

# Virtual screening with AutoDock Vina and the common pharmacophore engine of a low diversity library of fragments and hits against the three allosteric sites of HIV integrase: participation in the SAMPL4 protein–ligand binding challenge

Alexander L. Perryman · Daniel N. Santiago ·  
Stefano Forli · Diogo Santos-Martins ·  
Arthur J. Olson

Received: 15 November 2013 / Accepted: 11 January 2014 / Published online: 4 February 2014  
© Springer International Publishing Switzerland 2014

**Abstract** To rigorously assess the tools and protocols that can be used to understand and predict macromolecular recognition, and to gain more structural insight into three newly discovered allosteric binding sites on a critical drug target involved in the treatment of HIV infections, the Olson and Levy labs collaborated on the SAMPL4 challenge. This computational blind challenge involved predicting protein–ligand binding against the three allosteric sites of HIV integrase (IN), a viral enzyme for which two drugs (that target the active site) have been approved by the FDA. Positive control cross-docking experiments were utilized to select 13 receptor models out of an initial ensemble of 41 different crystal structures of HIV IN. These 13 models of the targets were selected using our new “Rank Difference Ratio” metric. The first stage of SAMPL4 involved using virtual screens to identify 62 active, allosteric IN inhibitors out of a

set of 321 compounds. The second stage involved predicting the binding site(s) and crystallographic binding mode(s) for 57 of these inhibitors. Our team submitted four entries for the first stage that utilized: (1) AutoDock Vina (AD Vina) plus visual inspection; (2) a new common pharmacophore engine; (3) BEDAM replica exchange free energy simulations, and a Consensus approach that combined the predictions of all three strategies. Even with the SAMPL4’s very challenging compound library that displayed a significantly lower amount of structural diversity than most libraries that are conventionally employed in prospective virtual screens, these approaches produced hit rates of 24, 25, 34, and 27 %, respectively, on a set with 19 % declared binders. Our only entry for the second stage challenge was based on the results of AD Vina plus visual inspection, and it ranked third place overall according to several different metrics provided by the SAMPL4 organizers. The successful results displayed by these approaches highlight the utility of the computational structure-based drug discovery tools and strategies that are being developed to advance the goals of the newly created, multi-institution, NIH-funded center called the “HIV Interaction and Viral Evolution Center”.

Alexander L. Perryman and Daniel N. Santiago have contributed equally to this work.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10822-014-9709-3) contains supplementary material, which is available to authorized users.

A. L. Perryman · D. N. Santiago · S. Forli · D. Santos-Martins ·  
A. J. Olson (✉)

Department of Integrative Structural and Computational  
Biology, The Scripps Research Institute, La Jolla, CA, USA  
e-mail: Olson@scripps.edu

## Present Address:

A. L. Perryman  
Department of Medicine, Rutgers Univ., NJ Medical School,  
Newark, NJ, USA

D. Santos-Martins  
REQUIMTE, Departamento de Química e Bioquímica,  
Faculdade de Ciências, Universidade do Porto, Porto, Portugal

**Keywords** Virtual screening · Ligand docking ·  
Computational pipeline · HIV integrase · Computational  
drug design · Docking challenge · AutoDock Vina

## Introduction

For an introduction to HIV infection, AIDS epidemics and integrase/LEDGF inhibition, we refer to the general introduction paper in this issue [1].

Our laboratory has been involved in HIV-related research for more than 20 years, with computational and

drug design efforts targeting HIV Protease (PR) and more recently HIV integrase (IN) [2–7]. This work has led to the identification of several potential new allosteric sites on HIV protease. Our main computational effort uses the FightAids@Home project, (<http://fightaidsathome.scripps.edu/>) in collaboration with IBM World Community Grid, where volunteers' computer power is used to perform high-throughput virtual screenings of millions of commercially available compounds versus HIV-related targets.

In 2012, we formed the HIV Interaction and Viral Evolution (HIVE) Center (<http://hive.scripps.edu/>) whose goal is to characterize at the atomic level the structural and dynamic relationships between interacting macromolecules in the HIV life cycle to understand the mechanistic evolution of drug resistance. The Center involves 13 individual research groups from six different institutions. One recent collaborative effort between the two HIVE Center computational groups, the Olson lab and the Levy lab, has been established to reduce the number of false positive results from large virtual screens. These virtual screens are used to recommend acquisition or synthesis of promising compounds for subsequent wet-lab analysis. Thus, the higher the rate of true positives from computational experiments, the less time, effort and money is wasted on testing compounds that are not true binders. To accomplish this, we are developing a procedure that takes the top hits selected from virtual screens using either AutoDock or AutoDock Vina (AD Vina) and evaluates their binding free energy using the replica exchange molecular dynamics computation, BEDAM, from the Levy Lab (described in detail in a companion paper in this issue [8]). Initial retrospective analysis of this procedure on HIV protease allosteric site-binders has shown to be very promising (paper in preparation). Fortunately, the SAMPL4 Challenge provided a useful blind data set with unpublished results upon which we could further test our methodology.

Moreover, the SAMPL4 Challenge organizers had chosen to use data from studies on three allosteric sites of the HIV IN [1]. While we had not previously worked on these HIV target sites, two other member labs in our HIVE Center, the Engelman and Kvaratskhelia groups, focus on allosteric inhibition of IN. The SAMPL4 Challenge presented an opportunity to initiate a computational effort in that area and promote further HIVE Center collaborations. Thus, we decided to participate in the SAMPL4 Challenge.

Before the Challenge began, the participants were informed that most of the SAMPL4 compounds were known to bind to the LEDGF site of IN, but some could bind to one or more additional allosteric sites of IN, which were referred to as the “FBP” site (for Fragment Binding Pocket) and the “Y3” site (see Fig. 1). Like the LEDGF site, the FBP site is also located at the dimer interface of the catalytic core domain (CCD) of IN. There are two LEDGF sites per IN CCD dimer, two FBP sites per IN

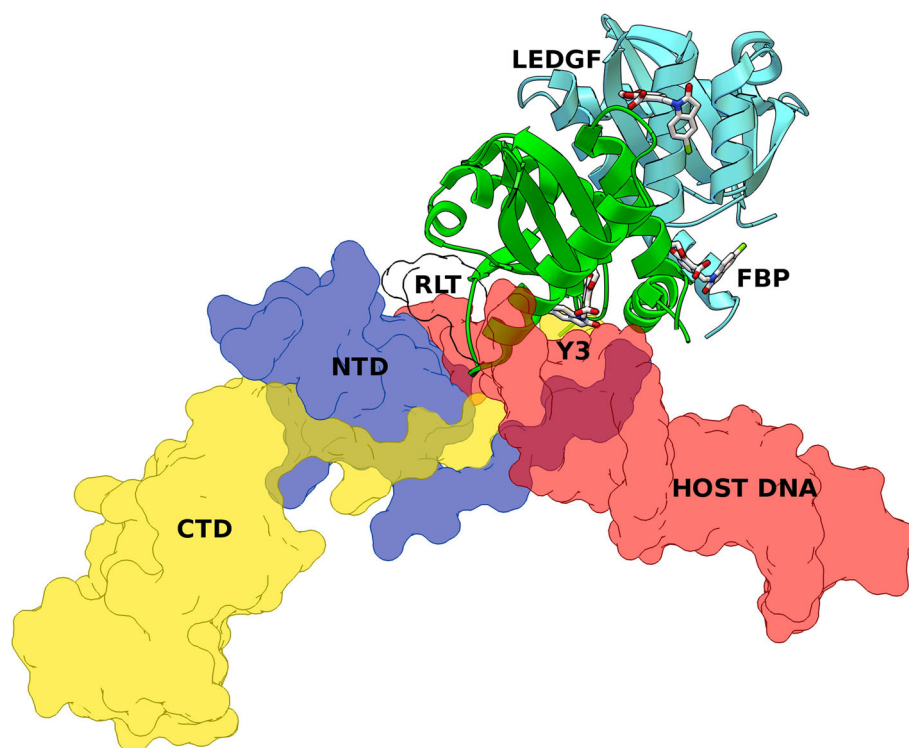
CCD dimer, and also two Y3 sites per IN CCD dimer. However, the Y3 site is entirely contained within each monomer of the core domain and is located underneath the very flexible 140s loop (i.e., Gly140-Gly149). The top of the 140s loop flanks the active site region, and the composition, conformation, and flexibility of the 140s loop is known to be critical to IN activity [9–11]. Since most of the inhibitors in the SAMPL4 library (i.e., the “true positives”) are LEDGF binders, and since the previously published inhibitors of the LEDGF site (e.g., the allosteric IN inhibitors or “ALLINIs”) are more advanced and more well-characterized [12–15], most of our effort focused on the LEDGF site. Most of this paper will focus on the results versus the LEDGF site, as well.

## Methods

### Positive control cross-docking studies

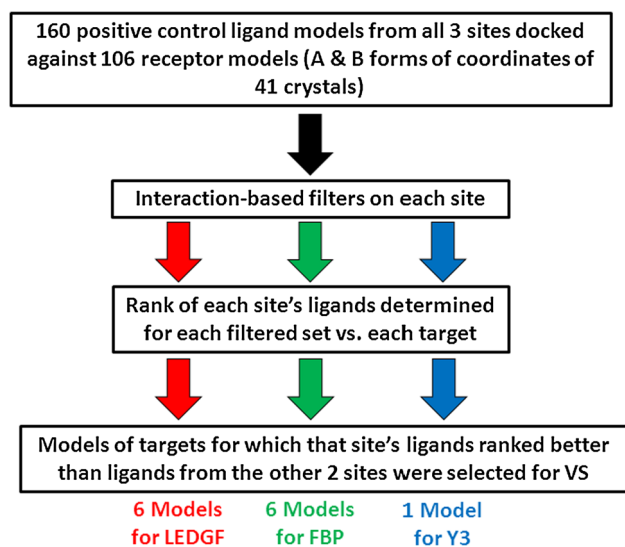
The challenge organizers gave as a reference structure an IN catalytic domain dimer with ligands bound to all three allosteric sites (PDB 3NF8). With myriad HIV IN structures available and 3 sites to consider, we inspected the interactions and rankings of known binders of the HIV IN allosteric sites to select an optimally informative subset of structures as targets for the virtual screening. Thus, co-crystallized ligands for each site were cross-docked from the collection of IN structures.

To prepare for the positive control cross-docking studies, the Protein Data Bank [16, 17] was searched for the available crystal structures of IN bound to an allosteric inhibitor. When a particular crystal structure displayed both “A” and “B” form coordinates for residues that were within one of the three allosteric sites of IN (or within the shell of residues that surround one of these sites), then that complex was split into two separate target files (i.e., *PDBID\_A* and *PDBID\_B*). If no “\_A” or “\_B” is listed, then that crystal structure had no alternate conformations in the regions surrounding the allosteric sites. The LEDGF site was represented by 64 different receptor models, the FBP was involved in 32 different receptor models, and the Y3 site was described by 10 different receptor models. These 106 receptor models of 41 crystallographic complexes were superimposed onto the coordinate reference frame provided by SAMPL4 (alignment by alpha carbons performed with PyMOL) and then organized according to which of the three allosteric sites had a ligand bound (using visual inspection). All hydrogen atoms were added to the proteins using the MolProbity server, which adjusts the pKa's of the titratable residues, optimizes the hydrogen bond network, and allows His, Gln, and Asn residues to flip if doing so lowers the energy of the system [18, 19]. All



**Fig. 1** Integrase functional structure and architecture. The three domain structure of an IN monomer bound to DNA is displayed. The SAMPL4 reference structure of the HIV IN Catalytic Core Domain dimer (CCD, PDB ID: 3NF8) was superimposed onto PFV IN crystal structure (PDB ID: 3OS1) to show the relative arrangement of the domain components. The CCD domain is represented as ribbon model with each monomer colored in green and cyan, respectively; the other domains are represented as semi-transparent surfaces: C-Terminal

Domain (CTD, yellow); N-Terminal Domain (NTD, light blue); host DNA (salmon). The CDQ allosteric ligand from 3NF8 is displayed as sticks (white) to highlight the three allosteric sites of HIV IN involved in the SAMPL4 challenge: LEDGF, Y3, and FBP. The IN inhibitor Raltegravir (RLT) bound in the PFV IN active site is shown with a black outline. The active catalytic state of HIV IN is a tetramer formed by a dimer of dimers (not shown)



**Fig. 2** Summary of the workflow used in the positive control cross-docking experiments with AutoDock Vina. This protocol was used to select the targets that were involved in the subsequent virtual screens with the SAMPL4 compounds

hydrogen atoms were added to the ligands using Avogadro [20]. Gasteiger-Marsili [21] charges were added to the models of both the ligands and the targets, and then the non-polar hydrogens were merged onto their respective heavy atoms using AutoDockTools and Raccoon [22, 23]. The docking studies were all performed using AD Vina 1.1.2 [24]. See Fig. 2 for a summary of the workflow employed in the positive control cross-docking experiments.

Only the small molecule allosteric inhibitors of IN were selected for the positive control cross-docking studies; the cyclic peptide inhibitor complexes being removed from the set. These positive control small molecules were utilized in the cross-docking experiments, with two ligand models for each ligand. One model started with a pose close to the crystallographic conformation and position, while the second model began with a randomized position, orientation, and conformation (generated by using the “randomize\_only” function in AD Vina) [24]. All 42 ligand models were known to be LEDGF binders, 33 ligand models were known to bind the FBP site, and 5 ligand models were

**Table 1** Common pharmacophore engine's training sets for each allosteric site

LEDGF	
Receptors (PDB ID)	Ligands (HET ID)
3ZCM	PX3
3ZSY	OM3
3ZSZ	OM2
3ZT1	OM1
3ZSQ	O4N
3ZSR	O3N
3ZSO	O2N
3ZSX	N44
3NF6	IMV
3NF8	CDQ
3NFA	CBJ
3VQ8	BCU
3VQ4	ONX
3ZT3	ZT4
3ZT4	ZT2
3ZT2	ZT2
3ZT0	ZT0
3ZSW	ZSW
3ZSV	ZSV
3VQ7	SNU
FBP	
Receptor (PDB ID)	Ligand (HET ID)
4AH9	OMB
3VQ4	ONX
3AO4	833
4AHS	AKH
3AO2	AVX
3AO5	BBY
3AO3	BMC
3AO3	BMC
3VQQ	BTE
3AO1 (top)	BZX
3AO1 (bottom)	BZX
3NF8	CDQ
3VQP	DBJ
3VQB	FBG
3VQE	FMQ
4AHR	I2E
4AHU	ICO
3VQ5	MMJ
3VQD	MOK
3VQC	MPK
3OVN	MPV
4AHT	Q6T
4AHV	Z5P

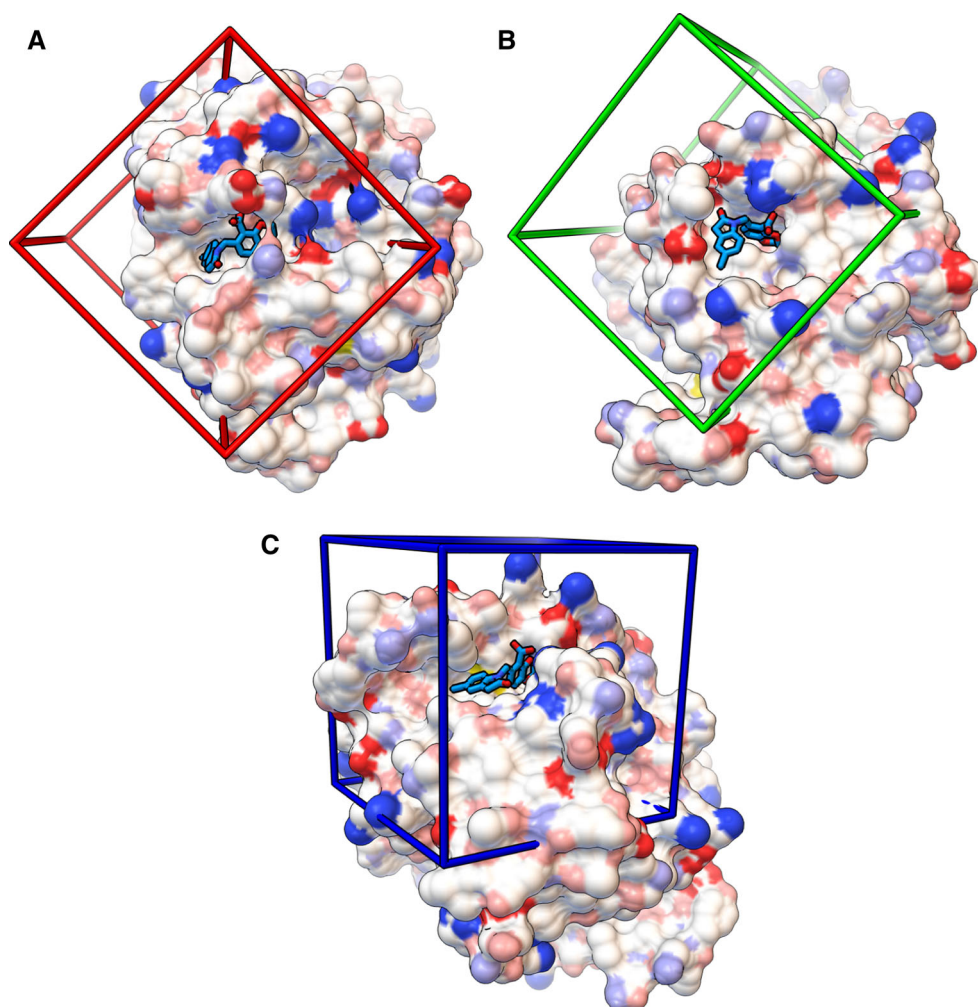
**Table 1** continued

FBP	
Receptor (PDB ID)	Ligand (HET ID)
3ZT2	ZT2
3ZT4	ZT2
Y3	
Receptor (PDB ID)	Ligand (HET ID)
3NF6	IMV
3NF7	CIW
3NF8	CDQ
3NF9	CD9
3NFA	CBJ

known to bind the Y3 site. These 80 ligand models and their 2 forms of randomized and non-randomized structures yielded 160 total ligand models for use in the positive control dockings. When a particular ligand crystallized in more than one type of allosteric site, it was included in the set of ligands for each of those sites. The aligned, crystallographic conformations of the positive control ligands from each site were used as inputs to train the common pharmacophore engine approach discussed below (see Table 1).

The following crystal structures had ligands that crystallized in both the LEDGF site and the FBP site: 3ZT2 (A and B forms), 3ZT4 (A and B forms), 3VQ4, and 3VQ7. Similarly, the following crystal structures had ligands that crystallized in both the LEDGF site and the Y3 site: 3NF6 (A and B forms) and 3NFA (A and B forms). For 3NF8 (which also had A and B forms), the fragment CDQ crystallized in all three allosteric sites. (See Fig. 3 for images that display the crystallographic binding mode of CDQ with each site and for a depiction of the grid boxes that were used in the AD Vina calculations against each site.) For both the positive controls and the virtual screens of the SAMPL4 compounds, 4 CPU's were used per docking calculation (on the TSRI Linux cluster), the grid box used for each site had a size of  $30 \times 30 \times 30 \text{ \AA}^3$ , and the grids were centered on an atom in CDQ from the 3NF8 reference provided by SAMPL4. For the FBP site, the nitrogen atom of CDQ was used for the center ( $x, y, z = 3.426, -21.239, -1.559$ ); for the LEDGF site, the C11 atom was used for the center ( $x, y, z = 11.095, -46.191, 0.368$ ); for the Y3 site, the N atom of CDQ was used for the center ( $x, y, z = 9.137, -25.675, -22.414$ ). Since large grid boxes were used, the "exhaustiveness" setting in AD Vina was increased to 20.

**Fig. 3** Grid boxes ( $30 \times 30 \times 30 \text{ \AA}^3$ ) utilized in the positive control cross-docking studies and in the virtual screens of the SAMPL4 compounds. Each image contains the solvent-excluded molecular surface of the HIV integrase CCD (colored by David Goodsell convention in AutoDock Tools [22]) from the crystal structure PDB ID 3NF8, in which the fragment CDQ (shown as *sticks* with turquoise carbon atoms) is bound. Shown are the LEDGF (A), FBP (B), and Y3 (C) sites

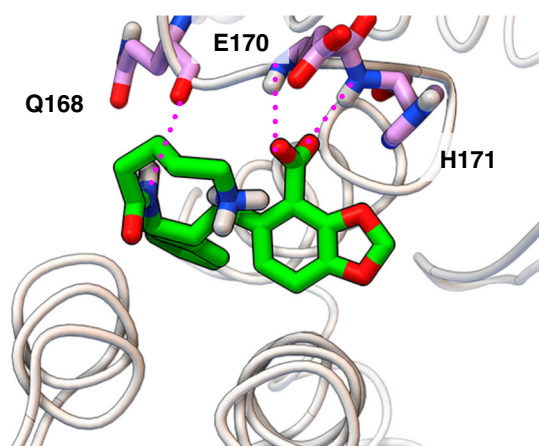


#### Selecting the models for each of the targets

The combination of the three sets of positive control compounds were cross-docked against all of the models for each of the 3 allosteric sites (see Fig. 2 for a summary of the workflow). Interaction-based filters, based on key hydrogen bonds displayed by many of the ligands in the crystal structures for each particular site, were applied to the AD Vina results against each allosteric site (see Fig. 4). The filtering was done using in-house Python scripts that were created during the development of Raccon2 and Fox [23]. The filters were only applied to the top-ranked AD Vina mode per model of each compound (for both the positive control experiments and the subsequent virtual screens of the SAMPL4 compounds). For example, about a dozen different sets of interaction-based filters were investigated for the LEDGF results, and two particular filters were selected, since they harvested a reasonable number of docked modes per target for visual inspection. For the LEDGF site, the filter requirements consisted of a minimum of 2 predicted hydrogen bonds to IN, and either:

(Fa) a hydrogen bond with Glu170; or (Fb) a hydrogen bond to the backbone amino group of His171, similar to the ALLINIs (see Fig. 5).

The method used for choosing receptors for virtual screening of the IN models for the LEDGF site described here is the same for the FBP site. For each LEDGF target, the top-ranked docked mode of the positive control ligands from all 3 sites that passed a particular filter were sorted according to the estimated free energy of binding as calculated by AD Vina. This sorting process determined the *absolute rank* for each compound whether they be known LEDGF-site binders or decoys (observed to bind at the FBP or Y3 sites). The ligands that crystallized in the LEDGF site were then extracted from that sorted list, and their order in that LEDGF-site specific list determined their *relative rank*. Receptor models were chosen from statistics and visual analysis of the rankings of the site-specific binders (Supplemental Information, Fig. 1). Subsequently, this ad hoc visualization process formalized into our “Rank Difference Ratio” (RDR) procedure where the relative ranking of the appropriate positive control ligands was



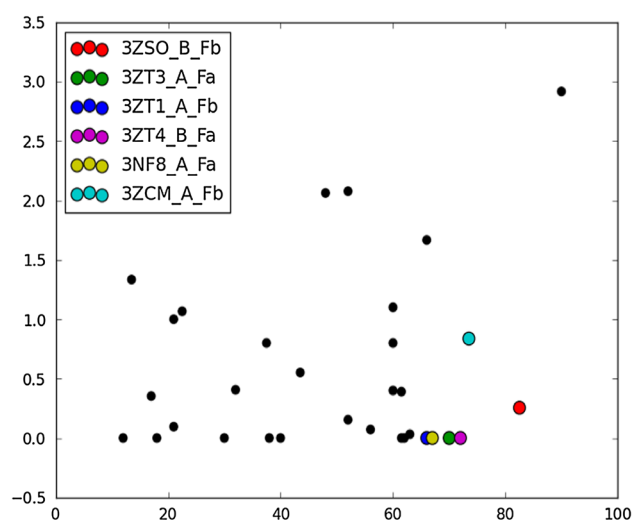
**Fig. 4** Hydrogen bond interactions of AVX17561 (sticks with green carbon atoms) docked with the LEDGF site (white backbone tube) of HIV integrase. Three residues (sticks with pink carbon atoms) are shown with hydrogen bonds (magenta dotted lines) to the ligand model. Two hydrogen bonds (Glu170 and His171) were required by the interaction filters

used with the corresponding absolute ranking as the RDR metric for a receptor  $i$ , as follows:

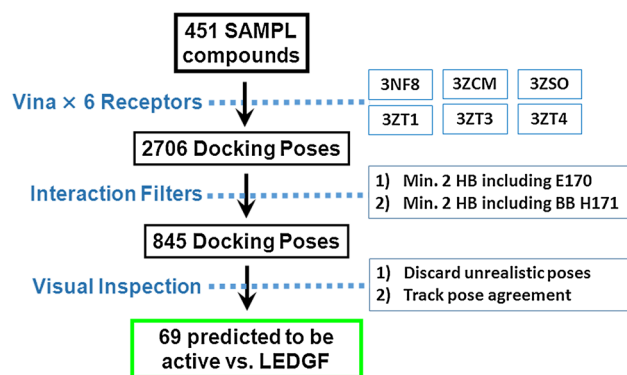
$$RDR_i = \frac{\sum_j R_{abs,j} - R_{rel,j}}{N_i}$$

where  $R_{abs,j}$  and  $R_{rel,j}$  are the absolute and relative rankings, respectively, for ligand  $j$ . For example, a LEDGF target model for which all of the LEDGF ligands ranked better than the ligands from the other two sites, the absolute and relative rankings for all ligands would be the same and the receptor would thus have an RDR value of 0. In Fig. 5 we plot RDR versus total number of site-specific hits. The receptor models having a low RDR value and a high number of LEDGF-site hits demonstrates that this formalized RDR metric matches the receptors chosen by the original ad hoc procedure. For targets that displayed similar RDR values using a particular filter, receptor models were selected to maximize structural diversity. Further, if two forms of a structure had a difference of total number of LEDGF-site hits less than 2, the receptor model with the lower RDR value was chosen. A similar strategy was used to select the FBP targets. For the small set of Y3 receptor models, the median statistic of rankings was sufficient to choose the best receptor model, 3NF8\_B.

The set of 106 receptor models representing the FBP, LEDGF, and Y3 sites enabled the selection of the following number of targets per site: 6 crystal structures of the LEDGF site, 6 structures of the FBP site, and 1 crystal structure of the Y3 site. The LEDGF targets selected were: 3ZSO\_B, 3ZT4\_B, 3ZT1\_A, 3ZT3\_A, 3NF8\_A, and 3ZCM\_A [25, 26]. The FBP targets selected were: 3AO1, 3AO2, 3VQD, 3VQE\_A, 3VQ4, and 3VQ7 [27, 28]. The Y3 target selected was 3NF8\_B [26].



**Fig. 5** Selection of the 6 crystal structures that were used as targets for the LEDGF site during the virtual screen of the SAMPL4 compounds. The X-axis indicates the number of docked models of LEDGF ligands that passed through a particular filter during the positive control cross-docking experiments, while the Y-axis plots the “Rank Difference Ratio” metric that quantifies how well the LEDGF ligands ranked, in relation to the FBP and Y3 ligands. All 84 models of small molecule LEDGF ligands were docked to 66 LEDGF-site receptor models and then 2 interaction filters (Fa: a hydrogen bond with Glu170; Fb: a hydrogen bond to the backbone amino group of His171) were applied. After normalization of data between the two sets of filtered data, the best for each receptor source is visualized here



**Fig. 6** Summary of the workflow used in the in the SAMPL4 challenge for the LEDGF site

Virtual screen of the SAMPL4 compounds using AD Vina

AutoDock Vina was used to screen the SAMPL4 compound library against these 13 receptor models of IN [24]. See Fig. 6 for a summary of the workflow used for the LEDGF targets. A similar strategy was utilized for the FBP and Y3 sites. The same grid box size, location, and settings for AD Vina from the positive controls were also used in these virtual screens (see Fig. 3). The 321 compounds

provided by SAMPL4 were used as inputs by the Levy Lab for LigPrep and Epik [29, 30] at a pH of  $7 \pm 2$ , which generated additional tautomers and protonation states of some compounds to produce a final set of 451 models of the SAMPL4 compounds. All 451 ligand models of the compounds were docked against the 6 LEDGF targets, 6 FBP targets, and 1 Y3 target, using the TSRI Linux cluster. The same filters used in the positive control dockings were applied to the results of the SAMPL4 dockings.

The sets of docked modes harvested by either of these two filters per target were combined and visually inspected. The hydrogen bond Python filter used only the donor–acceptor distance as a criterion (the donor–hydrogen–acceptor angles were manually measured during the visual inspection process). Positive factors considered during visual inspection included the hydrogen bonds exceeding the number specified in the filters, potentially favorable interactions of nearby side-chains if they flexed, and agreement of position between a set of stereoisomers. Negative factors included solvent-exposure of non-polar atoms or large portions of the ligand. Though somewhat subjective in nature, these criteria were used to make final choices of hits and later determine a confidence level for each ligand as required by SAMPL. The range of confidence was 1–5 with 5 being the highest confidence. The docked modes that passed the visual inspection process against any of the LEDGF targets were combined, to generate a set of 69 unique compounds that were predicted to be LEDGF binders.

For FBP, two filters were also combined to harvest ligands for visual inspection. In one filter, the top-ranked docked mode needed to have a minimum of two hydrogen bonds and a ligand efficiency value less than  $-0.30$  kcal/mol/heavy atom. In the other filter, the compound had to display a minimum of three hydrogen bonds and a ligand efficiency value better than  $-0.25$  kcal/mol/heavy atom. The interaction-based filters from the positive control dockings that used key hydrogen bonds to the FBP site were too stringent; very few or no docked compounds passed that filter. Twenty-five compounds passed the visual inspection process and were predicted to bind the FBP site.

For the Y3 receptor model, two different filters were combined to choose docking modes for visual inspection. One filter required the docked compounds to have one hydrogen bond with Lys188 and one additional unspecified hydrogen bond. The second filter only required the docked compounds to possess 4 hydrogen bonds. A set of seven compounds passed the visual inspection process and were predicted to bind to the Y3 site of IN.

The predicted binding modes of all of our candidate inhibitors that passed the visual inspection process against the 13 target models were sent to the Levy lab, to be used as inputs for their BEDAM replica exchange simulations

[31]. (See the companion paper on the BEDAM method for results and a discussion of the Consensus approach [8].) In addition, the filtered binding modes predicted by AD Vina were also re-ranked by the new common pharmacophore engine method (below), which identified an alternate set of candidate inhibitors.

#### Common pharmacophore engine

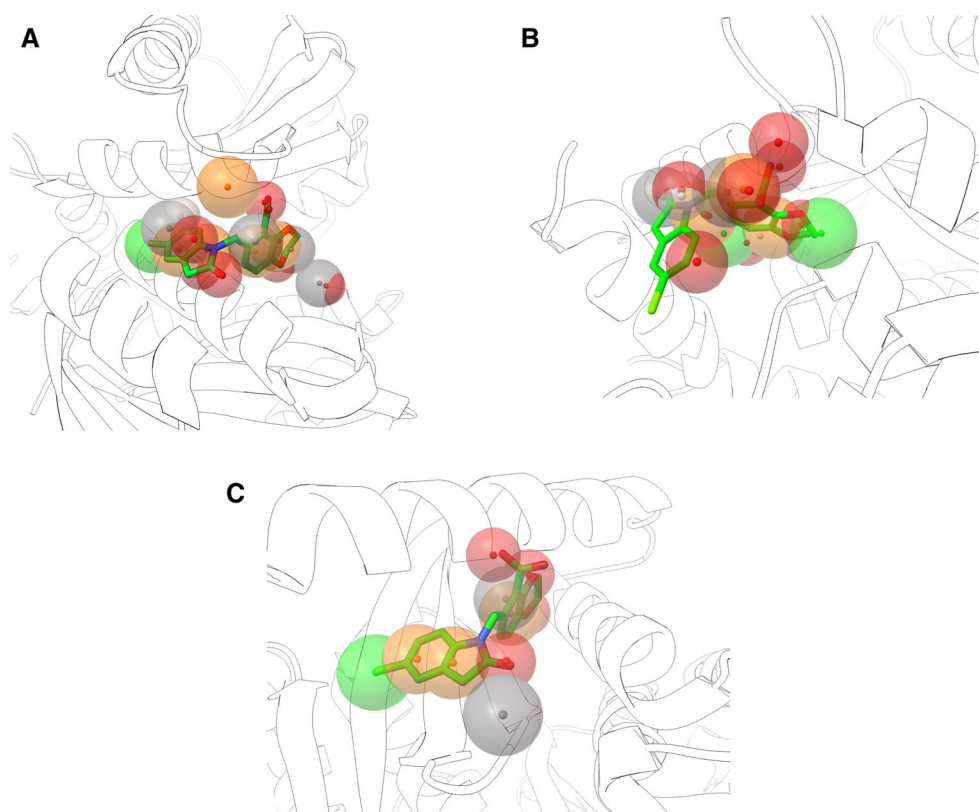
We have developed a 3D pharmacophore model [32, 33] based on the AutoDock forcefield/atom type set. The pharmacophore is based on the conversion of explicit chemical groups into basic chemical features: hydrogen bond acceptors and donor, aromatic rings, aliphatic carbons, and halogens. Each feature is represented by a combination of the three-dimensional location of a given feature with respect to the ligand structure and a series of properties specific to each feature (*i.e.*, hydrogen bond direction vector, aromatic ring plane). Each feature can also include a tolerance setting for its properties and location. This simplifies the comparison of chemical structures and provides a rapid method for quantitatively scoring structural similarity. By using pharmacophore representations it is possible to generate a common pharmacophore model that recapitulates the most representative set of chemical features of a series of ligands bound to the same site. The choice of features to represent the pharmacophore is performed after geometrical clustering. A common pharmacophore engine generates similar feature clusters that are processed to produce an average representation. Isolated features are discarded. In addition to the aforementioned chemical features, the engine also generates special sets (*i.e.*, an “honorable mention” set) representing recurrent atom type clusters present in the ligand set, like halogens or sulfur.

The common pharmacophore can then be used to re-score ligands docked to the binding site, providing a new quantitative score based on the similarity of binding patterns between docked results and known ligands. This pharmacophore model has been successfully applied to identify high micromolar allosteric inhibitors of HIV-1 protease, including several crystallographic hits (unpublished data). Further details about the method and the pharmacophore engine will be described in a future publication.

#### Generation of common pharmacophores

In order to generate the common pharmacophores for the three different sites (LEDGF, FBP, and Y3), the three different training sets of ligand–receptor complexes were aligned (by alpha carbons in PyMOL) to the 3NF8

**Fig. 7** Common pharmacophore engine representation for each of the three allosteric sites of HIV integrase. The SAMPL4 reference structure 3NF8 is represented as white cartoons with the crystallographic ligand CDQ represented as. Pharmacophore features (see Table 1 for full listing) are shown as semi-transparent spheres, with solid sphere centroids: aromatic features (*orange*); aliphatic carbon (*gray*); hydrogen bond (HB) acceptor (*red*), HB donor (*white*); and halogen (*green*). Shown are the pharmacophore models of the LEDGF (a), FBP (b), and Y3 (c) sites



structure provided as an input for the challenge. For each site, ligand structures were extracted and their overlapped conformations were processed with the common pharmacophore engine. Features present in at least two ligands were considered, directionality was disabled for the hydrogen bond and aromatic ring features, no maximum features count was used, and all features had the same weight and radius tolerance (1.5 Å). These settings were used to increase the tolerance of the common pharmacophores and to provide a generic filter to prioritize docked poses of ligands that displayed an interaction pattern similar to known ligands. A summary of the different training sets used for the generation of each common pharmacophore is shown in Table 1. The number and nature of the features generated for each pharmacophore is strictly correlated with the number of ligands available for each site and their structural diversity.

#### LEDGF common pharmacophore

The LEDGF common pharmacophore was calculated from 19 different ligands (see Table 1) and contains the following features: 5 aromatic rings, 2 hydrogen bond donors, 6 hydrogen bond acceptors, 5 aliphatic carbons, and 1 halogen (see Fig. 7a).

#### FBP common pharmacophore

The FBP common pharmacophore was calculated from 22 different ligands (see Table 1) and contains the following features: three aromatic rings, three hydrogen bond donors, nine hydrogen bond acceptors, three aliphatic carbons, and two halogens (see Fig. 7b).

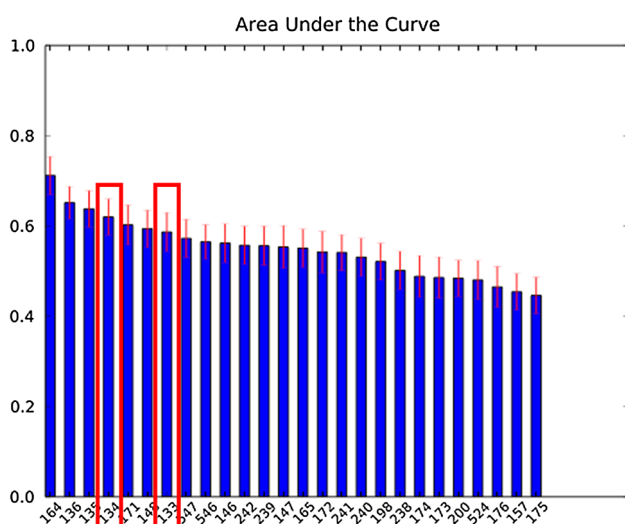
#### Y3 common pharmacophore

The Y3 common pharmacophore was calculated from 5 different ligands (see Table 1) and contains the following features: three aromatic rings, four hydrogen bond acceptors, two aliphatic carbons, and one halogen (see Fig. 7c).

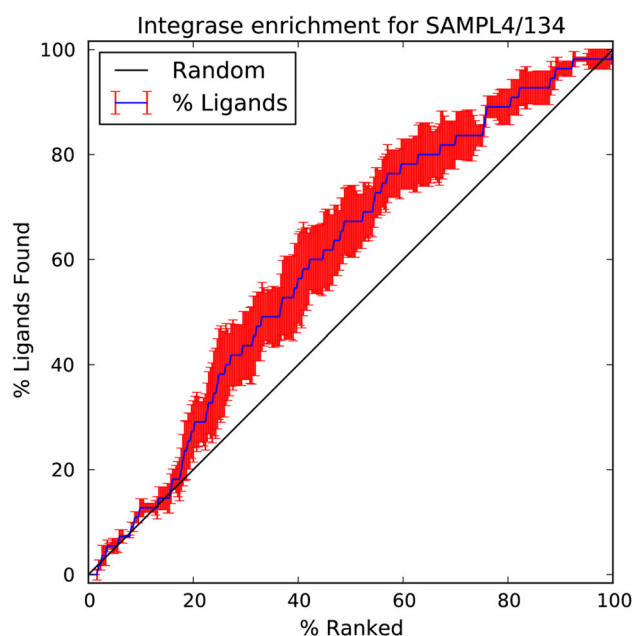
### Results and discussion

Since this was a blind challenge, the official classification of the success of these methods was determined by the SAMPL4 organizers, using metrics that they selected and applied. (See [1] for details used to calculate these metrics.) The results presented here were plotted by the SAMPL4 organizers, and the graphs in Figs. 8, 9, 10, 11, 12, 13, 14 were generated by them. For these graphs in Figs. 8, 9, 10,

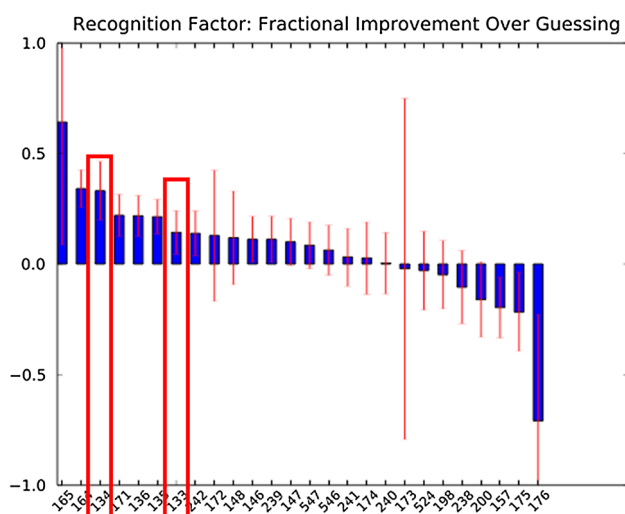




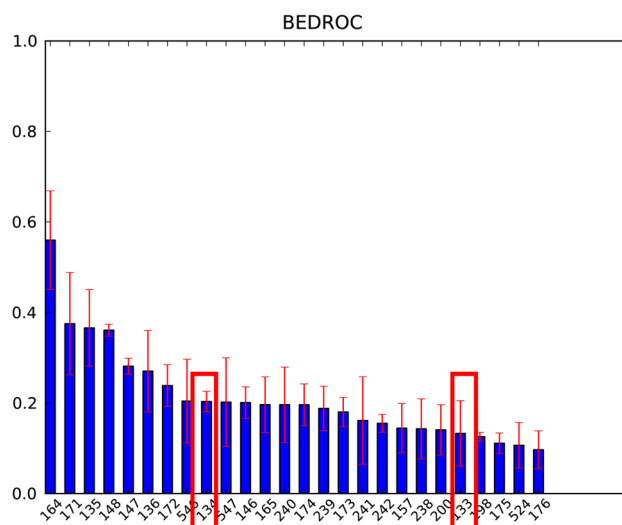
**Fig. 8** Area under the curve (AUC) performance of the Phase 1 entries in the SAMPL4 challenge provided by the SAMPL organizers. *Red boxes* highlight the entries for the common pharmacophore engine (134) and AutoDock Vina plus visual inspection (133)



**Fig. 10** Enrichment factor of SAMPL4 entry 134 corresponding to the common pharmacophore engine's score



**Fig. 9** Recognition Factor performance of the Phase 1 entries in the SAMPL4 challenge provided by the SAMPL organizers. *Red boxes* highlight the entries for the common pharmacophore engine (134) and AutoDock Vina plus visual inspection (133)



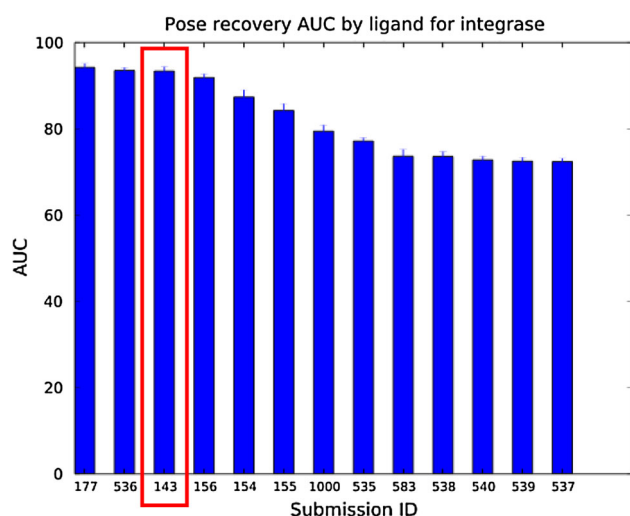
**Fig. 11** BEDROC score of the Phase 1 entries in the SAMPL4 challenge provided by the SAMPL organizers. *Red boxes* highlight the entries for the common pharmacophore engine (134) and AutoDock Vina plus visual inspection (133)

11, 12, 13, 14, our different predictions had the following submission ID numbers: (A) AD Vina plus visual inspection, entry 133; (B) common pharmacophore engine (without visual inspection), entry 134; (C) BEDAM replica exchange binding free energy simulations (of the AD Vina docked modes that passed the visual inspection process in A), entry 135; and (D) Consensus approach that combined A–C, entry 136. See the companion paper on the BEDAM

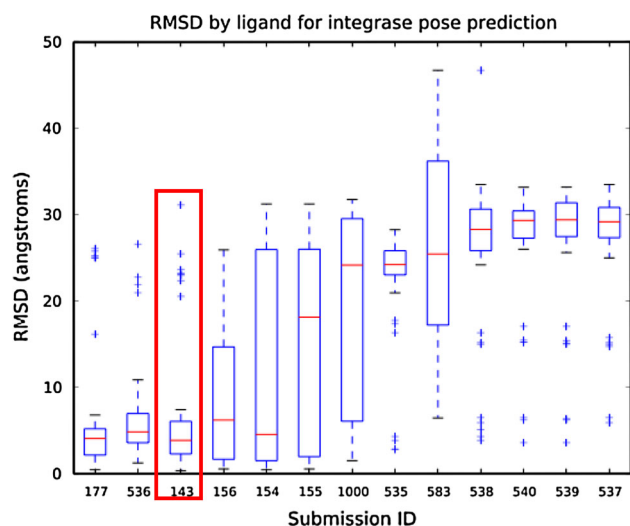
entry by the Levy Lab [8] to learn the details regarding the methods and results for C and D.

Virtual screens using AD Vina with visual inspection (Phase 1 of Challenge): identifying IN binders

Phase 1 involved predicting which of the SAMPL4 compounds actually bind to any of the three allosteric sites of

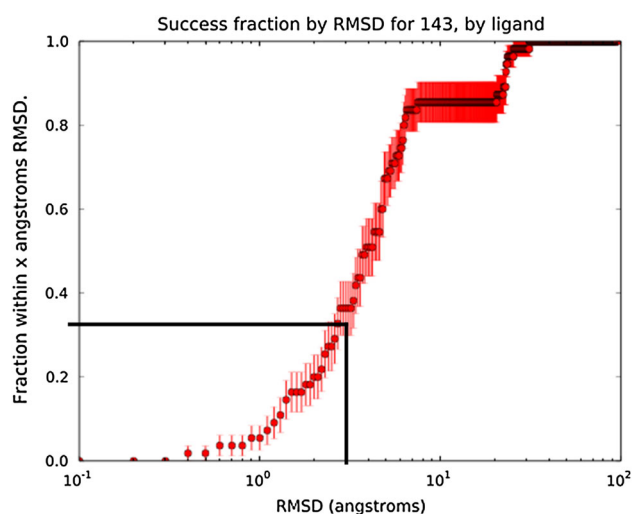


**Fig. 12** Pose Recovery Area Under the Curve performance of the Phase 2 entries in the SAMPL4 challenge provided by the SAMPL organizers. The AutoDock Vina plus visual inspection process (143, red box) was our only entry for Phase 2



**Fig. 13** RMSD per pose performance of the Phase 2 entries in the SAMPL4 challenge provided by the SAMPL organizers. The AutoDock Vina plus visual inspection process (143, red box) was our only entry for Phase 2

HIV IN. The common pharmacophore engine's predictions (without human intervention or visual inspection) are listed as entry number 134 and ranked 4th place, while the predictions from AD Vina with visual inspection are listed as entry number 133 and ranked 7th place according to the area under the curve (AUC) metric. Although the common pharmacophore engine's predictions ranked better than the results from AD Vina with visual inspection, their respective AUC values were within the standard deviations



**Fig. 14** Fraction of ligand poses that were successfully predicted by AutoDock Vina. RMSD data (in Å) provided by the SAMPL organizers

of each other. However, using the Recognition Factor metric, the performance of the common pharmacophore engine (which ranked 3rd place) was clearly superior to the predictions from AD Vina with visual inspection (which ranked 7th place) when considering the standard deviations.

For the general metric used to judge most prospective virtual screens in the published literature, the common pharmacophore engine had a hit rate of 24.85 % (42/169), and AD Vina with visual inspection had a similar hit rate of 23.76 % (24/101). Given the intrinsic approximations of empirical scoring functions, docking methods are usually less sensitive to congeneric compounds, while they are more effective in identifying potential hits in chemically diverse libraries [34]. Therefore, in the context of the low diversity library provided in this challenge, the overall success rate of these docking-based methods is high.

Pharmacophore results (Phase 1 of challenge):  
identifying IN binders

Every set of docking results generated with AD Vina for each binding site in the 13 targets were re-scored with the corresponding common pharmacophore. The pharmacophore score ( $ps$ ) ranged from 100.0 (all features matched) to 0.0 (no feature matched). For the LEDGF and FBP sites, where six different receptor models were targeted for each site, the pharmacophore score,  $ps$ , was calculated separately for each set. The 6 sets (per site) were then combined by weighting the score of each ligand versus a given receptor pose by its positional ranking in the set ( $ps_{tot} = \sum ps_i / \text{rank}_i$ ). The pharmacophore score for each

binding site was then used to rank the ligands and identify the binders. Depending on their ps ranking, the ligands were subdivided into three classes, with high (top 10), medium (10–50) and low (> 50) confidence, respectively. Accordingly to the challenge guidelines, ligands in the high confidence class were considered binders with a confidence value of 5, ligands in the medium confidence class were considered as binders with a confidence value of 3, and ligands in the low confidence class were considered as non-binders with a confidence value of 1. The overall success rate in hit identification for binders versus non-binders using the pharmacophore score was 25 %. Results for the common pharmacophore engine approach (*i.e.*, entry 134), using different metrics calculated by the SAMPL4 organizers, are presented: AUC (see Fig. 8), Recognition Factor (see Fig. 9), enrichment factor (see Fig. 10), and BEDROC (see Fig. 11).

Virtual screens with AD Vina plus visual inspection (Phase 2 of Challenge): identifying the binding site and binding pose within the three allosteric sites of IN

Phase 2 of the SAMPL4 challenge involved predicting the binding site (or sites) and binding mode (or modes) of the 57 SAMPL4 compounds that crystallized in an allosteric site of HIV IN. Since this was a blind contest, our team submitted all of our predictions for Phase 1 before we obtained the list of compounds that were involved in Phase 2.

AutoDock Vina with visual inspection was the only entry that we submitted for Phase 2, and it is listed as entry number 143. Generally, AD Vina placed between 2nd and 5th places over six different metrics (see Fig. 4 in [1]). According to the “Pose recovery AUC by ligand” metric (see Fig. 12), AD Vina ranked 3rd place. But considering the standard deviations, it was nearly identical to the 2nd place entry and similar to the 1st place entry. Using the “RMSD by ligand” box plot data (see Fig. 13), AD Vina again ranked 3rd place. However, when comparing the medians, AD Vina was actually 1st place (lowest median). Similarly, if only the data distributions within the 1st and 3rd quartiles are compared and outliers disregarded, AD Vina ranked 2nd place (with 1st and 3rd quartile values lower than those of 2nd place Entry 536). Thus, the binding modes predicted by AD Vina with visual inspection against these three allosteric sites of IN were relatively accurate, which contributed to the success of the BEDAM replica exchange re-scoring entry for Phase 1 [8].

Using the “Success fraction by RMSD, by ligand” plot (see Fig. 14), approximately 33 % of the ligands were docked within an RMSD of 3 Å of the crystallographic binding mode(s). For a visual comparison between the docked modes of true positives identified in Phase 1 against

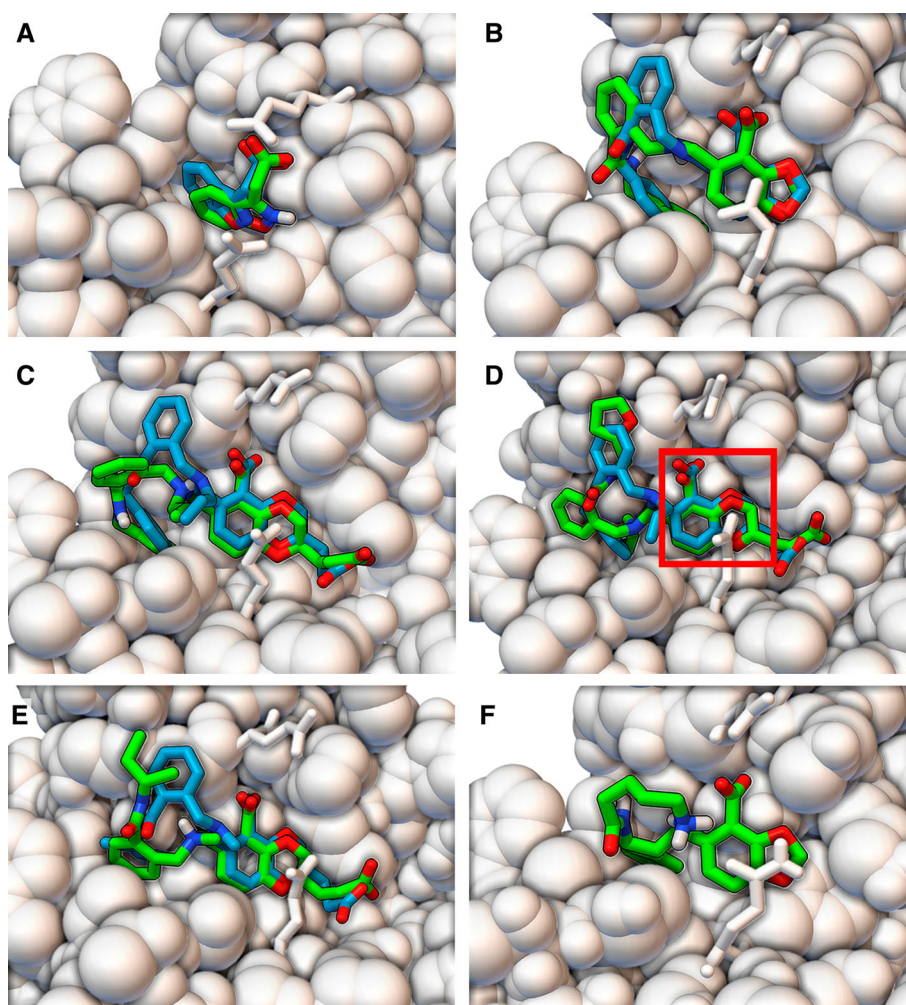
the LEDGF site and their crystallographic binding modes (provided after we submitted our entry for Phase 2), see Fig. 15. In these molecular images (created with PMV 1.5.6 release candidate 3) [22, 35], some of the best predictions are displayed as well as some representative results. In all of these molecular images in Fig. 15, the fragment hit or the scaffold region (within the larger derivatives based on extending the fragment hits) docked accurately with AD Vina and displayed some of the same interactions that the published ALLINIs display. The fragment AVX17679 docked within 1.3 Å of its crystallographic pose, while the larger compounds AVX17684 m and AVX38753 docked within 1.7 and 2.1 Å of their crystallographic binding modes, respectively. Even when AVX38783 and AVX38784 docked with larger RMSD values of 4.5 Å and 3.4 Å, respectively, (see Fig. 15d, e), their scaffolds superimposed well (RMSD of 1.1 and 1.4 Å) with the crystallographic pose. AVX17561 is an interesting case where the docked scaffold matches crystal data (RMSD is 1.2 Å without the amino alkyl group, co-crystallized ligand structure not shown), but the hydrophobic tail is solvated. The binding modes predicted by AD Vina earned 3rd place in the pose prediction with a relative performance that was very close to the 1st and 2nd place entries, using the metrics provided by SAMPL4.

## Conclusion

The positive control cross-docking experiments performed with AutoDock Vina indicated that at least 13 different crystal structures of HIV IN (6 LEDGF models, 6 FBP models, and 1 Y3 model) displayed reasonable predictive power for identifying the appropriate ligands that are known to bind each of the three sites. Our new RDR metric appears helpful when selecting targets out of a large ensemble of different receptor models and should be further investigated as an alternative and/or complementary way of selecting snapshots of targets for Relaxed Complex Scheme experiments [10, 36–39] and virtual screens.

The binding modes calculated by AD Vina (when docking the 451 models of the 321 SAMPL4 compounds against these 13 targets) were accurate enough to enable the following: (1) achieving a hit rate of 24 % using visual inspection of the poses predicted by AD Vina; (2) obtaining a hit rate of 24 % by re-ranking the docked poses using our new common pharmacophore engine (without any visual inspection); and (3) by using the docked poses that passed the visual inspection process as inputs for BEDAM replica exchange re-scoring calculations, the hit rate was further improved to 34 %.

**Fig. 15** Comparison of the AutoDock Vina docking poses to the crystallographic binding modes of ligands that bound to the LEDGF site of HIV integrase. The crystallographic binding modes of the ligands are displayed as sticks with turquoise carbon atoms, while the binding modes predicted by AutoDock Vina are rendered with green carbon atoms. HIV integrase is shown as white CPK spheres, with residues Glu170 and Gln95 displayed as white sticks for clarity. AVX17679 (a), AVX17684 m (b), AVX38753 (c), AVX38783 (d), and AVX38784 (e) is displayed with RMSD values compared to the crystallographic poses of 1.3, 1.7, 2.1, 4.5, and 3.4 Å, respectively. RMSD values for only the scaffold (red box in d) were 1.4 Å for AVX38784 and 1.1 Å for AVX38784. The docking pose of AVX17561 (f) is also shown (without the crystallographic structure) having an RMSD value of 2.3 Å; however, the RMSD value is 0.8 Å if the amino alkyl tail and amide group are disregarded



These results have helped to validate the efficacy of the virtual screening pipeline that we are developing. In particular we have established the following conclusions: (1) The common pharmacophore engine approach was more efficient than the visual inspection process in terms of the amount of human time involved and definitely merits further exploration and development. (2) The binding modes predicted by AD Vina were of sufficient accuracy to serve effectively as input to the free energy calculation of BEDAM. (3) The BEDAM post-processing of virtual screening results provided a significant improvement in false positive reduction.

Considering that SAMPL4 involved (A) three different allosteric sites of a flexible enzyme and (B) a very challenging library of compounds that displayed a low amount of structural diversity and that contained ligands that could bind to more than one type of allosteric site, the hit rates that our collaborative, multi-institution HIVE team achieved were impressive.

We have applied our new knowledge of structural detail about these three allosteric sites of HIV IN motivating us to test and hone our tools and protocols. These aspects will help

advance the goals we are pursuing as part of the new HIVE center, which is devoted to understanding and defeating the multidrug-resistant strains of HIV that are constantly evolving and spreading. In addition, the 13 models of these allosteric sites of HIV IN that we identified in our positive control cross-docking experiments are currently being used as targets for the FightAIDS@Home project.

**Acknowledgments** We thank the I.T. staff at The Scripps Research Institute (especially Jean-Christophe Ducom, and Lisa Dong) for maintaining a great Linux cluster and for giving the Levy lab access to it for their BEDAM calculations. This research was funded by the HIVE center Grant (P50 GM103368) and by the AutoDock grant (R01 GM069832).

## References

- Mobley DL, Liu S, Lim NM, Wymer KL, Perryman AL, Forli S, Deng N, Su J, Branson K, Olson AJ (2014) J Comput Aided Mol Des. doi:10.1007/s10822-014-9723-5
- Tiefenbrunn T, Forli S, Happer M, Gonzalez A, Tsai Y, Soltis M, Elder JH, Olson AJ, Stout CD (2013) Chem Biol Drug Des 83(2):141

3. Tiefenbrunn T, Forli S, Baksh MM, Chang MW, Happer M, Lin YC, Perryman AL, Rhee JK, Torbett BE, Olson AJ, Elder JH, Finn MG, Stout CD (2013) *ACS Chem Biol* 8(6):1223
4. Lin YC, Perryman AL, Olson AJ, Torbett BE, Elder JH, Stout CD (2011) *Acta Crystallogr D Biol Crystallogr* 67(Pt 6):540
5. Perryman AL, Zhang Q, Soutter HH, Rosenfeld R, McRee DE, Olson AJ, Elder JE, Stout CD (2010) *Chem Biol Drug Des* 75(3):257
6. Vajragupta O, Boonchoong P, Morris GM, Olson AJ (2005) *Bioorg Med Chem Lett* 15(14):3364
7. Perryman AL, Forli S, Morris GM, Burt C, Cheng Y, Palmer MJ, Whitby K, McCammon JA, Phillips C, Olson AJ (2010) *J Mol Biol* 397(2):600
8. Gallicchio E, Deng N, He P, Perryman AL, Santiago DN, Forli S, Olson AJ, Levy R (2014) *J Comput Aided Mol Des* 28(1)
9. Greenwald J, Le V, Butler SL, Bushman FD, Choe S (1999) *Biochemistry* 38(28):8892
10. Perryman AL, Forli S, Morris GM, Burt C, Cheng Y, Palmer MJ, Whitby K, McCammon JA, Phillips C, Olson AJ (2010) *J Mol Biol* 397(2):600
11. Dewdney TG, Wang Y, Kovari IA, Reiter SJ, Kovari LC (2013) *J Struct Biol* 184(2):245
12. Kessl JJ, Jena N, Koh Y, Taskent-Sezgin H, Slaughter A, Feng L, de Silva S, Wu L, Le Grice SFJ, Engelman A, Fuchs JR, Kvaratskhelia M (2012) *J Biol Chem* 287(20):16801
13. Tsiang M, Jones GS, Niedziela-Majka A, Kan E, Lansdon EB, Huang W, Hung M, Samuel D, Novikov N, Xu Y, Mitchell M, Guo H, Babaoglu K, Liu X, Geleziunas R, Sakowicz R (2012) *J Biol Chem* 287(25):21189
14. Christ F, Shaw S, Demeulemeester J, Desimmie BA, Marchand A, Butler S, Smets W, Chaltin P, Westby M, Debyser Z, Pickford C (2012) *Antimicrob Agents Chemother* 56(8):4365
15. Jurado KA, Wang H, Slaughter A, Feng L, Kessl JJ, Koh Y, Wang W, Ballandras-Colas A, Patel PA, Fuchs JR, Kvaratskhelia M, Engelman A (2013) *Proc Natl Acad Sci* 110(21):8690
16. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) *Nucleic Acids Res* 28(1):235
17. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C (2002) *Acta Crystallogr D* 58(6 Part 1):899
18. Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) *Acta Crystallogr D* 66(1):12
19. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC (2007) *Nucleic Acids Res* 35(suppl 2):W375
20. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR (2012) *J Cheminform* 4(1):17
21. Gasteiger J, Marsili M (1980) *Tetrahedron* 36(22):3219
22. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) *J Comput Chem* 30(16):2785
23. Forli S. Raccoon (2010) Molecular Graphics Laboratory, The Scripps Research Institute, La Jolla. <http://autodock.scripps.edu/resources/raccoon>. Accessed 2013
24. Trott O, Olson AJ (2010) *J Comput Chem* 31(2):455
25. Peat TS, Rhodes DI, Vandegraaff N, Le G, Smith JA, Clark LJ, Jones ED, Coates JAV, Thienthong N, Newman J, Dolezal O, Mulder R, Ryan JH, Savage GP, Francis CL, Deadman JJ (2012) *PLoS One* 7(7):e40147
26. Rhodes DI, Peat TS, Vandegraaff N, Jeevarajah D, Le G, Jones ED, Smith JA, Coates JA, Winfield LJ, Thienthong N, Newman J, Lucent D, Ryan JH, Savage GP, Francis CL, Deadman JJ (2011) *Antivir Chem Chemother* 21(4):155
27. Wielens J, Headey SJ, Deadman JJ, Rhodes DI, Le GT, Parker MW, Chalmers DK, Scanlon MJ (2011) *Chem Med Chem* 6(2):258
28. Wielens J, Headey SJ, Rhodes DI, Mulder RJ, Dolezal O, Deadman JJ, Newman J, Chalmers DK, Parker MW, Peat TS, Scanlon MJ (2013) *J Biomol Screen* 18(2):147
29. Greenwood JR, Calkins D, Sullivan AP, Shelley JC (2010) *J Comput Aided Mol Des* 24(6–7):591
30. LigPrep. 2.6. (2013) Schrödinger LLC, New York
31. Gallicchio E, Lapelosa M, Levy RM (2010) *J Chem Theory Comput* 6(9):2961
32. Güner OF (1999) Pharmacophore perception, development, and use in drug design. International University Line, La Jolla, CA
33. Langer T, Hoffman RD (2006) Pharmacophores and pharmacophore searches. Wiley-VCH, Weinheim, Germany
34. Zhu T, Cao S, Su P-C, Patel R, Shah D, Chokshi HB, Szukala R, Johnson ME, Hevener KE (2013) *J Med Chem* 56(17):6560
35. Sanner MF (1999) *J Mol Graph Model* 17(1):57
36. Lin J-H, Perryman AL, Schames JR, McCammon JA (2002) *J Am Chem Soc* 124(20):5632
37. Amaro RE, Baron R, McCammon JA (2008) *J Comput Aided Mol Des* 22(9):693
38. Nichols SE, Baron R, Ivetac A, McCammon JA (2011) *J Chem Inf Model* 51(6):1439
39. Lin JH, Perryman AL, Schames JR, McCammon JA (2003) *Biopolymers* 68(1):47