

HPPD: Ligand- and Target-Based Virtual Screening on a Herbicide Target

Miriam López-Ramos* and Francesca Perruccio*

Syngenta Crop Protection, Muenchwilen AG, WST-820.1.15, Schaffhauserstrasse, CH-4332 Stein, Switzerland

Received December 22, 2009

Hydroxyphenylpyruvate dioxygenase (HPPD) has proven to be a very successful target for the development of herbicides with bleaching properties, and today HPPD inhibitors are well established in the agrochemical market. Syngenta has a long history of HPPD-inhibitor research, and HPPD was chosen as a case study for the validation of diverse ligand- and target-based virtual screening approaches to identify compounds with inhibitory properties. Two-dimensional extended connectivity fingerprints, three-dimensional shape-based tools (ROCS, EON, and Phase-shape) and a pharmacophore approach (Phase) were used as ligand-based methods; Glide and Gold were used as target-based. Both the virtual screening utility and the scaffold-hopping ability of the screening tools were assessed. Particular emphasis was put on the specific pitfalls to take into account for the design of a virtual screening campaign in an agrochemical context, as compared to a pharmaceutical environment.

INTRODUCTION

Virtual screening (VS) is a technique now commonly used in drug discovery programs for lead finding and optimization and for scaffold hopping.¹ In such an approach, a collection of potential candidate compounds is screened against a target protein or a reference molecule in order to select a subset of compounds for effective experimental screening. The selection can be done using a wide range of VS methods, either ligand- or target-based when the three-dimensional (3D) structure of the target protein is available. Therefore, it is important to choose the VS method that performs best in the particular case considered. As a consequence, the comparison of VS tools has been the subject of many studies and debates.^{2–14}

Although VS has received great attention as a major tool for lead finding and scaffold hopping in the pharmaceutical context, very little has been published in this field concerning the agrochemical world. In agrochemistry, the concept underlying this approach remains the same as in drug discovery: active compounds are usually identified from biological screens, competitor's patents, or literature follow-up, and new active molecules are needed to scaffold-hop or to enrich the structure–activity relationship analysis. VS of the corporate collection or commercial libraries is, therefore, also in this case a useful means of selecting relevant compounds for biological tests. The VS tools available are the same as in drug discovery, although agrochemistry presents some particularities. For example, the fact that the activity of compounds is mainly evaluated directly *in vivo* and does not reflect the intrinsic activity of the molecule on its target. Also, “agro-like” compounds differ from “drug-like” ones;^{15,16} for example, log *P* values are generally lower, and hydrogen-bond (H-bond) donor atoms are less frequent in herbicidal compounds than in drugs. Some functional

groups, like alcohols and amines, are less represented, whereas heteroaromatic cycles, sulfonamides, and carboxylic acids are more common. Last, the content and characteristics of corporate databases are not the same in both fields, and no “agro-like” public databases are available for VS. Therefore, one has to be cautious when applying tools optimized for drug research to an agrochemical problem, in particular when designing a VS experiment. Within this context, the goal of our study is not to develop a new VS method or to perform an exhaustive comparison of the current VS tools, as has already been reported in several studies, but rather to evaluate the performance of the current industry standard VS tools available in house on an agrochemical target and to highlight any caveats found when applying such methodology in an agrochemical environment, as compared to a pharmaceutical one.

Several criteria were applied in selecting the biological system to use for this study. The possibility from an intellectual property point of view of disclosing and discussing in house data in a publication is always a fundamental request. Also, the project should correspond to a validated mode of action for which an *in vitro* assay is available, and it must contain a significant number of compounds active *in vitro*. In addition, a crystal structure of the target protein has to be accessible. After considering these requirements, the protein chosen for our analysis is the iron-containing enzyme 4-hydroxyphenylpyruvate dioxygenase (HPPD), the inhibition of which is a well-validated mode of action for selective herbicides. HPPD is involved in the tyrosine degradation pathway,¹⁷ and as such, it is a target for the treatment of hypertyrosinemia in humans.^{17,18} In plants, HPPD is a key enzyme for the production of plastoquinone and tocopherol. Its inhibition provokes a unique bleaching (whitening), followed by necrosis, and death.^{19,20} HPPD inhibitors have been a very active field of research across the agrochemical industry for many years,¹⁹ which has resulted in a considerable number of commercialized herbicides.²⁰ Mesotrione (Figure 1), sold by Syngenta under the

* Corresponding authors. Telephone: +41 613239334 (F.P.); +41 613237705 (M.L.-R.). E-mail: francesca.perruccio@syngenta.com (F.P.); miriam.lopez-ramos@m4x.org (M.L.-R.).

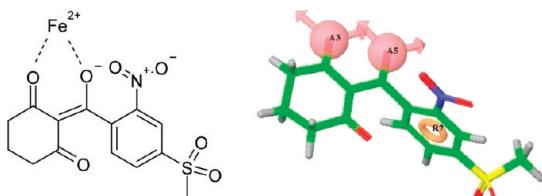


Figure 1. Mesotrione ligand (left) and its associated pharmacophore (right).

brand name Callisto, is the main representative of one of the chemical classes of such herbicides, the triketone family.

In order to evaluate the performance of different ligand- and target-based VS tools, we carried out a retrospective work in which we searched over a database consisting of compounds active on HPPD and decoy compounds. We analyzed the ability of different methods to discriminate between actives and inactives and also to successfully perform scaffold hopping. As in the pharmaceutical world, in the agrochemical scenario, the mode of action is often unknown or the corresponding 3D protein structure is unavailable, making difficult the use of a target for a structure-based VS approach. Therefore, besides target-based approaches (Gold^{21–24} and Glide^{25–27}), we also applied ligand-based 2D (Scitegic's extended connectivity fingerprints²⁸), 3D (Phase-shape,²⁹ ROCS,^{30–32} and EON³⁰), and pharmacophore-based (Phase²⁹) methods. It is worth emphasizing that, whereas several studies involving Gold, Glide, or ROCS have already been published,^{3,6,33–38} very few analyses include Phase as a ligand-based screening tool³⁹ or apply the target-based methods to metalloproteins.^{37,38} Also, the comparisons between ligand- and target-based methods are not so frequent,^{11,35,38,40,41} and very few cover 2D, 3D ligand- and target-based approaches,⁵ such as the study we report here.

METHODS

Database. Although the modeling community increasingly encourages the use of benchmarked data sets to reduce bias in the comparison of methods, to make the results comparable across studies and to allow the calculations to be reproduced,^{1,42,43} it appeared difficult to do so in an agrochemical context. Indeed, no benchmarked databases containing “agro-like” compounds or decoys for relevant agrochemical targets, which would be the equivalents of the MDDR or the DUD database, are available. Therefore, we chose to create our own database starting from Syngenta’s corporate collection.

Selection of the Active Set. The set of HPPD active compounds consisted of 216 molecules that had shown a good *in vitro* HPPD inhibitory activity ($\text{IC}_{50} < 1 \mu\text{M}$).

Decoy set selection. The first step in building a decoy set for a VS campaign from the corporate collection is to ensure that no false negatives are selected. Therefore, all the compounds related to the HPPD project were removed. In addition, it is important to select the decoy compounds in such a way that there are no obvious systematic differences between the actives and the decoy set that could bias the results,^{4,38} such as $\log P$, molecular weight, or other simple descriptors of molecular properties. After applying Pipeline Pilot’s²⁸ HTS and organic filters, several other filters were then used to narrow down the collection to the compounds displaying the same molecular properties as the set of known

HPPD actives (molecular weight between 250 and 550 and $\text{Alog } P$ between 0.35 and 6). Other constraints based on property counts calculated with Pipeline Pilot (the numbers of H-bond acceptors and donors, the number of rotatable and bridge bonds, and the number of spiro atoms and ring assemblies) were also applied.

Even after applying these filters, the collection still contained a vast number of compounds and, therefore, clustering based on molecular descriptors (FCFP_4 in the Pipeline Pilot software) was carried out to create a decoy of ~50 000 molecules. However, this approach did not lead to an evenness of cluster sizes; when selecting the centers of the clusters to build our decoy, the result was a collection of singletons rather than of compounds representative of chemical series within the Syngenta corporate file. Therefore, this clustering method was discarded. Our strategy consisted then in using the Pesticide Manual,⁴ a compendium of 1373 compounds commercialized for agricultural use, as a starting point, and in selecting the compounds from the filtered database that were most similar to each of the molecules referenced in this manual. The similarity measure used was Tanimoto FCFP_4 (as implemented in Pipeline Pilot), and 50 000 molecules were requested in total, keeping for each compound in the Pesticide Manual the molecules with the top 30 similarity values (a different similarity measure would result in a slightly different decoy set, but this should not affect the outcome of the virtual screening, since the measure is related to the compounds of the Pesticide Manual and not to the query ligand). After removal of the compounds that failed during subsequent ligand preparation or conformer generation, we obtained a decoy set consisting of 49 549 compounds.

Such an approach guarantees a sufficient chemical diversity, because the compounds from the Pesticide Manual belong to different chemical classes. This is important for ligand-based approaches, since the so-called “analogue bias” can lead to artificially high enrichments when similarity-based VS techniques, such as 2D fingerprints, are validated with a database lacking enough chemical diversity.⁴¹ In addition, the fact that the Pesticide Manual compounds are used in different indications (as insecticides, pesticides, herbicides, safeners, plant growth regulators, etc.) means that they also have different physicochemical properties. As can be seen from the distribution of molecular weight, $\text{Alog } P$ and number of H-bond donors and acceptors in Figure 2, the distribution of molecular properties for the compounds in the Pesticide Manual is rather wide. The profile of the decoy set is the truncated equivalent of that of the Pesticide Manual compounds. As compared to the decoy set, active compounds have a narrower distribution of properties, especially for the number of H-bond donors and acceptors. For $\text{Alog } P$ the mean of actives is well centered compared to the decoy and only slightly shifted toward heavier compounds for the molecular weight. The decoy compounds for which some properties are out of the range of values observed in the active set might be discarded more easily by the VS tools. Although this could lead to artificially high enrichments, the comparison of different methods, which is the aim of our work, should not be affected.

Preparation of the Database. All compounds (actives and decoys) were treated in the same manner to avoid any bias toward the selection of actives:⁴² no additional conformers

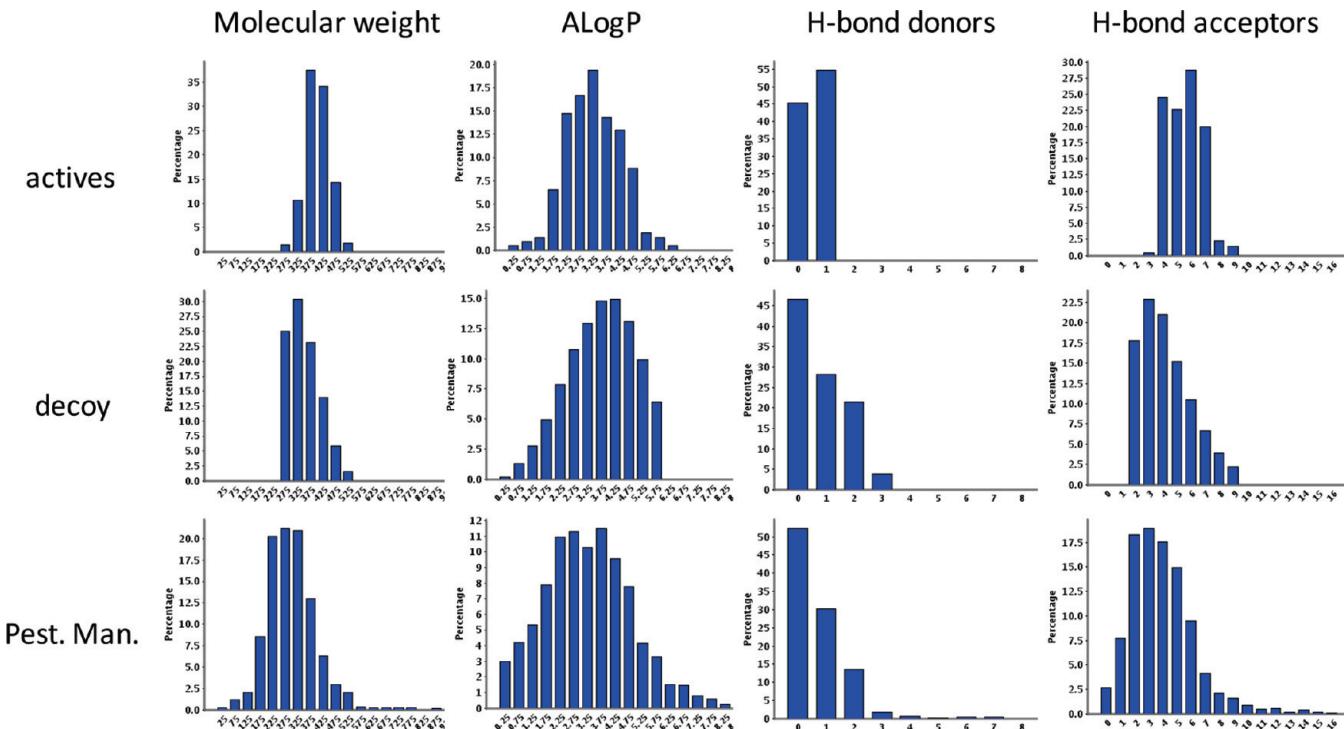


Figure 2. Histograms of selected simple molecular descriptors for the set of active compounds, the decoy set, and the compounds present in the Pesticide Manual, which were used to build the decoy set. All descriptors were calculated with Pipeline Pilot.

were used for actives, and tautomeric and protonation states were not modified by hand.

For docking with Gold and Glide, database compounds were prepared with the Schrödinger⁴⁴ Ligprep routine (version 2.3.108). Ionizer was used to generate all possible protonation states at pH = 7, and the compounds were desalting. Tautomers were generated, and the already specified chiralities were retained. All other parameters were used with their default values, the default force-field being OPLS_2005. The database obtained was then used to generate the pharmacophore points to be stored in the Phase database, for the search for matches to the pharmacophore (hypothesis) with Phase (the structures were not further prepared and all parameters were kept at their default values).

For ligand-based VS runs with ROCS and Phase-shape, Omega⁴⁵ version 2.3.2 (Openeye) was used to further prepare the database by generating sets of conformers. The force-field mmff_Trunc, which excludes both Coulombic interactions and the attractive part of the van der Waals interactions from the Merck molecular force-field (MMFF), was chosen for the conformer generation. Up to 100 conformers were saved for each compound, and default values were used for all other parameters.

Target Protein and Ligand Query. The choice of the crystal structure for the structure-based approaches and the ligand query for the ligand-based ones can have a great influence on the outcome of the VS.⁷ Therefore, particular care was taken to choose representative structures.

Target Protein. The structures of several HPPD-inhibitor complexes were solved in-house, most of them with inhibitors bearing the triketone pharmacophore typical of HPPD inhibitors. Among all these complexes, only the ones containing syncarpic acids were considered because they show a better ordering of the loops around the active site. A representative crystal structure solved with 2.25 Å resolution

was chosen for this study, keeping the iron atom and the water molecule that participates in iron coordination. For the docking studies with Gold, the protein was prepared using the Hermes⁴⁶ v1.0.1 interface. Hydrogen atoms were added, and the iron coordination geometry was set to octahedral. For the docking runs with Glide, the *protein preparation wizard* tool from the Schrödinger suite was used to add hydrogens and to set the iron charge to +2; then a protein grid was generated with the grid generation utility of Glide. In both cases, the active site was defined using the native ligand and the cavity detection option from Gold or the “dock ligands with length ≤20 Å” option in Glide.

Ligand Query. As already mentioned in the literature,⁷ for ligand-based approaches it is not necessary to choose as a probe ligand the compound crystallized with the same target protein that was used in target-based methods. Among all the ligands that were cocrystallized in complex with HPPD, we chose mesotriione as the query (Figure 1). In addition to being a commercial success, this compound is very representative of the triketone class of HPPD inhibitors and possesses only the minimal features necessary for HPPD-inhibition activity. The ligand was extracted as a mol2 from the crystal structure after adding all hydrogen atoms with the *protein preparation wizard* tool of the Schrödinger suite. The mesotriione structure so prepared was used as query for ROCS, EON, Phase-shape, and 2D screenings. The pharmacophore (hypothesis) used to find matches in the database with Phase was generated using the *create hypothesis* tool from Phase. Two H-bond acceptors corresponding to the carbonyl oxygen atoms involved in iron chelation as well as an aromatic ring present in all compounds of the active set were kept to define the pharmacophore hypothesis (Figure 1).

Table 1. Fingerprints Used and Their Associated Parameters^a

x	z	n	fingerprint	description
F	P	2	FCFP2	functional class based, extended connectivity fingerprint (<i>fingerprinting</i>), diameter of feature = 2
F	P	4	FCFP4	functional class based, extended connectivity fingerprint (<i>fingerprinting</i>), diameter of feature = 4
F	P	6	FCFP6	functional class based, extended connectivity fingerprint (<i>fingerprinting</i>), diameter of feature = 6
E	P	2	ECFP2	atom-type based, extended connectivity fingerprint (<i>fingerprinting</i>), diameter of feature = 2
E	P	4	ECFP4	atom-type based, extended connectivity fingerprint (<i>fingerprinting</i>), diameter of feature = 4
E	P	6	ECFP6	atom-type based, extended connectivity fingerprint (<i>fingerprinting</i>), diameter of feature = 6
F	C	2	FCFC2	functional class based, extended connectivity counts (<i>counts</i>), diameter of feature = 2
F	C	4	FCFC4	functional class based, extended connectivity counts (<i>counts</i>), diameter of feature = 4
F	C	6	FCFC6	functional class based, extended connectivity counts (<i>counts</i>), diameter of feature = 6
E	C	2	ECFC2	atom-type based, extended connectivity counts (<i>counts</i>), diameter of feature = 2
E	C	4	ECFC4	atom-type based, extended connectivity counts (<i>counts</i>), diameter of feature = 4
E	C	6	ECFC6	atom-type based, extended connectivity counts (<i>counts</i>), diameter of feature = 6

^a Adapted from Bender et al.⁴⁹

Proteins and Ligands for Assessment of Docking Accuracy.

Thirteen additional HPPD-inhibitor complexes solved in house with resolutions ranging from 1.9 to 2.5 Å were used to assess the docking accuracy in HPPD of Gold and Glide. The proteins were prepared in the same way as above, keeping the iron atom and the water molecule that participates in iron coordination and deleting all other water molecules, the ligand and eventual cofactors.

The extracted ligands were examined for bond orders and protonation states and written out as Maestro .mae and mol2 files for their use as “reference” ligands in root-mean-square deviation (rmsd) calculations. They were also converted to smiles strings and prepared using LigPrep, as described for the database molecules in order to be docked.

Ligand-Based Computational Methods. *2D Fingerprints.* Extended connectivity fingerprints were generated with SciTegic’s Pipeline Pilot software version 7.0. The circular fingerprints used are based on molecule-specific features, which are substructures centered on each atom of the molecule. The features of a whole molecule are recorded as an integer string,^{13,47,48} and the similarity of database compounds to the ligand query is then quantified using the Tanimoto coefficient between strings. Twelve different extended connectivity fingerprints of the type xCFz_n were used (Table 1),⁴⁹ in which x represents the atom typing (F for functional-class and E for atom-types atom typing); z indicates whether only the presence/absence of a feature is taken into account (P, hereafter referred to as *fingerprinting* to distinguish the value of this parameter from the whole fingerprint) or whether the number of features is also considered (C, hereafter referred to as *counts*); n is the diameter of the substructure representing a feature, counted in number of bonds (2-, 4-, and 6-bond long features have been used in this study).

MDL public keys were also used. MDL⁵⁰ keys are a set of 960, mostly substructural features, developed for rapid

substructural searching of ISIS databases. Molecular Design released the definition of 166 of the full set of 960. These are referred to as the MDL public keys.

ROCS. ROCS is a shape-based superposition method. Molecules are aligned by maximizing the overlap volume between them, and for every molecule in the database, the conformer with the best overlap with the query ligand is selected. Atom types can be taken into account using a color force-field, and color force-field forces and gradients can be used as part of the overlay optimization with the optchem option. Version 2.4.1 of ROCS was used, and three different VS runs were performed: without color force-field (shape_only), with ImplicitMillsDean color force-field (implicit), and with ImplicitMillsDean color force-field with forces and gradients (impl_optchem). When the color force-field was used, both the shape fit and the correspondence between the atom types of the database molecule and the query were considered (“combo score”). Otherwise, only the shape fit was taken into account. All other parameters were used with their default values.

EON. In EON, the Poisson–Boltzmann equation is solved to calculate the electrostatic potential of the query and database molecules. The score that takes into account the shape fit, as assessed by ROCS, as well as the similarity of the electrostatic potential around the database molecule and the query (“ET combo score”) were used to rescore all the alignments obtained with ROCS. The corresponding rankings will be hereafter referred to as shape_only_eon, implicit_eon, and impl_optchem_eon. Version 2.0.1 of EON was used with all the default parameters.

Phase-Shape. In the Phase-shape program, each conformer of a given molecule is aligned to the query. The shape search can treat all atoms as equivalent (hereafter Phase_shape_only), or it can favor alignments that superimpose atoms of the same type. There are three possibilities for atom typing in a shape search: macromodel types (Phase_shape_macro-

model), element types (Phase_shape_elements), or Phase QSAR types (Phase_shape_pharmacophore, the types used are among others: H-bond donor hydrogen, hydrophobic/nonpolar, negative and positive ionics, and electron withdrawing). Phase version 3.1.108 was used with the default parameters and with all available atom types.

Phase (Pharmacophore Search). Phase is a tool to find matches to a given pharmacophore (hypothesis) in a database. The search was performed using the conformers stored in the previously prepared Phase database, without rejecting any hit on the basis of its score and making sure that all molecules in the database were kept in the hit list. All other parameters were kept at their default values.

Target-Based Computational Methods. Gold. The docking program Gold version 4.0 was used with the native scoring function Goldscore, which includes the energy terms accounting for H-bonding, van der Waals interactions between the ligand and the protein, and ligand internal energy. Ring corners, amides, and carboxylic acids were allowed to flip, and early termination was allowed. All other parameters were kept at their default values. For the assessment of docking accuracy, two different speed settings were used: the standard genetic algorithm settings and the preset genetic algorithm settings corresponding to 10 000 genetic operations performed (hereafter referred to as Gold FS, for Fast Settings), which are the fastest preset parameters. In both cases, the ligand extracted from the crystal structure was used as a reference ligand, and rmsd calculations were requested. For the VS itself, only Gold FS was used.

Glide. Glide version 5.5.110 was run with the GlideScore function, a modified version of the ChemScore⁵¹ function that includes terms for steric clashes and buried polar interactions. The previously prepared protein grid was used, and all docking parameters were kept at their default values. For the assessment of docking accuracy, the three possible speed settings were used: standard precision (SP), extra precision (XP), and high-throughput mode (HTVS). The ligand extracted from the crystal structure was used as a core for rmsd calculations, taking exclusively heavy atoms into account. For the VS, which was set up with the VS Workflow version 2.1, distributed by Schrödinger, we decided to use not only the high-throughput mode, which is more appropriate for database searches, as in some reported studies,^{5,37} but also the standard precision mode, as other groups have described.^{6,33}

Analysis of the Results. Goodness Measures. There has been much debate recently about which metric is the most suitable to assess the relative performances of VS tools.^{5,42,52,53} In this study, we decided to use enrichment plots to compare the different methods used because we found that they were simple and well adapted to a typical agrochemical scenario in which VS is used as a prefilter for experimental biological screening. Indeed, an important criterion to take into account in such an approach is also the capacity of the testing facility. The main question is, therefore, given a certain number of compounds that can be effectively tested, to determine which VS method will give the highest number of actives. The answer to this question can be read at a glance from enrichment plots.

Enrichment curves represent the percentage of active compounds selected relative to the total number of active compounds, as a function of the proportion of the database

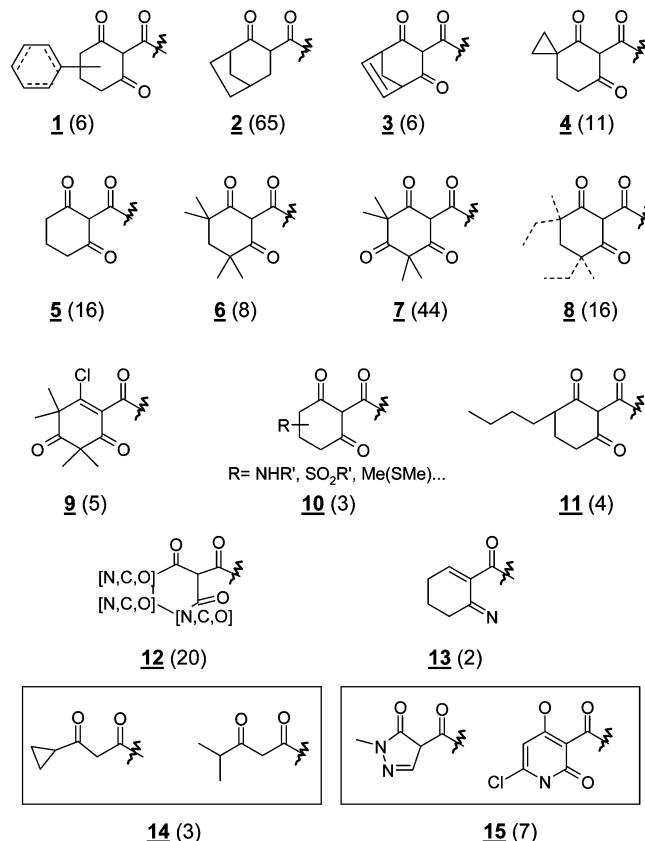


Figure 3. Scaffolds of the active compound set. The number of compounds in each class is indicated in brackets. Scaffold 4 contains 11 compounds with different sizes for the spiro ring. Scaffolds 1 and 8 are general depictions of clusters of very closely related compounds.

screened. The most significant retrieval rate in order to assess the performance in terms of screening utility was considered to be the one obtained when the top 1% of the database is selected. This would represent roughly 500 compounds in our case, which is already a rather large number of molecules to submit for biological screens. Indeed, in Syngenta's herbicide discovery process, the first and most significant biological test is performed directly on weeds grown in pots under controlled conditions in a glasshouse (only a small number of compounds of particular interest are then tested *in vitro*, if an *in vitro* assay is available). The screening capacity is, therefore, much lower than in the pharmaceutical industry, where the initial screen is a high-throughput *in vitro* test. However, looking at a smaller proportion of compounds would not have been statistically meaningful, so we used the top 1% of the database as a compromise.⁵ Although its relevance to assess the screening utility of the tools is lower, the retrieval rate of active compounds at 10% of the database was also calculated in order to allow an exhaustive analysis of the results.

For the target-based approaches, prior to processing the enrichment plots, an arbitrary bad score (zero in our case) was attributed to the compounds that could not be docked in the active site of the protein to place them at the bottom of the ranked list, as already described in other studies.^{5,33,37}

Assessment of Scaffold Diversity. The ability of VS methods to successfully perform scaffold hopping has received much attention in the past years, and many descriptors have been designed to measure scaffold-hopping

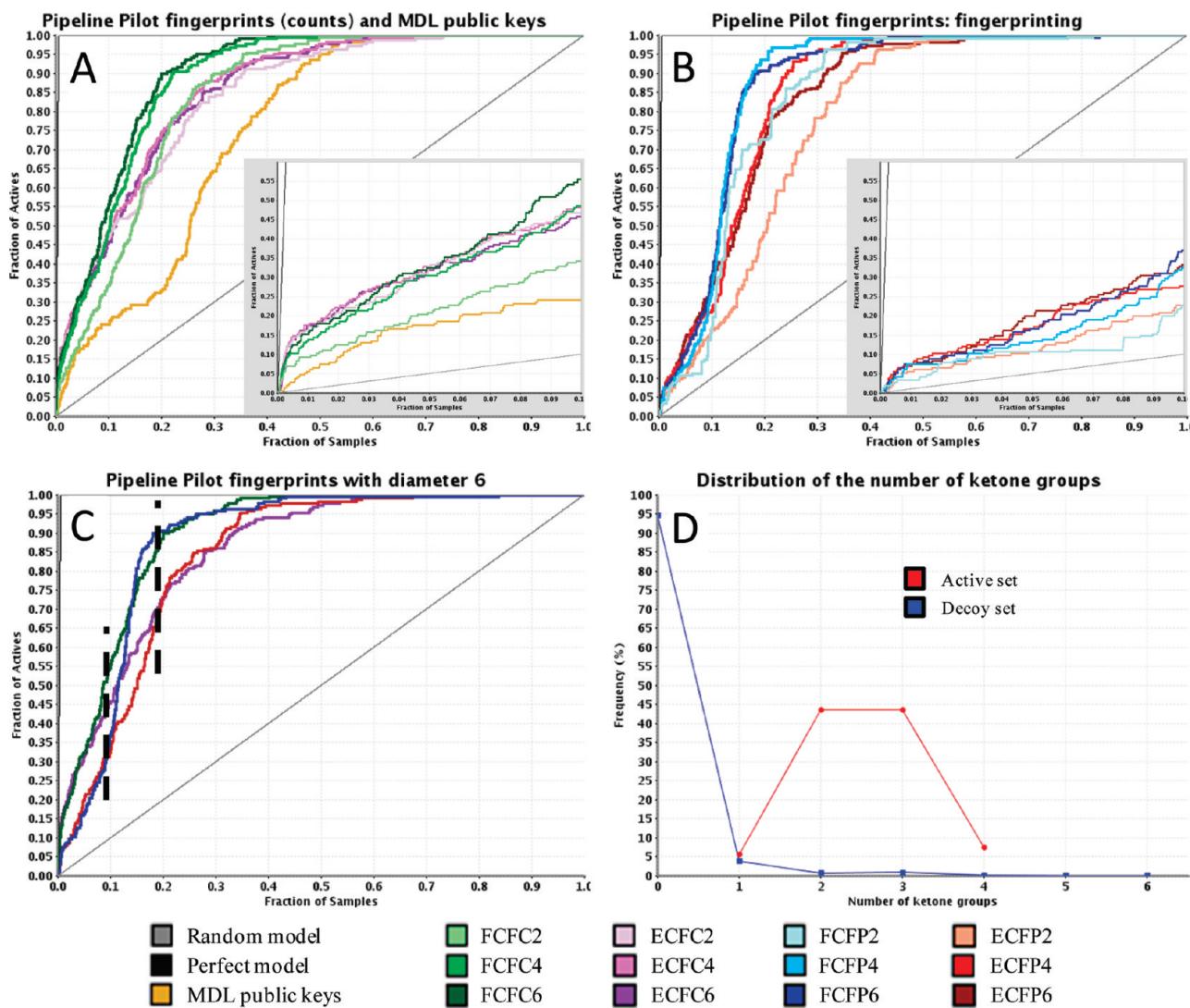


Figure 4. Enrichment plots obtained in 2D ligand-based VS using Pipeline Pilot’s extended connectivity fingerprints and MDL public keys. The embedded panels represent a zoom on the top 10% of the database. (A) counts-based fingerprints of the type (E,F)CFC with diameter values of 2, 4, and 6 and MDL public keys. (B) fingerprint-based fingerprints of the type (E,F)CFP with diameter values of 2, 4 and 6. (C) FCFP, FCFC, ECFP, and ECFC fingerprints with diameter 6. The dashed lines delimit the fractions of the database for which a different behavior of the fingerprints was observed. (D) Distribution of the number of ketone substructures found among the active and inactive compounds.

efficiency.¹ In our case, we decided to analyze by eye the chemical diversity of the active set and to assess how early the different scaffolds were retrieved by the various screening techniques. This kind of analysis method has already been reported.³⁴

The structure of all the compounds in the active set can be divided in two parts, connected by a ketone bridge: a “head” containing the most important pharmacophoric part, which would correspond to the cyclohexane-1,3-dione in mesotriione (Figure 1) and a “tail”, the 2-nitro-4-methanesulfonylphenyl part. Since the “tail” part of the compounds of the active set is always a substituted phenyl or pyridine, we considered only the “head” in order to evaluate the diversity. Using this approach, we defined 15 scaffold categories into which all compounds could be distributed (Figure 3). The first scaffolds (1–8), built around a cyclohexane and bearing only ketones or carbon substituents, are the most similar in structure to mesotriione, whereas the others (scaffolds 9–15) present more dissimilarities.

To analyze the diversity (from a chemist’s point of view) of the compounds retrieved by the various screening tools,

the most straightforward method would be to count the number of compounds of each category present in the best 1 and 10% of the database. However, Hert et al.¹³ pointed out that the relative performance in retrieving chemotypes mirrors closely that of active compounds when this kind of analysis is used, and indeed, this was the case with our data set. Therefore, we calculated instead the lowest percentage of the processed database that needed to be considered in order to find the first compound of each class (Tables 1–3 in Supporting Information). This approach allowed us to minimize the influence of the retrieval rate of actives on the retrieval of diversity, so as to evaluate both aspects as independently as possible.

RESULTS AND DISCUSSION

2D Ligand-Based Methods. Thirteen VS runs of the database with mesotriione as the reference ligand were performed using the 12 different extended connectivity fingerprints and the MDL public keys. The results obtained are represented as enrichment plots in Figure 4 (see also

Table 2. Retrieval Rates of Active Compounds for All Screening Methods When the Top 1 or 10% of the Database Is Selected

	model	% actives retrieved in best 1% of the collection	% actives retrieved in best 10% of the collection
2D Ligand-Based			
fingerprinting	FCFP2	3	23
	FCFP4	7	34
	FCFP6	7	37
	ECFP2	6	23
	ECFP4	7	28
	ECFP6	7	34
counts	FCFC2	9	34
	FCFC4	14	49
	FCFC6	15	56
	ECFC2	16	47
	ECFC4	18	49
	ECFC6	17	46
	MDL	6	24
3D Ligand-Based			
shape	Phase_shape_only	38	75
	ROCS_shape_only	28	62
shape + atom types	Phase_shape_elements	22	68
	Phase_shape_macromodel	21	70
	Phase_shape_pharmacophore	22	49
	ROCS_implicit	27	93
	ROCS_impl_optchem	23	93
	ROCS_shape_only_eon	40	78
	ROCS_implicit_eon	41	74
	ROCS_impl_optchem_eon	46	77
	pharmacophore_search	88	94
	Target-Based		
Gold FS		24	75
	Glide HTVS	17	50
	Glide SP	23	66

Figures 1 and 2 in Supporting Information), and the retrieval rates of active compounds when the top 1 or 10% of the database was selected are reported in Table 2. The MDL public keys clearly underperformed the extended connectivity fingerprints (Figure 4a) and will not be commented on any further. For the Pipeline Pilot fingerprints used in this study, the results show that all three parameters defining a fingerprint have an impact on the VS outcome: the fingerprint diameter (2-, 4- or 6-bonds long substructure), the atom typing (functional class (F) or atom type (E)), and the use of either fingerprinting (F) or counts (C) for each considered feature. The influence of these parameters is analyzed in detail below.

Influence of the Diameter. In all cases, fingerprints with diameters 4 and 6 performed better and in a similar manner when retrieving active compounds, whereas those with diameter 2 were significantly worse (Figure 4a and b and Table 2). This trend is less pronounced for the series of ECFC fingerprints, for which the difference between all diameters is less significant. This finding is consistent with the results of another recent study.⁴⁹ For the other models (FCFC, ECFP, and FCFP), the retrieval curves of the fingerprints with diameters 4 and 6 often intersect each other, the retrieval rates at 1% are very similar, and none of the methods is consistently better than the other one, except for FCFC6, which is always slightly better at retrieving actives than FCFC4.

Although the absence of significant performance differences between fingerprints with different diameters has been reported,⁴⁹ the underperformance of diameter 2 as compared with diameters 4 and 6 can be correlated in this case with

the amount of information contained in the fingerprint. Indeed, the analysis of the compounds in the active set shows that what can discriminate the active compounds from the decoys from a structural point of view is the presence of a di- or triketone substructure. Such a structure has a length of four bonds and can, therefore, be captured by a 4- or a 6-diameter fingerprint but not by a 2-diameter one. In other words, in the particular case of this study, a 2-diameter fingerprint cannot contain enough information to discriminate the active compounds from the decoys. On the contrary, the additional information included in a 6-diameter fingerprint with respect to a 4-diameter one does not make any difference in our test system, since the key information is already contained in the 4-diameter fingerprint. These considerations based on the structure of the active compounds could explain the influence of the diameter that we have observed here. Although they are very specific to our test system, they could be extended to other cases in which very specific chemical patterns are absolutely required for activity (a simple substructure search would of course in such situations provide a good enrichment).

Influence of Atom Types (ECFz or FCFz) and the Use of Fingerprinting (xCFP) versus Counts (xCFC). The analysis of the plots representing the retrieval rate of active compounds for the four types of fingerprints FCFP, ECFP, FCFC, and ECFC with a diameter of 6 (Figure 4c) shows that, when a small fraction of the database (up to roughly 10%) is considered, using counts (xCFC) instead of fingerprinting (xCFP) brings a significant improvement in the retrieval rates of active compounds, as recently reported by Bender et al.⁴⁹ (retrieval rates at 1% of 15% and 17% for

FCFC and ECFC compared to 7% for both FCFP and ECFP). In this early part of the curve, however, the kind of atom typing has only a slight influence (E fingerprints performing better than F). On the contrary, when large parts of the database are considered (larger than 20%), functional-class fingerprints (FCFz) show significantly better retrieval rates than those of atom-type fingerprints (ECFz) (retrieval rates at 30% of roughly 90% for FCFP and FCFC versus 72% for ECFP and ECFC), and the use of either fingerprinting (xCFP) or counts (xCFC) has a negligible impact on retrieval rates. The same trend was also observed for a diameter of 4 (data not shown), which led us to conclude that, for small fractions of the database, the retrieval rates are governed by the use of fingerprinting or counts (counts being better), whereas once 20% of the database has been collected, the methods are grouped by atom typing (functional-class atom types perform better).

These observations could, in our particular case, be partially explained by examining the kind of scaffolds retrieved at each stage of the database selection. The analysis of the smallest fraction of the database that had to be considered to find the first compound from each scaffold category (Table 1 in Supporting Information) showed that 10 out of 15 scaffolds were significantly more easily found by fingerprints using counts (of type xCFC) than by those based on the presence/absence of features (type xCFP). These scaffolds all corresponded to the ones having exactly three ketone groups, two in the head and an additional one in the bridge, like in particular the reference ligand mesotrione. Although the presence of ketones in drug- and agro-like compounds is very common, three ketone groups in the same molecule is a rather particular feature of the test system used. As can be seen in the distribution of the number of ketone groups in the active and decoy set shown in Figure 4d, the requirement for the presence of the feature *ketone* would already rule out most of the compounds from the decoy set (95%), but it would still leave 5% of the decoy set as potentially active compounds (so around 11-fold more compounds than in the active set). However, restricting the decoy set to only the compounds with three ketones would further significantly reduce the number of potential actives. This idea would explain why counts-based fingerprints, which could account for the presence of exactly three ketones, performed better than the fingerprints based on fingerprinting, which could only assess the presence or absence of the ketone group. It would also explain why counts (xCFC) were better than fingerprinting (xCFP) only when small fractions of the database were considered, since this rationale is only valid until all the compounds with exactly the same number of ketones as the reference compound have been selected.

The scaffolds that are most dissimilar to the reference molecule are found quite late in the database selection (see Table 1 in Supporting Information with the fraction of the database at which each scaffold appeared for the first time) but earlier by the functional-class fingerprints (F), which are based on a classification of the properties of atoms, rather than by the atom-type fingerprints (E), which takes into account the exact nature of the atoms. Such scaffold categories are the chloro- and imine-containing ones (9 and 13) as well as the ones with long chains and small rings (scaffolds 11 and 14). The greater ability of FCFz-type

fingerprints to retrieve more dissimilar compounds reveals their higher potential to scaffold hop and could explain why this kind of fingerprint performs best when large parts of the database have already been selected; once all triketone compounds have been retrieved, scaffold hopping becomes crucial to pick new active compounds.

Conclusions. The 2D ligand-based method used in this study that shows the best VS performance is the extended connectivity fingerprint FCFC6, with functional-class atom typing and based on feature counts within a diameter of 6 bonds. A diameter of 6 ensures in our test case that the key information is considered, and this fingerprint combines the best early retrieval rate of active compounds, which is characteristic of counts-based fingerprints in our test system, with the increased scaffold-hopping abilities of functional-class fingerprints. Overall, the chemical diversity appeared quite late along the database selection, probably because the selection of active compounds was biased toward the ones containing a triketone substructure. One might wonder to what extent the presence of the characteristic diketone/triketone group in the active compounds influenced the outcome of the VS. Unfortunately, no HPPD inhibitor lacking the diketone moiety has been reported to date, which makes it impossible to repeat the VS with a query ligand or a set of active compounds belonging to a different chemical series. A partial answer could be obtained by reducing the decoy set to the compounds containing one ketone to cancel the effect of this group. The only consequence of this approach was that ECFP fingerprints performed slightly better than FCFP (all the corresponding graphs can be found in Supporting Information), but globally the performance of all the other methods relative to each other was not altered. A more definite answer could be provided by a decoy set containing only diketone compounds or at least a significant proportion of them. But in that case, the VS methods would be compared in a particularly challenging situation in which the differences between the structures of active and inactive compounds would be rather subtle. In addition, that situation would not be representative any more of the everyday use of VS tools as a support for research projects in industry.

3D Ligand-Based Methods. 3D ligand-based VS methods (Schrödinger's Phase and Openeye's tandem ROCS and EON) were used in this study to screen our database against the reference ligand mesotrione. First, the shape of the compounds alone was considered, using Phase-shape and ROCS with its shape-only option. Then the amount of information exploited by the screening tool was progressively increased by taking into account the characteristics of the atoms present in the molecule structure, using more and more specific atom types in Phase-shape, introducing a color force-field in ROCS, and finally combining ROCS with the electrostatics comparison tool EON. Last, we took advantage of our knowledge of the mode of action of the reference ligand to implement a pharmacophore approach and to scan the database for matches to the pharmacophore with Phase. The results obtained are summarized in the enrichment plots of Figure 5 and in Table 2 (see also Figures 3 and 4 in Supporting Information).

Shape-Only Methods. Among the methods based only on shape similarity (Phase_shape_only and ROCS_shape_only), the Phase-shape program constantly outperformed ROCS (Figure 5a and Table 2); when considering the top 1% of

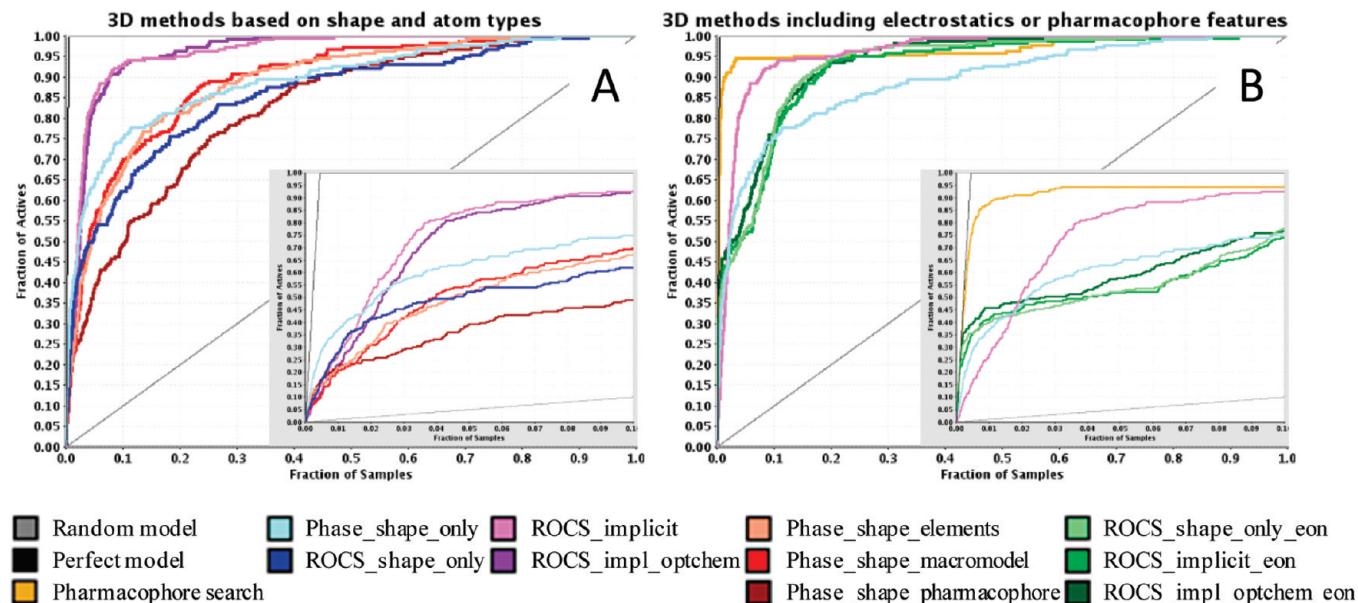


Figure 5. Enrichment plots obtained in 3D ligand-based VS using ROCS/EON and Phase. The embedded panels represent a zoom on the top 10% of the database. (A) Phase-shape, Phase-shape with atom types (elements, Macromodel, or pharmacophore atom types), ROCS in shape-only mode or combined with a color force-field (optimized or not). (B) Phase-shape without atom types, ROCS with a color force-field, reranking by EON of the three ROCS screenings, and search for matches to the pharmacophore with Phase.

the database, Phase-shape retrieved 38% of the actives while ROCS picked 28%. Although they are both based on the same feature of the compounds, the programs estimate the shape of a molecule in different ways: Phase-shape computes the overlap between hard-sphere volumes, whereas ROCS approximates shape with Gaussian functions. ROCS is faster than Phase-shape, however, the use of Gaussian functions might involve “smoother” shapes than the hard-sphere model, and thus, the reference ligand could be found to match with many compounds with very close scores. If no refining by atom types is used, then the prioritization of compounds might be driven by insignificant differences in score, leading to a loss of discrimination power. This could explain the differences in the retrieval of actives observed between both shape-based tools. No significant variation was found concerning the scaffold-hopping possibilities brought by each method (Table 2 in Supporting Information).

Shape and Atom Types. The addition of atom-type features to Phase-shape (Phase_shape_elements, Phase_shape_macromodel, and Phase_shape_pharmacophore in Figure 5a and Table 2) led to a decrease in the retrieval rate for all considered atom types and, in particular, for the pharmacophore one when large parts of the database were considered (75% decreased to 49% actives in the top 10% of the collection). This is consistent with the fact that the pharmacophore atom typing is the most general and least stringent one. Element and macromodel atom types performed comparably and slightly underperformed the selection by shape alone (retrieval rates of 22% and 21% respectively at 1% of the database, compared to 38% for shape alone). In the case of ROCS, the addition of a color force-field to take into account atom types (ROCS_implicit and ROCS_Impl_optchem) did not improve the initial retrieval rate; at 1% of the database, instead of 28% of the active compounds found by shape alone, this method was able to find 27% of the active set using the color force-field and 23% if color gradients and forces were applied. However, the use of a color force-field yielded a much higher enrichment when

10% of the collection was considered (93% of actives retrieved instead of 62%). This was indeed the case even without the use of gradients and forces in the color force-field, which did not seem to make any significant difference. It is interesting to note that the effects of taking the nature of atoms into account when comparing the shape of molecules were very different in both Phase-shape and ROCS. For Phase-shape not only did it not bring any improvement but also the results were slightly worse, whereas in the case of ROCS, the increase in the retrieval rate was quite spectacular. This could be due to the fact that the models underlying the use of atom types are different. In the case of Phase-shape, the atom types of overlapping atoms are compared, whereas in ROCS the comparison does not involve the atoms themselves but the probability-weighted H-bond maps around them (at each point of the ligand, all the possible positions of complementary H-bond donor or acceptor atoms that could be found in the receptor to which the ligand would bind are represented).^{4,55} This approach might be much more efficient in cases where there is no strict atom-to-atom superimposition of the atoms with the same type, although the compounds do bind in the same way.^{54,56}

It could be expected that methods that produce similar enrichment curves for the retrieval of active compounds would also show a very similar profile in scaffold selection. This was indeed the case for ROCS when using a color force-field, since the use of the force-field optimization (optchem) did not change the scaffold distribution. However, although Phase-shape used with elements and Macromodel atom types yielded very similar enrichment plots, there were substantial differences in the ability to identify the scaffolds containing a ring architecture different to that of the query ligand. Macromodel atom types induced a greater difficulty in retrieving the compounds in which the core 6-carbon cycle of mesotrione is replaced by a cyclopropane or an open aliphatic chain (scaffold 14, found at 28.18% of the database instead of 4.93% for element atom types, Table 2 in

Supporting Information). The compounds with an aromatic bridge were also selected much later (scaffold 3 at 7.67% instead of 0.78% for element atom types), whereas those with aliphatic bridges were more easily retrieved (scaffold 2, 21 compounds instead of 12 in the best 1% of the database, data not shown). This could be due to the fact that Macromodel atom types make a distinction between different hybridization states of carbon atoms and also take into account their degree of substitution. Thus, a carbon in a cycle (in the mesotriione head), a methyl carbon (in open aliphatic chains), and an sp² carbon (in the aromatic bridges) are not perceived in the same way using Macromodel types, while they are perfectly equivalent with a simple element typing. This could explain the differences observed between the Phase-shape screening methods.

Shape, Atom Types, and Electrostatics. The reranking of ROCS's output with EON, which is based on the electrostatic similarity with the query (ROCS_shape_only_eon, ROCS_implicit_eon, and ROCS_impl_optchem_eon, Table 2 and Figure 5b), suppressed almost all differences that could be due to the introduction of atom types in ROCS. Indeed, after reranking, the results obtained with an initial screening that takes into account the shape alone or includes a color force-field are almost equivalent. This is not surprising given the fact that EON performs just a rescoring of the compounds in the conformation selected by ROCS and optimizes only the terminal torsions. The differences between ROCS_shape_only_eon, ROCS_implicit_eon, and ROCS_impl_optchem_eon can then only arise from a different conformation choice induced by the color force-field at the superimposition stage done by ROCS. Given that the superimposition is mainly driven by volume overlap, this is not very likely to occur, which explains the uniform results observed with EON. Rescoring with EON considerably improved the retrieval rates when a small fraction (less than 2%) of the database was considered (40–46% of actives retrieved within the top 1% of the collection compared to 27% for the ImplicitMillsDean color force-field alone). However, when considering a larger fraction of the database, EON ranking clearly and constantly underperformed the ImplicitMillsDean model, retrieving only 74–78% of the actives within the top 10% of the database instead of the 93% recovered with the color force-field.

The enrichment curves obtained for the rerankings with EON showed a peculiarity, which is that the slope had a dramatic decrease between 1 and 6% of the database, leading to a plateau (Figure 5b). The analysis of the size of the smallest fraction of the database that needed to be considered, in order to retrieve the first compound of each scaffold (Table 2 in Supporting Information), showed a significant gap between the group of scaffolds that appeared in the first 1% of the database and the rest, which started to be identified after 5.6% of the database was selected. No such gap between the appearance of the various scaffolds existed for the other methods, and this could explain the plateau observed for EON. Among the late-discovered scaffolds were the compounds with aromatic bridges or with cyclopropane and aliphatic chains replacing the 6-carbon cycle (scaffolds 3 and 14, Table 2 in Supporting Information), already mentioned as being more difficult to retrieve when analyzing the results of Phase-shape using Macromodel atom types. In the case of EON, there are in addition chloro-, amino-, and sulfone-

substituted compounds (scaffolds 9 and 10) as well as molecules with a long-chain substituent or containing a conjugated imine in the core (scaffolds 11 and 13). Both the shape and electrostatic properties of these categories of compounds are the most distant from the ones of the query ligand mesotriione, which might be the reason for them being identified so late during the screening. This fact reveals that, although reranking ROCS's results with EON is a very efficient way of enhancing the initial retrieval of active compounds, this method is not very prone to scaffold hopping.

Pharmacophore. Considering the shape and atom properties of the reference ligand to search a database only exploits the information contained in the chemical structure itself. A further step in refining the model before screening could consist in including what is already known about the binding mode of the compound and, in particular, in defining its pharmacophoric features. In the case of HPPD inhibitors, the ligand has to fill the two free chelation sites of the iron II metal center of the protein to have the desired potency. Therefore, any active compound must contain two close H-bond acceptors (Figure 1), in an accurately defined geometry. A rough preliminary SAR had also shown that most of the active compounds had an aromatic ring in their "tail" part. Using this pharmacophore, a search in the database ran with Phase clearly outperformed all the other screening methods, with 88% of the active compounds retrieved in the best 1% of the database (Figure 5b and Table 2).

However, although close to the origin of the curve, this model almost fitted the perfect model, a plateau at 94% of the actives was reached very soon (after selection of 3% of the database), indicating a great difficulty in identifying the last 6% of the active set. It was striking to notice that, at that stage and until around 13% of the database, no imine-containing compounds were found (scaffold 13) and that only half of the pyrazolone structures (scaffold 15) had been identified (data not shown). This was in contrast with most of the other methods. In the case of imines, the imine nitrogen might not be considered a good H-bond acceptor; for scaffold 15, all but one compound were pyrazolones, with a five- instead of a six-membered ring, which led to a slightly different geometry of the diketone. Therefore, these scaffolds did not match perfectly the pharmacophore hypothesis, which could explain why the corresponding active compounds were not easily selected.

The performance of this pharmacophore search was extremely good, both in terms of retrieval rates and scaffold diversity. However, the fact that the pharmacophore was very clearly defined and present in almost all active molecules under exactly the same substructure probably made this test system very easy for a pharmacophore search. Therefore, this result might not be applicable to other cases.

Comparison between the Different Categories of 3D Ligand-Based Screening. Excluding the pharmacophore search, which gave the best results but is probably not general enough, the best 3D ligand-based method in our study appeared to be ROCS with the ImplicitMillsDean color force-field, which retrieved 93% of actives within the 10% best-scoring compounds of the collection. For an enhanced screening utility, it could be used with color force-field forces and gradients and combined with reranking with EON to

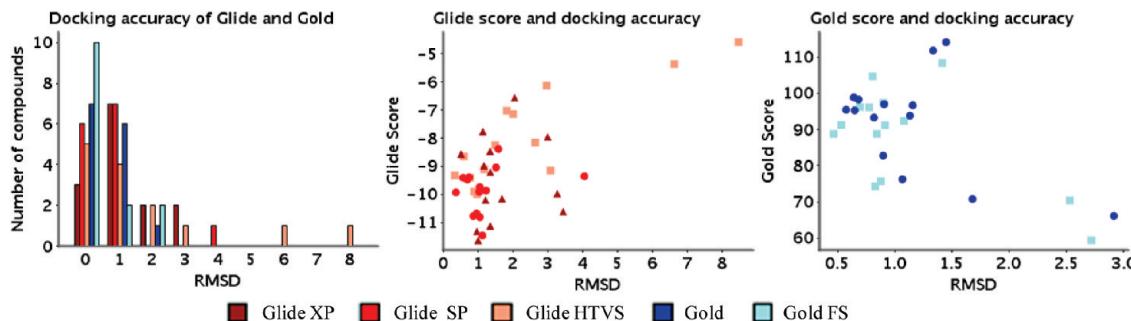


Figure 6. Assessment of the docking accuracy of Glide and Gold with several speed settings for 14 compounds docked in their native crystal structures: extra precision (XP), standard precision (SP), and high-throughput (HTVS) modes for Glide; default settings and 10 000 GA operations (FS) for Gold. The left panel shows the distribution of rmsd values, and the middle and right panels show the correlation between docking scores and rmsd values.

increase the retrieval rate at 1% of the database up to 46% of the active compounds (instead of 27% with the color force-field alone).

The fact that some categories of compounds are more dissimilar to the query ligand defines them as being more difficult to identify by methods based on similarity to the query. Indeed, they are the last ones to be selected not only by EON but also by Phase-shape when taking atom types into account. However, they appear much earlier in the screening with ROCS with both color force-field and gradients, and even slightly earlier when no optimization of the color force-field is used. And in all cases, the first methods to identify compounds from these classes are the ones based on shape only (scaffolds 9, 13, and 14 are representative examples of this trend, see Table 2 in Supporting Information). It seems, therefore, that, for the set of active compounds used in this study, the more strongly defined the atom types are, the less efficient at scaffold hopping the method is (such observation has already been reported in the literature).³⁴

Target-Based Methods. Since crystal structures of HPPD were available, target-based VS methods have also been applied to screen our database. Before proceeding to the VS itself, we evaluated the ability of Gold and Glide to reproduce the native ligand pose in the crystal structure used for VS and in 13 additional crystal structures of HPPD in complex with different inhibitors of the active set. These assessments are commonly performed using the default genetic algorithm settings for Gold, and standard (SP) and extra-precision (XP) docking for Glide.³⁷ However, since the VS was performed using the fastest available settings in both docking software, we found it interesting to include in our analysis of docking accuracy the results obtained with the fast presets for Gold (Gold FS) and the high-throughput mode of Glide (HTVS) as well. In order to evaluate the true predictive capability of these methods, only the best-ranked pose was considered, as already done in other studies.³⁸ The distribution of rmsd values between the docked pose and the position of the ligand in the crystal as well as the correlation between the docking score and the rmsd for all five methods (Gold, Gold FS, Glide XP, Glide SP, and Glide HTVS) are presented in Figure 6.

When used with the standard low-throughput settings, Gold and Glide SP performed in a very comparable manner and were able to reproduce the ligand pose found in the crystal in most of the cases (only one ligand out of 14 with rmsd > 2 Å). However, Glide XP underperformed Glide SP in this series of compounds, with 4 out of 14 structures

showing rmsd > 2 Å. Such weaker performance of the XP mode as compared to SP has already been reported³⁷ and was attributed to a lower ability of Glide XP to accommodate slight steric overlaps in the active site. Surprisingly, performing 10-fold fewer operations in the genetic algorithm of Gold (Gold FS) did not yield worse results, since only an additional ligand was found with rmsd > 2 Å, and most of the structures were found within 1 Å of the crystal pose. On the contrary, the high-throughput mode of Glide (Glide HTVS) produced 5 out of 14 structures with rmsd > 2 Å, among which two were very inaccurately docked (rmsd > 6 Å). The fact that Glide HTVS leads to a significantly less accurate docking has already been observed³⁷ and interpreted as the consequence of the reduced conformational sampling that is used in the HTVS method in order to increase docking speed. For this set of ligands and crystals structures, it appears, therefore, that the best methods for docking accuracy are Glide SP and Gold (with either standard or FS settings).

For docking runs with Gold and Gold FS, although the correlation between the score values and the rmsd is almost nonexistent, a cutoff of the score value at 71 discriminates between the ligands positioned with rmsd < 1.5 Å and the others. This would also be the case for Glide SP (cutoff at -9.35). However, such discrimination between accurate and inaccurate poses based on the score is not possible for Glide XP and Glide HTVS (except for the poses with rmsd around 6 and 8 Å). Here again, although no correlation exists between score and rmsd values, Gold, Gold FS, and Glide SP show the highest discrimination power between correct and incorrect ligand poses.

Given the conclusions drawn from the assessment of docking and scoring accuracy, it seemed reasonable to perform the VS using not only the fastest settings for each program (Gold FS and Glide HTVS) but also the standard ones for Glide (Glide SP). The results, summarized in Table 2 and Figure 7a, show that Gold FS and Glide SP clearly outperformed Glide HTVS. Gold FS also performed better than Glide SP once 6% of the database had been selected (75% of actives retrieved by Gold FS in the top 10% of the database compared to 66% for Glide SP). And, more importantly, Gold FS also yielded the best results as far as the initial retrieval rate is concerned (within the top 1% of the database). Such results are in disagreement with the fact that Glide, in its SP mode, has been claimed several times to perform better than Gold.^{6,33} The underperformance of Glide's HTVS mode as compared to Gold FS and Glide SP was to be expected, according to the conclusions drawn

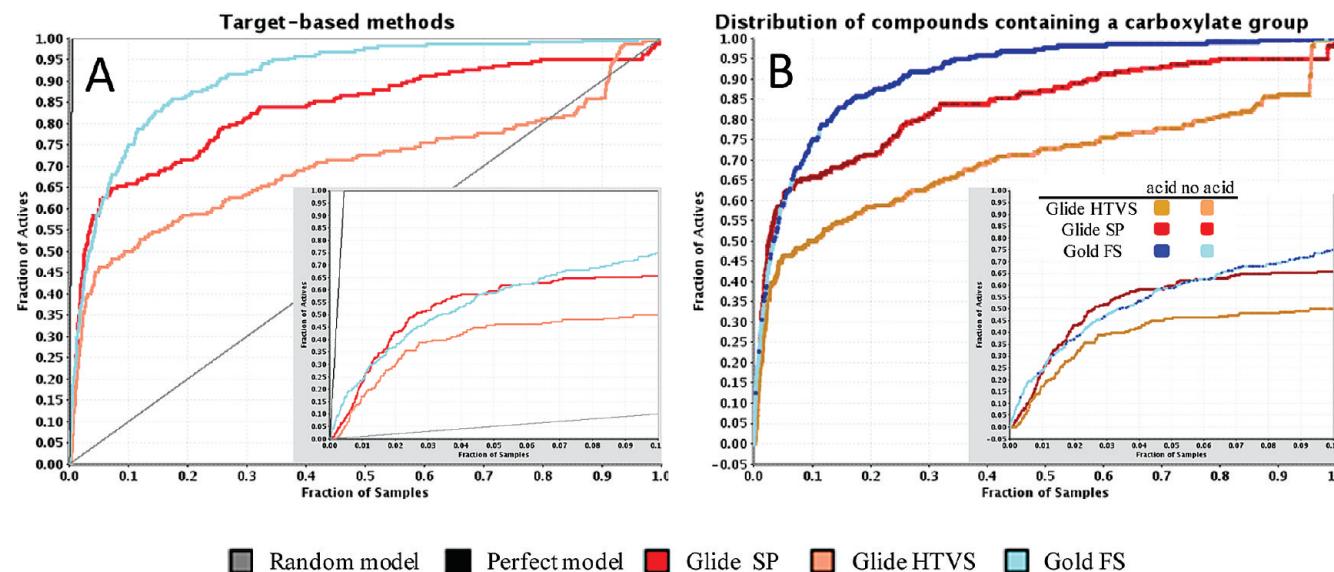


Figure 7. (A) Enrichment plots obtained in target-based VS using Gold and Glide. (B) Distribution of the carboxylic acid-containing compounds according to their ranking with Gold and Glide. In (A) and (B) the embedded panel represents a zoom on the best 10% part of the database.

above from the analysis of the correlation between the score values and the rmsd. In such correlation, Glide HTVS showed indeed a lower ability to correctly dock the active compounds in the active site, and 6 out of 14 actives were docked with high rmsd and low score (<-8), which means that this method could hardly identify them as active. It is not very surprising then that active compounds were also retrieved with difficulty during the VS.

A difference between Gold and Glide that might influence the outcome of the VS in our study is related to the particular nature of the HPPD protein, which contains a metal center. In Glide, for proteins in which the net metal charge in the *apo* protein is positive, only the interactions with anionic acceptor atoms of the ligand are included in the score. This is the case of HPPD, in which the iron is under the form Fe^{2+} . Visual inspection of Glide's best scored compounds showed that most of them contained a carboxylic acid moiety, which led us to analyze the relationship between the ranking of compounds and the presence of such a functional group in both Gold and Glide results (Figure 7b). It is quite obvious that the compounds retrieved in the best 1% of the database with Glide are mostly carboxylic acids, whereas with Gold the frequency of this functional group is much lower and its distribution more dispersed. This advantage given by the Glide scoring function to negatively charged compounds, and in particular to carboxylic acids, explains why this class of molecules is so heavily represented among the best-ranked compounds. It is not clear to what extent this different treatment of anions and neutral molecules influenced the outcome of the VS. Such a difficulty had not been mentioned either by other VS studies in which Glide was used for metalloproteins.^{37,38} A posteriori, we could think of modifying the scoring function or imposing strong constraints on the geometry around the iron to penalize the binding of acids. However, we chose not to adapt the screening methods to the targets, in order to place ourselves in a more realistic situation in which the user does not have prior experience or knowledge of the target or the software.^{5,37,42}

As far as the scaffold-hopping ability of these target-based methods is concerned, the analysis of the smallest fraction

of the database that has to be considered in order to select the first compound from each of the scaffolds defined in Figure 3 (Table 3 in Supporting Information) shows that some scaffolds are more easily retrieved by Gold (5, 7, and 9), whereas others appear sooner when using Glide SP (6, 8, 10, 11, and 15). Although no rationale could be found for these differences, in this case Glide SP seems to be slightly more prone to scaffold-hopping than Gold FS. However, this trend is not marked enough for this conclusion to be generalized to other systems.

Comparison of Ligand- And Target-Based Methods. *Retrieval Rates.* The increased efficiency of 3D versus 2D methods as well as target- versus ligand-based approaches is an often addressed issue.^{5,10} Indeed, although the amount of information used and the computational requirements of the different VS tools increase from 2D and 3D ligand-based to target-based methods, their outcome does not seem to be always correlated with the additional complexity they involve. In order to assess how the various categories of methods performed relatively to each other in our case, it was useful to represent in the same graph the retrieval rates of active compounds for the best method of each family (Figure 8).

The best method overall was the pharmacophore search using the Phase software, but as we have already pointed out, this test case was probably not representative enough for this approach. Among the other methods, 3D ligand-based VS performed with ROCS using a color force-field was the best choice, combined with a color force-field gradient and a reranking with EON for the first fragment of the database (until 20%) or on its own for larger fractions. The best target-based method was the docking program Gold, which performed less well than 3D ligand-based screening; Gold retrieved 24% of active compounds within the top 1% of the database, instead of 46% selected using ROCS with force-field gradients and EON. In the best 10% of the database, 75% of the active set was found by Gold, instead of 93% with ROCS and the color force-field. The 2D fingerprint was the least appropriate technique to retrieve active compounds from the database.

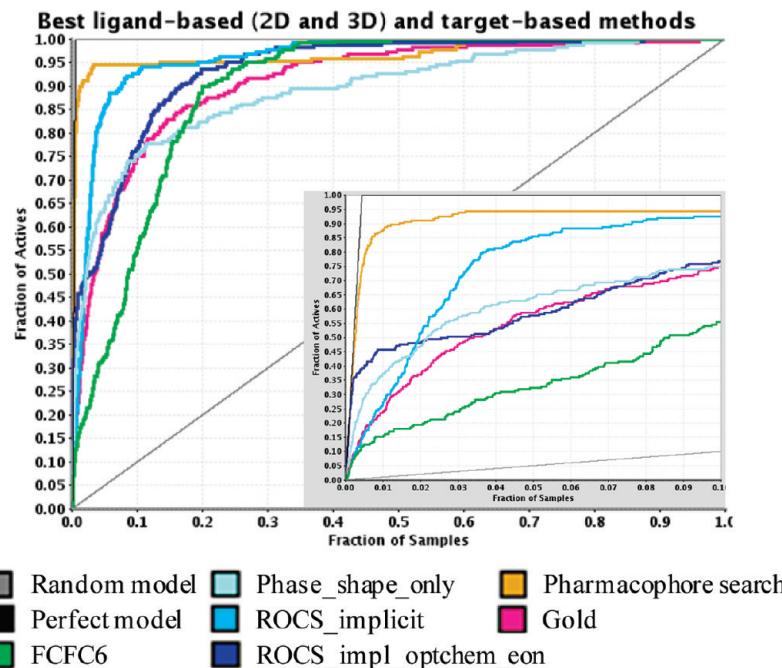


Figure 8. Enrichment plots obtained with the best 2D and 3D ligand-based and target-based methods. The embedded panel represents a zoom on the best 10% part of the database.

Therefore, in our study, much more CPU-demanding and time-consuming calculations, such as docking, did not result in an improvement of the retrieval rates. Other comparison studies had already led to the same conclusion that 3D ligand-based methods were more efficient than docking for database screening.^{5,35,38,40} On the contrary, among ligand-based methods, the increase in complexity and information used correlated well with an increase in the retrieval of actives, since an improvement was seen from 2D tools to 3D shape-based methods and from these to 3D techniques that take into account atom types. This is also in agreement with other studies.³⁴

Scaffold Hopping. The scaffolds that are closer in structure to the reference ligand were very easily retrieved by all methods, whereas those more chemically diverse were found much later during the database selection. Although within both 2D and 3D ligand-based methods a more accurate definition of atom types seemed to reduce the scaffold-hopping possibilities, no global trend in the ability of selecting diverse scaffolds could be seen between 2D, 3D, and target-based tools, which appeared to be equivalent in terms of scaffold hopping.

CONCLUSION

The virtual screening (VS) of an “agro-like” database for inhibitors of the herbicidal target HPPD provided a good opportunity for us to compare the VS methods that we have available in house and to adapt these techniques and workflows, already well tested in pharma, to a real research situation in the agrochemical industry. The results obtained showed that, in our particular case in which the pharmacophore of the reference ligand was very well defined, a pharmacophore search performed with the Phase pharmacophore tool gave the best results in the retrieval of active compounds. This method was followed by 3D shape- and atom-type similarity screening using ROCS combined with EON, which outperformed docking with Gold and screening

with 2D extended connectivity fingerprints. The very little studied software Phase-shape, which is also a program based on shape similarity, outperformed ROCS on shape-only screenings and showed good early retrieval rates. No significant differences in the scaffold-hopping abilities of 2D, 3D, and target-based methods could be highlighted, although within ligand-based methods, it could be noticed that increasing the distinction between atom types led to a decrease in the diversity of the scaffolds retrieved. The strong characteristic features of the reference ligand and the active set might have led to conclusions that are not general enough for the pharmacophore and the 2D fingerprint approaches. However, the results obtained with target-based and 3D shape-based tools highlighted the specificities of each method and their different suitability depending on the aim of the screening. These conclusions are in agreement with other published results and could probably be generalized by extending the study to other targets.

The major difficulty we faced when performing such a comparison of VS tools in an agrochemical context was the design of a suitable and agro-like decoy set. This database and the workflows we have implemented can now be used again internally; the decoy used for this publication can be further exploited as a basis to build models for current projects or as a training set to validate new computational procedures or software.

ACKNOWLEDGMENT

We thank Dr. Jane Wibley and Dr. Russell Viner from Syngenta Crop Protection, U. K., for providing the HPPD crystal structures and their analysis.

Supporting Information Available: Tables containing the smallest percentage of the database to be considered to select the first compound from each scaffold category for the analysis of the diversity of the retrieved active compounds

and full-size views of enrichment graphs. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- (1) Venhorst, J.; Núñez, S.; Terpstra, J. W.; Kruse, C. G. *J. Med. Chem.* **2008**, *51*, 3222–3229.
- (2) Bissantz, C.; Folkers, G.; Rognan, D. *J. Med. Chem.* **2000**, *43*, 4759–4767.
- (3) Schulz-Gasch, T.; Stahl, M. *J. Mol. Model.* **2003**, *9*, 47–57.
- (4) Cummings, M. D.; DesJarlais, R. L.; Gibbs, A. C.; Mohan, V.; Jaeger, E. P. *J. Med. Chem.* **2005**, *48*, 962.
- (5) McGaughey, G. B.; Sheridan, R. P.; Bayly, C. I.; Culberson, J. C.; Kreatsoulas, C.; Lindsley, S.; Maiorov, V.; Truchon, J.-F.; Cornell, W. D. *J. Chem. Inf. Model.* **2007**, *47*, 1504–1519.
- (6) Zhou, Z.; Felts, A. K.; Friesner, R. A.; Levy, R. M. *J. Chem. Inf. Model.* **2007**, *47*, 1599–1608.
- (7) Sheridan, R.; McGaughey, G.; Cornell, W. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 257–265.
- (8) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.* **2005**, *49*, 5912–5931.
- (9) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.* **2006**, *49*, 5912–5931.
- (10) Hawkins, P. C. D.; Skillman, A. G.; Nicholls, A. *J. Med. Chem.* **2006**, *50*, 74–82.
- (11) Moffat, K.; Gillet, V. J.; Whittle, M.; Bravi, G.; Leach, A. R. *J. Chem. Inf. Model.* **2008**, *48*, 719–729.
- (12) Hristozov, D.; Oprea, T.; Gasteiger, J. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 617–640.
- (13) Hert, J.; Willett, P.; Wilton, D. J.; Acklin, P.; Azzaoui, K.; Jacoby, E.; Schuffenhauer, A. *Org. Biomol. Chem.* **2004**, *2*, 3256–3266.
- (14) Raymond, J. W.; Willett, P. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 59–71.
- (15) Clarke, E. D.; Delaney, J. S. *Chimia* **2003**, *57*, 731–734.
- (16) Colin, M. T. *Pest Manage. Sci.* **2001**, *57*, 3–16.
- (17) Yang, C.; Pfleiderer, J. W.; Camper, D. L.; Foster, M. L.; Pernich, D. J.; Walsh, T. A. *Biochemistry* **2004**, *43*, 10414–10423.
- (18) Dayan, F. E.; Duke, S. O.; Sauldubois, A.; Singh, N.; McCurdy, C.; Cantrell, C. *Phytochemistry* **2007**, *68*, 2004–2014.
- (19) Beaudegnies, R.; Edmunds, A. J. F.; Fraser, T. E. M.; Hall, R. G.; Hawkes, T. R.; Mitchell, G.; Schaetzer, J.; Wendeborn, S.; Wibley, J. *Bioorg. Med. Chem.* **2009**, *17*, 4134–4152.
- (20) van Almsick, A. *Outlooks Pest Manage.* **2009**, *20*, 27–30.
- (21) Gold, version 4.0; The Cambridge Crystallographic Data Centre: Cambridge, U.K., 2010.
- (22) Marcel, L. V.; Jason, C. C.; Michael, J. H.; Christopher, W. M.; Richard, D. T. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 609–623.
- (23) Jones, G.; Willett, P.; Glen, R. C. *J. Mol. Biol.* **1995**, *245*, 43–53.
- (24) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (25) Glide, version 5.0; Schrödinger, LLC: New York, NY, 2008.
- (26) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (27) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. *J. Med. Chem.* **2004**, *47*, 1750–1759.
- (28) SciTegic, version 7.5; Accelrys Inc.: San Diego, CA, 2010.
- (29) Phase, version 3.1; Schrödinger, LLC: New York, NY, 2009.
- (30) 30 ROCS, version 2.4.1. and EON, version 2.0.1; OpenEyeScientific Software, Inc.: Santa Fe, NM, 2010.
- (31) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. *J. Comput. Chem.* **1996**, *17*, 1653–1666.
- (32) Rush, T. S.; Grant, J. A.; Mosyak, L.; Nicholls, A. *J. Med. Chem.* **2005**, *48*, 1489–1495.
- (33) Hevener, K. E.; Zhao, W.; Ball, D. M.; Babaoglu, K.; Qi, J.; White, S. W.; Lee, R. E. *J. Chem. Inf. Model.* **2009**, *49*, 444–460.
- (34) Tresadern, G.; Bemporad, D.; Howe, T. *J. Mol. Graphics Modell.* **2009**, *27*, 860–870.
- (35) Perez-Nueno, V. I.; Ritchie, D. W.; Rabal, O.; Pascual, R.; Borrell, J. I.; Teixido, J. *J. Chem. Inf. Model.* **2008**, *48*, 509–533.
- (36) Kirchmair, J.; Ristic, S.; Eder, K.; Markt, P.; Wolber, G.; Lagner, C.; Langer, T. *J. Chem. Inf. Model.* **2007**, *47*, 2182–2196.
- (37) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humbert, C. *J. Chem. Inf. Model.* **2009**, *49*, 1455–1474.
- (38) Chen, H.; Lyne, P. D.; Giordanetto, F.; Lovell, T.; Li, J. *J. Chem. Inf. Model.* **2006**, *46*, 401.
- (39) Wang, Z.; Lu, Y.; Seibel, W.; Miller, D. D.; Li, W. *J. Chem. Inf. Model.* **2009**, *49*, 1420–1427.
- (40) Hawkins, P. C. D.; Skillman, A. G.; Nicholls, A. *J. Med. Chem.* **2007**, *50*, 74–82.
- (41) Tiikkainen, P.; Markt, P.; Wolber, G.; Kirchmair, J.; Distinto, S.; Poso, A.; Kallioniemi, O. *J. Chem. Inf. Model.* **2009**, *49*, 2168–2178.
- (42) Jain, A. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 201–212.
- (43) Cole, J. C.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 325.
- (44) Schrödinger, www.schrodinger.com.
- (45) Omega, version 2.3.2; Openeye Scientific Software, Inc.: Santa Fe, NM, 2010.
- (46) Hermes; The Cambridge Crystallographic Data Centre: Cambridge, U.K., 2010.
- (47) Hu, Y.; Lounkine, E.; Bajorath, J. *ChemMedChem* **2009**, *4*, 540–548.
- (48) Rogers, D.; Brown, R. D.; Hahn, M. *J. Biomol. Screening* **2005**, *10*, 682–686.
- (49) Bender, A.; Jenkins, J. L.; Scheiber, J.; Sukuru, S. C. K.; Glick, M.; Davies, J. W. *J. Chem. Inf. Model.* **2009**, *49*, 108–119.
- (50) MDL keys, MDL Information Systems: San Leandro, CA, 2010.
- (51) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425–445.
- (52) Truchon, J.-F.; Bayly, C. I. *J. Chem. Inf. Model.* **2007**, *47*, 488–508.
- (53) Triballeau, N.; Acher, F.; Brabet, I.; Pin, J.-P.; Bertrand, H.-O. *J. Med. Chem.* **2005**, *48*, 2534–2547.
- (54) Mills, J. E. J.; Dean, P. M. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 607–622.
- (55) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (56) Perkins, T. D. J.; Mills, J. E. J.; Dean, P. M. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 479–490.

CI900498N