JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Comparative Analysis of Pharmacophore Screening Tools

Marijn P. A. Sanders,[†,#] Arménio J. M. Barbosa,[‡] Barbara Zarzycka,[§] Gerry A.F. Nicolaes,[§] Jan P.G. Klomp,[||] Jacob de Vlieg,[†,⊥] and Alberto Del Rio*[,‡]

[†]Computational Drug Discovery Group, CMBI, Radboud University Nijmegen, Geert Grooteplein Zuid 26-28, 6525 GA, Nijmegen, The Netherlands

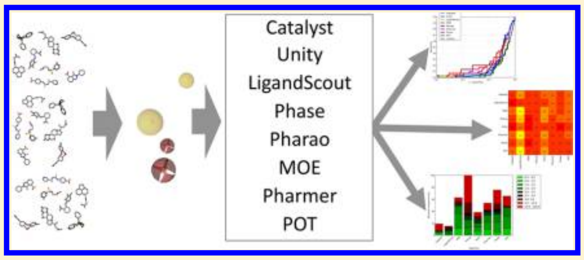
[‡]Department of Experimental Pathology, Alma Mater Studiorum, University of Bologna, Via S. Giacomo 14, 40126 Bologna, Italy

[§]Department of Biochemistry, Cardiovascular Research Institute Maastricht, Maastricht University, Universiteitssingel 50, 6229 ER, The Netherlands

[||]Lead Pharma Medicine, Kapittelweg 29, 6525 EN, Nijmegen, The Netherlands

[⊥]Netherlands eScience Center, Science Park 140, 1098XG, Amsterdam, The Netherlands

Ⓢ *Supporting Information*



ABSTRACT: The pharmacophore concept is of central importance in computer-aided drug design (CADD) mainly because of its successful application in medicinal chemistry and, in particular, high-throughput virtual screening (HTVS). The simplicity of the pharmacophore definition enables the complexity of molecular interactions between ligand and receptor to be reduced to a handful set of features. With many pharmacophore screening softwares available, it is of the utmost interest to explore the behavior of these tools when applied to different biological systems. In this work, we present a comparative analysis of eight pharmacophore screening algorithms (Catalyst, Unity, LigandScout, Phase, Pharao, MOE, Pharmer, and POT) for their use in typical HTVS campaigns against four different biological targets by using default settings. The results herein presented show how the performance of each pharmacophore screening tool might be specifically related to factors such as the characteristics of the binding pocket, the use of specific pharmacophore features, and the use of these techniques in specific steps/contexts of the drug discovery pipeline. Algorithms with rmsd-based scoring functions are able to predict more compound poses correctly as overlay-based scoring functions. However, the ratio of correctly predicted compound poses versus incorrectly predicted poses is better for overlay-based scoring functions that also ensure better performances in compound library enrichments. While the ensemble of these observations can be used to choose the most appropriate class of algorithm for specific virtual screening projects, we remarked that pharmacophore algorithms are often equally good, and in this respect, we also analyzed how pharmacophore algorithms can be combined together in order to increase the success of hit compound identification. This study provides a valuable benchmark set for further developments in the field of pharmacophore search algorithms, e.g., by using pose predictions and compound library enrichment criteria.

■ INTRODUCTION

In the field of drug design, high-throughput virtual screening (HTVS) methods encompass a valuable set of computational approaches for the analysis of large chemical structure libraries with the purpose of identifying hit compounds capable of interacting with a biological target of interest.[1] While over the last few decades combinatorial chemistry and high-throughput screening (HTS) represented an important step in drug discovery to accumulate large amount of data, the global importance of in silico techniques is vice versa ascribable to reduced costs and the increased time-efficiency to unveil new potential active compounds.[2,3] Among all computational approaches that can help guide drug discovery, the so-called structure-based (SB) design approaches, which use the three-dimensional information of the biological target, are among the most popular.[4−6] However, despite the existence of large numbers of apparently different computational approaches, recent studies emphasize the usage of simple and already established techniques for the successful disclosure of important information toward the selection of novel bioactive compounds.[7,8] In this context, some seemingly old concepts, such as that of the pharmacophore, have proven to be extremely useful over the past 30 years, and it is surprising that many of these concepts are regaining momentum.[9−11] For pharmacophore modeling, in particular, this *renaissance* has also been fostered by the current possibility to generate hypotheses directly from crystallographic, NMR, or computational models of protein−ligand complexes.[9,10,12] Together with the fact that nowadays three-dimensional structures of biological receptors and enzymes become available much more frequently than in the past, one would theoretically need few steps in order to

Received: November 4, 2011
Published: May 30, 2012

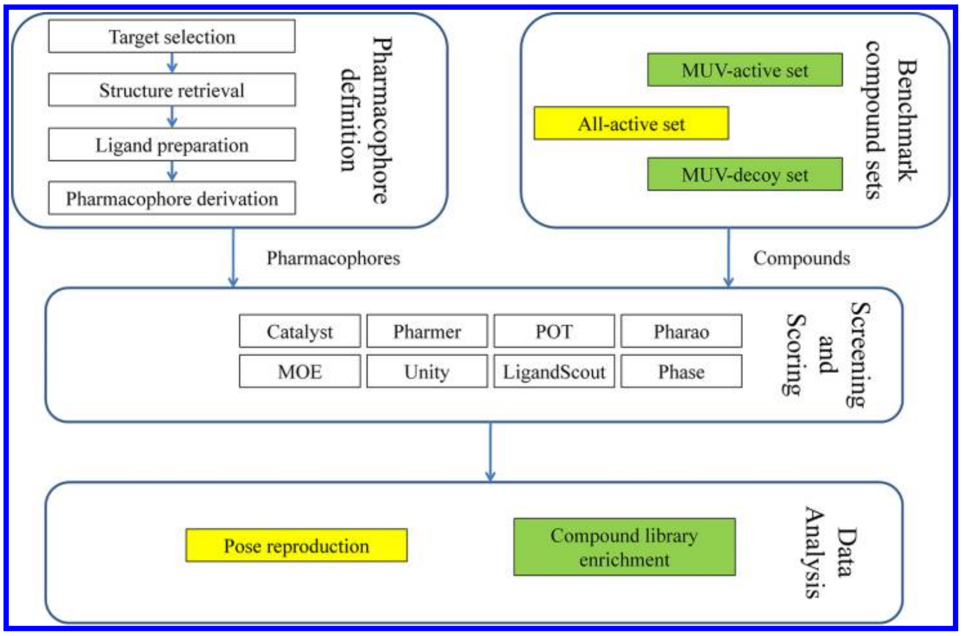


Figure 1. Flow diagram of the data set preparation protocol and computational procedures. Yellow boxes indicate procedures in which all cocrystallized ligands are used. Green boxes indicate procedures where the unbiased active and decoy set is used.

rapidly setup pharmacophore screening campaigns toward the selection of novel molecular entities for biological testing and/or lead optimization purposes.

Besides the increased availability of structural information on pharmacological targets, advances in computing power and improvements in pharmacophore screenings algorithms are also further stimulated by the increasing number of academic services and chemical vendors that offer large databases of commercially available compounds and virtual libraries for this purpose.[13−17] Some of these databases allow the circumvention of typical preparation steps such as hydrogen addition, tautomer and stereoisomers enumeration, and, most importantly, conformer generation.[17,18] From the practical point of view, pharmacophores can be used to screen millions of high quality compound structures within a reasonable amount of time, particularly when approximations such as rigid pharmacophore fitting procedures are used. Like molecular docking, pharmacophore search algorithms should not only discriminate between active and inactive compounds but should also correctly orient the ligand in the protein-binding region.[19,20]

While in the last year many studies focusing on the reproduction of binding modes and compound library enrichment have been published to assess docking screening algorithms,[5,21,22] the pharmacophore concept, introduced above, has been the object of few comparative studies assessing the performance of pharmacophore screening tools.[9,12,20,23,24]

Here, we present an assessment of eight free and commercial software packages for pharmacophore screening starting with the hypothesis that default/typical settings are used, which is often the case when campaigns on new uncharacterized biological targets are pursued. In order to ensure an unbiased comparison of the screening algorithms, we have chosen to manually curate pharmacophores extracted from X-ray structures and to perform the virtual screens on rigid compound structures of precalculated conformers using default software settings. Four biological targets were analyzed, for which the locations of a large number of ligands elucidated by X-ray crystallography were collected from the literature, and were used to derive structure-based pharmacophores and evaluate pharmacophore screening performance.

The accuracy of the predicted binding modes as well as library enrichments for the different pharmacophore search algorithms was analyzed to elucidate how different factors, e.g., pharmacophore hypotheses and conformational states, may influence the outcome of high-throughput screenings.

■ METHODS

Figure 1 shows the computational protocol that was applied to each biological target. Full details of each step are discussed in the next paragraphs.

Pharmacophore Definition. *Target Selection.* Four data sets corresponding to different biological targets were used in this study, namely, CDK2 (Cyclin-dependent kinase 2), Chk-1 (Checkpoint kinase 1), PTP-1B (Protein tyrosine phosphatase 1B), and Urokinase. The last three protein targets were chosen in order to take advantage of published data from Brown et al.,[25] while the CDK2 data set was created from crystal structures retrieved from the Protein Data Bank.[26] All the biological targets have been considered for their important roles in different biological processes. CDK2 is a well-known protein kinase involved in the control of the cell cycle.[27] Chk-1 is a kinase required for checkpoint-mediated cell cycle arrest in response to DNA damage or the presence of unreplicated DNA.[28] PTP1B is a regulator involved in insulin signaling and has been implicated as a potential therapeutic target for treatment of type II diabetes.[29] Urokinase is a serine protease that circulates in plasma and has been implicated in a number of tumor-related activities.[30]

Structure Retrieval. Ligand coordinates for PTP1B, CHK1, and Urokinase data sets were obtained from Brown et al.[25] as being aligned with the PDB IDs 1NZ7, 2YWP, and 1OWK, respectively. CDK2 was selected because of the high number of complexes with different cocrystallized ligands available for the biological target in the Protein Data Bank.[26,27] In this case, we collected reference X-ray structures from the PDB (accessed

Table 1. Overview of Data Sets Included in the Present Study[a]

Target	All-actives	All-actives Conformations	MUV-decoys	MUV-decoys Conformations	MUV-actives	MUV-actives Conformations
CDK2	80	992	15000	248250	30	352
CHK1	123	1913	15000	287203	30	457
PTP1B	110	4634	15000	405398	30	1123
Urokinase	75	703	15000	268646	30	192

[a] All files in single- and multi-conformation are available in Supporting Information.

November 2010)[26] by using the human CDK2 Uniprot[31], Accession ID: P24941. The collected CDK2 complexes were filtered in order to retrieve proteins with bound ligands and no modified residues resulting in a set of 107 CDK2 complexes. These complexes were visually inspected and discarded in cases where ligand atoms were missing or multiple ligand conformations were present in a single PDB entry. Duplicate entries were removed that resulted in a nonredundant set of 80 complexes (the full list is available in the Supporting Information). Next, all complexes were superimposed on the basis of the protein backbones with the Biopolymer tool *Fit Monomers*, from SYBYL-X.[32] Finally, the ligands were extracted from the aligned complexes.

Ligand Preparation. After manual correction of ligands for all targets (e.g., bond orders, aromaticity, and charges), Epik software[33] was used to precalculate the possible physiological protonation states. All ligands were subsequently visually inspected in the original protein PDB structure so as to assign correct protonation and tautomeric forms consistently with ligand-protein binding interactions.

Pharmacophore Derivation. Ligand structures were given appropriate atom-types using an in-house rule-based classification system developed at Organon NV. Pharmacophore features were generated by application of Renner's fuzzy pharmacophore algorithm with an Rc-value of 2.0 Å[34] for all available atom types (hydrophobic, donor, acceptor, negative ionizable, and positive ionizable). Selections of combinations of features to comprise a pharmacophore for screening were based on the visual inspection of all ligand binding modes with the help of literature descriptions. In particular, manual inspection and combination of features shared by most ligands resulted in the final set of features that are depicted in the Results and Discussion section and in the Supporting Information (Figure 1S-4S). In case multiple binding modes were found, i.e., for the PTP1B data set, we defined additional pharmacophore hypotheses.

It should be noted that all pharmacophore hypotheses consisted of standard features, i.e., donor, acceptor, hydrophobic, and excluded volumes, that are contemplated by any software we considered in this study. Pharmacophore files were converted to all formats by using some of the software described and custom in-house scripts. In case multiple pharmacophore schemes were available, e.g., for MOE, the least restrictive definition of donor, acceptor, and hydrophobic features was chosen.

Volume features were added by an iterative procedure and manually reviewed. Starting with the addition of an excluded volume feature at the position of the closest atom in the protein to a cocrystallized ligand atom, the algorithm continues to add excluded volumes until all protein atoms are considered. This addition is recursively performed for atom distances between 3.0−6.0 Å from the cocrystallized ligands and are at least 1.0 Å from the nearest excluded volume feature. In contrast to most algorithms, which use the atom center to evaluate if a pose is inside the excluded volume, Phase software rejects ligand poses with an overlap of the ligand VDW radius with the excluded volumes. To be more consistent with the other algorithms, we therefore reduced the excluded volume radii of Phase pharmacophores with 1.7 Å (approximately the VDW radius of a carbon atom).

Benchmark Compound Sets. *All-Actives Set.* The compound sets comprising all active compounds of each respective target (Table 1) were used to assess how well each individual algorithm performs in the reproduction of the crystal structure pose by means of the rmsd calculation (Results and Discussion section).

Maximum Unbiased Validation Sets of Active and Decoys. In order to assess the compound library enrichment of the different pharmacophore screens, data sets of actives and decoys were designed to avoid analogue bias (overrepresentation of certain scaffolds or chemical entities) and artificial enrichment (classification caused by differences in simple physicochemical descriptors like molecular weight, number of bonds and acceptors, and donors, rather than correct representation of the protein−ligand interactions). First, for each target, a set of assumed inactive compounds was prepared from the CoCoCo database[14,35] by selecting compounds with a BCI−Tanimoto fingerprint similarity[36] of 0.5 or less to at least one ligand reported in the ChEMBL[37,38] for that particular target. Second, the maximum unbiased validation (MUV) sets protocol[39] was used to select 30 actives and 15000 decoys (consistent with the set sizes available on the MUV Web site[40]) per biological target from each subset of the CoCoCo database. This protocol ensures an unbiased validation set by maximizing "active−active distances" $G(\mathrm{t})$ and "active-decoy distances" $F(\mathrm{t})$ (Supporting Information). Numerical integration of both distribution functions enables computation of global figures for data set self-similarity ($\sum G$) and the separation between active and decoys ($\sum F$). A parameter describing the data set clumping $\mathrm{S(t)}$ and its numerical integral $\sum S$ can be calculated by subtraction of $\mathrm{G(t)}$ from $F(\mathrm{t})$. Negative values of $\sum S$ indicate clumping of actives, while positive values indicate dispersion of actives and clumping of small clusters of decoys with single active compounds, and values near zero indicate a spatially random distribution of actives and decoys.[39] The data sets generated with the MUV protocol will be used for the compound library enrichment studies (Results and Discussion section) and will be referred in the text as MUV-active or MUV-decoys. The output of the MUV algorithm is provided in Table 1 and Table 1S of the Supporting Information.

For all compound data sets (All-actives, MUV-actives, and MUV-decoys) of the four biological targets, three-dimensional conformations were generated with the Confgen software by application of the *comprehensive* mode that insures the good accuracy to reproduce experimental bioactive conformations taken from the Protein Data Bank.[26,41] This setting employs extended criteria to search steps per rotatable bond, collect total number of conformers, explore a wider energy window,

Table 2. List of Pharmacophore Screening Algorithms Used in This Study

Tool	Version	Scoring algorithm[a]	Scoring method[b]	Best[c]	Tool availability	Reference
Catalyst	Discovery Studio v2.5.5.9350	FitValue[d]	Overlay	High	Commercial	44
Pharmer	–	rmsd	rmsd	Low	Open-source	45
POT	–	rmsd	rmsd	Low	Open-source[e]	46
Pharao	3.0.3	Tversky_ref[f]	Overlay[f]	High	Open-source	42
MOE	2010.10 (date)	rmsd	rmsd	Low	Commercial	47
Unity	Sybyl X1.1.1	QFIT[g]	Overlay	High	Commercial	32
LigandScout	3.02	Pharmacophore Fit[h]	Overlay	High	Commercial	48
Phase	3.3	Fitness[i]	rmsd[i]	High	Commercial	49

[a] Scoring algorithm refers to the tool routine that was used in the present study to score ligand poses. [b] Scoring method refers to the methodology class of scoring algorithm. The rmsd-based methods check the distance of the feature group of the compound to the pharmacophore feature center. Overlay-based methods take the radii of the features and/or atoms into account and use this to assess how well a feature is matched. [c] Best refers to the score value (high/low) that is used to select the best match. [d] FitValue evaluates for each molecular feature the distance to the center of the pharmacophore feature with respect to the pharmacophore feature radius. [e] Will be made available soon. [f] Tversky_ref evaluates the volume-overlap of features of the pharmacophore and compound with respect to the volume of the pharmacophore. Pharao uses Gaussian overlaps of pharmacophore features and ligand atoms and does not require that all features are matched. Poses with a lower number of matched features are able to obtain as high scores as poses with more matched features. [g] QFIT is intended to compare alternate mappings of a single compound to the query and choose the best match. While Unity reports hits as soon as they fulfill the query, post-processing options to relax and tighten hits to more closely match the query are available within Sybyl package. [h] Pharmacophore Fit is a simple geometric scoring function that takes into account only chemical feature overlap. Other scoring functions are available within LigandScout package. [i] Fitness score of Phase is based on an rmsd term, vector term and a term describing the overlay of the produced pose with a reference pose. No vector features and no reference pose were considered in the present study, thus the resulting score is purely rmsd-based.

perform effective duplicate conformer elimination, obtain accurate ring sampling, and minimize the final structures.

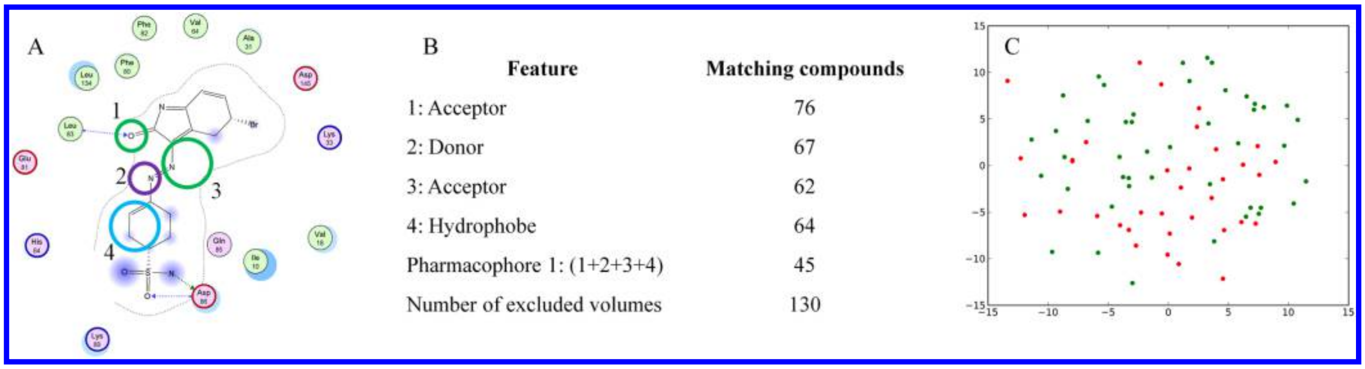
Screening and Scoring. *Screening.* Screens were performed using eight different software tools as shown in Table 2 below. To mimic the scenario in which new biological target are studied or nonexperts apply pharmacophore screens and in order to avoid artificial bias toward certain algorithms, we ran all algorithms with default parameters with the exception of Pharao and Pharmer. For Pharao, we used the *tversky_ref* score instead of the Tanimoto score as the reference article indicated that this score was most suitable for scenarios where it is important that as many features of the pharmacophore are matched as possible.[42] Because Pharmer cannot handle excluded volume features, we postprocessed these poses with POT with use of only the excluded volume definition and without a fitting. Because virtual screens were performed in different laboratories and with different hardware and software systems, we decided to focus entirely on the quality of the produced results of the different algorithms and disregard the CPU-timing that is required for the presented pharmacophore searches.

Scoring. Internal molecular symmetries were considered by calculation of the Root Mean Square Deviations (rmsd) of all possible structural matches of fitted conformations and corresponding cocrystallized reference poses using RDkit[43] functionality. The lowest rmsd value, corresponding to the best fit, was reported for each pose and used for all further analysis. Receiver Operator Characteristic (ROC) curves, Area Under the ROC Curves (AUC), and enrichment factors were calculated after ranking compounds from the MUV-active and MUV-decoys set on the basis of the score values as reported in Table 2. Enrichment factors (EF) after x% of the library screened were calculated according to the following formula ($N_{\text{experimental}}$ = number of experimentally found active structures in the top x% of the sorted database, N_{expected} = number of expected active structures, and N_{active} = total number of active structures in database).

$$\text{EF} = \frac{N^{x\%}_{\text{experimental}}}{N^{x\%}_{\text{expected}}} = \frac{N^{x\%}_{\text{exprimental}}}{N_{\text{active}\cdot x\%}} \qquad (1)$$

For each software package, we considered the methodology class of the scoring algorithm (Table 2). We divided scoring functions in two classes, i.e., rmsd-based and overlay-based methods. The rmsd-based methods check the distance of the feature group of the compound to the pharmacophore feature center. The overlay-based methods take the radii of the features and/or atoms into account and use this to assess how well a feature is matched. While these two concepts are related, as discussed in the Results and Discussion section, they do not entail same results. For instance, rmsd scoring does not take the radii of the features into account and does not apply a weighting of the different features (e.g., by considering the radii of both pharmacophore features and ligand features). The reduction of ligand and pharmacophore features to single points is one of the factors leading to different results in a typical compound library enrichment study despite the fact that a low rmsd is likely to have a good overlap. This classification of scoring functions will be conveniently used to discuss the various pharmacophore screening algorithms (see following).

Data Analysis. *Pose Reproduction.* Pharmacophoric poses were generated by application of the eight screening algorithms to the four data sets, and the results were analyzed with respect to the accuracy of experimental binding mode reproduction and compound library enrichments. For each molecule, both the pose with the lowest rmsd to the reference structure and the rmsd of the pose with the best score were reported. The cumulative percentage of poses below a certain rmsd was calculated for both X-ray structures and the multiconformational data sets of actives previously generated. For CHK1, PTP1B and Urokinase multiple pharmacophores are defined that recognize different subsets of active molecules (Results and Discussion section). To evaluate the combined performance for CHK1, PTP1B and Urokinase we also merged the outcomes of the individual pharmacophore searches for these targets. The success rate is reported at rmsd thresholds for both the top scoring and best-predicted poses according to recommenda-



Feature	Matching compounds
1: Acceptor	76
2: Donor	67
3: Acceptor	62
4: Hydrophobe	64
Pharmacophore 1: (1+2+3+4)	45
Number of excluded volumes	130

Figure 2. CDK2 data set. (a) Pharmacophore depiction as used in this study on top of PDB entry: 1FVT (note that 1FVT with its cocrystallized ligand is used as a reference and does not contain the donor feature that is present in most ligands cocrystallized with CDK2). (b) List of pharmacophore features with corresponding matching compounds in the set of actives. (c) Two-dimensional illustration of active compound similarities created using stochastic proximity embedding (SPE) with euclidean distances of BCI fingerprints.[59−61] Green dots represent compounds that match the pharmacophore according to the observed ligand alignment in the crystal structures; red dots are the compounds that do not match the pharmacophore.

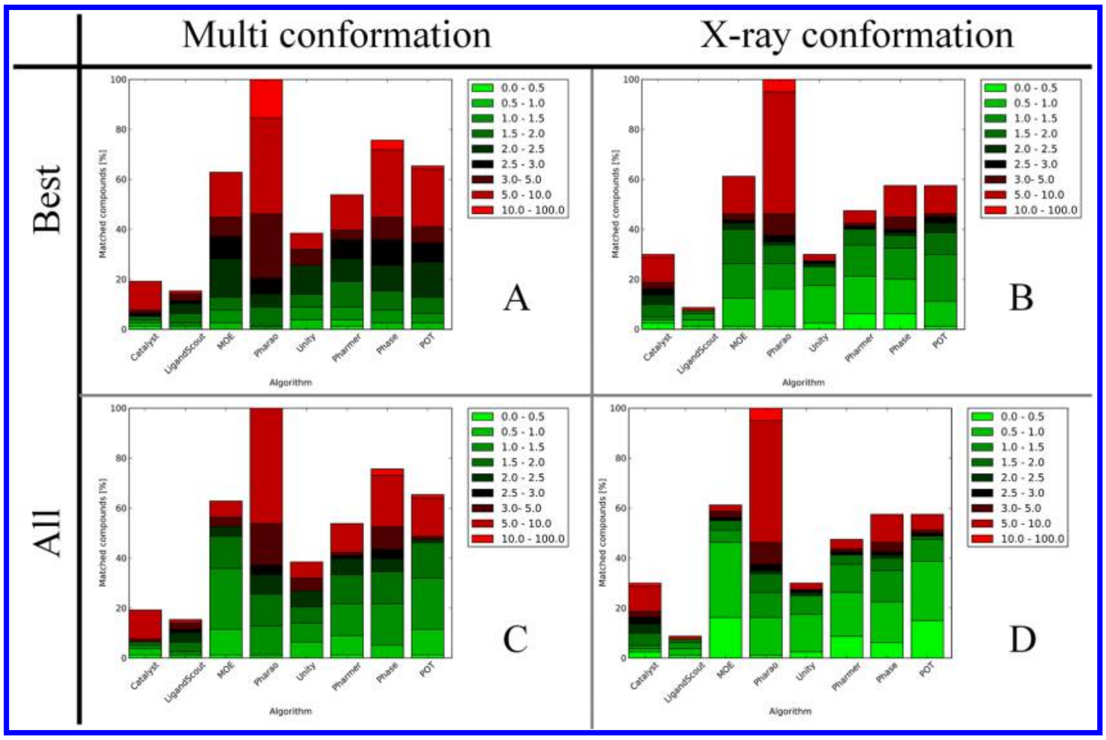
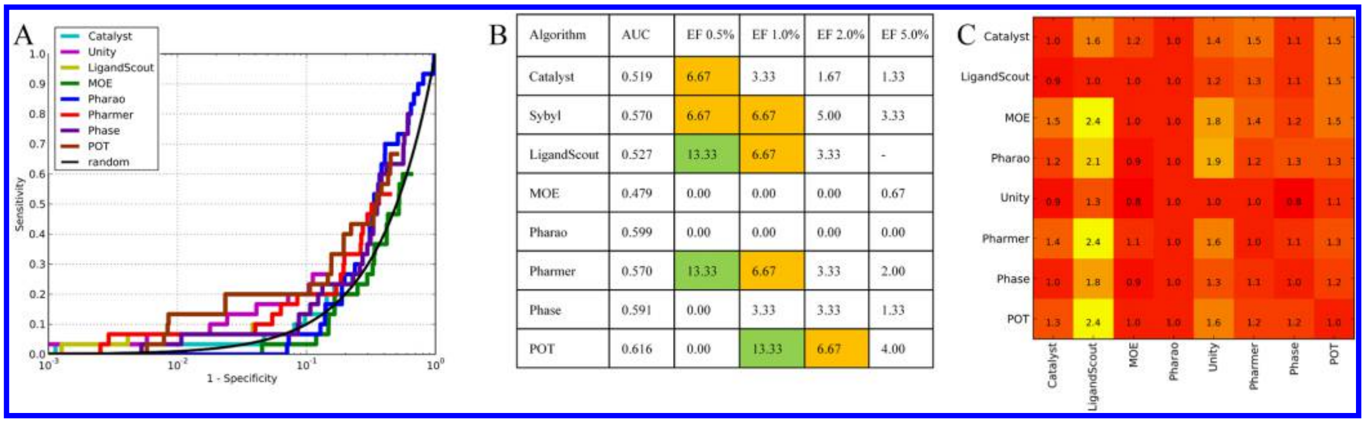


Figure 3. Rmsd ranges of matched compounds from the cocrystallized ligand in four different scenarios. (A) Best ranked pose from the ligand set in their multiconformational format. (B) Best ranked pose from the ligand set in X-ray conformation. (C) Lowest rmsd from all poses from the ligand set in multiconformational format. (D) Lowest rmsd from all poses in the ligand set in their X-ray conformation.

tions for reporting results on pose predictions.[50] In order to ensure that the conformational sampling did not significantly impact on the accuracy of the binding mode reproduction, we performed a conformational analysis that is included in the Supporting Information (Figure 5S and discussion)[51,52] highlighting that conformational sampling is accurate for all data sets because the experimentally determined ligand conformation is reproduced within 1.5 Å for the majority of compounds of all data sets.

Compound Library Enrichment. The performance of each algorithm with respect to compound library enrichment, i.e., the percentage of retrieved actives (sensitivity) versus the percentage of retrieved decoys (specificity), was calculated using the ranking of MUV data sets (actives and decoys) as deduced from their respective score values and visualized in receiver operator curves (ROC). The corresponding area under the ROC curves (AUC) and enrichment values at 0.5%, 1.0%, 2.0%, and 5.0% are calculated and reported.

To assess the cooperative behavior of different algorithms, we calculated the improvement in enrichment factors for each pharmacophore algorithm when combined with another algorithm. This is calculated as the enrichment of the set of compounds matching algorithm XY (Y after X) divided by the enrichment of the set of compounds that match algorithm Y. Such analysis can be useful not only to identify the combination of algorithms that results in the best enrichment but also might suggest how fast algorithms can be used as a prefilter for the slower, but more accurate, algorithms. The heatmaps generated from this analysis (Results and Discussion section) do not show the necessarily algorithms that have the best enrichment but



Algorithm	AUC	EF 0.5%	EF 1.0%	EF 2.0%	EF 5.0%
Catalyst	0.519	6.67	3.33	1.67	1.33
Sybyl	0.570	6.67	6.67	5.00	3.33
LigandScout	0.527	13.33	6.67	3.33	-
MOE	0.479	0.00	0.00	0.00	0.67
Pharao	0.599	0.00	0.00	0.00	0.00
Pharmer	0.570	13.33	6.67	3.33	2.00
Phase	0.591	0.00	3.33	3.33	1.33
POT	0.616	0.00	13.33	6.67	4.00

	Catalyst	LigandScout	MOE	Pharao	Unity	Pharmer	Phase	POT
Catalyst	1.0	1.6	1.2	1.0	1.4	1.5	1.1	1.5
LigandScout	0.9	1.0	1.0	1.0	1.2	1.3	1.1	1.5
MOE	1.5	2.4	1.0	1.0	1.8	1.4	1.2	1.5
Pharao	1.2	2.1	0.9	1.0	1.9	1.2	1.3	1.3
Unity	0.9	1.3	0.8	1.0	1.0	1.0	0.8	1.1
Pharmer	1.4	2.4	1.1	1.0	1.6	1.0	1.1	1.3
Phase	1.0	1.8	0.9	1.0	1.3	1.1	1.0	1.2
POT	1.3	2.4	1.0	1.0	1.6	1.2	1.2	1.0

Figure 4. Enrichment analysis of CDK2 MUV-data set. (A) ROC curves showing the enrichment of CDK2 actives/decoys (data set created with MUV, Table 1). (B) AUC and enrichment values at 0.5%, 1.0%, 2.0%, and 5.0% false positive rate; green indicates an enrichment factor (EF) above 10.0 and orange indicates EF above 5.0. (C) Heatmap showing the improvement in enrichment factor of the algorithms on the *Y*-axis if prescreening with the algorithm on the *X*-axis is performed. Values greater than 1.0 show that the reapplication of the algorithms in the *X*-axis improves the results of algorithms in the *Y*-axis. Conversely, values smaller than 1.0 show that the reapplication of a second algorithm worsens the overall results. A value of 1.0 denotes no influence of the second algorithm.

illustrate the gain in performance achieved by prescreening with another algorithm. In particular, values below 1.0 indicate that prescreening with another algorithm will result in a worse enrichment, while values above 1.0 indicate that prescreening is beneficial for the final enrichment value.

■ RESULTS AND DISCUSSION

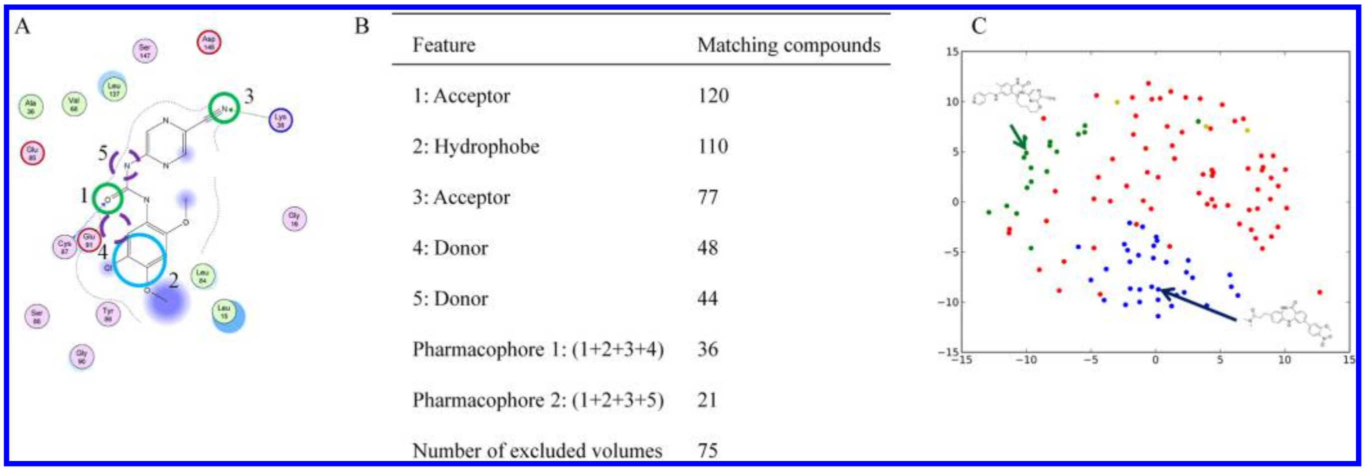
CDK2 Data Set. *Pharmacophore Perception.* CDK2 is a protein kinase whose pharmacophore features, delineating ligands that target the ATP-binding site, are well described in the literature.[53] Ligand sites typically include a hydrogen bond donor (HBD) and a hydrogen bond acceptor (HBA) representing a pair of key intermolecular interaction occurring with the *hinge* backbone (feature 1 and 2, Figure 2a). The importance of these features is exemplified by the fact that almost all the active compounds match those features (Figure 2B). A second HBA feature is common to most of the active compounds and represents the interaction with the *gatekeeper* residue Glu81 or for bridging water molecules with catalytic Lys33. The hydrophobic feature labeled 4 in Figure 2a usually matches halogen-substituted aromatic rings that occupy the hydrophobic pocket of the CDK2 ATP binding site.

Retrospective Compound Set Analysis. The 45 compounds statisfying all pharmacophore constraints, in the conformations and positions observed in the crystal structure, are scattered over the chemical space represented by the 80 compounds included in the reference set of actives (Figure 2c). Most compounds match three of the four features, including the acceptor and donor features required for the hydrogenbonding to the *hinge region* of the CDK2.

Prospective Binding Mode Reproduction. The percentage of compounds for which a binding mode is predicted exceeds the 45 compounds (56%) that were found to fulfill the pharmacophore criteria for all algorithms except Catalyst, LigandScout, and Unity. Scoring seems to be problematic for most algorithms as only ~20−40% of compounds have a rmsd to the cocrystallized reference pose below 3.0 Å for the best scored pose (Figure 3a). This behavior appears to be partly caused by the number of conformers that are used as inputs for benchmarking the different algorithms, as performance is better if the best scored pose among the X-ray conformation active set is evaluated (Figure 3b,d).

An analysis of the best reported rmsd of all matched poses for each molecule reveals that ~40% of compounds are predicted with a rmsd below 2.5 Å, with the majority of those even below 2.0 Å (Figure 3c). MOE even reaches almost 56% percent. Again, it is indicated that pose prediction is more accurate if only the cocrystallized conformation of the ligand is used for the pharmacophore search (Figure 3d). However, the increase in compounds predicted below rmsd 2.5 Å is only minor (Figure 3c,d), indicating that the multiple conformation generation protocol used includes at least one conformation representative for the cocrystallized conformer.

Compound Library Enrichment. Receiver Operator Characteristic curves (ROC) are generated from the results of the pharmacophore searches of the MUV-actives and MUV-decoys sets. Most algorithms retrieve less than 70% and 50% of actives and decoys, respectively, as indicated by the endpoints of the lines in Figure 4a. The AUC calculation returns values between 0.5 and 0.6 for most algorithms, indicating that the overall enrichment is only slightly better than could be expected from a random selection (Figure 4a,b). This seems to be mainly because of the relatively low number of active compounds retrieved by the pharmacophore. In particular, LigandScout retrieves only 3 out of 30 actives by matching those actives among the first 0.5% of the screened database and reaching an enrichment factor (EF) of 13.33 with AUC of 0.527 (Figure 4b). Most algorithms show decreasing enrichment factors at higher false positive retrieval rates, indicating that the score measures to rank all compounds that match a pharmacophore are relatively successful for CDK2 ligands in a compound library enrichment experiment (Figure 4a,b). The analysis of algorithm combinations (Figure 4c) shows that LigandScout, Unity, and POT are capable of improving enrichment of other algorithms. In particular, LigandScout and POT seem to be complementary as there is an improvement of both enrichment factors if they are used in a consecutive screening pipeline. Without any prescreening, LigandScout matches 3 out of 30 actives and 711 out of 15000 decoys resulting in an enrichment factor of 2.1, while POT matches 20 out of 30 actives and 7571 out of 15000 decoys resulting in an enrichment factor of 1.3. When combined, these algorithms retrieve 3 actives and 471 decoys and have an enrichment factor of 3.2. This is an improvement in enrichment of a factor 1.5 for LigandScout and



Feature	Matching compounds
1: Acceptor	120
2: Hydrophobe	110
3: Acceptor	77
4: Donor	48
5: Donor	44
Pharmacophore 1: (1+2+3+4)	36
Pharmacophore 2: (1+2+3+5)	21
Number of excluded volumes	75

Figure 5. CHK1 data set. (a) Pharmacophore depiction as used in this study on top of PDB entry: 2YWP. Features used in either pharmacophore 1 or 2 are visualized with dashed lines. (b) List of pharmacophore features with corresponding matching compounds in the set of actives. (c) Two-dimensional illustration of active compound similarities created using stochastic proximity embedding (SPE) with euclidean distances of BCI fingerprints.[59−61] Blue dots represent compounds that match pharmacophore 1; green dots represent compounds that match pharmacophore 2; yellow dots represent compounds that match both pharmacophores according to the observed ligand alignment in the crystal structures; and red dots are the compounds not satisfying the pharmacophore.

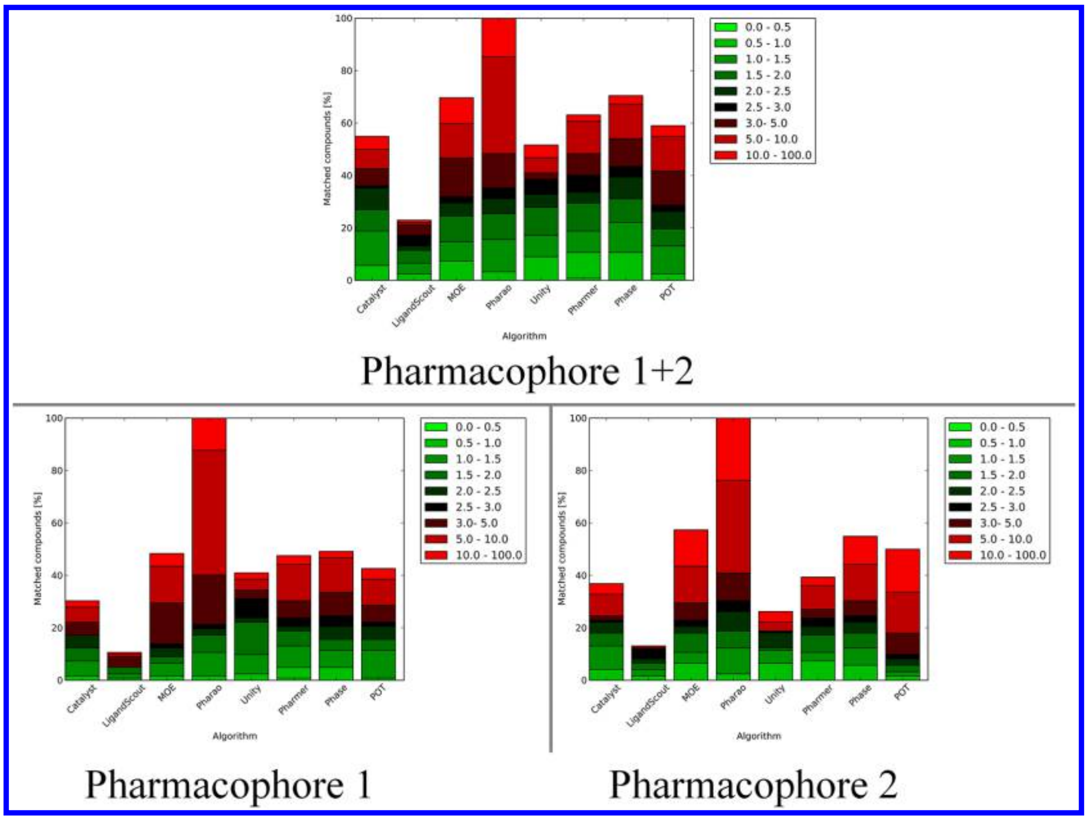
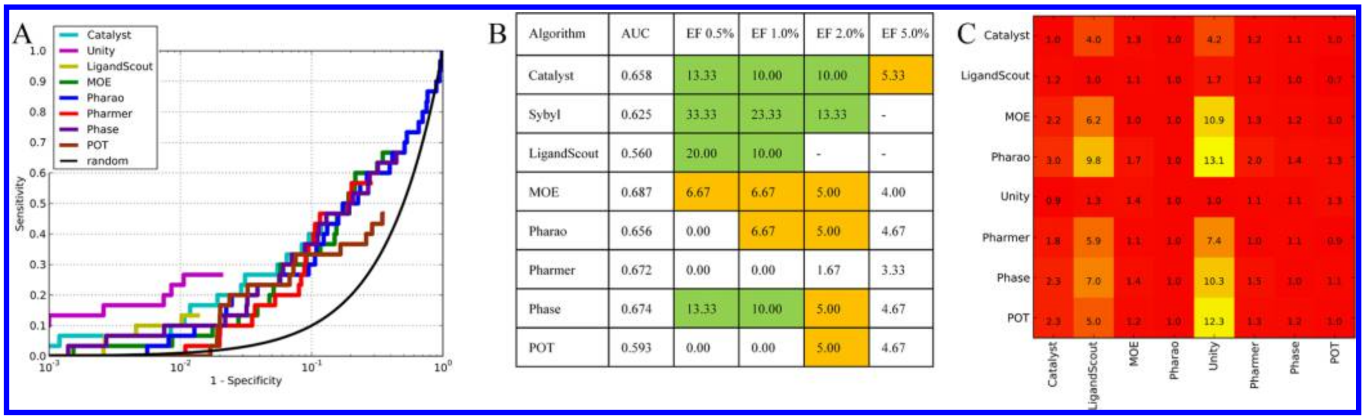


Figure 6. Rmsd ranges of matched compounds from the cocrystallized ligand for pharmacophore 1 and 2 (A), pharmacophore 1 (B), and pharmacophore 2 (C). Both figures refer to the best matching pose of multiconformational data sets.

2.4 for POT (Figure 4c). Only few values below 1.0 are observed, indicating that most algorithms can be combined with others without adversely affecting enrichments. This is an important observation that suggests the use of fast algorithms as prefiltering steps for large compound collections before more accurate and computationally expensive algorithms.

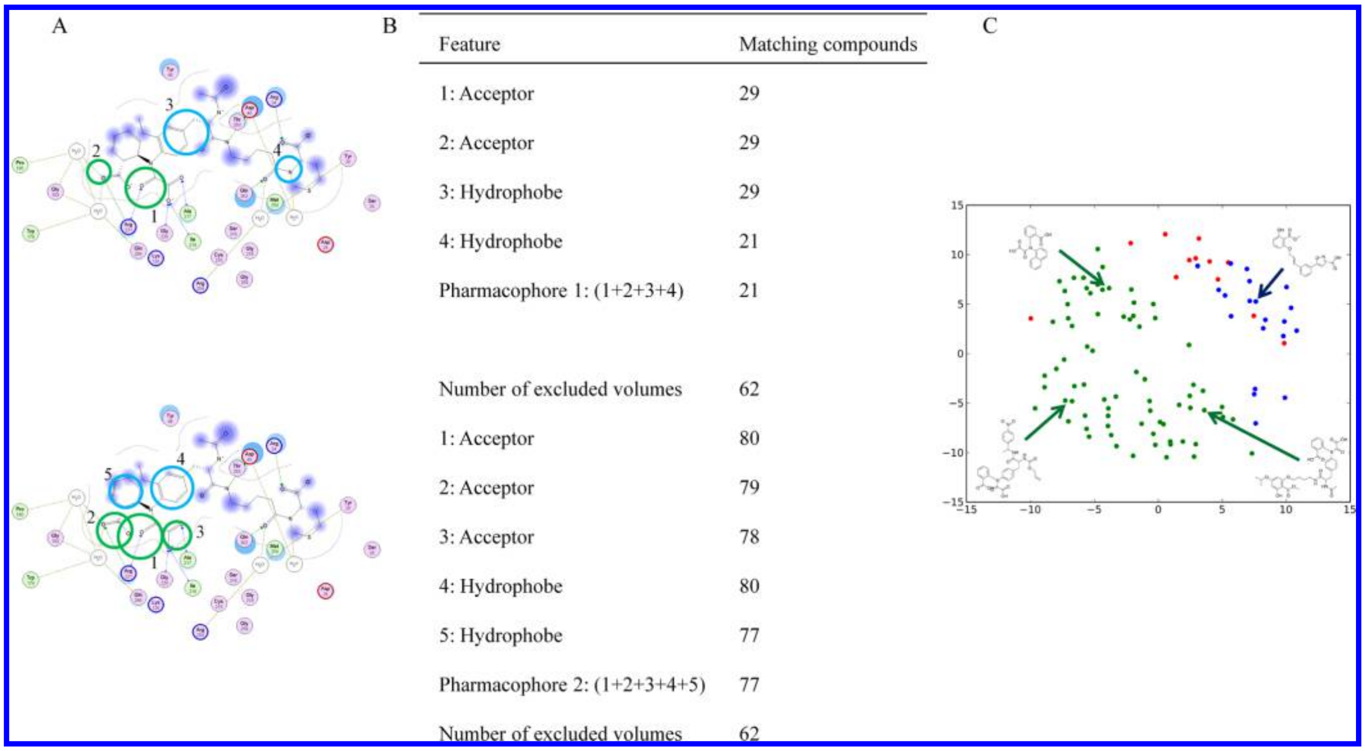
CHK1 Data Set. *Pharmacophore Perception.* The pharmacophore features of protein kinase CHK1 are well described in the literature.[54] Similarly to CDK2, a hydrogen bond donor and acceptor pair represents the key interactions for binding the *hinge region* of the kinase (features 1, 4, and 5 of Figure 5a). However, in this case, two locations of the HBD feature are possible for CHK1 ligands, and the analysis of matching compounds indicates their exclusive behaviors (Figure 5b,c). Other differences with respect to the CDK2 pharmacophore are that the hydrophobic feature 2 (Figure 5a) is located in the solvent-exposed region of the binding site and is conserved for almost all ligands of the data set (Figure 5b)



Algorithm	AUC	EF 0.5%	EF 1.0%	EF 2.0%	EF 5.0%
Catalyst	0.658	13.33	10.00	10.00	5.33
Sybyl	0.625	33.33	23.33	13.33	-
LigandScout	0.560	20.00	10.00	-	-
MOE	0.687	6.67	6.67	5.00	4.00
Pharao	0.656	0.00	6.67	5.00	4.67
Pharmer	0.672	0.00	0.00	1.67	3.33
Phase	0.674	13.33	10.00	5.00	4.67
POT	0.593	0.00	0.00	5.00	4.67

	Catalyst	LigandScout	MOE	Pharao	Unity	Pharmer	Phase	POT
Catalyst	1.0	4.0	1.3	1.0	4.2	1.2	1.1	1.0
LigandScout	1.2	1.0	1.1	1.0	1.7	1.2	1.0	0.7
MOE	2.2	6.2	1.0	1.0	10.9	1.3	1.2	1.0
Pharao	3.0	9.8	1.7	1.0	13.1	2.0	1.4	1.3
Unity	0.9	1.3	1.4	1.0	1.0	1.1	1.1	1.3
Pharmer	1.8	5.9	1.1	1.0	7.4	1.0	1.1	0.9
Phase	2.3	7.0	1.4	1.0	10.3	1.5	1.0	1.1
POT	2.3	5.0	1.2	1.0	12.3	1.3	1.2	1.0

Figure 7. Enrichment analysis of CHK1 MUV-data set. (A) ROC curves showing the enrichment of CHK1 actives/decoys (data set created with MUV, Table 1). (B) AUC values and enrichment values at 0.5%, 1.0%, 2.0%, and 5.0% false positive rate; green indicates an enrichment factor (EF) above 10.0 and orange indicates EF above 5.0. (C) Heatmap showing the improvement in enrichment factor of the algorithms on the *Y*-axis if prescreening with the algorithm on the *X*-axis is performed. Values greather than 1.0 indicate that the reapplication of the algorithms in the *X*-axis improves the results of algorithms in the *Y*-axis; values smaller than 1.0 indicate that reapplication of a second algorithm worsens the results; a value of 1.0 denotes no influence of the second algorithm.



Feature	Matching compounds
1: Acceptor	29
2: Acceptor	29
3: Hydrophobe	29
4: Hydrophobe	21
Pharmacophore 1: (1+2+3+4)	21
Number of excluded volumes	62
1: Acceptor	80
2: Acceptor	79
3: Acceptor	78
4: Hydrophobe	80
5: Hydrophobe	77
Pharmacophore 2: (1+2+3+4+5)	77
Number of excluded volumes	62

Figure 8. PTP1B data set. (a) Pharmacophore depiction as used in this study on top of PDB entry: 1NZ7. (b) List of pharmacophore features with the number of matching compounds in the set of actives. c) Two-dimensional depiction of active compound similarities created using stochastic proximity embedding (SPE) with euclidean distances of BCI fingerprints.[59−61] Blue dots represent compounds that match pharmacophore 1; green dots represent compounds that match pharmacophore 2 according to the observed ligand alignment in the crystal structures; and red dots are the compounds not satisfying the pharmacophore.

and that an additional HBA that is located deep in the binding pocket that represents interactions with the catalytic lysine of the CHK1.

Retrospective Compound Set Analysis. For the reasons provided above, two different pharmacophores were defined for this CHK1 data set differing in the position of the donor feature (Figure 5a). As shown in Figure 5c, the pharmacophores correspond to two topologically distinct clusters of compounds. Among the whole data set of 123 actives, only three compounds match all five pharmacophore features that are common to pharmacophore 1 and 2 (Figure 5c). Remarkably those compounds have a relatively high dissimilarity to compounds from both clusters. Most compounds, however, do not match either of the pharmacophores (Figure 5c, red dots).

Prospective Binding Mode Reproduction. The searches with pharmacophores 1 and 2 retrieve an approximately equal number of compounds and match ~20% of compounds with a rmsd of below 3.0 Å (~25 compounds) with respect to the cocrystallized reference structure (Figure 6b,c). This is in agreement with the retrospective analysis in which pharmacophore 1 and pharmacophore 2 are both derived from ~20% of

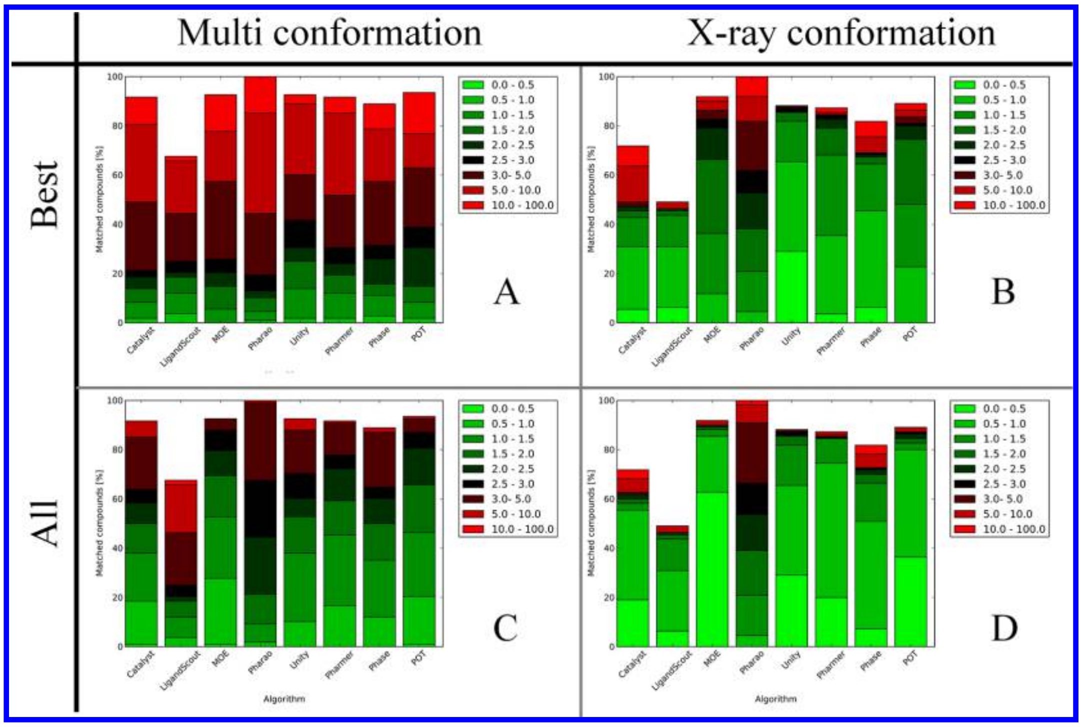
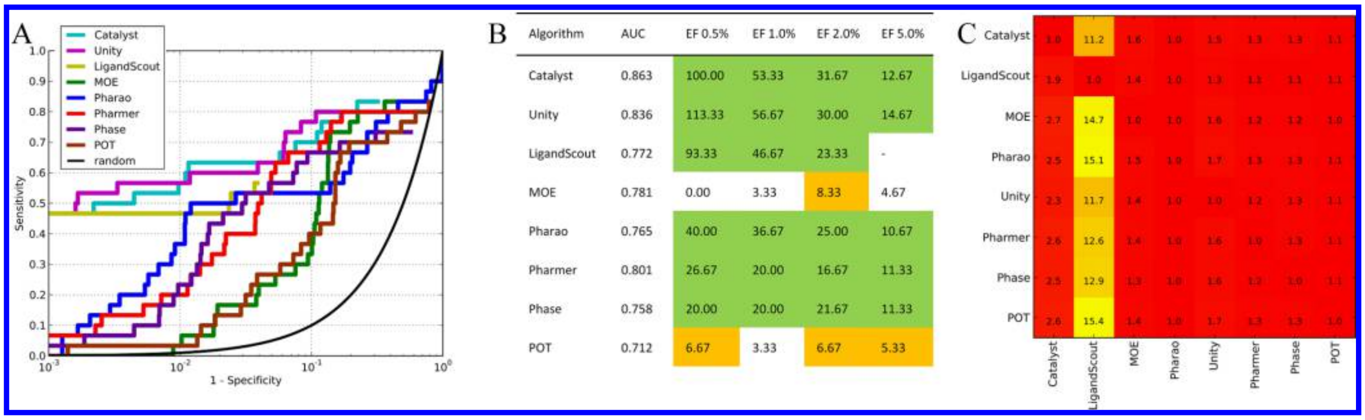


Figure 9. Rmsd ranges of matched compounds from the cocrystallized ligand in two different scenarios. (A) Best ranked pose from the ligand set in their multiconformational format. (B) Best ranked pose from the ligand set in X-ray conformation. (C)Lowest rmsd from all poses from the ligand set in multiconformational format. (D) lowest rmsd from all poses in the ligand set in their X-ray conformation.



Algorithm	AUC	EF 0.5%	EF 1.0%	EF 2.0%	EF 5.0%
Catalyst	0.863	100.00	53.33	31.67	12.67
Unity	0.836	113.33	56.67	30.00	14.67
LigandScout	0.772	93.33	46.67	23.33	-
MOE	0.781	0.00	3.33	8.33	4.67
Pharao	0.765	40.00	36.67	25.00	10.67
Pharmer	0.801	26.67	20.00	16.67	11.33
Phase	0.758	20.00	20.00	21.67	11.33
POT	0.712	6.67	3.33	6.67	5.33

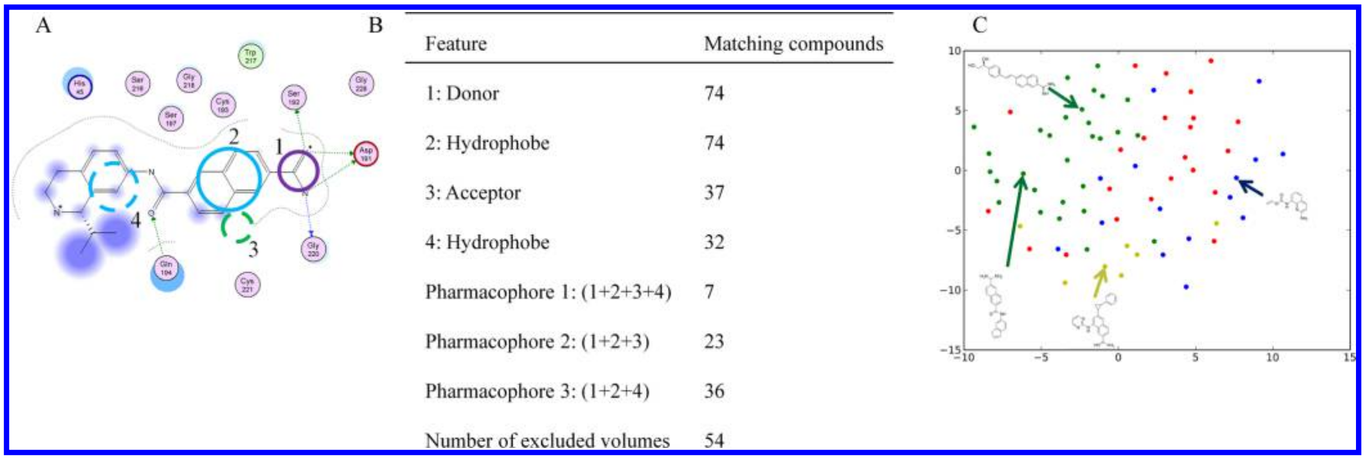
	Catalyst	LigandScout	MOE	Pharao	Unity	Pharmer	Phase	POT
Catalyst	1.0	11.2	1.6	1.0	1.5	1.3	1.3	1.1
LigandScout	1.9	1.0	1.4	1.0	1.3	1.1	1.1	1.1
MOE	2.7	14.7	1.0	1.0	1.6	1.2	1.2	1.0
Pharao	2.5	15.1	1.5	1.0	1.7	1.3	1.3	1.1
Unity	2.3	11.7	1.4	1.0	1.0	1.2	1.3	1.1
Pharmer	2.6	12.6	1.4	1.0	1.6	1.0	1.3	1.1
Phase	2.5	12.9	1.3	1.0	1.6	1.2	1.0	1.1
POT	2.6	15.4	1.4	1.0	1.7	1.3	1.3	1.0

Figure 10. Enrichment analysis of PTP1B MUV-data set. (A) ROC curves showing the enrichment of PTP1B actives/decoys (data set created with MUV, Table 1). (B) AUC values and enrichment values at 0.5%, 1.0%, 2.0%, and 5.0% false positive rate; green indicates an enrichment factor (EF) above 10.0 and orange indicates EF above 5.0. (C) Heatmap showing the improvement in enrichment factor of the algorithms on the *Y*-axis if prescreening with the algorithm on the *X*-axis is performed. Values greather than 1.0 indicate that reapplication of the algorithms in the *X*-axis improves the results of algorithms in the *Y*-axis; values smaller than 1.0 indicate that reapplication of a second algorithm worsen the results; values of 1.0 denotes no influence of the second algorithm.

the compounds. Notably, LigandScout retrieves a low number of actives, while the percentage of compounds in the correct conformations is equal to or even better than those generated by other algorithms.

Compound Library Enrichment. The stricter matching criteria of LigandScout is reflected in the search of the decoys as just over 1% of compounds are retrieved, while the same analysis shows at least 10% for all other algorithms, except Unity (Figure 7). These tighter criteria may explain the improved early enrichment of Unity and LigandScout, for which enrichment factors exceeded 20.0 in the top 0.5% of the ranked database (Figure 7b). The rmsd-based scoring methods, like POT, Pharmer, MOE, and Phase, have also relatively good (≥5.0) enrichments at 2.0% percent of the searched database but do not achieve enrichment factors (EF) of 10.0 or higher. The consecutive screening of compounds with MOE and Unity results in the best enrichment (data not shown), mainly due to the strong performance of Unity that retrieves 8 out of 30 actives and only 306 out of 15000 decoys. The consecutive application of both algorithms results in an enrichment of 18.7 that is 1.4 times the enrichment factor of the Unity search and 10.9 times the enrichment factor of MOE search. (Figure 7c).

PTP1B Data Set. *Pharmacophore Perception.* PTP1B is a protein tyrosine phosphatase and is characterized by a highly conserved and positively charged active-site.[55] The core of the binding features is characterized by a dyad of hydrogen bond



Feature	Matching compounds
1: Donor	74
2: Hydrophobe	74
3: Acceptor	37
4: Hydrophobe	32
Pharmacophore 1: (1+2+3+4)	7
Pharmacophore 2: (1+2+3)	23
Pharmacophore 3: (1+2+4)	36
Number of excluded volumes	54

Figure 11. Urokinase data set. (a) Pharmacophore depiction as used in this study on top of PDB entry: 1OWK. Features used in either pharmacophore 2 or 3 are visualized with dashed lines. (b) List of pharmacophore features with corresponding matching compounds in the set of actives. (c) Two-dimensional illustration of active compound similarities created using stochastic proximity embedding (SPE) with euclidean distances of BCI fingerprints.[59−61] Yellow dots represent compounds that match in pharmacophore 1, 2 and 3; blue dots represent compounds that match pharmacophore 2; green dots represent compounds that match pharmacophore 3 according to the observed ligand alignment in the crystal structures; and red dots are the compounds not satisfying the pharmacophore.

acceptors (HBAs), often represented by acid moieties, that ensure several interactions with arginine and histidine residues present in the binding pocket. Additional features include a HBA (pharmacophore 2, feature 3, Figure 3a) and hydrophobic sites that occupy the binding site region at different locations.

Retrospective Compound Compound Set Analysis. Like CHK1, we defined two distinct pharmacophores for PTP1B (Figure 8a). These pharmacophores relate to compounds from distinct chemical moieties. For instance, pharmacophore 2, which contains 5 features, matches 77 compounds that vary in size but all contain the N-substituted oxamic acid moiety, while pharmacophore 1 relates only to 21 compounds (Figure 8c). Twelve compounds do not match either pharmacophore 1 or 2.

Prospective Binding Mode Reproduction. The pharmacophore search based on X-ray conformations (Figure 9b) shows that 22 of the compounds (~20%) are retrieved by pharmacophore 1 with an rmsd ≤2.5 Å and ~70% (77 compounds) by pharmacophore 2 (data in the Supporting Information). Notably, MOE is able to retrieve ~50% (55 compounds) of the compounds with an rmsd <2.5 Å with a search of pharmacophore definition 1. This is most likely the result of a less stringent feature definition that also explains the relatively high number of retrieved decoys and the moderate performance of MOE in compound library enrichment across all targets (Figure 10a,b, and following). Conformation generation and appropriate scoring seems, however, to be a problem for nearly all algorithms. Only 20−30% of compounds are matched with an rmsd <2.5 Å (Figure 9a), while many more are correctly positioned in cases where only the X-ray conformation is used (Figure 9b) or where the pose with the lowest rmsd is picked among all matched poses (Figure 9c).

Compound Library Enrichment. Overlay-based scoring functions (Catalyst, LigandScout, and Unity) seem to perform very well with respect to compound library enrichment and exhibit high early enrichment values (Figure 10a,b). Nonetheless, one should keep in mind that for all algorithms scoring is often based on predicted binding modes that do not match the true binding modes. For example, in the cases of Catalyst, LigandScout, and to a lesser extent Unity, ~45% of active compounds are ranked above all decoys (Figure 10a), while only ~20−30% of compounds have rmsd values <2.5 Å (Figure 9a). A possible explanation for this discrepancy might be the fact that only a fraction of the ligand is represented by pharmacophore features, especially in case of pharmacophore 2 (Figure 8a). Yet, enrichments are good, especially for overlay-based (Catalyst, LigandScout, and Unity) scoring functions (Figure 10b), and the output of nearly every algorithm can be improved by prescreening with another algorithm as indicated by the values above 1.0 in Figure 10c. Combinations with Catalyst, LigandScout, Unity, and MOE show the largest enrichment improvement factors (on average >1.4) in Figure 10c. The best enrichment factor results from the combination of Catalyst with LigandScout that together retrieve 17 out of 30 actives and 308 out of 15000 decoys resulting in an enrichment factor of 27.6.

Urokinase Data Set. *Pharmacophore Perception.* Despite its name, Urokinase is a serine protease that is clinically used for therapy of thrombolytic disorders and whose small-molecule inhibitors have already been shown to inhibit cancer growth.[56] The pharmacophore created from the list of active compounds (Figure 11a) shows that two features are always present, specifically a hydrogen bond donor (HBD) (feature 1) and a hydrophobic feature (feature 2). The rest of the pharmacophore consists of one hydrogen bond acceptor (HBA) and another hydrophobic feature that seems to be mutually exclusive because matching compounds are almost complementary (Figure 11b,c).

Retrospective Compound Set analysis. Urokinase compounds show relatively diverse features with respect to the other data sets. This is exemplified by the fact that the best four-feature pharmacophore (pharmacophore 1) only satisfied seven compounds (Figure 11), as deduced from the ligand overlay of cocrystallized ligands. For this reason, we generated two additional three-feature pharmacophores of which both are comprised of three of the features from the original four-feature pharmacophore (Figure 11a,b). Pharmacophore 2 contains a donor, hydrophobic, and acceptor feature and matched 23 compounds that show little similarity (Figure 11c). Pharmacophore 3 contains a donor and two hydrophobic features and matches 36 compounds in the cocrystallized overlay. Those compounds are more similar to each other than the compounds

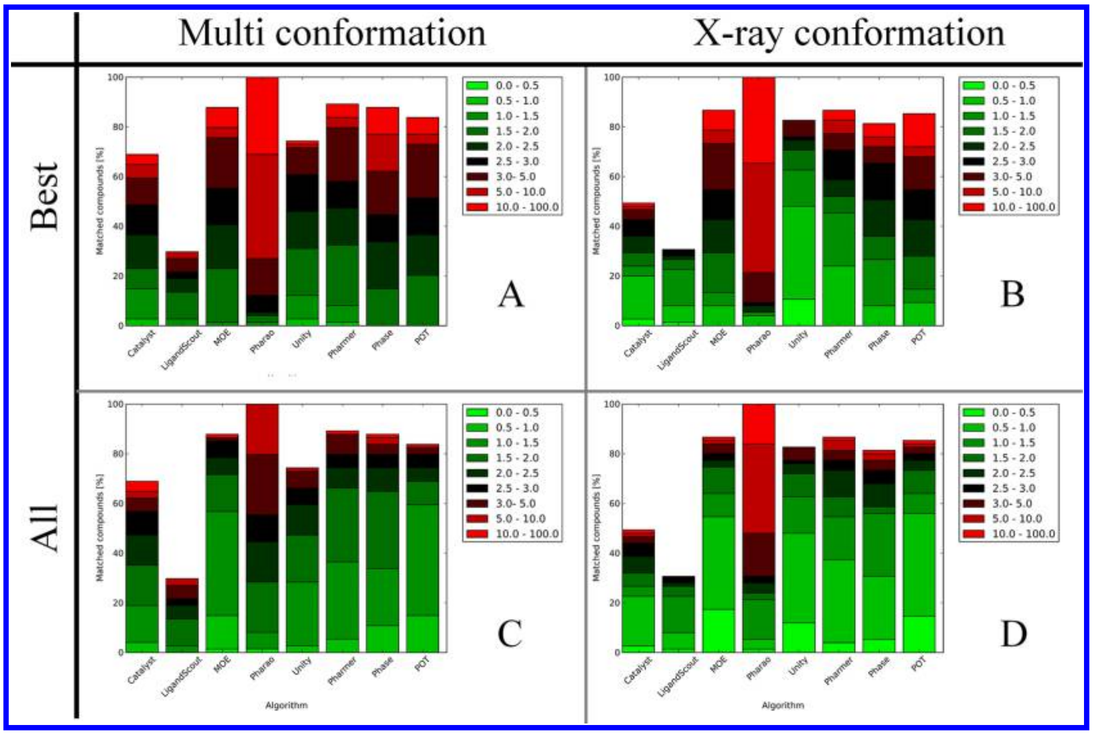
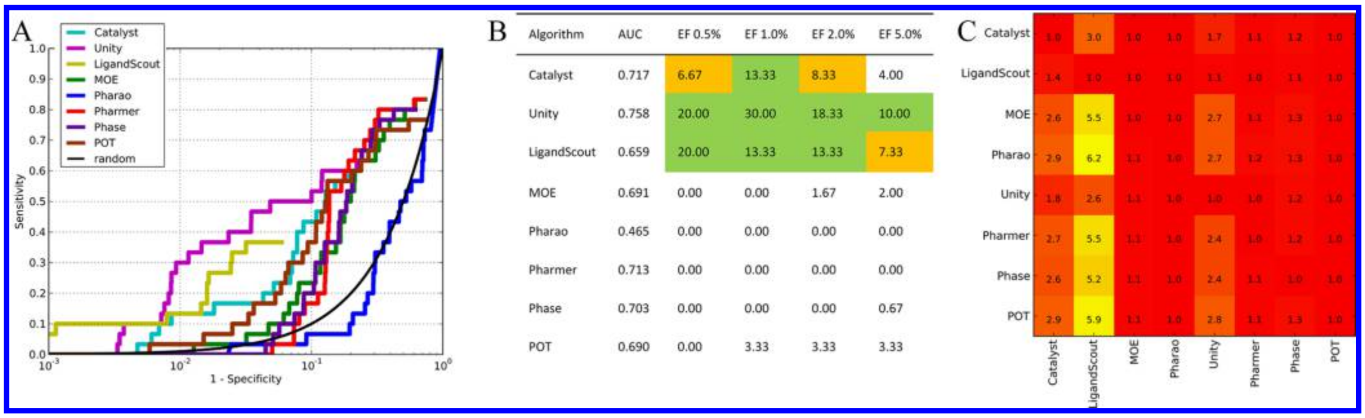


Figure 12. Rmsd ranges of matched compounds from the cocrystallized ligand in two different scenarios. (A) Best ranked pose from the ligand set in their multiconformational format. (B) Best ranked pose from the ligand set in X-ray conformation. (C) Lowest rmsd from all poses from the ligand set in multiconformational format. (D) lowest rmsd from all poses in the ligand set in their X-ray conformation.



Algorithm	AUC	EF 0.5%	EF 1.0%	EF 2.0%	EF 5.0%
Catalyst	0.717	6.67	13.33	8.33	4.00
Unity	0.758	20.00	30.00	18.33	10.00
LigandScout	0.659	20.00	13.33	13.33	7.33
MOE	0.691	0.00	0.00	1.67	2.00
Pharao	0.465	0.00	0.00	0.00	0.00
Pharmer	0.713	0.00	0.00	0.00	0.00
Phase	0.703	0.00	0.00	0.00	0.67
POT	0.690	0.00	3.33	3.33	3.33

Figure 13. Enrichment analysis of Urokinase MUV-data set. (A) ROC curves showing the enrichment of Urokinase actives/decoys (data set created with MUV, Table 1). (B) AUC values and enrichment values at 0.5%, 1.0%, 2.0%, and 5.0% false positive rate; green indicates an enrichment factor (EF) above 10.0 and orange indicates EF above 5.0. (C) Heatmap that illustrates component importance when combining two pharmacophore algorithms. Values greather than 1.0 indicate that reapplication of the algorithms on the *X*-axis improves the results of algorithms on the *Y*-axis; values smaller than 1.0 indicate that reapplication of a second algorithm worsens the results; a value of 1.0 denotes no influence of the second algorithm.

matching pharmacophore 2 and are clustered together in topological structure space (Figure 11c).

Prospective Binding Mode Reproduction. Retrieval rates for the separate pharmacophores correspond to what is observed in the overlay of cocrystallized ligands with ~10% (7 compounds), 40–60% (30–45 compounds), and 50–70% (37–52) matching to pharmacophores 1, 2, and 3, respectively (Supporting Information). The scoring of poses could, however, be improved for most algorithms because more accurate poses are usually in the ensemble of solutions (Figure 12c) but are not scored as being best (Figure 12a). Although rmsd-based scoring methods (MOE, Pharmer, POT) have a comparable performance in pose prediction for Urokinase, they perform poorly in compound library enrichment (Figure 13a,b).

Compound Library Enrichment. Similarly to the other targets, overlay-based scoring algorithms outperform rmsd methods in compound library enrichment (Figure 13a, b). It is also notable that the combination of two knowledge-based scoring algorithms (Catalyst, Unity, and LigandScout) improves enrichment values (Figure 13c), while such trend is not observed for the combination of two rmsd-based methods (Figure 13c). For instance, larger values in Figure 13c are obtained if Catalyst, Unity, and LigandScout are combined as consecutive screens with Pharmer, Phase, POT, and MOE. The best enrichment is obtained by a combination of Catalyst with

LigandScout: together both algorithms retrieve 10 out of 30 actives and 577 out of 15000 decoys and have an enrichment factor of 8.7 (Methods section, eq 1).

General Discussion. The compound sets considered in this study allowed us to explore the different characteristics of a range of pharmacophore screening algorithms in terms of compound retrieval and pose prediction. Different pharmacophores were derived from the ligand alignments observed in 80 CKD2, 120 CHK1, 110 PTP1B, and 74 Urokinase crystal structures. For CDK2, one pharmacophore was defined from 45 actives, while 36 and 21 actives defined two CHK1 pharmacophores, 21 and 77 actives defined two PTP1B pharmacophores, and 7, 23, and 36 actives defined three Urokinase pharmacophores. Some of these pharmacophores match well-defined clusters of active molecules (CHK1 and PTP1B), while others match a more diverse range of chemical structures. In this regard, the ability to identify different scaffolds by capturing the chemistry of different ligands, which is one of the most important capabilities of virtual screening methods, shows that all algorithms are generally good in recognizing different scaffolds (Figure 6S of the Supporting Information and subsequent discussion). In fact, scaffolds that do not match the pharmacophore in the cocrystallized pose have been successfully retrieved by the various algorithms emphasizing the advantages of using low-resolution methods like pharmacophores in drug design. Most algorithms retrieve a greater number of compounds, as one would expect after analysis of the ligand poses in the available crystal structures, and this might indicate that several active compounds are matched in conformations that do not correspond to the experimental one. The ability of the scoring functions to identify the correct experimental-bound pose is limited, as this can depend on (i) the ligand input structures, (ii) the pharmacophore's definition, or (iii) the scoring method applied by the pharmacophore screening tool. For instance, for CDK2, PTP1B and to a lesser extent CHK1 running the pharmacophore searches against X-ray conformations results in better pose reproductions than when searching against the precalculated conformational ensembles. The pharmacophore definition can also be responsible for poor pose reproduction as illustrated by the PTP1B pharmacophore 2, which describes only a small part of the interaction patterns of the cocrystallized ligands and generates only poses with relatively high rmsd values. Scoring methods are also suboptimal, as they frequently fail to identify the best pose from the full ensemble of matched poses (see CDK2 and Urokinase data sets). Compound retrieval based on poses dissimilar to the biophysical binding mode suggests that hit identification by serendipity does still frequently occur in pharmacophore search strategies. It should be noted here that while the definition of hydrophobic features may entail differences in screening results, as reported by Wolber et al.[57] and Spitzer et al.,[58] in our case we did not observe discrepancies. For example, molecules with heteroatomic rings at positions of hydrophobic pharmacophore features are predicted with similar accuracies by algorithms which type these rings as either hydrophobic or non-hydrophobic. This might be a result of the fairly large hydrophobic features definition in our pharmacophores that allow fitting of a compound in the pharmacophore with nearby hydrophobicity typed groups, resulting in almost equal rmsd values. In this context, it should be noted that while the use of different pharmacophore search algorithms in their default settings may lead to comparable results, deriving and searching a pharmacophore within the same software may render the setup of a pharmacophore screening more straightforward toward optimum performances.[58]

Compound library enrichment seems not so much related to the pharmacophore search algorithm used but more to the compound sets and corresponding pharmacophores used in this study. PTP1B, particularly, showed very good early enrichments that might be related to the N-substituted oxamic acid moiety present in most active compounds.

Combining the strength of several algorithms seems possible by screening compound libraries with different algorithms in a consecutive order. By comparison of the enrichment factor of all compounds matching algorithm pairs and enrichment factor of all compounds matching a single algorithm, we showed that improvements over a factor of 1.5 are possible.

■ CONCLUSIONS

We carried out a comparative study of eight pharmacophore screening tools by describing their different ability to retrieve active compounds for four biological targets of interest in their default settings. Several analyses allowed us to better elucidate advantages and drawbacks of algorithms when they are applied with their default settings for high-throughput pharmacophore screening purposes.

Our analysis shows that the correct reproduction of experimental binding poses is generally better with algorithms using rmsd-based methods (MOE, Pharmer, Phase, and POT) than with overlay-based methods (Catalyst, LigandScout, Pharao, and Unity). However, because many compounds match pharmacophore hypotheses without reproducing the experimental pose, it is also important to assess the ratio of correctly predicted compounds on the overall number of matched compounds in a given data set. In this respect, the performance of overlay-based algorithms is slightly better than the rmsd-based methods. Thus, while rmsd-based algorithms generally return "more shots on goal" due to the high number of poses, overlay-based methods seem to provide the best chance of retrieving the relevant biophysical binding mode. As a whole, these findings suggest that one may prefer certain algorithms depending on the research application stage. For example, when optimizing lead compounds, it may be necessary to collect a large number of binding modes to explore the conformational space thoroughly. In such a case, one may prefer an rmsd-based method. Vice versa, if only a single binding mode is desired, one may prefer overlay-based methods. In fact, the stricter criteria of overlay-based methods returns better results in compound library enrichment studies, as they are better at discriminating between active and inactive compounds. As this is most likely due to the stricter fitting criteria (which retrieve subsets of rmsd-based methods) and better scoring, it seems feasible to prescreen large compound databases with rmsd-based pharmacophore screening methods, which are typically faster, to obtain the same results in a less time-consuming manner.

Overall ,we can conclude that (i) the more advanced overlay-based scoring algorithms do result in more enriched compound libraries than rmsd-based algorithms, (ii) compound library enrichment dependends on the biological target and not strictly from the choice of a given algorithm, and (iii) the use of different pharmacophore search algorithms may lead, in their default settings, to comparable results.

We acknowledge that our findings could be extended and corroborated with further analyses with other biological targets.

However, the result of this work may be already of practical use in the planning of more efficient high-throughput pharmacophore screens.

ASSOCIATED CONTENT

Supporting Information

Preparation details of the CDK2 data set, selection of active and decoys data sets with MUV, data set files, pharmacophore definition details; Conformational analysis and its discussion; Scaffold analysis and its discussion. This information is available free of charge via the Internet at http://pubs.acs.org

AUTHOR INFORMATION

Corresponding Author

*Phone: +39 051 2094004. Fax: +39 051 2094746. E-mail: alberto.delrio@gmail.com.

Present Address

#Netherlands eScience Center, Science Park 140, 1098XG, Amsterdam, The Netherlands

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dave Wood for critical reading of the manuscript and all software companies and developers for their support. This work was supported by AIRC Emilia Romagna Start-up grant 6266 (A.J.M.B. and A.D.); Bayer Haemophilia Awards programme (G.A.F.N); Dutch Top Institute Pharma, project number: D1-105 (M.P.A.S.); and Cyttron II - FES0908 (B.Z.)

REFERENCES

(1) *Molecular Design: Concepts and Applications*; Schneider, G., Baringhaus, K., Kubinyi, H., Eds.; Wiley-VCH: Weinheim, 2008.

(2) *Chemoinformatics Approaches to Virtual Screening*; Varnek, A., Tropsha, A., Eds.; RSC Publishing: Cambridge, 2008.

(3) *Chemoinformatics*; Gasteiger, J., Engel, T., Eds.; Wiley-VCH: Weinheim, 2003.

(4) Caporuscio, F.; Rastelli, G.; Imbriano, C.; Del Rio, A. Structure-Based Design of Potent Aromatase Inhibitors by High-Throughput Docking. *J. Med. Chem.* **2011**, *54* (12), 4006−4017.

(5) Kolb, P.; Ferreira, R. S.; Irwin, J. J.; Shoichet, B. K. Docking and chemoinformatic screens for new ligands and targets. *Curr. Opin. Chem. Biol.* **2009**, *20* (4), 429−436.

(6) Villoutreix, B. O.; Renault, N.; Lagorce, D.; Sperandio, O.; Montes, M.; Miteva, M. A. Free resources to assist structure-based virtual ligand screening experiments. *Curr. Protein Pept. Sci.* **2007**, *8* (4), 381−411.

(7) Ballester, P. J.; Westwood, I.; Laurieri, N.; Sim, E.; Richards, W. G. Prospective virtual screening with Ultrafast Shape Recognition: The identification of novel inhibitors of arylamine N-acetyltransferases. *J. R. Soc., Interface* **2010**, *7* (43), 335−342.

(8) Schneider, G. Virtual screening: An endless staircase? *Nat. Rev. Drug Discovery* **2010**, *9* (4), 273−276.

(9) Leach, A. R.; Gillet, V. J.; Lewis, R. A.; Taylor, R. Three-dimensional pharmacophore methods in drug discovery. *J. Med. Chem.* **2010**, *53* (2), 539−558.

(10) Langer, T. Pharmacophores in drug research. *Mol. Inf.* **2010**, *29* (6−7), 470−475.

(11) Caporuscio, F.; Tafi, A. Pharmacophore modelling: A forty year old approach and its modern synergies. *Curr. Med. Chem.* **2011**, *18* (17), 2543−2553.

(12) Sun, H. Pharmacophore-based virtual screening. *Curr. Med. Chem.* **2008**, *15* (10), 1018−1024.

(13) Del Rio, A.; Barbosa, A.; Caporuscio, F., Use of large multiconformational databases with structure-based pharmacophore models for fast screening of commercial compound collections. *J. Cheminf.* [online] **2011**, *3* (Suppl 1), P27, http://www.jcheminf.com/content/3/S1/P27 (accessed February 2, 2012).

(14) Del Rio, A.; Barbosa, A. J.; Caporuscio, F.; Mangiatordi, G. F. CoCoCo:A free suite of multiconformational chemical databases for high-throughput virtual screening purposes. *Mol. BioSyst.* **2010**, *6* (11), 2122−2128.

(15) Irwin, J. J.; Shoichet, B. K. ZINC: A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177−182.

(16) Masciocchi, J.; Frau, G.; Fanton, M.; Sturlese, M.; Floris, M.; Pireddu, L.; Palla, P.; Cedrati, F.; Rodriguez-Tome, P.; Moro, S. MMsINC: A large-scale chemoinformatics database. *Nucleic Acids Res.* **2009**, *37* (Database issue), D284−290.

(17) Barbosa, A. J.; Del Rio, A. Freely accessible databases of commercial compounds for high-throughput virtual screenings. *Curr. Top. Med. Chem.* **2012**, *12* (8), 866−877.

(18) Oellien, F.; Cramer, J.; Beyer, C.; Ihlenfeldt, W. D.; Selzer, P. M. The impact of tautomer forms on pharmacophore-based virtual screening. *J. Chem. Inf. Model.* **2006**, *46* (6), 2342−2354.

(19) Kolb, P.; Irwin, J. J. Docking screens: Right for the right reasons? *Curr. Top. Med. Chem.* **2009**, *9* (9), 755−770.

(20) Chen, Z.; Li, H. L.; Zhang, Q. J.; Bao, X. G.; Yu, K. Q.; Luo, X. M.; Zhu, W. L.; Jiang, H. L. Pharmacophore-based virtual screening versus docking-based virtual screening: A benchmark comparison against eight targets. *Acta Pharmacol. Sin.* **2009**, *30* (12), 1694−1708.

(21) Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection—What can we learn from earlier mistakes? *J. Comput.-Aided Mol. Des.* **2008**, *22* (3−4), 213−228.

(22) ten Brink, T.; Exner, T. E. Influence of protonation, tautomeric, and stereoisomeric states on protein−ligand docking results. *J. Chem. Inf. Model.* **2009**, *49* (6), 1535−1546.

(23) Patel, Y.; Gillet, V. J.; Bravi, G.; Leach, A. R. A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP. *J. Comput.-Aided Mol. Des.* **2002**, *16* (8−9), 653−681.

(24) Peach, M. L.; Nicklaus, M. C., Combining docking with pharmacophore filtering for improved virtual screening. *J. Cheminf.* [online] **2009**, *1* (1), 6, http://www.jcheminf.com/content/1/1/6 (accessed February 2, 2012).

(25) Brown, S. P.; Muchmore, S. W. Large-scale application of high-throughput molecular mechanics with Poisson−Boltzmann surface area for routine physics-based scoring of protein−ligand complexes. *J. Med. Chem.* **2009**, *52* (10), 3159−3165.

(26) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235−242.

(27) Morgan, D. O. Cyclin-dependent kinases: Engines, clocks, and microprocessors. *Annu. Rev. Cell Dev. Biol.* **1997**, *13*, 261−291.

(28) Sanchez, Y.; Wong, C.; Thoma, R. S.; Richman, R.; Wu, Z.; Piwnica-Worms, H.; Elledge, S. J. Conservation of the Chk1 checkpoint pathway in mammals: Linkage of DNA damage to Cdk regulation through Cdc25. *Science* **1997**, *277* (5331), 1497−1501.

(29) Elchebly, M.; Payette, P.; Michaliszyn, E.; Cromlish, W.; Collins, S.; Loy, A. L.; Normandin, D.; Cheng, A.; Himms-Hagen, J.; Chan, C. C.; Ramachandran, C.; Gresser, M. J.; Tremblay, M. L.; Kennedy, B. P. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science* **1999**, *283* (5407), 1544−1548.

(30) Andreasen, P. A.; Kjoller, L.; Christensen, L.; Duffy, M. J. The urokinase-type plasminogen activator system in cancer metastasis: A review. *Int. J. Cancer* **1997**, *72* (1), 1−22.

(31) 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* (Database issue), D115−119.

(32) *Sybyl X*; Tripos: : St. Louis, MO.

(33) *Epik*, version 2.1107; Schrödinger, LLC: Portland, OR, 2011.

(34) Renner, S.; Schneider, G. Fuzzy pharmacophore models from molecular alignments for correlation-vector-based virtual screening. *J. Med. Chem.* **2004**, *47* (19), 4653−4664.

(35) *CoCoCo Database*. http://www.cococo-database.it (accessed date February 2, 2012)

(36) Wild, D. J.; Blankley, C. J. Comparison of 2D fingerprint types and hierarchy level selection methods for structural grouping using Ward's clustering. *J. Chem. Inf. Comput. Sci.* **2000**, *40* (1), 155−162.

(37) *ChEMBL*. https://www.ebi.ac.uk/chembldb/ (accessed date February 2, 2012)

(38) Warr, W. ChEMBL. An interview with John Overington, team leader, chemogenomics at the European Bioinformatics Institute Outstation of the European Molecular Biology Laboratory (EMBL-EBI). *J. Comput.-Aided Mol. Des.* **2009**, *23* (4), 195−198.

(39) 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* (2), 169−184.

(40) *Maximum Unbiased Validation (MUV)*, Datasets for Virtual Screening. http://www.pharmchem.tu-bs.de/lehre/baumann/MUV.html (accessed date February 2, 2012)

(41) Watts, K. S.; Dalal, P.; Murphy, R. B.; Sherman, W.; Friesner, R. A.; Shelley, J. C. ConfGen: A conformational search method for efficient generation of bioactive conformers. *J. Chem. Inf. Model.* **2010**, *50* (4), 534−546.

(42) Taminau, J.; Thijs, G.; De Winter, H. Pharao: pharmacophore alignment and optimization. *J. Mol. Graphics Modell.* **2008**, *27* (2), 161−169.

(43) Landrum, G. *RDkit*. http://rdkit.org (accessed date February 2, 2012).

(44) Accelrys Inc. http://accelrys.com (accessed date Feb 2, 2012).

(45) Koes, D. R.; Camacho, C. J. Pharmer: Efficient and exact pharmacophore search. *J. Chem. Inf. Model.* **2011**, *51* (6), 1307−1314.

(46) Sanders, M. P.; Verhoeven, S.; de Graaf, C.; Roumen, L.; Vroling, B.; Nabuurs, S. B.; de Vlieg, J.; Klomp, J. P. Snooker: A structure-based pharmacophore generation tool applied to class A GPCRs. *J. Chem. Inf. Model.* **2011**, *51* (9), 2277−2292.

(47) Chemical Computing Group. http://www.chemcomp.com (accessed date February 2, 2012).

(48) Wolber, G.; Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* **2004**, *45* (1), 160−169.

(49) *Phase*, version 3.3; Schrödinger, LLC: Portland, OR, USA, 2011.

(50) Jain, A. N.; Nicholls, A. Recommendations for evaluation of computational methods. *J. Comput.-Aided Mol. Des.* **2008**, *22* (3−4), 133−139.

(51) Agrafiotis, D. K.; Gibbs, A. C.; Zhu, F.; Izrailev, S.; Martin, E. Conformational sampling of bioactive molecules: A comparative study. *J. Chem. Inf. Model.* **2007**, *47* (3), 1067−1086.

(52) Griewel, A.; Kayser, O.; Schlosser, J.; Rarey, M. Conformational sampling for large-scale virtual screening: accuracy versus ensemble size. *J. Chem. Inf. Model.* **2009**, *49* (10), 2303−2311.

(53) Zou, J.; Xie, H. Z.; Yang, S. Y.; Chen, J. J.; Ren, J. X.; Wei, Y. Q. Towards more accurate pharmacophore modeling: Multicomplex-based comprehensive pharmacophore map and most-frequent-feature pharmacophore model of CDK2. *J. Mol. Graphics Modell.* **2008**, *27* (4), 430−438.

(54) Chen, X. M.; Lu, T.; Lu, S.; Li, H. F.; Yuan, H. L.; Ran, T.; Liu, H. C.; Chen, Y. D. Structure-based and shape-complemented pharmacophore modeling for the discovery of novel checkpoint kinase 1 inhibitors. *J. Mol. Model.* **2010**, *16* (7), 1195−1204.

(55) Zhang, S.; Zhang, Z. Y. PTP1B as a drug target: Recent developments in PTP1B inhibitor discovery. *Drug Discovery Today* **2007**, *12* (9−10), 373−381.

(56) Shui, L.; Bharatham, N.; Bharatham, K.; Lee, K. W., Urokinase inhibitor design based on pharmacophore model derived from diverse classes of inhibitors. *Interdiscip. Bio Cent.* [online] **2009**, http://www.ibc7.org/article/e_archive_v.php?sid=67 (accessed February 2, 2012)

(57) Wolber, G.; Seidel, T.; Bendix, F.; Langer, T. Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discovery Today* **2008**, *13* (1−2), 23−29.

(58) Spitzer, G. M.; Heiss, M.; Mangold, M.; Markt, P.; Kirchmair, J.; Wolber, G.; Liedl, K. R. One concept, three implementations of 3D pharmacophore-based virtual screening: Distinct coverage of chemical search space. *J. Chem. Inf. Model.* **2010**, *50* (7), 1241−1247.

(59) Agrafiotis, D. K. Stochastic proximity embedding. *J. Comput. Chem.* **2003**, *24* (10), 1215−1221.

(60) Agrafiotis, D. K.; Xu, H. A self-organizing principle for learning nonlinear manifolds. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (25), 15869−15872.

(61) BCI Fingerprints. Barnard Chemical Information Ltd.. http://www.bci1.demon.co.uk/ (accessed date February 2, 2012).