JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Comparative Modeling and Benchmarking Data Sets for Human Histone Deacetylases and Sirtuin Families

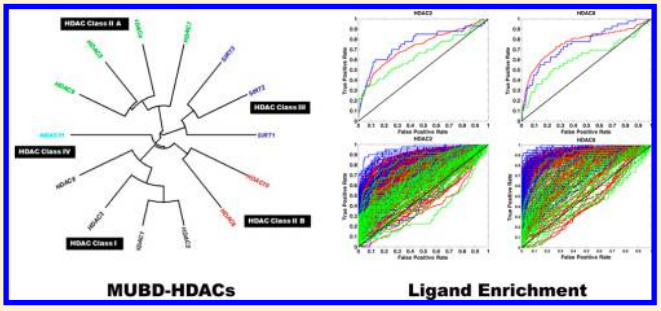
Jie Xia,[†,§] Ermias Lemma Tilahun,[§] Eyob Hailu Kebede,[§] Terry-Elinor Reid,[§] Liangren Zhang,*[,†] and Xiang Simon Wang*[,§]

[†]State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, 38 Xueyuan Road, Beijing 100191, China

[§]Molecular Modeling and Drug Discovery Core for District of Columbia Developmental Center for AIDS Research (DC D-CFAR); Laboratory of Cheminformatics and Drug Design, Department of Pharmaceutical Sciences, College of Pharmacy, Howard University, 2300 Fourth Street NW, Washington, D.C. 20059, United States

Supporting Information

ABSTRACT: Histone deacetylases (HDACs) are an important class of drug targets for the treatment of cancers, neurodegenerative diseases, and other types of diseases. Virtual screening (VS) has become fairly effective approaches for drug discovery of novel and highly selective histone deacetylase inhibitors (HDACIs). To facilitate the process, we constructed maximal unbiased benchmarking data sets for HDACs (MUBD-HDACs) using our recently published methods that were originally developed for building unbiased benchmarking sets for ligand-based virtual screening (LBVS). The MUBD-HDACs cover all four classes including Class III (Sirtuins family) and 14 HDAC isoforms, composed of 631 inhibitors and 24 609 unbiased decoys. Its ligand sets have been validated extensively as chemically diverse, while the decoy sets were shown to be property-matching with ligands and maximal unbiased in terms of "artificial enrichment" and "analogue bias". We also conducted comparative studies with DUD-E and DEKOIS 2.0 sets against HDAC2 and HDAC8 targets and demonstrate that our MUBD-HDACs are unique in that they can be applied unbiasedly to both LBVS and SBVS approaches. In addition, we defined a novel metric, i.e. NLBScore, to detect the "2D bias" and "LBVS favorable" effect within the benchmarking sets. In summary, MUBD-HDACs are the only comprehensive and maximal-unbiased benchmark data sets for HDACs (including Sirtuins) that are available so far. MUBD-HDACs are freely available at http://www.xswlab.org/.



INTRODUCTION

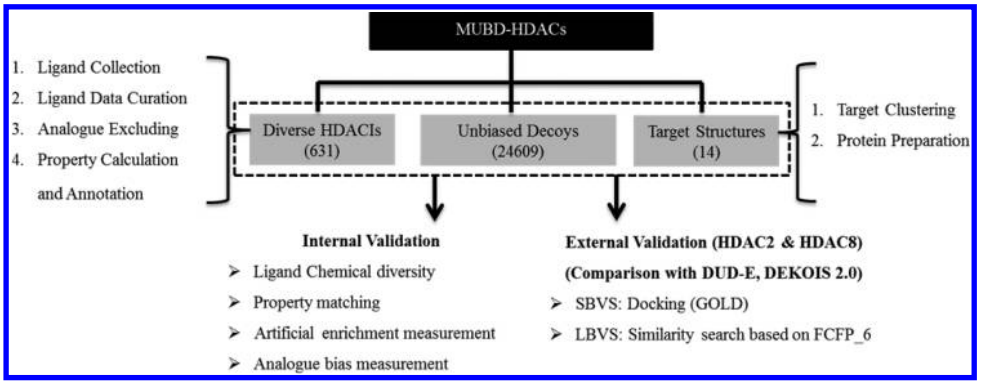
Nucleosomal histone acetylation and deacetylation are the major forms of posttranslational modifications in mammals, which play a pivotal role in the regulation of gene transcription.[1,2] Normally, the acetylation levels, i.e. hyperacetylation vs hypoacetylation, are well balanced by histone acetyltransfereases (HATs) and histone deacetylases (HDACs).[3] HATs transfer acetyl groups from acetyl-CoAs to generate acetylated lysine residues in histone tails. The presence of these acetylated lysine residues, an indication of hyperacetylation, is normally associated with relaxed chromatin status and transcriptional activation.[4] By contrast, HDACs remove acetyl groups from acetylated lysine residues thus lead to the hypoacetylation, an increase of deacetylated lysine termini.[5,6] The histones with deacetylated lysine termini are positively charged thus can bind tightly to the negatively charged DNA phosphate groups, resulting in chromatin compaction.[2] The compact chromatin structure reduces accessibility of transcription factors to DNA,[7] which hence causes transcriptional repression.[8−10] However, when genes that encode HATs, HDACs, or their binding partners are mutated or their expression is altered, the homeostasis of histone acetylation levels can be disturbed.[8] These types of imbalances are involved in the pathology of a wide range of diseases, e.g. cancers[2,8,11,12] and neurodegenerative diseases such as Alzheimer's disease (AD) and Parkinson's disease (PD).[13,14]

Up to now, 18 mammalian HDAC isoforms have been identified and are classified into 4 classes (I, II, III, and IV) according to their sequence homology to yeast HDACs.[2] Different classes of HDACs have their unique characteristics such as cofactors, subcellular localization, and tissue distribution. As for the cofactors, Classes I, II, and IV of HDACs are zinc-dependent enzymes, whereas class III is nicotinamide adenine dinucleotide (NAD^+) dependent.[13] For subcellular localization, Class I of HDACs, comprised of HDAC1, HDAC2, HDAC3, and HDAC8, primarily localizes in the nucleus where they regulate histone acetylation process. Class II of HDACs is further divided into two subclasses: IIA of

Received: September 10, 2014
Published: January 29, 2015

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

Scheme 1. Workflow for the Construction and Validation of MUBD-HDACs



HDAC4, HDAC5, HDAC7, and HDAC9 and IIB of HDAC6 and HDAC10. Unlike Class I, Class II HDACs are able to shuttle between the nucleus and cytoplasm in response to certain cellular signals.[1] HDAC11 is the only member in the Class IV HDACs, which mainly localizes in the nucleus as well[15] and is homologous to HDAC3 and HDAC8.[16] Class III HDACs are homologous to the yeast Sir2 family of proteins, thus were named as Sirtuins (SIRT1-SIRT7). Contrary to Class I and II HDACs, they are insensitive to Trichostatin A (TSA) or other hydroxamate inhibitors.

HDACs are involved in a wide range of key physiological functions and constitute an important class of modern drug targets. First, HDACs have been validated extensively as druggable targets for the treatment of various cancers.[1,6] On one hand, the upregulation of HDACs activity or over-expression induces the aberrant transcriptional repression of key genes that regulate important cellular functions,[2,17,18] which is linked to tumorigenesis.[19,20] On the other hand, it has been proven that HDAC inhibitors (HDACIs) could increase acetylation level thus to inhibit aberrant cellular proliferation by interfering with cell cycle regulation.[6,12,21] Many HDACIs had been approved for clinical use to treat cancers such as vorinostat,[22] mocetinosat (MGCD0103),[23,24] entinostat (MS-275),[25,26] and romindepsin (cyclodepsipeptide or FK228).[27,1,6] Second, HDACs also serve as promising targets for modern drug discovery for the treatment of AD. The pathology of neurodegenerative diseases such as AD has been proposed in recent years to be associated with imbalances in histone acetylation levels and transcriptional dysfunctions.[13,14,28] For example, HDAC2 was found to negatively regulate memory formation and synaptic plasticity.[29] HDAC3 is also a negative regulator of long-term memory formation.[30] HDAC6, a nonhistone deacetylase, is a target for protection and regeneration following injury in the central nervous system.[31] HDACIs could interfere with the activities of proteins which play an important role in AD, e.g. Aβ, GSK-3β, and tau protein, thus are able to enhance synaptic plasticity and improve learning and memory.[32] In addition, SIRT1, SIRT2, and SIRT3 were also proposed to play a role in the pathology of AD.[13]

For drug discovery targeting various HDACs, virtual screening (VS) approaches including both structure-based virtual screening (SBVS)[33−39] and ligand-based virtual screening (LBVS)[1,40−42] have been widely used. In these studies, it is indicated clearly that the appropriate selection of VS approach is essential to the successful identification of novel HDACIs.[1,33,38,39,43] Nowadays, there are a variety of VS approaches available to use, whereas the performance of each one when applied to a specific target is not known. In order to choose the optimal method for future VS campaigns against certain HDAC isoform, it is necessary to evaluate their potential performances beforehand. Normally, the evaluation aims to access the potential ligand enrichment[44,45] in prospective VS based on the retrospective small-scale VS against benchmarking data sets that mimic the real-world chemical libraries.[46] So far, the ready-to-apply benchmarking data sets for HDACs are rather limited and do not cover the whole panel of HDAC isoforms. In the gold-standard series of benchmarking sets, i.e. directory of useful decoys (DUD),[44] DUD clusters,[47] charge matched DUD,[48] DUD LIB V1.0,[49] and DUD-enhanced (DUD-E),[45] only two HDACs targets of HDAC2 and HDAC8 were recently covered by DUD-E. In addition to that, the WOMBAT database[47] and virtual decoy sets (VDS)[50] cover the same set of isoforms with DUD; demanding evaluation kits for objective in silico screening (DEKOIS)[51] and DEKOIS 2.0[52] include only HDAC2, HDAC8, and SIRT2. Besides those, other major benchmarking sets such as G protein-coupled receptors (GPCRs) ligand library (GLL), GPCR decoy database (GDD),[53] nuclear receptors ligands and structures benchmarking database (NRLiSt BDB),[54] database of reproducible virtual screens (REPROVIS-DB), and maximum unbiased validation (MUV) do not contain any target from HDACs. A viable solution to supplement insufficient coverage of HDAC isoforms can be the accessibility of decoy-making tools, which are normally web-based in this case.[45,55] However, other issues arise concerning the application domain of such tools, especially for LBVS, SBVS, or both. Current decoy makers such as DecoyFinder[55] or the DUD-E server (http://dude.docking.org/generate)[45] are not able to generate decoys for LBVS. The only tool for LBVS, i.e. MUV,[56] is restricted by the targets and sufficient experimental decoys (chemical space of decoys) in the PubChem database. In fact, the ready-to-apply data sets from DUD-E and DEKOIS 2.0 were also specifically designed for SBVS, i.e. molecular docking.[45,52] To the best of our knowledge, benchmarking sets for whole HDACs family with application domains for LBVS approaches are not available. There are thus huge demands to build unbiased benchmarking data sets for LBVS approaches, ideally for both SBVS and LBVS, to assist the ongoing efforts of drug discovery for HDACs.

In this study, we applied our new algorithms[57] to select ligands and generate decoys for the whole family of HDAC isoforms and built benchmarking data sets named "Maximal Unbiased Benchmarking Data Sets for HDACs" (MUBD-HDACs), for both LBVS and SBVS approaches. Each data set incorporates carefully curated diverse HDACIs, unbiased

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

decoys, and well-prepared crystal structures of HDACs. For the purpose of comparison, we chose our data sets of HDAC2 and HDAC8 and evaluate their performance with the corresponding data sets in DUD-E and DEKOIS 2.0. We anticipate that MUBD-HDACs can be a powerful tool to assist the discovery of new HDACIs.

METHODS

Workflow to Construct MUBD-HDACs. For the purpose of benchmarking both LBVS and SBVS approaches, we designed MUBD-HDACs to contain three parts: diverse HDACIs, unbiased decoys, and target structures (cf. Scheme 1). All diverse HDACIs were originally obtained from the ChEMBL18 database (https://www.ebi.ac.uk/chembl/, accessed Oct. 2014),[58] with the only exception for SIRT3 whose data were taken directly from recent scientific literatures. All the data were further refined by consecutive steps: ligand collection, data curation, analogue excluding, and property calculation and annotation. Unbiased decoys were generated using our in-house workflow and algorithms.[57] Target structures were also curated and prepared as ready-to-apply. Every benchmarking set in MUBD-HDACs has been validated extensively, including evaluation of ligand chemical diversity, property matching, artificial enrichment bias, and analogue bias. We also conducted comparative studies with other published benchmarking sets that cover HDACs, i.e. DUD-E[45] and DEKOIS 2.0,[52] in terms of ligand enrichment by docking (GOLD) and similarity search based on the function class fingerprints of maximum diameter 6 (FCFP_6) fingerprint. Only HDAC2 and HDAC8 were covered at this time because of their availabilities in both DUD-E and DEKOIS 2.0.

Generation of Diverse HDACIs Data Sets. *Ligand Collection.* Initially, all the ligands for each HDAC isoform were downloaded from ChEMBL18 database.[58] The following criteria were then applied to the collection of ideal ligands: (1) To obtain ligands annotated with ChEMBL confidence score ≥4. The confidence score reflects both the type of target assigned to a particular assay and the level of confidence that the target assigned is the correct target for that assay.[68] In order to make sure that the assays were truly for assigned protein targets, the ChEMBL confidence score of 4 was set as the cutoff, which was previously applied to DUD-E as well.[45] As a result, this criterion did not exclude any ligand as the confidence scores of all ligands for each isoform are 8 for homologous proteins, 9 for direct proteins, or 5 for multiple direct proteins such as HDAC1. (2) To obtain ligands annotated with quantitative activity data of $IC_{50} \leq 1$ μM, preferably.[45] As Jahn et al. advised, the number of chemotypes must be sufficient to reduce potential bias thus the minimal size greater than 17 was ensured in their study.[49] Likewise, when no more than 17 diverse HDACIs (after ligand data curation and analogue excluding, see below) were obtained for certain targets in our study, the cutoff of IC_{50} for that target would be increased to 10 or 100 μM in this case. For instance, IC_{50} = 10 μM was used as the final cutoff for HDAC5, HDAC7, HDAC10, SIRT1, and SIRT2, while IC_{50} = 100 μM was the final cutoff for HDAC9 and SIRT3 (cf. Table S1). (3) If the number of collected ligands was still limited after the applying of criterion 2, other ligands in ChEMBL annotated with inhibition rates (%) or activities (%) would be considered. If needed, ligands that were published but had not been included in ChEMBL at the time of data collection were also included. To be noted, this criterion was only applied to the ligand compilation for SIRT3 (cf. Table S1). (4) Among the ligands that met the above three criteria, some ligands were tested by different types of bioassays or reported by multiple sources thus had various activities. We named those ligands as the "duplicates" and their number for each HDACs isoform is shown in Table S1. Because the benchmarking data set is mostly designed for ligand enrichment study, the only information required for each compound in the data sets is its activity state, i.e. active or inactive, rather than the exact activity values and/or types. Therefore, only the chemical structure and its activity state were retained for each ligand inside the data sets. But for the clarification purpose, we listed an example of HDAC1 potent inhibitors with inconsistent activity values from scientific literatures in Table S2. For example, CHEMBL271741 was tested by three different HDAC1 bioassays and its IC_{50} values were 52, 60, and 410 nM, respectively. As a result, only the chemical structure of CHEMBL271741 and its activity state were included in the ligand set. The whole collection of ligands after "duplicates" excluding was named to be "raw ligands" (cf. Table 1 and Table S3).

Ligand Data Curation. We applied extensive and thorough manual-checking procedures to make certain that the chemical structure and activity state of each ligand in the "raw ligands" set and the sequential ligand sets are correct in a very strict sense. First, the source (i.e., scientific literature) of each ligand was retrieved and its chemical structure in the "raw ligands" set, in particular the stereochemistry, was carefully inspected by referring to its original literature. In the meantime, all sources of activities data for the collected ligands were checked in order to make certain that the activity state we annotated in the "raw ligands" set and the sequential ligand sets are correct. Second, components of "Strip salts", "Standardize molecule" inside Pipeline Pilot (version 7.5, Accelrys Software, Inc.) were applied and a property filter of "RBs > 20" or "MW ≥ 600" originating from the chemical curation workflow for DUD-E was followed.[45] At last, the charged ligands at pH range of 7.3−7.5 were generated and constituted the set of "ligands after curation" (cf. Table 1).

Analogue Excluding. Analogue excluding remains to be the main task in the step of "Ligand Processing" inside our in-house workflow and plays a pivotal role in reducing the analogue bias.[57] First, the pairwise topological similarities in terms of Tanimoto coefficient (Tc) based on MACCS structural keys[59] between those charged ligands were calculated. Next, ligands with Tc over 0.75 were excluded. As mentioned in our methodology publication,[57] the use of Tc = 0.75 as a cutoff is a well-established criterion for ligand selection instead of an arbitrary one. It is highly associated with the cutoff of Tc applied during the decoy building to reducing "false negative" rate. The remaining ligands were named as "diverse HDACIs" (cf. Table 1 and Table S3).

Property Calculation and Annotation. Six physicochemical properties, i.e. LogP, molecular weight (MW), number of hydrogen bond acceptors (HBAs), number of hydrogen bond donors (HBDs), number of rotatable bonds (RBs), and net formal charge (nFC) were calculated for all the diverse HDACIs using Pipeline Pilot. At the same time, the original IDs of those HDACIs (i.e., CMPD_CHEMBLID) were retrieved from ChEMBL18 database and added to each diverse structure for the purpose of reference annotation.[52,54] In this way, the data sets of diverse HDACIs are ready to use.

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

Table 1. Information of Ligands, Decoys, and Protein Structures for All Targets in MUBD-HDACs

HDAC class	HDAC isoform	UniProt accession code	PDB codes	no. raw ligands	no. ligands after curation	no. diverse HDACIs	scaffolds	compound/scaffold ratio	no. decoys
HDAC Class I	HDAC1	Q13547	4BKX	1284	1251	180	147	1.22	7020
	HDAC2	Q92769	4LXZ, 4LY1, 3MAX	215	205	63	50	1.26	2457
	HDAC3	O15379	4A69	158	140	39	34	1.15	1521
	HDAC8	Q9BY41	1T64, 1T67, 1T69, 1VKG, 1W22, 2 V5W, 2 V5X, 3EW8,3EWF, 3EZP, 3EZT,3F06, 3F07, 3F0R, 3MZ7, 3RQD, 3SFF, 3SFH	231	201	39	34	1.15	1521
HDAC Class IIA	HDAC4	P56524	4CBT, 4CBY, 2VQJ, 2VQM, 2VQO, 2VQQ, 2VQV, 2VQW	135	122	39	36	1.08	1521
	HDAC5	Q9UQL6		53	52	22	19	1.16	858
	HDAC7	Q8WUI4	3C0Y, 3C0Z, 3C10, 3ZNR, 3ZNS	52	50	24	22	1.09	936
	HDAC9	Q9UKV0		65	61	28	25	1.12	1092
HDAC Class IIB	HDAC6	Q9UBN7		495	467	96	74	1.30	3744
	HDAC10	Q969S8		90	90	23	19	1.21	897
HDAC Class III	SIRT1	Q96EB6	4I5I, 4IF6, 4IG9, 4KXQ	103	91	21	17	1.24	819
	SIRT2	Q8IXJ6	4L3O, 1J8F, 3ZGO, 3ZGV	123	118	24	21	1.14	936
	SIRT3	Q9NTG7	3GLR, 3GLS, 3GLT, 3GLU, 4BN4, 4BN5, 4BV3, 4BVB, 4BVE, 4BVF, 4BVG, 4BVH, 4C78, 4C7B, 4FVT, 4FZ3, 4HD8, 4JSR, 4JT8, 4JT9	95	99	16	15	1.07	624
HDAC Class IV	HDAC11	Q96DB2		29	29	17	15	1.13	663

Generation of Unbiased Decoy Data Sets. Our new algorithms and the whole workflow were implemented in Pipeline Pilot and applied to generate unbiased decoys for all targets in MUBD-HDACs.[57] The steps of "Preliminary Filtering" and "Precise Filtering" were carried out first to choose ideal decoys from the "All-Purchasable Molecules" subset (~18 million) of ZINC (http://zinc.docking.org/, downloaded Dec. 2012).[48,60] In next step of "Preliminary Filtering", ZINC compounds were reduced using two preliminary filters, one of which was defined by the range of physicochemical properties for HDACIs and another was mutual topological similarities within the ligand sets, i.e. the "similarity in structure" ("sims") measured by MACCS, in a range from the minimum value to 0.75. Because the decoy compounds whose sims values to any ligand over 0.75 were likely to be actives, we followed the GLL/GDD[53] to set the upper bound to be 0.75 in order to greatly reduce the risk of being "false negatives" among decoys. The sims mentioned above is in essence the Tc, the percentage of shared bits in all bits in the fingerprints of target compound i and reference compound j, as can be described in eq 1:

$$\mathrm{sims}_{i,j} = \mathrm{Tc}_{i,j} = \frac{N_{i,j}}{N_i + N_j - N_{i,j}} \tag{1}$$

For "Precise Filtering", its aim is to select 39 suitable decoys for each HDACI that have (1) physicochemical properties (LogP, MW, HBAs, HBDs, RBs, and nFC) that are highly similar to their corresponding ligand and (2) random spatial distribution in comparison to other ligands. To this end, a precise filter based on "similarity in properties" ("simp") was applied first and its output was further filtered using the "sims difference" ("simsdiff") based precise filter. The cutoff for the simp-based filter can be lowered automatically from 0.95 to 0.50 until a sufficient number of final decoys were obtained. Finally, 39 final decoys after the $\overline{\text{simsdiff}}$-based filter were chosen according to their low $\overline{\text{simsdiff}}$ values. The precise filters of simp and $\overline{\text{simsdiff}}$ are defined in eqs 2 and 3. The term simp is derived from Euclidean distance and calculated according to eq 2. In this equation, p represents the scaled value of physicochemical property, n is the total number of physicochemical properties, and i is the index for individual property. T denotes the target compound, and R denotes the reference compound:

$$\mathrm{simp}_{T,R} = 1 - \sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_{i,\mathrm{T}} - p_{i,\mathrm{R}})^2} \tag{2}$$

The term $\overline{\text{simsdiff}}$ is used to record the average difference between two topological similarities measured by MACCS, one is between the potential decoy k and remaining ligands j, i.e. $\mathrm{sims}_{k,j}$, and another is between the query ligand i and the remaining ligands j, i.e. $\mathrm{sims}_{i,j}$. (cf. eq 3) In this equation, m is a constant that represents the number of the ligands. The indexes of potential decoys and query ligands, i.e. k and i respectively, are also constants taken from the individual data set. However, the index of remaining ligands j iterates from 1 to $m - 1$ by a step of 1:

$$\overline{\mathrm{simsdiff}}_{i,k} = \frac{1}{m-1}\sum_{j=1}^{m-1}|\mathrm{sims}_{k,j} - \mathrm{sims}_{i,j}| \tag{3}$$

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

Similar to those diverse HDACIs in the ligand sets, the final decoy data sets were annotated with ZINC IDs, values of simp, simsdiff, and six physicochemical properties.

Target Clustering and Preparation. *Generate Polar Dendrograms.* In order to give an overview of all the targets covered in MUBD-HDACs, we generated a polar dendrogram based on the protein sequences. First, the protein sequences of all the targets were retrieved from UniProt (http://www.uniprot.org)[61] using the annotated UniProt accession codes shared from the ligands in ChEMBL. We then employed ClusterX2 to align the sequences.[62] Next, the package "ape" for phylogenetics in R[63] was applied to calculate genetic distances and perform hierarchical clustering to generate the polar dendrogram. Because of the lack of enough diverse HDACIs for SIRT4, SIRT5, SIRT6, and SIRT7, these isoforms were not included in the final target list for MUBD-HDACs.

Protein Preparation. For the purpose of benchmarking of SBVS approaches, high-quality and ready-to-apply protein structures are an essential part of the database as well. With UniProt accession codes, those targets in MUBD-HDACs were mapped to protein data bank (PDB) to obtain X-ray structures for *Homo sapiens* using cross-references tool in http://www.uniprot.org/docs/pdbtosp.txt (accessed Oct. 2014). All currently available structures were downloaded and their related information was retrieved from the literature. Those structures with incomplete or no ligand binding pocket were discarded due to their limitations for future docking studies. For those remaining protein structures, the identical protein chains, cocrystallized water molecules, nonessential cofactors except for zinc ions, NAD^+, and its analogues in the active sites were removed. The protein structures were then prepared using the "Clean Protein" module of Discovery Studio (version 2.5, Accelrys Software, Inc.) to add hydrogens, modify chain termini, correct nonstandard names, alternate conformations, repair incomplete residues and atom order in amino acids, and protonate the whole protein at pH 7.0. For those targets with *holo* structures, the cognate ligand was stripped and retained. The cognate ligands contain much useful information on interactions within active site, thus can be a good starting point for enrichment study and method development. In the end, all the prepared 3D structures were provided in the way as recommended by Lagarde et al. in NRLiSt BDB.[54]

Extensive Validation of MUBD-HDACs. *Ligand Chemical Diversity.* Bemis-Murcko atomic frameworks[64] were generated to represent the unique scaffolds in the ligand set of our MUBD-HDACs. The ratio of compounds per scaffold was used as a metric to measure ligand diversity and calculated by the "Generate Fragments component" in Pipeline Pilot for this purpose.

Property Matching. Two groups of curves were plotted to evaluate property matching between ligands and decoys: property distribution curves and radar plots. The former curve has been widely used to measure overall property matching of ligands and decoys in different benchmarking systems, e.g. DUD,[44] DUD-E,[45] and GLL/GDD.[53] Herein, to validate property matching in MUBD-HDACs, we also plotted distribution curves of six physicochemical properties (LogP, MW, HBAs, HBDs, RBs, and nFC) for the whole set of diverse HDACIs and their unbiased decoys. The radar charts of physicochemical properties were recently applied to compare different data sets in NRLiSt BDB.[54] This type of plots considers the whole space of distribution in physicochemical properties of data sets compounds at the same time. Therefore, we plotted radio charts of mean, maximal and minimal values of those physicochemical properties for each data set in MUBD-HDACs to measure property matching.

Artificial Enrichment Bias Measurement. To this end, we applied a metric of the average AUC of ROC curves, i.e. mean(ROC AUCs), from simp-based VS using the format of leave-one-out (LOO) cross-validation (CV).[57] As mentioned before, in each cycle the values of simp between the query ligand and other compounds in the screening set (ligand subset and decoy subset) were calculated and sorted. Next the ROC curve was plotted based on the sorted simp values and its corresponding AUC was calculated. Finally, the AUCs from multiple cycles were averaged to obtain the value of mean(ROC AUCs). Its specific value of 0.5 represents the status of random distribution of ligands and decoys inside the chemical space of six physicochemical properties. In another word, it is unable to distinguish ligands and their corresponding decoys at this point. The closer to 0.5 the mean(ROC AUCs) value is, the smaller the "artificial enrichment" bias is. Because the mean(ROC AUCs) of simp-based similarity search is based on physicochemical properties of ligands and decoys, it can be considered as a metric for property matching as well.

Analogue Bias Measurement. The sims-based VS using the MACCS structural keys[59] was performed to check whether "analogue bias" had been reduced toward the ideal level of random distribution in chemical topology. The similar procedure of LOO CV was applied and the mean(ROC AUCs) was calculated. Unlike simp, the measurement of topological similarity is sims, i.e. Tc.[57,65]

Metric to Measure the Effect of Nearer Ligands (NLs). NLs are ligands that are more similar to the query in chemical structure than the nearest decoy when similarity search is conducted. The existence of NLs is one of the major reasons that causes enrichment bias for 2D LBVS approaches.[66] Herein, we designed a metric to quantify the effect of NLs based on the calculation of LOO CV. Similar to the LOO CV of MACCS sims-based similarity search, FCFP_6 sims-based search was conducted for this specific purpose. Let n be the number of ligands in the ligand set, then $n - 1$ is the total number of ligands in the ligand subset (index is k, $k = 1, 2, ..., n - 1$) for each iteration of LOO CV. Due to the nature of LOO CV, the similarity search is repeated n times. In each iteration i, the values of sims based on FCFP_6 are calculated for each ligand in the ligand set ($sims_{Q,L}$), and each decoy in the decoy set with the reference to the ligand query ($sims_{Q,D}$). The maximum of $sims_{Q,D}$ (i.e., $sims_{Q,D\_max}$) is located and set as a threshold. Next the number of ligands in the ligand set whose $sims_{Q,L}$ is greater than the threshold is counted, represented as $\sum\delta$ ($\delta = 1$ or 0)(cf. eq 6). These ligands are regarded as NLs, which are nearer to the query than the nearest decoys. The percentage of NLs (NL%) in the ligand collection of the benchmarking data sets is the metric to quantify the effect on the enrichment. In this way, the overall effect, represented as "NL bias score (NLBScore)", can be measured by the average value of NL% from n iterations. Both NL% and NLBScore are described by the formula below (eqs 4 and 5):

$$\text{NLBScore} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{k=1}^{n-1}\delta_k\right) \tag{4}$$

$$\text{NL\%} = \frac{1}{n-1}\sum_{k=1}^{n-1}\delta_k \tag{5}$$

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

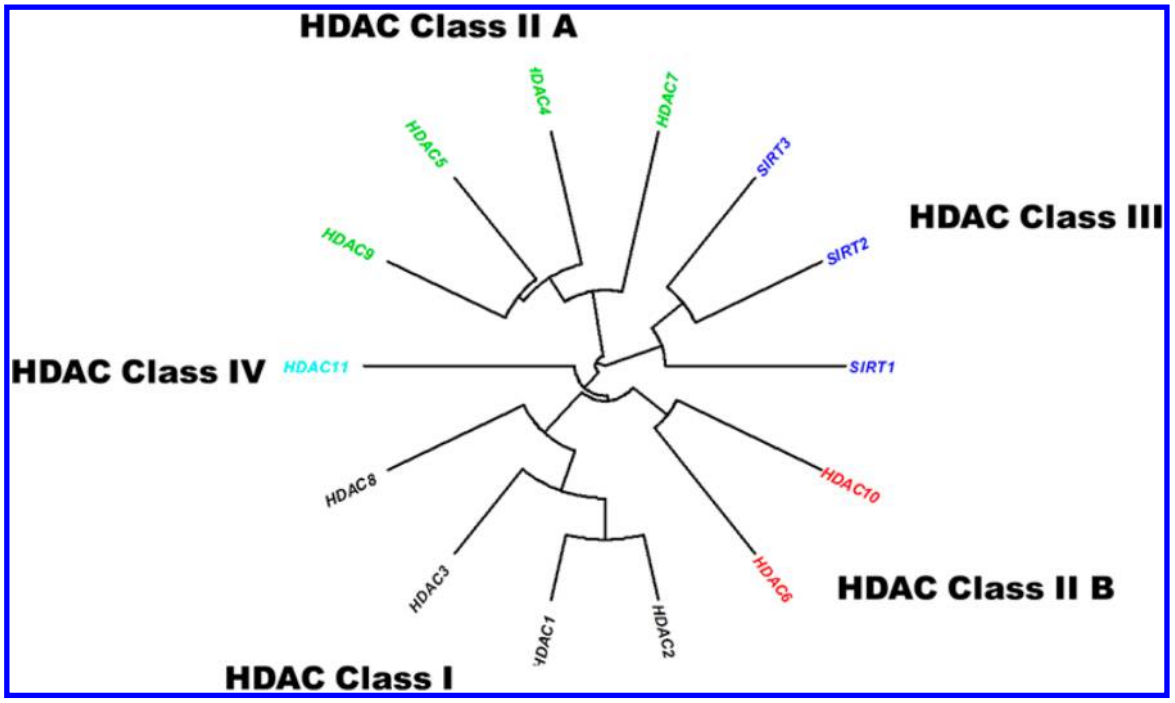


Figure 1. Clustering of MUBD-HDACs targets. Fourteen HDAC isoforms are clustered into five groups by hierarchical clustering according to genetic distances between protein sequences and depicted as a polar dendrogram. Different classes are represented by different colors.

$$\delta_k = \begin{cases} 1, \text{ if } & \text{sims}_{Q,k} > \text{sims}_{Q,D\_max} \\ 0, \text{ if } & \text{sims}_{Q,k} \leq \text{sims}_{Q,D\_max} \end{cases} \quad (6)$$

According to the above two equations, the lower NLBScore, the fairer the benchmarking evaluation.

Comparison with DUD-E and DEKOIS 2.0. As a measure of further validation, we compared MUBD-HDACs with two popular benchmarking sets, DUD-E and DEKOIS 2.0. It should be noted that this type of comparison is limited because only HDAC2 and HDAC8 data sets were included due to the few overlap. In this comparative study, the ligand diversity and the enrichment performance for both SBVS and LBVS approaches were explored. For SBVS approaches, molecular docking by GOLD (version 3.0) was performed on all three benchmarking sets and the protein structures for human HDAC2 (PDB code: 3MAX) and HDAC8 (PDB code: 3F07) prepared in MUBD-HDACs were used. The binding site residues of both structures were defined by a sphere centering on the coordinates of their cognate ligands and with a radius of 8 Å (HDAC2) or 9 Å (HDAC8). The ChemScore was employed as the scoring function for measuring the binding affinities. The docking mode was set as default and the iteration was allowed to terminate when RMSDs among the top three solutions were within 1.5 Å RMSD. The poses with the highest scores were retained for the overall ranking. For LBVS approaches, FCFP_6 based similarity search was performed in the form of LOO CV, similar to the MACCS sims-based similarity search. For the evaluation metrics, the ROC AUC was used for the SBVS approach while the mean(ROC AUCs) was applied for the LBVS approach.

RESULTS AND DISCUSSION

Overview of MUBD-HDACs. Our MUBD-HDACs database was designed for benchmarking of both SBVS and LBVS approaches for drug discovery targeting HDACs. It includes diverse HDACIs, unbiased decoys and protein structures, with the basic statistics listed in Table S3. As Jahn et al. mentioned before, a sufficient number of chemotypes are needed for ligand enrichment study.[49] Accordingly, we adopted 17 or so as the minimum number of diverse ligands throughout our targets. Because of this strict criterion, SIRT4, SIRT5, SIRT6, and SIRT7 were not included in the final target list due to the scarce of active compounds retrieved from ChEMBL. In the end, the final format of database contains 14 targets out of the whole panel of 18 HDACs isoforms. These targets were further clustered into five classes according to their protein sequence homology, which were consistent with current definition of HDACs subfamilies.[67] The polar dendrogram generated from target clustering clearly shows the target coverage of MUBD-HDACs. Notably, MUBD-HDACs covers all these subfamilies though it does not include all the targets of sirtuins in HDAC Class III (cf. Figure 1). Among them, nine targets have been supplemented with crystal structures to assist the SBVS benchmarking. There were a total of 3128 raw ligands for all the targets of MUBD-HDACs initially. After our extensive curation, the number of ligands including the charged form was reduced to 2976. Finally, there are 631 diverse HDACIs left, with an average of 45 for each target, a maximum of 180 and a minimum of 16. From the reductions of ligand numbers at each step, we conclude that it is the step of "analogue excluding" that plays the most pivotal role in suppressing ligand entries. In total, 24 609 compounds were chosen from approximately 18 million of "All-Purchasable Molecules" in ZINC database as our maximal unbiased decoys, with 39 decoys for each ligand (cf. Table S3). Table 1 shows the composition of benchmarking data sets for each target so as to give a detailed view of this collection. In this table, we provide the UniProt accession code, multiple PDB codes, number of raw ligands, ligands after curation, diverse HDACIs and unbiased decoys for each target. The crystal structures with the reference of PDB codes were extensively checked and prepared to be in the ready-to-apply

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

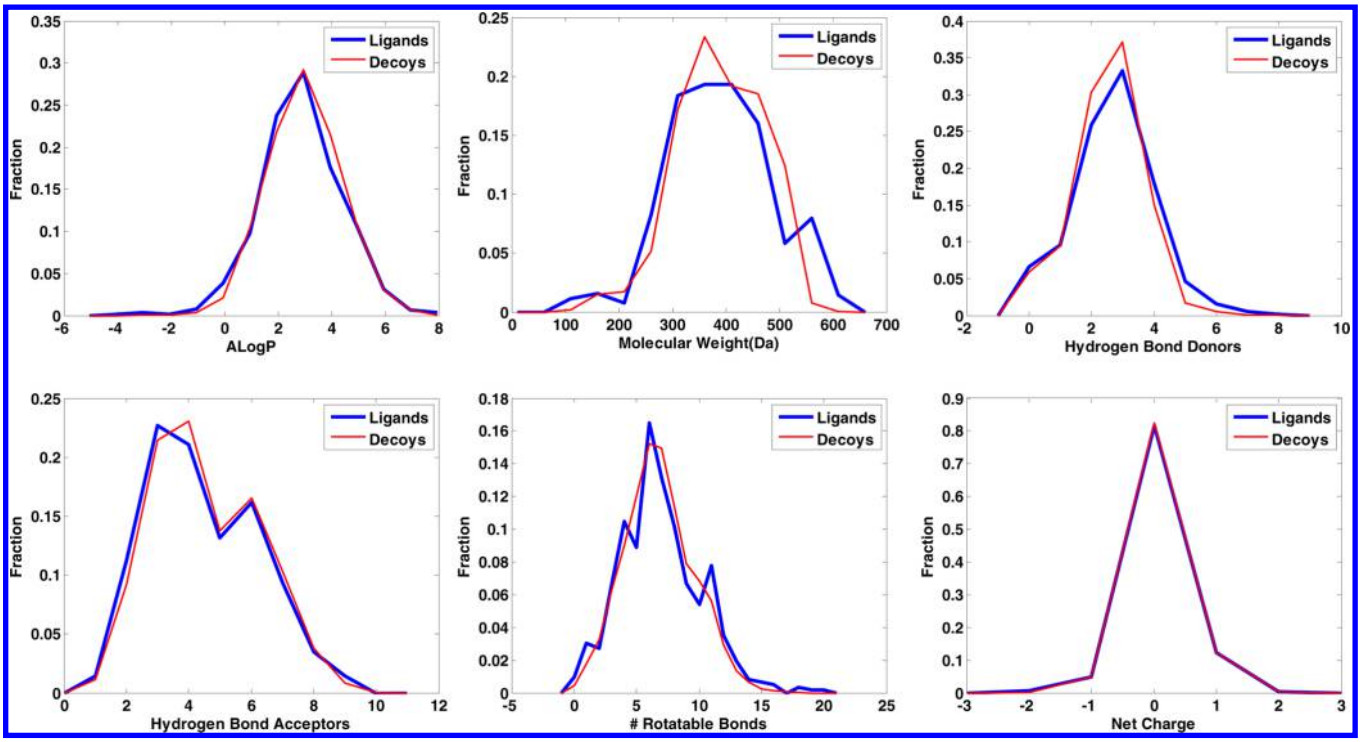


Figure 2. Physicochemical properties distribution curves for all diverse ligands and unbiased decoys in MUBD-HDACs.

format, which can be the powerful resources for ligand enrichment study of SBVS. In addition, the ratios of compound to scaffold taken from chemical diversity analysis are also listed in Table 1. The value for each ligand set is close to 1:1, with a maximum of 1.30 and a minimum of 1.07, which indicates that each compound in the ligand set represents nearly one unique scaffold (i.e., Bemis−Murcko atomic frameworks).

Property Matching in MUBD-HDACs. Figure 2 shows property distribution between all the ligands and decoys in MUBD-HDACs. The distribution curves match fairly well for each of the six properties, i.e. LogP, MW, HBAs, HBDs, RBs, and nFC. First, they span the similar range in terms of the maximum and minimum values for both ligands and decoys. Second, the peaks of curves are located at the same interval for most of six properties, especially the AlogP, HBD, RBs, and nFC. Figure 3 is a complete set of radar plots, showing the average, maximum, and minimum values of physicochemical properties for each target. It should be pointed out that all the properties of the ligands match well with those of decoys for the average values (middle circle). For the maximums (outer circle) and minimums (inner circle), the ligands are consistent with the decoys for most properties though there are minor gaps in one or two properties, e.g. the max of HBDs for HDAC1, the min of AlogP for HDAC2. To be specific, the matching of maximums and minimums indicates that the properties span similar ranges between ligands and decoys.

Artificial Enrichment Bias and Analogue Bias Corrections. Figure 4 and Table S4 show the profile of mean(ROC AUCs) from LOO CV based on similarity searches for each target in MUBD-HDACs. All the mean values are fairly close to the line of 0.5 (random distribution) for simp-based search, which indicates it is rather challenging to distinguish ligands and decoys based on six physicochemical properties. Therefore, all the data sets in MUBD-HDACs are deemed to be of no "artificial enrichment" thus ideal for SBVS benchmarking. It is consistent with the good property matching between ligands and decoys shown in Figures 2 and 3. Another curve in Figure 4 is about the mean(ROC AUCs) from LOO CV with MACCS sims-based similarity search. For the majority of the targets (9), the values of mean(ROC AUCs) are less than 0.6; for the other targets (5), the values are less than 0.7 with the maximal value of merely 0.689 for the target of HDAC10. These low values of mean(ROC AUCs) imply that it is also difficult to distinguish ligands and decoys based on MACCS structural keys. As expected, the analogue bias caused by the inherent advantage of analogues retrieval in ligand enrichment, has been reduced significantly by choosing the most ideal decoys according to our in-house criteria. In summary, the data sets in MUBD-HDACs appear to be the maximal-unbiased for both SBVS and LBVS approaches.

Comparison of MUBD-HDACs to DUD-E and DEKOIS 2.0. *HDACIs Chemical Diversity.* For HDAC2, the ratios of compounds per scaffold are 1.26, 1.11, and 1.33 for MUBD-HDACs, DUD-E, and DEKOIS 2.0, respectively. For HDAC8, these ratios were 1.15, 1.10, and 1.29, respectively (cf. Table 2). When protonated forms of the same ligand in DUD-E were taken into account as one compound, the ratios were lowered to 1.02 for both HDAC2 and HDAC8. Generally speaking, all three ligand sets are regarded as diverse since their ratios approximate to 1:1, which indicates that all three methods work in an effective way. Nevertheless, there is a small difference among the three methods. The ligands in DUD-E are the most diverse followed by MUBD-HDACs, while DEKOIS 2.0 ranks the last apparently. DUD-E ligand set applied the Bemis-Murcko atomic frameworks (scaffolds) for clustering the ligands, the same algorithm for scaffold analysis thus could show advantages in this case. For MUBD-HDACs and DEKOIS 2.0, MACCS structural keys and FCFP_6 were applied for selecting diverse ligands. Based on the current result, it seems that our strategy using MACCS structural keys

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

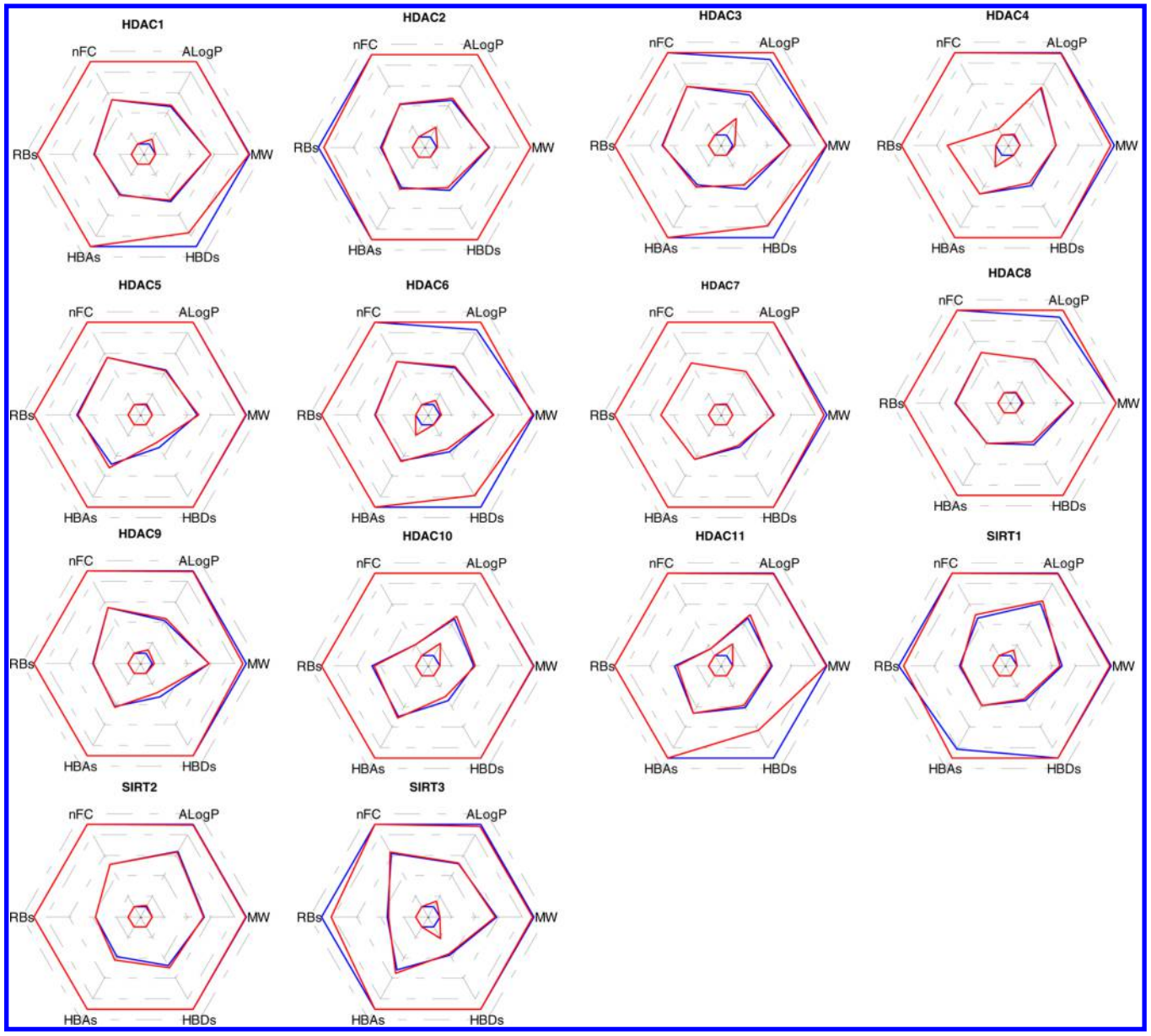


Figure 3. Radar chart representation of maximum (outer), average (mid), and minimum (inner) of six physicochemical properties for each target in MUBD-HDACs: (color codes) blue for ligands; red for decoys.

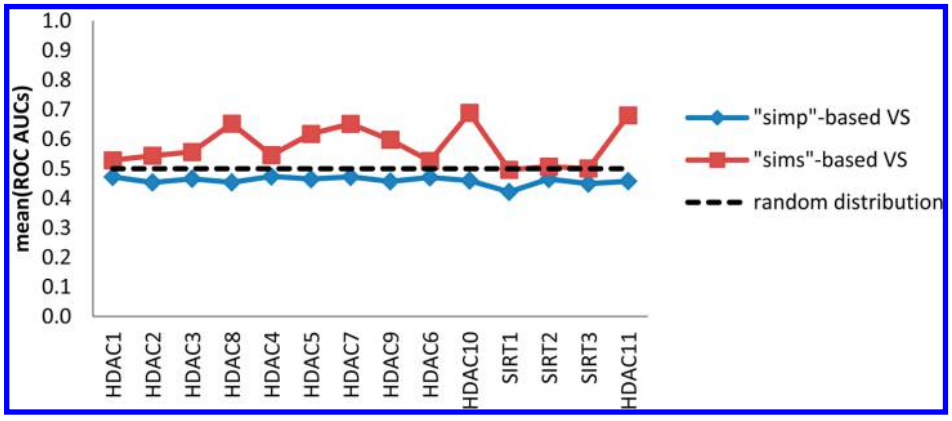


Figure 4. Performance of leave-one-out cross validation in the metric of mean(ROC AUCs) across all targets in MUBD-HDACs, by similarity search based on six physicochemical properties (blue) and MACCS structural keys (red).

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

Table 2. Comparison of MUBD-HDACs with DUD-E and DEKOIS 2.0 in Terms of Ligand Diversity, Ligand Enrichment, Nearer Ligand (NL) Biases for HDAC2 and HDAC8

		benchmarking sets		
target	items	MUBD-HDACs	DUD-E	DEKOIS 2.0
HDAC2	no. of compds	63	238 (185*)	40
	no. of scaffolds	50	214 (182*)	30
	compound/ scaffold ratio	1.260	1.11 (1.02*)	1.333
	ROC AUC (GOLD)	0.647	0.722	0.757
	mean(ROC AUCs) (FCFP_6)	0.623 ([b]0.083)	0.731([b]0.108)	0.813([b]0.110)
	NLBScore	0.052	0.091	0.160
	NL% (min)	0.000	0.000	0.000
	NL% (max)	0.210	0.257	0.538
HDAC8	no. of compds	39	234 (170*)	40
	no. of scaffold	34	213 (167*)	31
	compound/ scaffold ratio	1.147	1.10 (1.02*)	1.290
	ROC AUC (GOLD)	0.618	0.762	0.754
	mean(ROC AUCs) (FCFP_6)	0.666 ([b]0.118)	0.831([b]0.120)	0.903([b]0.086)
	NLBScore	0.040	0.183	0.178
	NL% (min)	0.000	0.000	0.000
	NL% (max)	0.211	0.536	0.538

*When different protonated forms of the same ligand are considered as one chemical entity. [b]Standard deviations for all AUCs from ROC curves.

works better in increasing ligand chemical diversity than DEKOIS 2.0.

Property Distribution. Figure 5 shows distribution curves of six physicochemical properties for ligands and decoys in HDAC2 and HDAC8 data sets from MUBD-HDACs, DUD-E, and DEKOIS 2.0. For both targets, MUBD-HDACs shows good property matching in all six properties, similar to the benchmarking sets of DUD-E. For the property of HBDs, MUBD-HDACs shows closer property matching than DUD-E in the distribution. DEKOIS 2.0 appears not to be as good as MUBD-HDACs and DUD-E in the property matching for HBDs, HBAs and nFC.

Ligand Enrichment. The ROC curves calculated from molecular docking by GOLD are shown in Figure 6A, while their values of ROC AUC are also listed in Table 2. For both targets of HDAC2 and HDAC8, the ROC curves for MUBD-HDACs are closer to the random distribution curve (diagonal line) than other two databases. Consistently, the ROC AUCs for MUBD-HDACs are 0.623 for HDAC2 and 0.618 for HDAC8, much less than those values with DUD-E of 0.722 for HDAC2 and 0.757 for HDAC8, and DEKOIS 2.0 of 0.762 for HDAC2 and 0.754 for HDAC8. These results indicate that the ligands in MUBD-HDACs are more challenging to be enriched from the background of our unbiased decoys by docking with GOLD. There seems to be no significant difference between DUD-E and DEKOIS 2.0 in this regard. Figure 6B shows the ROC curves generated from similarity search based on FCFP_6; Table 2 lists the values of mean(ROC AUCs)s of those curves for different data sets. Similarly, for both targets the ROC curves for MUBD-HDACs are the closest to random distribution curves, followed by DUD-E. DEKOIS 2.0 seems to be away from the diagonal lines somehow. The values of mean(ROC AUCs)s show the same trend to ROC curves. For both HDAC2 and HDAC8, the values are the lowest in MUBD-HDACs (e.g., 0.666 for HDAC8) and the highest in DEKOIS 2.0 (e.g., 0.903 for HDAC8) while DUD-E (e.g., 0.831 for HDAC8) ranks in the middle. These curves and values indicate that MUBD-HDACs data sets are more challenging than the other two benchmarking sets for similarity search using FCFP_6 fingerprint. In summary, the data sets of MUBD-HDACs for HDAC2 and HDAC8 are (1) as diverse as DUD-E while more diverse than DEKOIS 2.0; (2) comparable to DUD-E in property matching but better than DEKOIS 2.0; and (3) more challenging and unbiased than both DUD-E and DEKOIS 2.0 for enrichment by both types of VS approaches, i.e. docking using GOLD and similarity search using FCFP_6 fingerprint.

Higher Similarity of Decoys to Ligands in MUBD-HDACs. The Tc values based on FCFP_6 fingerprint, which measures structural similarities, were calculated between all ligands and decoys in MUBD-HDACs, DUD-E, and DEKOIS 2.0. The distribution curves of Tc were plotted for the purpose of comparison (cf. Figure 7). In fact, the curves for MUBD-HDACs shift to the right direction (i.e., higher similarity) for both HDAC2 and HDAC8. DUD-E's curves are located in the middle, while DEKOIS 2.0 is to the left. These data indicate the higher similarities of decoys to ligands in MUBD-HDACs followed by DUD-E and DEKOIS 2.0. In addition, the top five pairs of ligands and decoys based on the rank of Tc values were listed for three benchmarking sets in Tables S5 and S6. In those structures, we can observe higher resemblance between ligands and decoys in MUBD-HDACs compared to other two sets. In DUD-E, the most dissimilar compounds based on ECFP4 fingerprints were chosen as decoys, while DEKOIS 2.0 applied a stricter standard to select structurally divergent decoys, i.e. "avoidance of latent actives in the decoy set (LADS) score". Though both methods were valuable to avoid "false negatives", it may artificially increase the ligand enrichment.[45] In contrast, our method of selecting decoys for HDACs considered the degrees of structural similarity of decoys to ligands along with dissimilarity.[57] Theoretically, the more structurally similar of the decoys to the ligands, the more challenging it is to distinguish ligands and decoys. Therefore, the higher structural similarity in MUBD-HDACs could be one possible reason to the lower ligand enrichments in terms of ROC AUCs measured by both docking and FCFP_6 based similarity search (cf. Table 2).

Benchmarking Performance. The ligand enrichment of molecular docking and FCFP_6 based similarity search for the same two targets show varied outcomes with different benchmarking sets (cf. Table 2). For HDAC2, there is no significant difference (ΔROC AUC < 0.05) between two approaches for MUBD-HDACs and DUD-E, while FCFP_6 based similarity search is more efficient than molecular docking for DEKOIS 2.0. For HDAC8, the performances of two approaches are similar based on MUBD-HDACs, while the results based on DUD-E and DEKOIS 2.0 consistently show that the ROC AUC of FCFP_6 based similarity search is larger than molecular docking.

As discussed by Cleves and Jain,[66] a large amount of "obvious analogues" including the NLs in the benchmarking sets make ligand enrichment easier than the real-world screening practice,

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

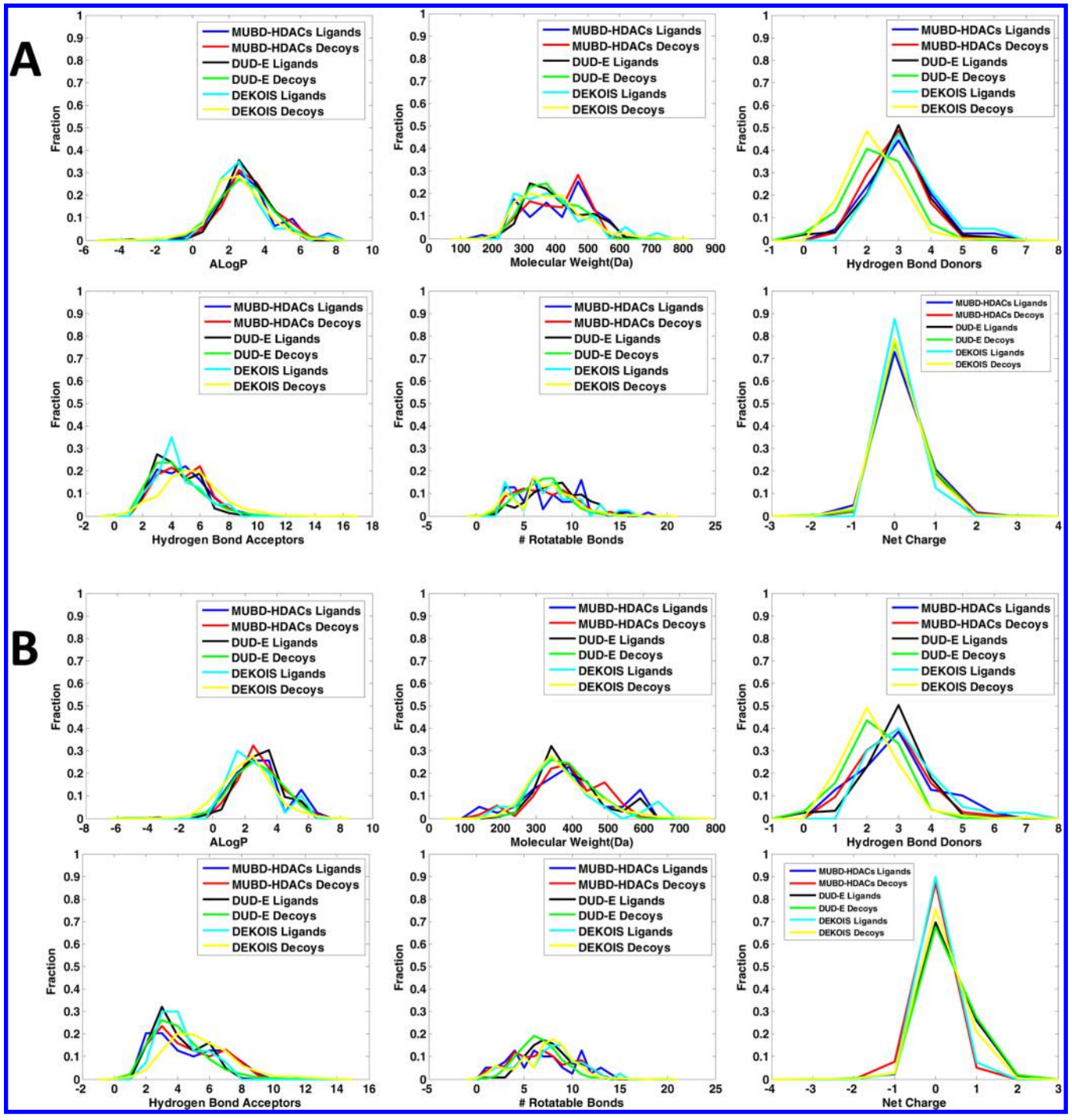


Figure 5. Physicochemical property distributions of ligands and decoys for HDAC2 (**A**) and HDAC8 (**B**) from MUBD-HDACs, DUD-E, and DEKOIS, respectively: (color codes) MUBD-HDACs ligands, blue; MUBD-HDACs decoys, red; DUD-E ligands, black; DUD-E decoys, green; DEKOIS ligands, light blue; DEKOIS decoys, yellow.

which was called the "analogue bias" in general. Because 2D similarity search method is more sensitive to ligand analogues than molecular docking, benchmarking set with obvious analogues brings forth a favorable enrichment bias during method evaluation. Such benchmarking sets are regarded as "2D biased", which could cause "LBVS favorable" outcome. In our study, we considered the high percentage of NLs as one major and obvious factor that causes 2D bias and thus LBVS favorable evaluation outcome. Therefore, we applied NLBScore to measure the percentages of NLs in three benchmarking sets. NLBScore, the average value of NL%, and the corresponding maximum and minimum values of NL% for each benchmarking set are summarized in Table 2. For HDAC2, the NLBScore is smaller for both MUBD-HDACs and DUD-E (0.052 vs 0.091) but larger for DEKOIS 2.0 (0.160). This indicates that DEKOIS 2.0 is more 2D biased than the other two thus the evaluation based on DEKOIS 2.0 is likely to be more LBVS favorable than the others. For the data sets of HDAC8, both NLBScores for DUD-E and DEKOIS 2.0 are larger than that of MUBD-HDACs, which brings up an alert these two data sets

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

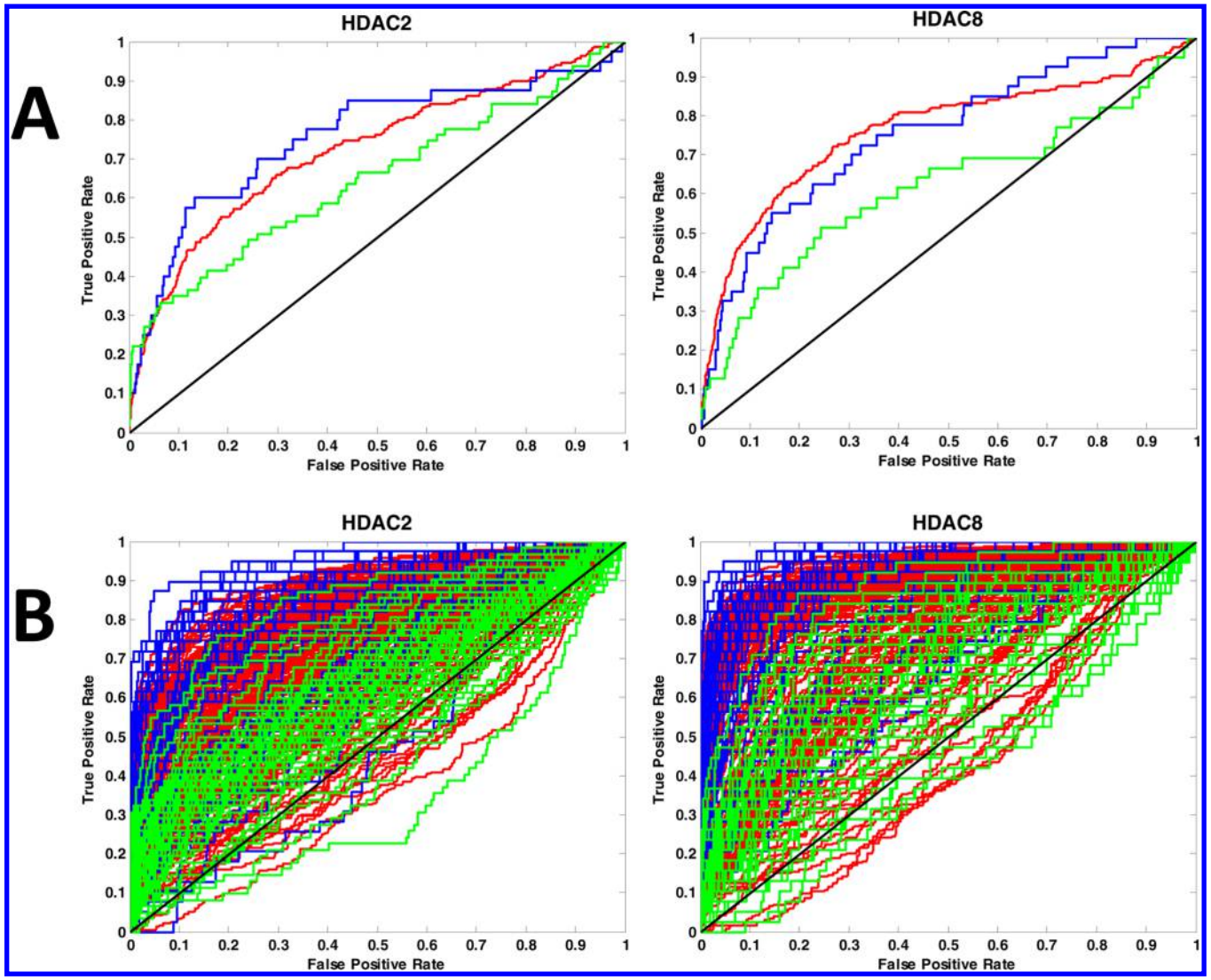


Figure 6. ROC curves from molecular docking by GOLD (**A**) and similarity search using FCFP_6 fingerprints (**B**) for HDAC2 and HDAC8 data sets in MUBD-HDACs, DUD-E, and DEKOIS, respectively: (color codes) MUBD-HDACs, green; DUD-E, red; DEKOIS decoys, blue.

might be LBVS favorable. Consistently, the ligand enrichments based on the benchmarking sets with larger NLBScores show that FCFP_6 based similarity search performs better than molecular docking. For example, for HDAC8/DEKOIS 2.0 the ROC AUCs is as high as 0.903 for FCFP_6 compared to 0.754 for GOLD (cf. Table 2); its corresponding values are 0.831 vs 0.762 for HDAC8/DUD-E and 0.666 vs 0.618 for HDAC8/MUBD-HDACs. Because of the undesired LBVS favorable effect caused by 2D bias in the benchmarking practices, NLBScore should be reduced to as small as possible for our purpose. Since our in-house methodology can ensure small NLBScore in MUBD-HDACs, the evaluation outcome appears to be more accurate and versatile, i.e. similar enrichments for both docking and FCFP_6 based methods, as shown for both targets of HDAC2 and HDAC8. In summary, MUBD-HDACs is an accurate and versatile benchmarking system that can be used for the evaluation of both LBVS and SBVS approaches. In addition, the NLBScore we defined can assist to detect the existence of 2D bias as well as LBVS favorable effect for benchmarking sets in general; the high NLBScore serves as an alert for biased evaluation outcomes.

CONCLUSIONS

In the current study, we have presented exhaustive benchmarking data sets for most isoforms in the superfamily of HDACs including sirtuins (Class III HDACs), i.e. MUBD-HDACs. It is dedicated to assist modern drug discovery for HDACs-associated diseases, e.g. cancers and neurodegenerative diseases. In the data sets, all ligands of HDACIs have been carefully curated, unbiased decoys have been generated by our new in-house method, and multiple crystal structures of HDACs been prepared for immediate use. From our extensive validations, the ligand sets in MUBD-HDACs have been proved to be structurally diverse and the decoy sets are shown as property-matching of ligands. In addition, MUBD-HDACs is deemed to be maximal-unbiased in terms of artificial enrichment and analogue bias, measured by the average value of ROC AUCs from simp-based VS and sims-based VS. The comparative studies with DUD-E and DEKOIS 2.0 sets against HDAC2 and HDAC8 targets show that (1) MUBD-HDACs has comparable chemical diversity to DUD-E and higher diversity than DEKOIS 2.0; (2) the ligand and decoy sets in MUBD-HDACs match well in terms of physicochemical properties, which is similar to DUD-E and better than DEKOIS

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

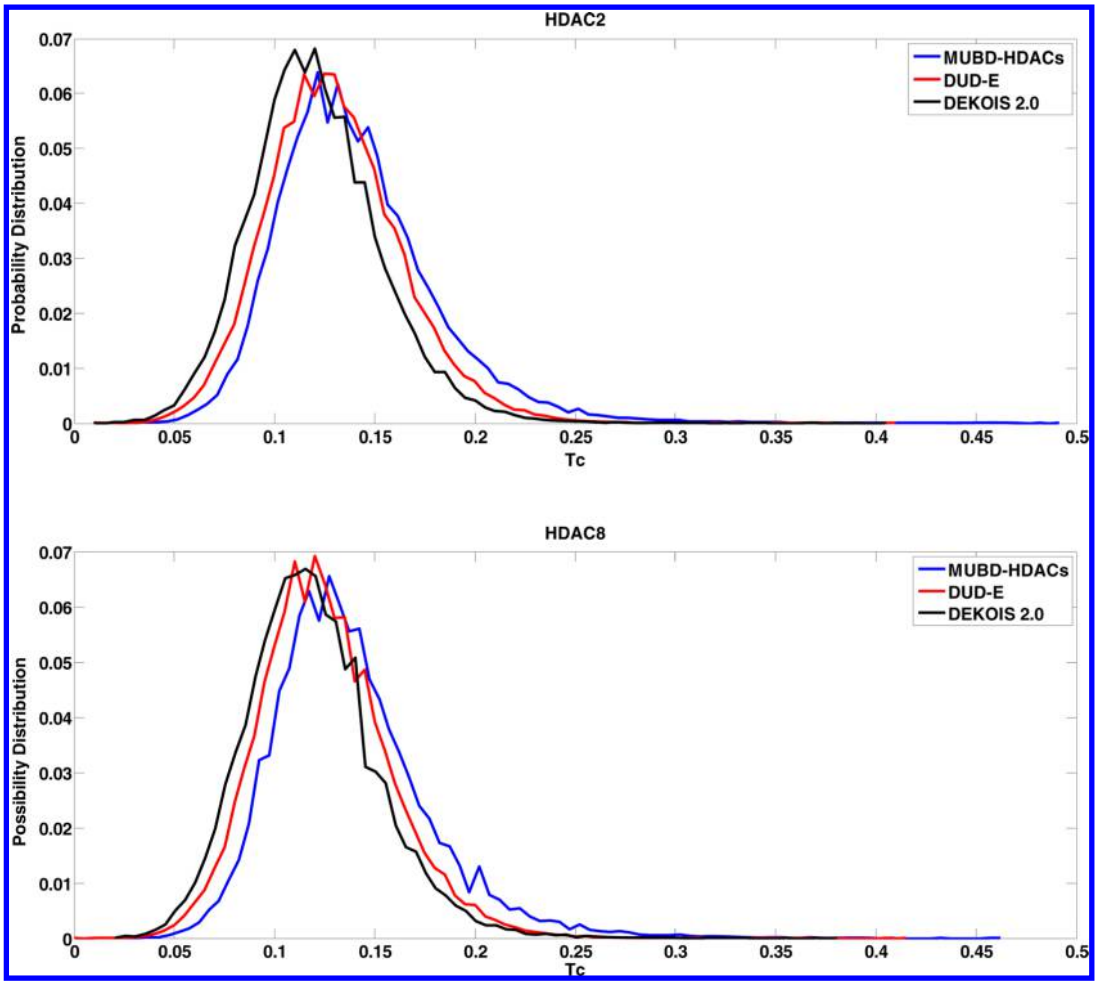


Figure 7. Distribution of topological similarity (Tanimoto coefficient, Tc) based on FCFP_6 fingerprints for HDAC2 and HDAC8 data sets: (color codes) MUBD-HDACs, blue; DUD-E, red; DEKOIS 2.0, black.

2.0; (3) in terms of enrichment performance, MUBD-HDACs seems to be more challenging than DUD-E and DEKOIS 2.0 for both SBVS and LBVS approaches, especially ideal for LBVS benchmarking purpose. Further analysis represented by Tc distribution and ligands/nearest decoys structural comparison indicates that decoys in MUBD-HDACs appear to be more similar to their corresponding ligands than those of DUD-E, while decoys in DEKOIS 2.0 are the most dissimilar to their ligands. We deem such a difference could be a reason for MUBD-HDACs being challenging and fair during benchmarking practice, which is essential to evaluate the merits and weaknesses of various VS methods. In addition, we defined a novel metric, i.e. NLBScore, to detect the 2D bias and LBVS favorable effect within the benchmarking sets and have proved that MUBD-HDACs is maximal-unbiased for benchmarking both LBVS (mainly similarity search) and SBVS approaches. To the best of our knowledge, MUBD-HDACs is the only comprehensive and maximal-unbiased benchmark system for HDACs (including sirtuins) that is available so far. In particular, MUBD-HDACs can be applied for the assessment of both SBVS and LBVS approaches in an unbiased manor. We anticipate that MUBD-HDACs can be versatile benchmarking data sets for the discovery of novel HDACIs.

■ ASSOCIATED CONTENT

Supporting Information

Number of "duplicate" ligands, the activity types/cutoff for each HDACs isoform, the examples of duplicate ligands, the mean(ROC AUCs) values from leave-one-out cross-validation based on simp-based similarity search and MACCS sims-based similarity search (i.e., data for Figure 4), and top five pairs of structurally similar ligands/decoys for HDAC2 and HDAC8 among MUBD-HDACs, DUD-E, and DEKOIS 2.0. This material is available free of charge via the Internet at http://pubs.acs.org. All the data sets of MUBD-HDACs are freely accessible at http://www.xswlab.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: x.simon.wang@gmail.com (X.S.W.).
*E-mail: liangren@bjmu.edu.cn (L.Z.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by District of Columbia Developmental Center for AIDS Research (P30AI087714), National Institutes of Health Administrative Supplements for U.S.-China Biomedical Collaborative Research (5P30AI0877714-02), and the National Institute on Minority

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

Health and Health Disparities of the National Institutes of Health under Award Number G12MD007597. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We are also grateful to the China Scholarship Council (CSC) (201206010076), National Natural Science Foundation of China (NSFC 81373272, 81172917).

■ ABBREVIATIONS USED

HATs, histone acetyltransfereases; HDACs, histone deacetylases; AD, Alzheimer's disease; PD, Parkinson's disease; NAD, nicotinamide adenine dinucleotide; TSA, Trichostatin A; HDACIs, histone deacetylases inhibitors; VS, virtual screening; SBVS, structure-based virtual screening; LBVS, ligand-based virtual screening; DUD, directory of useful decoys; DUD-E, DUD-enhanced; VDS, virtual decoy sets; GPCRs, G protein-coupled receptors; GLL, GPCR ligand library; GDD, GPCR decoy database; NRLiSt BDB, nuclear receptors ligands and structures benchmarking database; REPROVIS-DB, database of reproducible virtual screens; MUV, maximum unbiased validation; DEKOIS, demanding evaluation kits for objective in silico screening; MUBD-HDACs, maximal unbiased benchmarking data sets for histone deacetylases; FCFP_6, function class fingerprints of maximum diameter 6; Tc, Tanimoto coefficient; MW, molecular weight; HBAs, number of hydrogen bond acceptors; HBDs, number of hydrogen bond donors; RBs, number of rotatable bonds; nFC, net formal charge; sims, similarity in structure; simp, similarity in properties; simsdiff, sims difference; PDB, Protein Data Bank; LOO CV, leave-one-out cross-validation; NLs, nearer ligands; NLBScore, NL bias score; LADS, avoidance of latent actives in the decoy set

■ REFERENCES

(1) Tang, H.; Wang, X. S.; Huang, X. P.; Roth, B. L.; Butler, K. V.; Kozikowski, A. P.; Jung, M.; Tropsha, A. Novel Inhibitors of Human Histone Deacetylase (Hdac) Identified by Qsar Modeling of Known Inhibitors, Virtual Screening, and Experimental Validation. *J. Chem. Inf. Model.* **2009**, *49*, 461−476.

(2) Ropero, S.; Esteller, M. The Role of Histone Deacetylases (Hdacs) in Human Cancer. *Mol. Oncol* **2007**, *1*, 19−25.

(3) Yamagoe, S.; Kanno, T.; Kanno, Y.; Sasaki, S.; Siegel, R. M.; Lenardo, M. J.; Humphrey, G.; Wang, Y.; Nakatani, Y.; Howard, B. H.; Ozato, K. Interaction of Histone Acetylases and Deacetylases in Vivo. *Mol. Cell. Biol.* **2003**, *23*, 1025−1033.

(4) Hassig, C. A.; Schreiber, S. L. Nuclear Histone Acetylases and Deacetylases and Transcriptional Regulation: Hats Off to Hdacs. *Curr. Opin. Chem. Biol.* **1997**, *1*, 300−308.

(5) Wolffe, A. P. Histone Deacetylase: A Regulator of Transcription. *Science* **1996**, *272*, 371−372.

(6) Marson, C. M.; Matthews, C. J.; Yiannaki, E.; Atkinson, S. J.; Soden, P. E.; Shukla, L.; Lamadema, N.; Thomas, N. S. Discovery of Potent, Isoform-Selective Inhibitors of Histone Deacetylase Containing Chiral Heterocyclic Capping Groups and a N-(2-Aminophenyl)-Benzamide Binding Unit. *J. Med. Chem.* **2013**, *56*, 6156−6174.

(7) Strahl, B. D.; Allis, C. D. The Language of Covalent Histone Modifications. *Nature* **2000**, *403*, 41−45.

(8) Johnstone, R. W. Histone-Deacetylase Inhibitors: Novel Drugs for the Treatment of Cancer. *Nat. Rev. Drug Discov* **2002**, *1*, 287−299.

(9) Iizuka, M.; Smith, M. M. Functional Consequences of Histone Modifications. *Curr. Opin Genet Dev* **2003**, *13*, 154−160.

(10) Taunton, J.; Hassig, C. A.; Schreiber, S. L. A Mammalian Histone Deacetylase Related to the Yeast Transcriptional Regulator Rpd3p. *Science* **1996**, *272*, 408−411.

(11) Petrella, A.; Fontanella, B.; Carratu, A.; Bizzarro, V.; Rodriquez, M.; Parente, L. Histone Deacetylase Inhibitors in the Treatment of Hematological Malignancies. *Mini Rev. Med. Chem.* **2011**, *11*, 519−527.

(12) Minucci, S.; Pelicci, P. G. Histone Deacetylase Inhibitors and the Promise of Epigenetic (and More) Treatments for Cancer. *Nat. Rev. Cancer* **2006**, *6*, 38−51.

(13) Xu, K.; Dai, X. L.; Huang, H. C.; Jiang, Z. F. Targeting Hdacs: A Promising Therapy for Alzheimer's Disease. *Oxid Med. Cell Longev* **2011**, *2011*, 143269.

(14) Fischer, A. Targeting Histone-Modifications in Alzheimer's Disease. What Is the Evidence That This Is a Promising Therapeutic Avenue? *Neuropharmacology* **2014**, *80*, 95−102.

(15) Gao, L.; Cueto, M. A.; Asselbergs, F.; Atadja, P. Cloning and Functional Characterization of Hdac11, a Novel Member of the Human Histone Deacetylase Family. *J. Biol. Chem.* **2002**, *277*, 25748−25755.

(16) de Ruijter, A. J.; van Gennip, A. H.; Caron, H. N.; Kemp, S.; van Kuilenburg, A. B. Histone Deacetylases (Hdacs): Characterization of the Classical Hdac Family. *Biochem. J.* **2003**, *370*, 737−749.

(17) Zhang, Z.; Yamashita, H.; Toyama, T.; Sugiura, H.; Ando, Y.; Mita, K.; Hamaguchi, M.; Hara, Y.; Kobayashi, S.; Iwase, H. Quantitation of Hdac1Mrna Expression in Invasive Carcinoma of the Breast*. *Breast Cancer Res. Treat* **2005**, *94*, 11−16.

(18) Zhang, Z.; Yamashita, H.; Toyama, T.; Sugiura, H.; Omoto, Y.; Ando, Y.; Mita, K.; Hamaguchi, M.; Hayashi, S.; Iwase, H. Hdac6 Expression Is Correlated with Better Survival in Breast Cancer. *Clin. Cancer Res.* **2004**, *10*, 6962−6968.

(19) Mahlknecht, U.; Hoelzer, D. Histone Acetylation Modifiers in the Pathogenesis of Malignant Disease. *Mol. Med.* **2000**, *6*, 623−644.

(20) Gayther, S. A.; Batley, S. J.; Linger, L.; Bannister, A.; Thorpe, K.; Chin, S. F.; Daigo, Y.; Russell, P.; Wilson, A.; Sowter, H. M.; Delhanty, J. D.; Ponder, B. A.; Kouzarides, T.; Caldas, C. Mutations Truncating the Ep300 Acetylase in Human Cancers. *Nat. Genet.* **2000**, *24*, 300−303.

(21) Marson, C. M. Histone Deacetylase Inhibitors: Design, Structure-Activity Relationships and Therapeutic Implications for Cancer. *Anticancer Agents Med. Chem.* **2009**, *9*, 661−692.

(22) Kelly, W. K.; O'Connor, O. A.; Krug, L. M.; Chiao, J. H.; Heaney, M.; Curley, T.; MacGregore-Cortelli, B.; Tong, W.; Secrist, J. P.; Schwartz, L.; Richardson, S.; Chu, E.; Olgac, S.; Marks, P. A.; Scher, H.; Richon, V. M. Phase I Study of an Oral Histone Deacetylase Inhibitor, Suberoylanilide Hydroxamic Acid, in Patients with Advanced Cancer. *J. Clin Oncol* **2005**, *23*, 3923−3931.

(23) Zhou, N.; Moradei, O.; Raeppel, S.; Leit, S.; Frechette, S.; Gaudette, F.; Paquin, I.; Bernstein, N.; Bouchain, G.; Vaisburg, A.; Jin, Z.; Gillespie, J.; Wang, J.; Fournel, M.; Yan, P. T.; Trachy-Bourget, M. C.; Kalita, A.; Lu, A.; Rahil, J.; MacLeod, A. R.; Li, Z.; Besterman, J. M.; Delorme, D. Discovery of N-(2-Aminophenyl)-4-[(4-Pyridin-3-Ylpyrimidin-2-Ylamino)Methyl]Benzamide (Mgcd0103), an Orally Active Histone Deacetylase Inhibitor. *J. Med. Chem.* **2008**, *51*, 4072−4075.

(24) Fournel, M.; Bonfils, C.; Hou, Y.; Yan, P. T.; Trachy-Bourget, M. C.; Kalita, A.; Liu, J.; Lu, A. H.; Zhou, N. Z.; Robert, M. F.; Gillespie, J.; Wang, J. J.; Ste-Croix, H.; Rahil, J.; Lefebvre, S.; Moradei, O.; Delorme, D.; Macleod, A. R.; Besterman, J. M.; Li, Z. Mgcd0103, a Novel Isotype-Selective Histone Deacetylase Inhibitor, Has Broad Spectrum Antitumor Activity in Vitro and in Vivo. *Mol. Cancer Ther* **2008**, *7*, 759−768.

(25) Saito, A.; Yamashita, T.; Mariko, Y.; Nosaka, Y.; Tsuchiya, K.; Ando, T.; Suzuki, T.; Tsuruo, T.; Nakanishi, O. A Synthetic Inhibitor of Histone Deacetylase, Ms-27−275, with Marked in Vivo Antitumor Activity against Human Tumors. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 4592−4597.

(26) Knipstein, J.; Gore, L. Entinostat for Treatment of Solid Tumors and Hematologic Malignancies. *Expert Opin Investig Drugs* **2011**, *20*, 1455−1467.

(27) Piekarz, R. L.; Robey, R. W.; Zhan, Z.; Kayastha, G.; Sayah, A.; Abdeldaim, A. H.; Torrico, S.; Bates, S. E. T-Cell Lymphoma as a

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

Model for the Use of Histone Deacetylase Inhibitors in Cancer Therapy: Impact of Depsipeptide on Molecular Markers, Therapeutic Targets, and Mechanisms of Resistance. *Blood* **2004**, *103*, 4636−4643.
(28) Chuang, D. M.; Leng, Y.; Marinova, Z.; Kim, H. J.; Chiu, C. T. Multiple Roles of Hdac Inhibition in Neurodegenerative Conditions. *Trends Neurosci* **2009**, *32*, 591−601.
(29) Guan, J. S.; Haggarty, S. J.; Giacometti, E.; Dannenberg, J. H.; Joseph, N.; Gao, J.; Nieland, T. J.; Zhou, Y.; Wang, X.; Mazitschek, R.; Bradner, J. E.; DePinho, R. A.; Jaenisch, R.; Tsai, L. H. Hdac2 Negatively Regulates Memory Formation and Synaptic Plasticity. *Nature* **2009**, *459*, 55−60.
(30) McQuown, S. C.; Barrett, R. M.; Matheos, D. P.; Post, R. J.; Rogge, G. A.; Alenghat, T.; Mullican, S. E.; Jones, S.; Rusche, J. R.; Lazar, M. A.; Wood, M. A. Hdac3 Is a Critical Negative Regulator of Long-Term Memory Formation. *J. Neurosci.* **2011**, *31*, 764−774.
(31) Rivieccio, M. A.; Brochier, C.; Willis, D. E.; Walker, B. A.; D'Annibale, M. A.; McLaughlin, K.; Siddiq, A.; Kozikowski, A. P.; Jaffrey, S. R.; Twiss, J. L.; Ratan, R. R.; Langley, B. Hdac6 Is a Target for Protection and Regeneration Following Injury in the Nervous System. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19599−19604.
(32) Yu, C. W.; Chang, P. T.; Hsin, L. W.; Chern, J. W. Quinazolin-4-One Derivatives as Selective Histone Deacetylase-6 Inhibitors for the Treatment of Alzheimer's Disease. *J. Med. Chem.* **2013**, *56*, 6775−6791.
(33) Price, S.; Bordogna, W.; Bull, R. J.; Clark, D. E.; Crackett, P. H.; Dyke, H. J.; Gill, M.; Harris, N. V.; Gorski, J.; Lloyd, J.; Lockey, P. M.; Mullett, J.; Roach, A. G.; Roussel, F.; White, A. B. Identification and Optimisation of a Series of Substituted 5-(1h-Pyrazol-3-Yl)-Thiophene-2-Hydroxamic Acids as Potent Histone Deacetylase (Hdac) Inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 370−375.
(34) Park, H.; Kim, S.; Kim, Y. E.; Lim, S. J. A Structure-Based Virtual Screening Approach toward the Discovery of Histone Deacetylase Inhibitors: Identification of Promising Zinc-Chelating Groups. *ChemMedChem.* **2010**, *5*, 591−597.
(35) Tervo, A. J.; Kyrylenko, S.; Niskanen, P.; Salminen, A.; Leppanen, J.; Nyronen, T. H.; Jarvinen, T.; Poso, A. An in Silico Approach to Discovering Novel Inhibitors of Human Sirtuin Type 2. *J. Med. Chem.* **2004**, *47*, 6292−6298.
(36) Tervo, A. J.; Suuronen, T.; Kyrylenko, S.; Kuusisto, E.; Kiviranta, P. H.; Salminen, A.; Leppanen, J.; Poso, A. Discovering Inhibitors of Human Sirtuin Type 2: Novel Structural Scaffolds. *J. Med. Chem.* **2006**, *49*, 7239−7241.
(37) Neugebauer, R. C.; Uchiechowska, U.; Meier, R.; Hruby, H.; Valkov, V.; Verdin, E.; Sippl, W.; Jung, M. Structure-Activity Studies on Splitomicin Derivatives as Sirtuin Inhibitors and Computational Prediction of Binding Mode. *J. Med. Chem.* **2008**, *51*, 1203−1213.
(38) Huhtiniemi, T.; Suuronen, T.; Rinne, V. M.; Wittekindt, C.; Lahtela-Kakkonen, M.; Jarho, E.; Wallen, E. A.; Salminen, A.; Poso, A.; Leppanen, J. Oxadiazole-Carbonylaminothioureas as Sirt1 and Sirt2 Inhibitors. *J. Med. Chem.* **2008**, *51*, 4377−4380.
(39) Uciechowska, U.; Schemies, J.; Neugebauer, R. C.; Huda, E. M.; Schmitt, M. L.; Meier, R.; Verdin, E.; Jung, M.; Sippl, W. Thiobarbiturates as Sirtuin Inhibitors: Virtual Screening, Free-Energy Calculations, and Biological Testing. *ChemMedChem.* **2008**, *3*, 1965−1976.
(40) Vadivelan, S.; Sinha, B. N.; Rambabu, G.; Boppana, K.; Jagarlapudi, S. A. Pharmacophore Modeling and Virtual Screening Studies to Design Some Potential Histone Deacetylase Inhibitors as New Leads. *J. Mol. Graph Model* **2008**, *26*, 935−946.
(41) Zhao, L.; Xiang, Y.; Song, J.; Zhang, Z. A Novel Two-Step Qsar Modeling Work Flow to Predict Selectivity and Activity of Hdac Inhibitors. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 929−933.
(42) Kalyaanamoorthy, S.; Chen, Y. P. Energy Based Pharmacophore Mapping of Hdac Inhibitors against Class I Hdac Enzymes. *Biochim. Biophys. Acta* **2013**, *1834*, 317−328.
(43) Salo, H. S.; Laitinen, T.; Poso, A.; Jarho, E.; Lahtela-Kakkonen, M. Identification of Novel Sirt3 Inhibitor Scaffolds by Virtual Screening. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2990−2995.
(44) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49*, 6789−6801.
(45) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (Dud-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582−6594.
(46) Nicholls, A. What Do We Know and When Do We Know It? *J. Comput. Aided Mol. Des* **2008**, *22*, 239−255.
(47) Good, A. C.; Oprea, T. I. Optimization of Camd Techniques 3. Virtual Screening Enrichment Studies: A Help or Hindrance in Tool Selection? *J. Comput. Aided Mol. Des* **2008**, *22*, 169−178.
(48) Irwin, J. J.; Shoichet, B. K. Zinc−a Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45*, 177−182.
(49) Jahn, A.; Hinselmann, G.; Fechner, N.; Zell, A. Optimal Assignment Methods for Ligand-Based Virtual Screening. *J. Cheminform* **2009**, *1*, 14.
(50) Wallach, I.; Lilien, R. Virtual Decoy Sets for Molecular Docking Benchmarks. *J. Chem. Inf. Model.* **2011**, *51*, 196−202.
(51) Vogel, S. M.; Bauer, M. R.; Boeckler, F. M. Dekois: Demanding Evaluation Kits for Objective in Silico Screening−a Versatile Tool for Benchmarking Docking Programs and Scoring Functions. *J. Chem. Inf. Model.* **2011**, *51*, 2650−2665.
(52) Bauer, M. R.; Ibrahim, T. M.; Vogel, S. M.; Boeckler, F. M. Evaluation and Optimization of Virtual Screening Workflows with Dekois 2.0−a Public Library of Challenging Docking Benchmark Sets. *J. Chem. Inf. Model.* **2013**, *53*, 1447−1462.
(53) Gatica, E. A.; Cavasotto, C. N. Ligand and Decoy Sets for Docking to G Protein-Coupled Receptors. *J. Chem. Inf. Model.* **2012**, *52*, 1−6.
(54) Lagarde, N.; Ben Nasr, N.; Jeremie, A.; Guillemain, H.; Laville, V.; Labib, T.; Zagury, J. F.; Montes, M. Nrlist Bdb, the Manually Curated Nuclear Receptors Ligands and Structures Benchmarking Database. *J. Med. Chem.* **2014**, *57*, 3117−3125.
(55) Cereto-Massague, A.; Guasch, L.; Valls, C.; Mulero, M.; Pujadas, G.; Garcia-Vallve, S. Decoyfinder: An Easy-to-Use Python Gui Application for Building Target-Specific Decoy Sets. *Bioinformatics* **2012**, *28*, 1661−1662.
(56) Rohrer, S. G.; Baumann, K. Maximum Unbiased Validation (Muv) Data Sets for Virtual Screening Based on Pubchem Bioactivity Data. *J. Chem. Inf. Model.* **2009**, *49*, 169−184.
(57) Xia, J.; Jin, H.; Liu, Z.; Zhang, L.; Wang, X. S. An Unbiased Method to Build Benchmarking Sets for Ligand-Based Virtual Screening and Its Application to Gpcrs. *J. Chem. Inf. Model.* **2014**, *54*, 1433−1450.
(58) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. Chembl: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2012**, *40*, D1100−1107.
(59) *Maccs Structural Keys*; Mdl Information Systems Inc.: San Ramon, Ca, 2005.
(60) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. Zinc: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52*, 1757−1768.
(61) Apweiler, R.; Bairoch, A.; Wu, C. H.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; Martin, M. J.; Natale, D. A.; O'Donovan, C.; Redaschi, N.; Yeh, L. S. Uniprot: The Universal Protein Knowledgebase. *Nucleic Acids Res.* **2004**, *32*, D115−119.
(62) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G.; Clustal, W.; Clustal, X. Version 2.0. *Bioinformatics* **2007**, *23*, 2947−2948.
(63) *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing, Vienna, Austria, 2008.
(64) Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **1996**, *39*, 2887−2893.
(65) Tanimoto, T. *Ibm Internal Report*; Ibm Corp: Armonk, NY, 1957.

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388

(66) Cleves, A. E.; Jain, A. N. Effects of Inductive Bias on Computational Evaluations of Ligand-Based Modeling and on Drug Discovery. *J. Comput. Aided Mol. Des* **2008**, *22*, 147−159.
(67) Dokmanovic, M.; Clarke, C.; Marks, P. A. Histone Deacetylase Inhibitors: Overview and Perspectives. *Mol. Cancer Res.* **2007**, *5*, 981−989.
(68) https://www.ebi.ac.uk/chembl/faq#faq24 (accessed Oct 2014).

DOI: 10.1021/ci5005515
J. Chem. Inf. Model. 2015, 55, 374−388