand O3	47.12	3.226 (0.18)
compound **2**	Thr174(A) OH	ligand O2	98.72	2.699 (0.13)
	Glu170(A) NH	ligand O1	98.32	2.806 (0.11)
	Glu170(A) NH	ligand O2	52.08	3.236 (0.18)
	His171(A) NH	ligand O2	43.20	3.257 (0.16)
compound **3**	Glu170(A) NH	ligand O1	99.88	2.794 (0.10)
	Thr174(A) OH	ligand O2	99.80	2.712 (0.14)
	His171(A) NH	ligand O2	35.36	3.303 (0.15)
	Glu170(A) NH	ligand O2	32.52	3.301 (0.15)

[a]The listed donor and acceptor pairs satisfy the criteria for the hydrogen bond over 20.0% occupancy during the entire simulation. [b]The average distance with standard error (SE = standard deviation/$N^{1/2}$) in parentheses between hydrogen-acceptor atom and hydrogen-donor atom in the investigated time period.

Glu170 in chain A of IN and the OD1 atom of carboxylate of Asp366 in LEDGF (the OD1 atom of Asp366 formed a very stable hydrogen bond with Glu170, but after a turning point, the OD2 atom took the place of OD1, indicating that the two O atoms of Asp366 rotated about 180° during the MD simulations); (2) the hydrogen bond between the side chain OH of Thr174 in chain A of IN and the OD2 atom of carboxylate of Asp366 in LEDGF/p75; (3) the hydrogen bond between the backbone NH of Ile365 in LEDGF and the carbonyl oxygen of Gln168 in chain A of IN.

Other hydrogen bonds were found between the OD2 atom of Asp366 (LEDGF) and NH backbone of His171 (IN chain A); the carbonyl oxygen of Asp366 (LEDGF) and NH_2 of side chain of Gln95 (IN chain B); the carbonyl oxygen of Ile365 (LEDGF) and side chain OH of Thr125 (IN chain B); the side chain NH of Trp131 (IN chain B), and the carbonyl oxygen of Arg405 (LEDGF). Finally, the two oxygen atoms of the carboxylate in Glu170 (IN chain A) formed hydrogen bonds with the side chain NH_2 of Asn367 (LEDGF), together with a discontinuous salt bridge with NH_3 of Lys364 (data not shown). These results are in agreement with a previous study.[31]

For the three protein-inhibitor systems of **1−3** derivatives, the hydrogen bond interactions with key residues in the binding pocket are listed in Table 2. As observed for the IN−LEDGF complex, the three inhibitors formed stable hydrogen bonds with the Thr174 (IN chain A) OH and Glu170 (IN chain A) NH through the two oxygens of the carboxylate group, that after a turning point were observed to rotate about 180° during the MD simulations. Furthermore, lower occupancy was observed for the hydrogen bond with His171 (IN chain A) NH. The main difference in hydrogen bonds, between this class of inhibitors and LEDGF protein bound to IN was the lack of an interaction with Gln168 in chain A of IN that was very stable in the complex between the two proteins.

Looking at the indole derivatives, hydrogen bonds were less stable, if compared to LEDGINs (Table 3). In both complexes, the diketo acid moiety forms the same hydrogen bonds in the starting structure (one oxygen atom of the carboxylate interacts with the main chain nitrogen of residue His171 while the other one binds to the side chain of residues Thr174 and Gln95; O2 forms an hydrogen bond with the main chain nitrogen of residue Glu170), but the occupancy proved to be different

Table 3. Hydrogen Bond Analysis from the Results of MD Simulations for IN in Complex with Compounds CHIBA-3053 and CHI-1043

IN complex with	donor	acceptor	occupancy (%)[a]	distance (Å)[b]
CHIBA-3053	Thr174(A) OH	ligand O3	61.60	2.672 (0.12)
	His171(A) NH	ligand O4	52.48	3.083 (0.19)
	Glu170(A) NH	ligand O2	44.32	3.204 (0.17)
	Glu170(A) NH	ligand O3	37.12	2.955 (0.17)
	Gln95(B) NH_2	ligand O4	25.24	2.927 (0.19)
CHI-1043	Gln95(B) NH_2	ligand O3	30.24	2.929 (0.18)
	Thr174(A) OH	ligand O4	27.96	2.688 (0.17)

[a]The listed donor and acceptor pairs satisfy the criteria for the hydrogen bond over 20.0% occupancy during the entire simulation. [b]The average distance with standard error (SE = standard deviation/$N^{1/2}$) in parentheses between hydrogen-acceptor atom and hydrogen-donor atom in the investigated time period.

dx.doi.org/10.1021/ci300361e | J. Chem. Inf. Model. 2012, 52, 3245−3254

Table 4. Binding Free Energy and Its Components for the Six Complexes[a]

	IN-1	IN-2	IN-3	IN-CHIBA-3053	IN-CHI-1043	IN-LEDGF/p75
total nonpolar contributions	−29.15	−35.95	−37.88	−42.09	−41.86	−80.66
total electrostatic contribution	7.48	8.25	7.7	16.82	16.84	33.4
ΔG_{bind}	−21.68	−27.70	−30.19	−25.27	−25.02	−47.26
IC_{50} (μM)	12.2	1.37	0.58	3.5	36.16	

[a]Mean energies are in kcal/mol. Total nonpolar contributions = van der Waals energy term (ΔG_{VDW}) + nonpolar desolvation term ($\Delta G_{nonpolar}$). Total electrostatic contribution = electrostatic energy term (ΔG_{ELE}) + polar desolvation term (ΔG_{GB}).

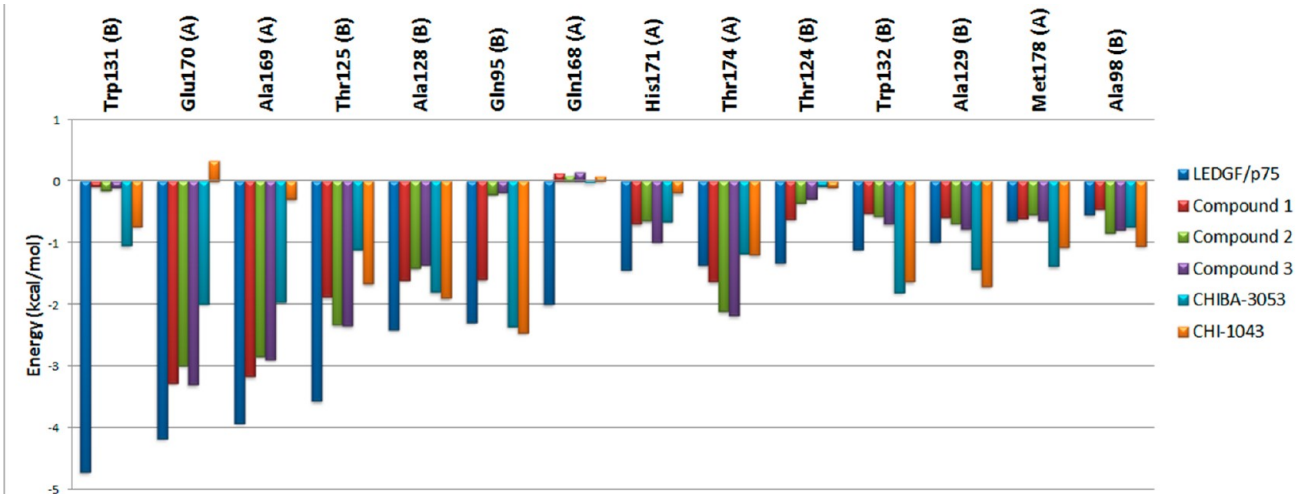


Figure 3. Comparison of the contribution of key residues to the binding energy of the six complexes.

during the simulation. In the IN_{CCD}−CHIBA-3053 complex, the O3 of the carboxylate group switched from Thr174 to Glu170 (taking the place of O2 that was interacting with Glu170) throughout the simulation, while O4 moved from His171 to Gln95. CHI-1043 lost several hydrogen bonds (Glu170 and His171) after the first nanoseconds of the simulation. Only two hydrogen bonds were somehow maintained, involving the two oxygen atoms of the carboxylate group (O3 and O4) and the side chain of residues Thr174 and Gln95.

Binding Free Energy. To quantitatively gain insight into different contributions to the affinity of ligands binding to IN and the main driving force for their binding, the MM-GBSA method[12] was used to calculate binding free energy for each complex with 100 snapshots taken from the last 4 ns of the MD trajectory.

For the IN−LEDGF complex, the total energy of binding was −47.26 kcal/mol. In agreement with previous studies,[31] the decomposition of the binding free energy showed that the nonpolar contribution dominated IN−LEDGF/p75 complex interactions. In contrast to the nonpolar contribution, the electrostatic contribution was not significant for the binding of IN with LEDGF/p75. Although the electrostatic contribution in the gas phase was favorable, it was not large enough to compensate the desolvation penalties associated with the binding event.

The binding free energies of the five IN−inhibitor complexes calculated by using the MM-GBSA method are shown in Table 4, together with experimental IC_{50}; as can be seen, the ranking of the predicted binding free energies are in good agreement with the experimental IC_{50} except for CHI-1043. The overestimation of binding energy of compound CHI-1043 is however counteracted by the lack of fundamental hydrogen bonding during the simulation and, as shown below in the text, by a different interaction energy with hot-spot residues.

In the five studied protein−inhibitor complexes, the van der Waals interactions and the nonpolar solvation energies responsible for the burial of the inhibitor's hydrophobic groups upon binding are the basis for favorable binding free energies, as observed for the IN−LEDGF complex. The favorable Coulomb interactions within the protein−inhibitor complexes are counteracted by the unfavorable electrostatics of desolvation. The resulting balance of the electrostatic interaction contributions in a vacuum and in solvent is unfavorable to binding in all of the systems. A particularly detrimental electrostatic contribution can be observed for the two molecules with the diketo acid moiety that is however compensated by a greater hydrophobic contribution.

All these binding free energy results seem to confirm that PPI inhibitors tend to be lipophilic molecules with few hydrogen bond features that are however essential to conferring target specificity. Indeed, in a drug discovery context, molecules with such a lipophilic profile could be considered potentially promiscuous and therefore unattractive.[32]

Residue Decomposition. The binding free energy was decomposed into the contribution of each residue to the binding of the five inhibitors to IN_{CCD} and compared to its cofactor, as can be seen in Figure 3. The major favorable energy contributions in the IN−LEDGF complex originate from residues Trp131, Glu170, Ala169, Thr125, Ala128, Gln95, Gln168, His171, Thr174, Thr124, Trp132, Ala129, Met178, and Ala98, whose contribution to ΔG_{bind} is lower than −0.50 kcal/mol. If compared to the IN−inhibitor complexes, one of the main differences is a significantly lower interaction with residue Trp131; indeed a relatively good interaction can be achieved only by CHIBA-3053 and CHI-1043. Almost all

dx.doi.org/10.1021/ci300361e | J. Chem. Inf. Model. 2012, 52, 3245−3254

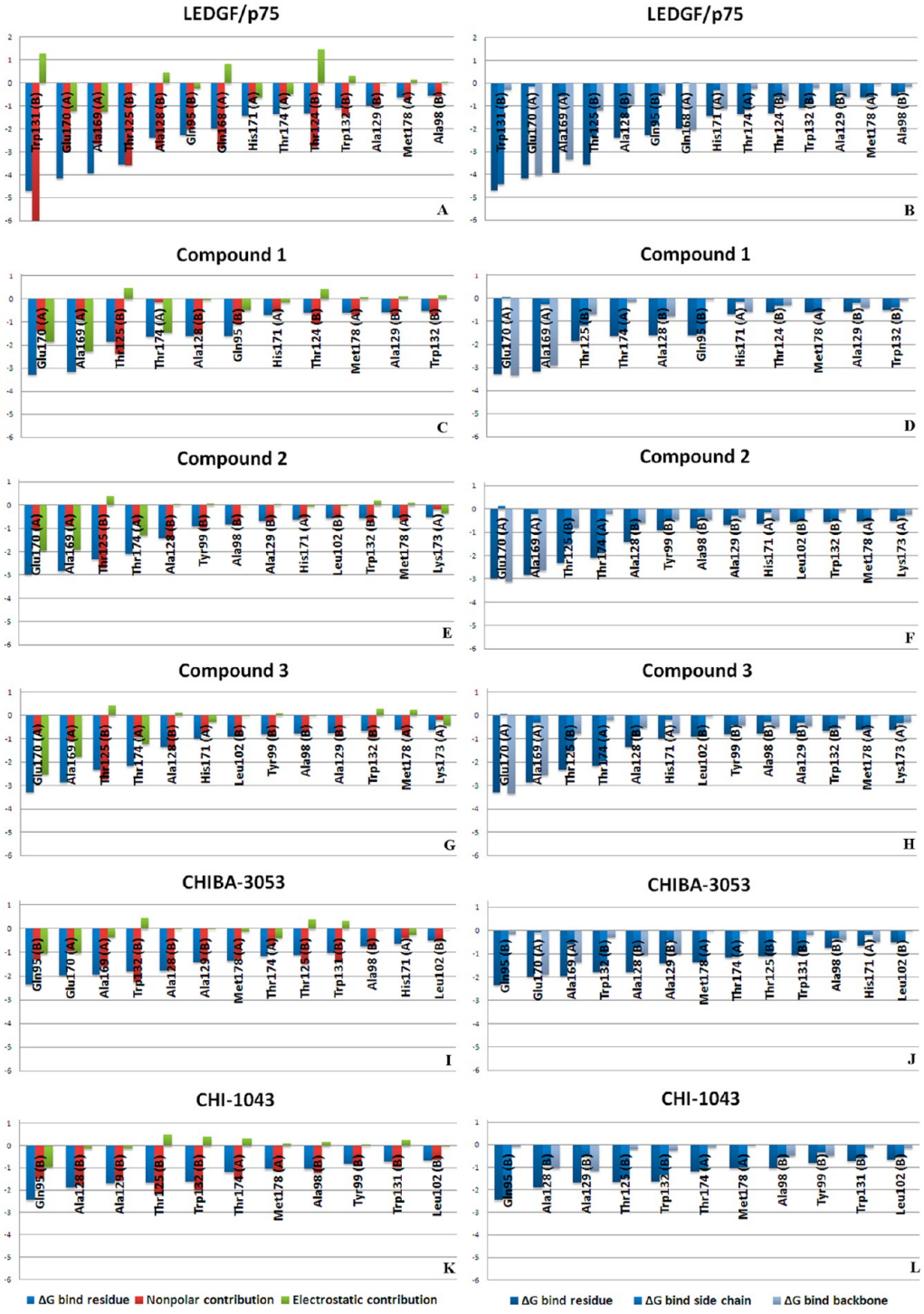


Figure 4. Total electrostatic contribution and total nonpolar contribution of each key residue in kcal/mol (panels A, C, E, G, I, K). Contributions from backbone atoms ($\Delta G_{\text{bind-backbone}}$) and side chain atoms ($\Delta G_{\text{bind-side-chain}}$) of each key residue in kcal/mol (panels B, D, F, H, J, L). Only residues showing $\Delta G_{\text{bind-res}} \leq -0.50$ kcal/mol were taken into account.

compounds show favorable interactions with residues Glu170, Ala169, Thr125, Ala128, Gln95; importantly CHI-1043 presents unfavorable energy with Glu170, thus justifying its lower activity. Finally, it is worth noting that none of these compounds is able to show good values of energy for residue Gln168 that seem to be essential for LEDGF binding.

Furthermore, for each complex the net electrostatic contribution and the total nonpolar contribution of each residue were evaluated, together with the contributions from backbone atoms ($\Delta G_{\text{bind-backbone}}$) and side chain atoms ($\Delta G_{\text{bind-side-chain}}$). Results are summarized in Figure 4, only residues showing $\Delta G_{\text{bind-res}} \leq -0.50$ kcal/mol were taken into account.

The main electrostatic contribution for the LEDGINs series derives from residues Glu170, Ala169, and Thr174, together with Gln95 and His171 for compound **1**. CHIBA-3053 shows a similar behavior even if values are relatively lower, and an important contribution comes from residue Gln95. On the

dx.doi.org/10.1021/ci300361e | J. Chem. Inf. Model. 2012, 52, 3245−3254

other hand, CHI-1043 has a relevant electrostatic contribution only from Gln95. These results are consistent with hydrogen bond formation, considering that H bonds were monitored throughout the complete MD simulation while in this case only the last 4 ns were examined. In fact, on the right panels, we can see that the main contribution for Glu170 and Ala169 derives from backbone atoms while Thr174 contributes through the side chain. Moreover, we can see that the side chain of Glu170 shows unfavorable energy; this may be due to the electrostatic repulsion with the carboxylate of inhibitors, while in the IN−LEDGF/p75 complex the Glu170 side chain remains involved in H bonds with Asn367 and salt bridge with Lys364 of $LEDGF_{IBD}$.

Nonpolar contributions originate from side chains of residues surrounding the hydrophobic pocket, that are Ala128, Thr125, Met178, Ala129, Trp132, and Leu102, together with Trp131 for CHIBA-3053 and CHI-1043. The propyl group of **2** and **3**, which seemed to determine the better activity of these two compounds with respect to **1**, is responsible for the additional nonpolar interactions with residues Tyr99 and Ala98; the same role is somehow maintained by the methoxy group of CHI-1043.

Implications for Inhibitor Design. Analyzing the results of this study, several tips for future inhibitor design could be suggested.

(1) Static structural analysis of IN−PPI inhibitor complexes should be supported by dynamic analysis. As seen for compounds CHI-1043 and CHIBA-3053, H bonds from the starting structures appear to be exactly the same while MD simulations suggest a different stability for the two starting models, justifying the different activity. Furthermore, compound RMSD gives information about pose stability.

(2) Targeting the hot-spot amino acids is important for activity. The PPI inhibitor should exhibit the same interactions as the LEDGF protein. Decomposing binding free energy into per residue contribution could be particularly useful.

(3) The main driving force for ligand binding is hydrophobic. Even if electrostatic contribution in the gas phase is favorable, it is not large enough to compensate for the desolvation penalties associated with the binding event. Increasing hydrophobic contributions could help to improve activity. As regards the examined inhibitors of the LEDGIN series, increasing nonpolar contribution resulted in a better activity. In particular, the extra propyl group of **2** and **3** is responsible for the additional nonpolar interactions with residues Tyr99 and Ala98. However, none of these compounds showed significant interaction energy with Trp131 except CHI-1043 and CHIBA-3053 derivatives. Merging these features together should result in an activity improvement.

(4) The inhibitor should form stable hydrogen bonds, in particular with the NH backbone of Glu170, OH the of side chain of Thr174, carbonyl oxygen of the backbone of Gln168, and the NH backbone of His171. The examined inhibitors lack the hydrogen bond donor to bind Gln168. Different occupancies during MD simulations should be taken into account as described above.

(5) The best way to avoid false positives would be to evaluate simultaneously multiple parameters and to avoid trusting one single result, such as the binding affinity alone.

CONCLUSIONS

In this study, docking, MD simulation, and free energy calculation were used to obtain detailed information on the key interactions and dynamic properties of five IN-LEDGF/p75 interaction inhibitors. The comparison between the behavior of the original IN−LEDGF/p75 complex and the five IN−inhibitor complexes highlighted the fundamental interactions related with biological activity, together with the success and pitfalls of the examined PPI inhibitors. On one hand, the LEDGINs compounds form more stable hydrogen bonds with Glu170, Thr174, and His171; on the other hand, CHIBA-3053 and CHI-1043 derivatives showed better hydrophobic interactions, in particular with IN residue Trp131. In both series, however, we observed a lack of a hydrogen bond with IN residue Gln168, which was one of the most important residues at the IN−LEDGF interface, in terms of energy contribution and hydrogen bond occupancy.

We can conclude that the use of docking calculation together with MD simulation could be a useful tool in the discovery and development of novel and more potent IN−LEDGF/p75 interaction inhibitors.

AUTHOR INFORMATION

Corresponding Author

*Phone: 00390906766464. Fax: 00390906766402. E-mail: ldeluca@unime.it (L.D.L.), fmorreale@unime.it (F. M.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the European Commission THINC project (HEALTH-F3-2008-201032).

ABBREVIATIONS

AIDS, acquired immunodeficiency syndrome; CCD, catalytic core domain; HAART, highly active antiretroviral therapy; HIV-1, human immunodeficiency virus type 1; IBD, integrase binding domain; IN, integrase; LEDGF, lens epithelium-derived growth factor; MD, molecular dynamics; MM-GBSA, molecular mechanics-generalized Born surface area; PDB, protein data bank; PPI, protein−protein interaction; PPIIs, protein−protein interaction inhibitors; RMSD, root-mean-square deviation

REFERENCES

(1) Mehellou, Y.; De Clercq, E. Twenty-six years of anti-HIV drug discovery: where do we stand and where do we go? *J. Med. Chem.* **2010**, *53*, 521−538.

(2) Ramkumar, K.; Serrao, E.; Odde, S.; Neamati, N. HIV-1 integrase inhibitors: 2007−2008 update. *Med. Res. Rev.* **2010**, *30*, 890−954.

(3) FDA approves raltegravir tablets. *AIDS Patient Care STDS.* **2007**, *21*, 889.

(4) Molina, J. M.; Lamarca, A.; Andrade-Villanueva, J.; Clotet, B.; Clumeck, N.; Liu, Y. P.; Zhong, L.; Margot, N.; Cheng, A. K.; Chuck, S. L. Efficacy and safety of once daily elvitegravir versus twice daily raltegravir in treatment-experienced patients with HIV-1 receiving a ritonavir-boosted protease inhibitor: randomised, double-blind, phase 3, non-inferiority study. *Lancet Infect. Dis.* **2012**, *12*, 27−35.

(5) van Lunzen, J.; Maggiolo, F.; Arribas, J. R.; Rakhmanova, A.; Yeni, P.; Young, B.; Rockstroh, J. K.; Almond, S.; Song, I.; Brothers, C.; Min, S. Once daily dolutegravir (S/GSK1349572) in combination therapy in antiretroviral-naive adults with HIV: planned interim 48 week results from SPRING-1, a dose-ranging, randomised, phase 2b trial. *Lancet Infect. Dis.* **2012**, *12*, 111−118.

(6) Al-Mawsawi, L. Q.; Neamati, N. Allosteric Inhibitor Development Targeting HIV-1 Integrase. *ChemMedChem* **2011**, *6*, 228−241.

(7) De Luca, L.; Ferro, S.; Morreale, F.; Chimirri, A. Inhibition of the interaction between HIV-1 integrase and its cofactor LEDGF/p75: a promising approach in anti-retroviral therapy. *Mini-Rev. Med. Chem.* **2011**, *11*, 714−727.
(8) De Luca, L.; Ferro, S.; Morreale, F.; De Grazia, S.; Chimirri, A. Inhibitors of the interactions between HIV-1 IN and the cofactor LEDGF/p75. *ChemMedChem.* **2011**, *6*, 1184−1191.
(9) Al-Mawsawi, L. Q.; Neamati, N. Blocking interactions between HIV-1 integrase and cellular cofactors: an emerging anti-retroviral strategy. *Trends Pharmacol. Sci.* **2007**, *28*, 526−535.
(10) Engelman, A.; Cherepanov, P. The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS Pathog.* **2008**, *4*, e1000046.
(11) Alonso, H.; Bliznyuk, A. A.; Gready, J. E. Combining docking and molecular dynamic simulations in drug design. *Med. Res. Rev.* **2006**, *26*, 531−568.
(12) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., 3rd. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(13) Cherepanov, P.; Ambrosio, A. L.; Rahman, S.; Ellenberger, T.; Engelman, A. Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17308−17313.
(14) *Discovery Studio*, version 2.5.5; Accelrys: San Diego, CA, 2009.
(15) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727−748.
(16) Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmie, B. A.; Marchand, D.; Bardiot, D.; Van der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; De Maeyer, M.; Chaltin, P.; Debyser, Z. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat. Chem. Biol.* **2010**, *6*, 442−448.
(17) Sechi, M.; Bacchi, A.; Carcelli, M.; Compari, C.; Duce, E.; Fisicaro, E.; Rogolino, D.; Gates, P.; Derudas, M.; Al-Mawsawi, L. Q.; Neamati, N. From ligand to complexes: inhibition of human immunodeficiency virus type 1 integrase by beta-diketo acid metal complexes. *J. Med. Chem.* **2006**, *49*, 4248−4260.
(18) Maignan, S.; Guilloteau, J. P.; Zhou-Liu, Q.; Clement-Mella, C.; Mikol, V. Crystal structures of the catalytic domain of HIV-1 integrase free and complexed with its metal cofactor: high level of similarity of the active site with other viral integrases. *J. Mol. Biol.* **1998**, *282*, 359−368.
(19) *Maestro*, version 9.0; Schrödinger LLC: New York, NY, 2009.
(20) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 11*; University of California: San Francisco, CA, 2010.
(21) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(22) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(23) Tintori, C.; Veljkovic, N.; Veljkovic, V.; Botta, M. Computational studies of the interaction between the HIV-1 integrase tetramer and the cofactor LEDGF/p75: insights from molecular dynamics simulations and the informational spectrum method. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 3396−3408.
(24) Tintori, C.; Demeulemeester, J.; Franchi, L.; Massa, S.; Debyser, Z.; Christ, F.; Botta, M. Discovery of small molecule HIV-1 integrase dimerization inhibitors. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 3109−3114.
(25) Al-Mawsawi, L. Q.; Christ, F.; Dayam, R.; Debyser, Z.; Neamati, N. Inhibitory profile of a LEDGF/p75 peptide against HIV-1 integrase: insight into integrase-DNA complex formation and catalysis. *FEBS Lett.* **2008**, *582*, 1425−1430.
(26) De Luca, L.; Barreca, M. L.; Ferro, S.; Christ, F.; Iraci, N.; Gitto, R.; Monforte, A. M.; Debyser, Z.; Chimirri, A. Pharmacophore-based discovery of small-molecule inhibitors of protein-protein interactions between HIV-1 integrase and cellular cofactor LEDGF/p75. *ChemMedChem.* **2009**, *4*, 1311−1316.
(27) De Luca, L.; Ferro, S.; Gitto, R.; Barreca, M. L.; Agnello, S.; Christ, F.; Debyser, Z.; Chimirri, A. Small molecules targeting the interaction between HIV-1 integrase and LEDGF/p75 cofactor. *Bioorg. Med. Chem.* **2010**, *18*, 7515−7521.
(28) De Luca, L.; Ferro, S.; Morreale, F.; Christ, F.; Debyser, Z.; Chimirri, A.; Gitto, R. Fragment hopping approach directed at design of HIV IN-LEDGF/p75 interaction inhibitors. *J. Enzyme Inhib. Med. Chem.* DOI: dx.doi.org/10.3109/14756366.2012.703184.
(29) De Luca, L.; Barreca, M. L.; Ferro, S.; Iraci, N.; Michiels, M.; Christ, F.; Debyser, Z.; Witvrouw, M.; Chimirri, A. A refined pharmacophore model for HIV-1 integrase inhibitors: Optimization of potency in the 1H-benzylindole series. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2891−2895.
(30) De Luca, L.; Gitto, R.; Christ, F.; Ferro, S.; De Grazia, S.; Morreale, F.; Debyser, Z.; Chimirri, A. 4-[1-(4-Fluorobenzyl)-4-hydroxy-1H-indol-3-yl]-2-hydroxy-4-oxobut-2-enoic acid as a prototype to develop dual inhibitors of HIV-1 integration process. *Antiviral Res.* **2011**, *92*, 102−107.
(31) Zhao, Y.; Li, W.; Zeng, J.; Liu, G.; Tang, Y. Insights into the interactions between HIV-1 integrase and human LEDGF/p75 by molecular dynamics simulation and free energy calculation. *Proteins: Struct., Funct., Bioinf.* **2008**, *72*, 635−645.
(32) Higueruelo, A. P.; Schreyer, A.; Bickerton, G. R.; Pitt, W. R.; Groom, C. R.; Blundell, T. L. Atomic interactions and profile of small molecules disrupting protein-protein interfaces: the TIMBAL database. *Chem. Biol. Drug Des.* **2009**, *74*, 457−467.
(33) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific LLC: San Carlos, CA, 2008.