JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

Improving VEGFR-2 Docking-Based Screening by Pharmacophore Postfiltering and Similarity Search Postprocessing

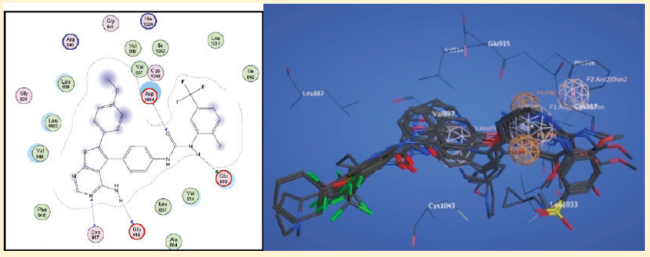
Jesús M. Planesas,[‡] Rosa M. Claramunt,[‡] Jordi Teixidó,[†] José I. Borrell,[†] and Violeta I. Pérez-Nueno[†,*]

[†]Grup d'Enginyeria Molecular, Institut Químic de Sarriá (IQS), Universitat Ramon Llull, Barcelona, Spain

[‡]Departamento de Química Orgánica y Bio-Orgánica, Facultad de Ciencias, UNED, Senda del Rey 9, E-28040 Madrid, Spain

Supporting Information

ABSTRACT: Conventional docking-based virtual screening (VS) of chemical databases is based on the ranking of compounds according to the values retrieved by a scoring function (typically, the binding affinity estimation). However, using the most suitable scoring function for each kind of receptor pocket is not always an effective process to rank compounds, and sometimes neither to distinguish between correct binding modes from incorrect ones. To improve actives from decoys selection, here we propose a three-step VS protocol, which includes the conventional docking step, a pharmacophore postfilter step, and a similarity search postprocess. This VS protocol is retrospectively applied to VEGFR-2 (Kdr-kinase) inhibitors. The resulting docking poses calculated using the Alpha HB scoring function implemented in MOE are postfiltered according to defined pharmacophore interactions (structure based). The selected poses are again ranked according to their molecular similarity (MACCS fingerprint) to the cognate ligand. Results show that both the overall and early VS performance improve the application of this protocol.



INTRODUCTION

The vascular endothelial growth factor (VEGF) is responsible for regulation of blood vessel growth processes starting from preexisting blood vessels (angiogenesis).[1] There is evidence and clinical observations which confirm that processes of abnormal angiogenesis can cause various diseases such as diabetic retinopathy, psoriasis, rheumatoid arthritis, macular degeneration, and cancer. In particular, it is considered that metastasis and growth of solid tumors largely depend on angiogenesis. Among the different kind of factors, angiopoietins[2] and VEGF are particularly interesting because they are specific factors of endothelial cells and they play important and coordinated roles in vascular development.[3]

Inside the family of VEGF factors we find their receptors, VEGFR. Those are protein kinases that modulate angiogenesis.[4] There are different kinds of kinase receptors: VEGFR-1, VEGFR-2 (or KDR/Flk1), VEGFR-3, and VEGF-A to VEGF-E. While VEGFR-1 is critical for the development of angiogenesis and VEGFR3 is restricted to lymphatic endothelial cells, VEGFR-2 seems to mediate in the majority of angiogenesis and pathological effects at high levels (processes like cell proliferation, differentiation of endothelial cells, invasion, or migration).[5,6] Therefore, the development of VEGFR-2 inhibitors[7,8] has been manifested as a potent treatment for metastatic tumor growth and cancer.[9]

Conventional structure-based docking[10] drug design aims to predict the preferred conformation (pose) of a molecule (ligand) in relation to a target receptor (protein) when both bind to form a stable complex. The function which measures and evaluates the goodness-of-fit for protein docking is the scoring function. Recently, several papers have been published[11,12] to evaluate, in a comparative manner, docking performance for different targets using different scoring functions. Authors evaluate these scoring functions using in silico databases of known ligands and decoys, such as DUD database[13] or MUV.[14] The purpose of these papers is to study how the different scoring functions can distinguish ligands from decoys (retrospective screening), in order to find the scoring function that should be used to perform prospective docking-based virtual screening (VS) with the highest performance.

This work focuses on the retrospective docking of VEGFR-2 protein with known ligands and decoys[15] obtained from the DUD database. The computational software MOE[16] has been used, and the most suitable scoring function for VEGFR-2 inhibitors has been selected. The objective is to improve VS when scoring functions themselves do not reach high specificity to distinguish actives from decoys, as it happens in many retrospective dockings when dealing with difficult targets, mainly the ones displaying large binding sites, such as VEGFR-2, where it is difficult to accommodate all ligand conformations, and in which multiple ligands can bind in different subsites. Some authors have reported docking-based VS experiments on VEGFR-2 showing poor performance; for example, Kirchmair et al. have reported

Received: July 17, 2010
Published: March 18, 2011

AUC values from 0.43 to 0.48[17] and Tawa et al. have performed[18] VEGFR-2 VS based in Rapid Overlay of Chemical Structures (ROCS) obtaining an AUC of 0.61. These low values are due to the difficulty in finding a common query matching all different VEGFR-2 ligands, given that they bind in multiple pocket subsites. Hence, here we propose a three-step VS protocol applied to VEGFR-2 inhibitors, which includes the conventional docking step, a pharmacophore postfilter step, and a similarity search postprocess to improve actives from decoys selection. Our protocol might also work for other difficult targets given that it is taking into account not only docking information but also receptor-based pharmacophore filter and ligand similarity searching, which can help to determine important features for pocket subsite binding.

Table 1. Ligands and Decoys Molecular Properties: Summary of the 1D Physicochemical Properties of Active and Decoy Molecules in the VEGFR-2 DUD Screening Database[a]

	ligands: 78	decoys: 2479
molecular weight:	238.3/692.6 (86.5)	325.4/474.2 (21.6)
rotable single bonds	0.000/0.294 (0.064)	0.057/0.286 (0.037)
hydrogen bond acceptors	1/7 (1.3)	0/7 (0.9)
hydrogen bond donors	0/5 (1.2)	0/4 (0.7)
hydrophobic atoms	12/32 (3.7)	11/25 (1.7)
octanol−water partition coeff	0.25/6.79 (1.17)	1.33/5.74 (0.72)
polar surface area	34.2/374.2 (54.2)	19.1/288.2 (37.6)

[a] Similar molecular values between ligands and decoys avoid bias in docking process. Standard deviation values in parentheses.

METHODS

Crystal Structure and Docking Scoring Function Selection. To date, all known protein kinase structures have been determined by protein crystallography.[19] In the RCSB Protein Databank[20] we can find more than 20 crystal structures (PDB) for protein kinase VEGFR-2, all determined by X-ray crystallography, with different resolutions and different ligands, and in some cases with some mutations in the protein sequences.

Table 2. Self-Docking Validation[a]

scoring function: ASE					scoring function: Affinity dG				
PDB[b]	Alpha Triangle	Alpha PMI	Proxy Triangle	Triangle Matcher	PDB	Alpha Triangle	Alpha PMI	Proxy Triangle	Triangle Matcher
1Y6A	2.4	5.1	4.6	4.6	1Y6A	4.7	5.1	4.5	3.6
1YWN	1.4	0.9	0.8	0.9	1YWN	2.5	0.9	3.3	3.3
2OH4	0.9	1.2	0.8	0.8	2OH4	0.9	1.2	0.8	0.8
2P2H	3.0	4.5	1.4	1.5	2P2H	3.2	1.4	1.4	1.4
2QU5	0.9	0.7	0.8	0.8	2QU5	1.3	0.7	0.8	0.8
2RL5	0.8	0.3	0.9	0.9	2RL5	1.1	0.2	1.0	0.9
3BE2	1.6	0.9	1.8	1.8	3BE2	3.8	0.9	1.6	1.6
3C7Q	2.4	5.5	2.2	5.5	3C7Q	1.7	5.5	2.7	2.7
3CJF	3.2	5.8	3.3	3.3	3CJF	4.1	6.0	1.2	2.4
3EWH	1.5	3.0	1.7	1.2	3EWH	2.3	3.0	2.8	2.8
3CJG	3.5	7.3	2.2	2.2	3CJG	2.5	7.5	2.9	2.9
aRMSD[c]	2.0	3.2	1.9	2.1	aRMSD	2.6	3.0	2.1	2.1
Inf. 2 Å[d]	6	5	7	7	Inf. 2 Å	4	6	6	5
Scoring function: Alpha HB					**Scoring function: London dG**				
PDB	Alpha Triangle	Alpha PMI	Proxy Triangle	Triangle Matcher	PDB	Alpha Triangle	Alpha PMI	Proxy Triangle	Triangle Matcher
1Y6A	1.3	5.1	3.6	3.6	1Y6A	4.5	5.1	3.8	3.4
1YWN	0.8	0.9	0.8	0.8	1YWN	1.8	0.9	0.9	1.6
2OH4	1.5	1.2	0.8	0.8	2OH4	0.9	1.2	0.8	0.8
2P2H	0.8	5.4	1.4	1.4	2P2H	1.2	5.4	2.8	1.4
2QU5	0.8	0.7	1.1	0.8	2QU5	1.3	0.7	0.8	0.8
2RL5	1.1	0.2	0.9	0.9	2RL5	1.1	0.2	1.0	1.0
3BE2	1.2	0.9	1.6	1.6	3BE2	2.1	0.9	1.8	1.6
3C7Q	1.3	5.5	2.4	2.1	3C7Q	4.2	5.5	2.4	2.2
3CJF	3.2	6.0	3.4	3.4	3CJF	3.1	6.0	3.2	3.2
3EWH	1.8	3.0	1.2	1.2	3EWH	2.5	2.8	2.6	2.6
3CJG	0.9	7.5	1.7	1.7	3CJG	1.1	7.1	1.7	1.7
aRMSD	1.3	3.3	1.7	1.7	aRMSD	2.2	3.3	2.0	1.9
Inf. 2 Å	10	5	8	8	Inf. 2 Å	6	5	6	7

[a] Best RMSD values obtained when the 11 cognate ligands were docked in their corresponding PDBs protein structures using the different placement functions for each scoring function. Scoring function Alpha HB and placement function Alpha Triangle obtain the lower aRMSD (1.3 Å) and the highest number of best poses whose RMSD value is below or equal 2.0 Å. All values are expressed in angstroms. [b] PDBs used do not contain allosteric ligands. [c] Average rmsd. [d] Number of best poses with a rmsd value equal or below 2 Å.

dx.doi.org/10.1021/ci1002763 |J. Chem. Inf. Model. 2011, 51, 777–787

The first VEGFR-2 PDB complex published was 1Y6A (07-06-05), with a crystal resolution of 2.2 Å. This protein structure contains as cognate ligand a 2-anilino-5-(2-aryl)oxazole. This ligand has been assayed and tested[21] to inhibit VEGFR-2 protein (IC_{50} = 0.022 μM). We used this structure, together with 10 more recent VEGFR-2 structures published to date, to select the most appropriate MOE scoring function for docking compounds into VEGFR-2 pocket. The 11 VEGFR-2 PDB complexes were downloaded, crystallized water molecules were eliminated in order to avoid possible hindrances during docking runs,[22] and the proteins were protonated. MOE allows working with several scoring and placement functions to perform protein−ligand docking such as ASE, Affinity dG, Alpha HB, and London dG. The reliability of these docking scoring functions was assessed by a self-docking and a cross-docking analysis.

From the 11 existent VEGFR-2 PDB complexes, we selected one for our VS experiments by also applying a self-docking analysis as a first step, and a cross-docking study as a second step. Both analyses used the best scoring and placement functions found in the previous scoring function validation. If the average rmsd (aRMSD) of the top 10 ranked poses was equal or below 2.0 Å, the PDB complex was considered suitable for VEGFR-2 docking-based VS.

Self-Docking Analysis. The ligands for each of the 11 VEGFR2 complexes used were docked back into their corresponding protein structures using ASE, Affinity dG, Alpha HB, and London dG as docking-scoring functions and Alpha PMI, Alpha Triangle, Proxy Triangle, and Triangle Matcher as docking-placement funcions implemented in MOE. The docking results were evaluated through comparison of the best docked ligands binding modes with the experimental ones. As a measure of docking reliability, the root-mean-square deviation (rmsd)[11] was used to compare differences between the atomic distances of the docked poses and the real cocrystallized pose. An evaluation of the most suitable scoring-placement function was assessed attending to the aRMSD and the number of best poses with rmsd values below or equal to 2.0 Å, obtained in both cases for the 11 PDBs.

Cross-Docking Analysis[23]. To perform cross-docking evaluation, all VEGFR-2 complexes were aligned with each other, and then all ligands were docked into all complex structures. The docking reliability was evaluated by calculating for each ligand and each complex structure the rmsd between the reference position of the ligand in the experimental protein−ligand complex and that predicted by the docking function in the various VEGFR-2 complex structures. The aRMSD comparison and the number of best poses with rmsd values below or equal to 2.0 Å were carried out equally as described above for the self-docking studies.

Finally, in order to select the working structure we analyzed the statistical values obtained when all ligands were docked into their corresponding structure using the selected docking function and also taking into account their crystallographic resolution.

Dockable Database Preparation. We used VEGFR-2 test set in the DUD database[11,12,22,24,25] for docking-based VS. This test set contains 78 ligands and 2479 decoys with similar physical properties to the actives, such as molecular weight, octanol−water partition coefficient (logP), polar surface area, number of molecular rotable single bonds, or minimum number of hydrogen bond donors and acceptors,[26] in order to avoid bias in docking results[27] (see Table 1). All ligands and decoys were energy minimized using the MMFF94x forcefield in MOE.

Table 3. Summary of the Cross-Docking Validation[a−d]

	aRMSD	rmsd std dev	no. of poses ≤2 Å
ASE−Alpha PMI	4.9	2.0	10
ASE−Alpha Triangle	3.7	2.0	29
ASE−Proxi Triangle	3.3	1.7	33
ASE−Triangle Matcher	3.3	1.8	40
Affinity dG−Alpha PMI	4.9	1.9	11
Affinity dG−Alpha Triangle	4.0	2.2	22
Affinity dG−Proxi Triangle	3.6	2.0	29
Affinity dG−Triangle Matcher	3.6	1.9	26
Alpha HB−Alpha PMI	5.1	1.9	5
Alpha HB−Alpha Triangle	3.0	1.6	45
Alpha HB−Proxi Triangle	3.1	1.7	40
Alpha HB−Triangle Matcher	3.1	1.8	44
London dG−Alpha PMI	4.9	1.9	6
London dG−Alpha Triangle	3.1	1.8	41
London dG−Proxi Triangle	3.0	1.6	43
London dG−Triangle Matcher	3.1	1.7	40

[a] PDB used do not contain allosteric ligands. [b] All values are expressed in angstroms. [c] Protein−ligand complexes were aligned prior to perform cross-dockings, and the resulting docked poses were compared with their respective cognate pose. [d] Averaged RMSD (aRMSD) values and number of best poses with a RMSD equal or below 2 Å (no. of poses ≤2 Å) obtained when cognate ligands were docked in the different PDBs protein structures using the different scoring functions (bold) and placement functions. See Table S-1 in Supporting Information for details.

Similarity Search and Receptor-Based Pharmacophore. The ligand similarity search was performed comparing 2D MACCS fingerprints, and the Tanimoto index was used as scoring function.

The receptor-based phamacophore was built using the "Pharmacophore consensus" tool implemented in MOE, which finds the most common interactions resulting from the alignment of the protein residues in the published VEGFR-2 PDB complexes. Pharmacophore feature radii were manually optimized by executing several dockings followed by pharmacophore postfiltering in order to maximize the ratio ligands/decoys.

Assessment of Docking-Based VEGFR-2 Virtual Screening. Following the docking calculations, all compounds were sorted into ranked lists based upon their docking score, pharmacophore postfilter score, and similarity postprocessing score. These lists were then used to plot the percentage of known actives found versus the percentage of the ranked database screened (enrichment plot), and to calculate the enrichment factor (EF) at 1%, 5%, and 10% of false positives selected.[28] The EF measures the number of known ligands in the top-ranked list, relative to a random selection.[29] Moreover, Receiving Operating Characteristic (ROC) plots were calculated. ROC plots represent sensitivity, or the fraction of true positives recovered, versus 1-specificity, or the percentage of false positives recovered. The area under the ROC curve (AUC) gives the probability of ranking a randomly selected active higher than a randomly chosen decoy. It varies between 0 and 1, where 1 represents a perfect ranking (all actives ranked above the decoys), while 0.5 corresponds to a completely random ranking.[30]

Table 4. PDB Selection: Statistical Values of Pose Prediction for VEGFR-2 PDBs with One Cognate Ligand Using Alpha HB as Scoring Function and Alpha Triangle as Placement Function. (a) RMSD Values Extracted from the Self-Docking Analysis. (b) RMSD Values Extracted from the Cross-Docking Analysis

(a)

PDB	resolution[a] (Å)	best rmsd pose (Å)	rmsd median (Å)	aRMSD (Å)	rmsd std dev (Å)	no. of poses[b]
1Y6A	2.10	1.3	4.0	3.8	0.9	10
1YWN	1.71	0.8	1.4	1.4	0.3	10
2OH4	2.05	1.5	2.6	2.5	0.6	10
2P2H	1.95	0.8	2.2	2.3	0.9	10
2QU5	2.95	0.8	1.4	1.3	0.4	10
2RL5	2.65	1.1	1.5	1.4	0.2	10
3BE2	1.75	1.2	2.9	2.6	0.9	10
3C7Q	2.10	1.3	3.4	3.4	0.9	10
3CJF	2.15	3.2	4.4	4.2	0.5	10
3EWH	1.60	1.8	3.0	2.9	0.9	10
3CJG	2.25	0.9	4.7	4.1	2.0	10

(b)

Alpha HB—Alpha Triangle

	Ligands											
PDB	1Y6A	1YWN	2OH4	2P2H	2QU5	2RL5	3BE2	3C7Q	3CJF	3EWH	3CJG	mean aRMSD
1Y6A	[c]	(5.3/**7.0**)	(6.4/7.7)	(1.2/**4.3**)	(7.0/**7.4**)	(5.5/**6.4**)	(5.6/**8.6**)	(3.0/**5.1**)	(4.9/**5.3**)	(5.8/**3.1**)	(1.1/**3.8**)	5.9
1YWN	(3.2/**5.3**)		(1.2/**2.6**)	(1.6/**4.1**)	(2.0/**7.4**)	(1.4/2.3)	(1.4/**5.3**)	(1.9/**6.7**)	(2.9/**5.6**)	(1.3/**9.3**)	(1.5/**3.7**)	5.3
2OH4	(3.3/**4.4**)	(4.9/**8.0**)		(1.6/**3.0**)	(1.9/**7.4**)	(0.9/**7.2**)	(1.6/**9.5**)	(2.2/**3.7**)	(4.3/**4.6**)	(2.8/**7.0**)	(1.1/**1.7**)	5.6
2P2H	(4.7/**5.2**)	(5.1/**8.0**)	(4.4/**8.5**)		(4.8/**7.8**)	(5.5/**6.5**)	(6.5/**9.3**)	(2.8/**3.1**)	(1.7/**2.3**)	(5.9/**8.2**)	(1.7/**3.0**)	6.2
2QU5	(4.4/**4.7**)	(5.8/**8.8**)	(1.7/**8.9**)	(3.3/**8.8**)		(1.7/**6.1**)	(3.0/**9.7**)	(3.4/**8.6**)	(4.6/**4.8**)	(3.0/**10.4**)	(1.9/**2.2**)	7.3
2RL5	(4.7/**4.3**)	(0.8/**7.0**)	(0.9/**7.5**)	(1.4/**4.6**)	(1.8/**7.3**)		(1.7/**8.7**)	(2.5/**2.9**)	(4.5/**4.8**)	(1.2/**7.8**)	(0.9/**2.1**)	5.7
3BE2	(1.8/**7.3**)	(1.8/**2.2**)	(2.4/**3.4**)	(1.0/**4.9**)	(1.3/**3.3**)	(1.3/**3.7**)		(4.0/**9.5**)	(3.1/**4.1**)	(2.7/**9.9**)	(1.3/**2.9**)	5.1
3C7Q	(4.4/**5.9**)	(2.1/**9.7**)	(2.9/**7.8**)	(4.8/**8.3**)	(3.0/**7.8**)	(1.6/**2.8**)	(1.7/**7.3**)		(1.9/**4.7**)	(2.7/**4.4**)	(4.5/**4.4**)	6.3
3CJF	(3.1/**5.7**)	(5.8/**7.1**)	(4.2/**7.3**)	(2.5/**3.8**)	(4.2/**7.6**)	(4.5/**4.3**)	(3.6/**4.3**)	(2.3/**6.3**)		(3.5/**6.3**)	(1.6/**3.0**)	5.6
3EWH	(3.2/**3.9**)	(1.2/**7.7**)	(1.0/**6.8**)	(1.3/**2.9**)	(2.2/**2.6**)	(1.4/**2.5**)	(1.3/**2.7**)	(3.4/**5.7**)	(4.1/**5.9**)		(1.6/**2.7**)	4.3
3CJG	(4.2/**5.7**)	(5.5/**6.6**)	(1.6/**7.0**)	(1.5/**5.1**)	(5.6/**6.9**)	(3.5/**5.8**)	(5.4/**8.5**)	(3.2/**4.2**)	(1.6/**3.9**)	(5.1/**8.9**)		6.3
aRMSD:		3.0										
std dev:		1.6										
no. of poses ≤2 Å:		45										

[a] PDB crystallographic resolution. [b] Number of poses used to calculate statistical rmsd values. [c] (best rmsd pose/***aRMSD***) for the best 10 poses.

RESULTS

Docking Function and Crystal Structure Selection. Regarding the docking function selection, among the different MOE scoring and placement functions, we tested 1936 dockings by self-docking and cross-docking validation. According to the self-docking validation (Table 2) the docking function Alpha HB—Alpha Triangle obtained the best results. The function returned 10 rmsd values below to 2.0 Å and also the lowest aRMSD value (1.3 Å) in comparison to the rest of docking functions. For the cross-docking validation, each cocrystallized ligand was docked into the rest of protein structures. Hence, for each cross-docking combination, a rmsd value was obtained, and therefore for each docking function, we computed an aRMSD value and also the number of poses ranked with a rmsd value below or equal to 2.0 Å. All cross-docking detailed results are available in the Supporting Information (Table S-1). Table 3 summarizes the aRMSD, standard deviation, and number of poses with a rmsd value equal or below 2 Å for all the 16 combinations of scoring and placement functions. All 16 combinations returned aRMSD values equal or above 3.0 Å. Regarding the number of best poses with a rmsd value equal or below 2.0 Å, 4 of the 16 combinations (Alpha HB—Alpha Triangle, Alpha HB—Triangle Matcher, London dG—Alpha Triangle, London dG—Proxy Triangle) were selected as the best functions, with more than 40 poses retrieved. All the aforementioned combinations of scoring and placement functions returned also the lowest aRMSD values (3.0 Å and 3.1 Å). Combining these results with the self-docking analysis, Alpha HB—Alpha Triangle was chosen as the most suitable docking function to work with our set of VEGFR-2 PDBs.

Hence, we used Alpha HB—Alpha Triangle docking function for the PDB selection. Table 4a and 4b shows the rmsd values of pose prediction for all the available VEGFR-2 PDBs extracted from the self-docking (Table 4a) and the cross-docking (Table 4b) analyses. Table 4a shows that the best aRMSD values

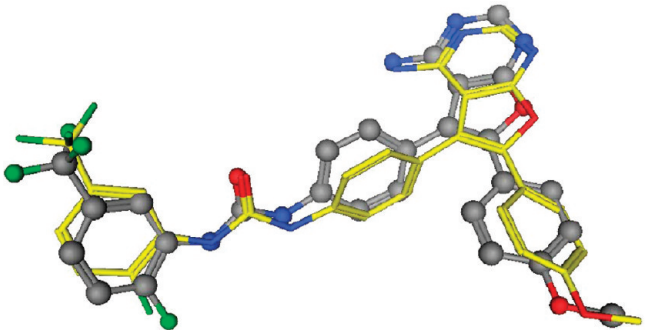


Figure 1. Superposition of the best rmsd docked pose onto the cocrystallized ligand in the VEGFR-2 complex (PDB code: 1YWN). Cocrystallized ligand from PDB 1YWN (carbon atoms in balls and gray color), and best rmsd docked pose (0.8 Å) (carbon atoms in sticks and yellow color) using Alpha HB scoring function and Alpha Triangle placement function implemented in MOE.

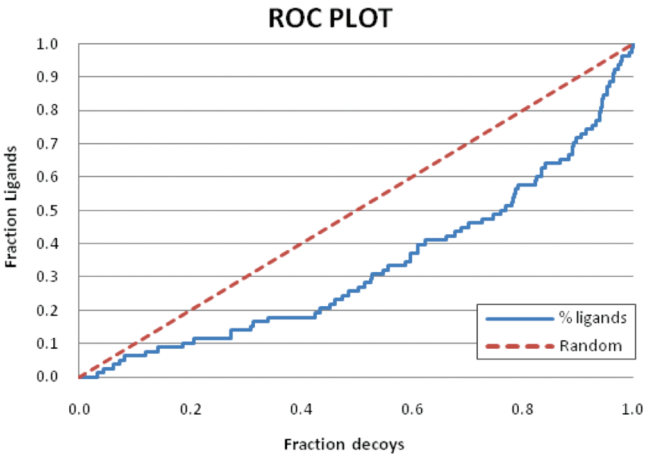
were obtained for the PDBs 2RL5 (1.4 Å), 2QU5 (1.3 Å), and 1YWN (1.4 Å), the three having similar average values for the top 10 ranked poses. PDB 1YWN and 2QU5 retrieved also the best rmsd poses: 0.8 Å (1YWN) (Figure 1) and 0.8 Å (2RL5). Nevertheless, crystallographic resolution was poorer for PDB 2RL5 (2.95 Å) and 2RL5 (2.65 Å) than for PDB 1YWN (1.71 Å). Table 4b shows that in the cross-docking analysis PDBs 3EWH, 3BE2, and 1YWN, respectively, retrieved the best mean aRMSD for the top 10 ranking poses. PDB 1YWN also retrieved the best rmsd pose together with PDBs 2RL5, 3BE2, and 3EWH.

Therefore, taking into account the crystallographic resolution and the analysis of the docking results obtained for each PDB, we defined structure 1YWN as our working structure[31] to use for VEGFR-2 docking-based VS. This structure has also been used in docking experiments applied to an in-house database by Enyedy et al.[32] This VEGFR-2 complex (1YWN) has a substituted 4-aminofuro[2,3-*d*]pyrimidine as cocrystallized ligand (IC_{50} = 0.003 μM[33]). Hence, PDB 1YWN protein structure with Alpha HB function and Alpha Triangle placement function were used for the subsequent docking-based VS.

Retrospective Docking of DUD VEGFR-2 Data Set. We docked the VEGFR-2 test set in the DUD database using Alpha HB and Alpha Triangle MOE placement function. All compounds were ranked from higher to lower Alpha HB value. Figure 2 shows the resulting ROC plot. The overall AUC was worse than the random selection (AUC = 0.33), with a very low early recovery performance (% true positives at 1% of false positives: 0%, and % true positives at 5% of false positives: 2.56%).

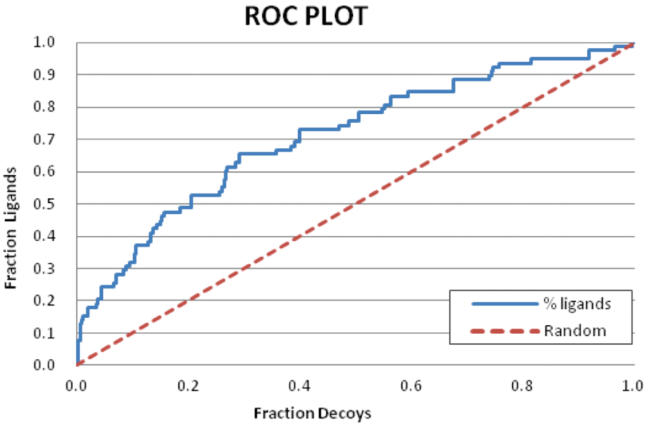
An alternative commonly used to postprocess docking poses is through molecular similarity analysis, using coefficients such as the Tanimoto index.[34] Similarity between two molecules is measured, comparing fingerprints or physicochemical properties. In this case, we used the 2D MACCS fingerprints although it is also possible to use 3D properties such as three-dimensional shape of molecules, due to their important role in bioactivity.[35] The underlying assumption is that similar structures have similar biological activity,[36] taking into account that there is a wide range of different similarity methods.[37]

Thus, it is expected that similar molecules (molecules with Tanimoto coefficient values close to 1) should have similar activities when they bind to a biological target and they inhibit it.[38]



% True Pos at 1% of False Pos	% True Pos at 5% of False Pos	% True Pos at 10% of False Pos.
0.00%	2.56%	6.41%

Figure 2. ROC plot obtained when docked poses are ranked by Alpha HB scoring function. AUC = 0.33. The table below the ROC plot shows early recognition values when poses are ranked by Alpha HB energy scoring function.



% True Pos at 1% of False Pos	% True Pos at 5% of False Pos	% True Pos at 10% of False Pos.
14.10%	24.36%	32.05%

Figure 3. ROC plot obtained when docked poses are ranked by MACCS similarity to the congnate ligand (Tanimoto index). AUC = 0.70. The table below the ROC plot shows early recognition values when poses are ranked by MACCS similarity to the cognate ligand.

Jahn et al.[24] successfully achieved discrimination between VEGFR-2 ligands and decoys using ligand-based VS methods. The results processed by Tanimoto MACCS similarity fingerprints improved early recognition in comparison with results obtained using only standard scoring functions for ranking compounds (e.g., awROC enrichments at 1% false positive fraction of 9.4 ± 2.2 versus 3.2 ± 1.5 applying similarity postprocess and a docking standard scoring function, respectively). Figure 3 shows VEGFR-2 docking-based VS results ranked by MACCS (Tanimoto coefficient) using the cocrystallized ligand from PDB 1YWN as the query. In our case, it can be

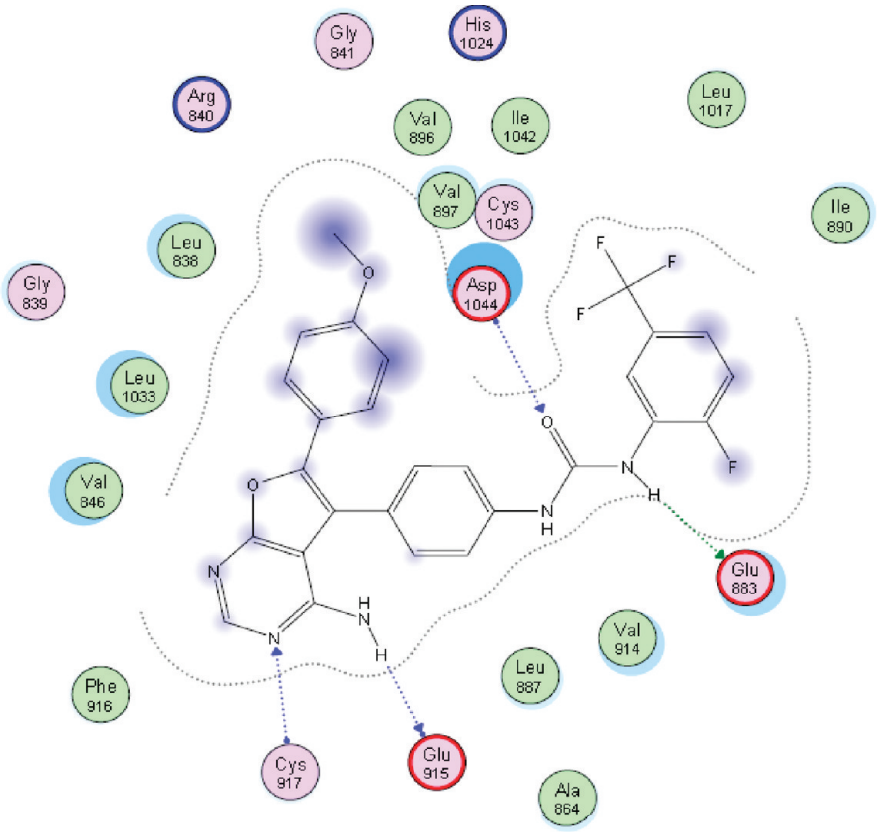


Figure 4. Key interactions between VEGFR-2 binding site residues and the cocrystallized ligand (PDB code: 1YWN). The cognate ligand makes two hydrogen bond donor interactions with residues Glu915 and Glu883 and two hydrogen bond acceptor interactions with residues Cys917 and Asp1044. The dashed line shows the ligand proximity contour. Residues in lilac color indicate polar interactions, and residues in green indicate nonpolar interactions. Ligand atoms with blue contour and residues with turquoise discs surrounding them (Leu838, Val846, Val914, Glu883, Leu1033, and Asp1044) indicate solvent accessibility.

observed that the overall area under the curve (AUC = 0.70) improves significantly the AUC obtained using only Alpha HB scoring function (AUC = 0.33). The early performance also improves substantially (% true positives at 1% of false positives: 0.0% using only Alpha HB scoring function, and % true positives at 1% of false positives: 14.1% with similarity postprocessing).

Pharmacophore Postfiltering Using Protein−Ligand Interaction Information. To improve VEGFR-2 docking-based VS, we used pharmacophore[39] postfiltering, that is to say, we filtered all docked poses, taking into account relevant protein−ligand interactions[35] (Figure 4). Often ligand-based pharmacophore filtering techniques are used as predocking screening methods. However, we pose-filtered docking results using a receptor-based pharmacophore which captures the essential protein−ligand interactions information.

Docking poses were filtered according to a series of protein−ligand interactions defined as features. All molecules whose docking poses did not satisfy all required interactions were eliminated.

Usually a pharmacophore query is defined following two strategies. One strategy is to perform dockings of several cocrystallized ligands and align these poses. Common interactions found for the different cognate ligands are analyzed and grouped defining the pharmacophore[25] query. The second strategy is to use tools for analyzing protein−ligand interactions for several PDB complexes and build a map-histogram for all of them. Common and repeated interactions are used as pharmacophore features (e.g., PLIF tools implemented in MOE). In this work, we used protein alignment to superpose all cognate ligands. Ligand's superposition generated from the alignment of PDB complexes is more accurate than superposition of docked ligands, e.g., by flexible alignment or other ligand-based superposition methods, where the position (and interacting information) of ligand in reference to the receptor binding pocket is not taken into account. Nevertheless, these strategies are useful approximations[40] when their PDB complexes are not available for a protein target. Therefore, we aligned protein chains for all VEGFR-2 PDB complexes with crystallographic resolution lower than 3 Å (see Figure 5). Common ligand functional groups and common protein−ligand interactions located inside a threshold volume were used as pharmacophore features. Pharmacophore Consensus implemented in MOE was used for searching the common interactions between the overlapped ligands. The resulting pharmacophore query (Figure 6) was therefore defined on the basis of a mixed strategy, structure based and ligand based. Pharmacophore Consensus search returned five pharmacophore features, some of them showing multiple interactions (features F1, F2, and F5 in Figure 6). Features F1 and F2 indicate hydrogen bond acceptor[41−43] and hydrogen bond donor[44] interaction sites with residue Cys 917, since this residue admits both interactions.[45−47] Features F3, F4, and F5 show both hydrophobic and nonaromatic π-system rings interaction sites, mainly with residues Leu 838, Val 897, Leu 887, Leu 1033, and Cys1043.[47]

dx.doi.org/10.1021/ci1002763 |J. Chem. Inf. Model. 2011, 51, 777–787

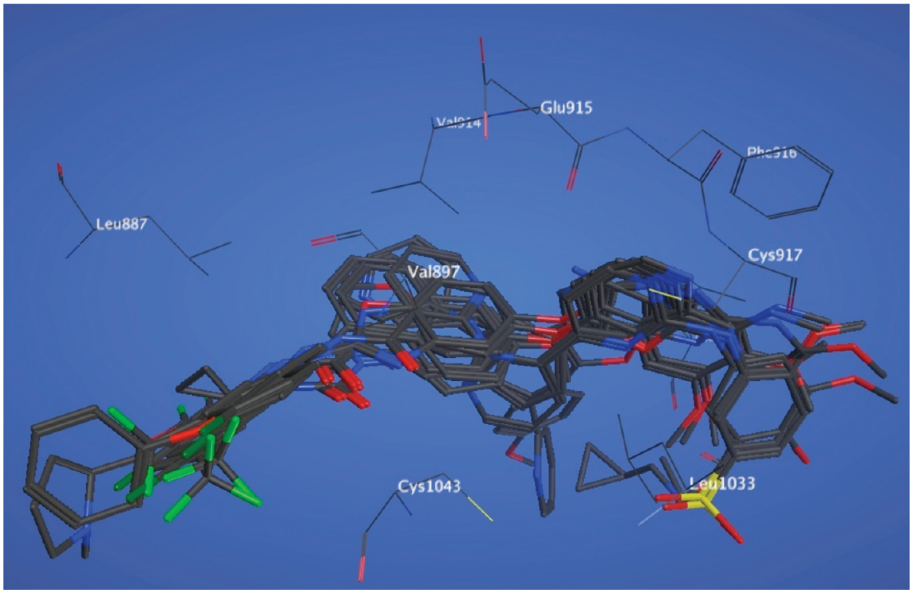


Figure 5. Superposition of cocrystallized ligands by aligning VEGFR-2 PDB complexes with crystallographic resolution lower than 3 Å.

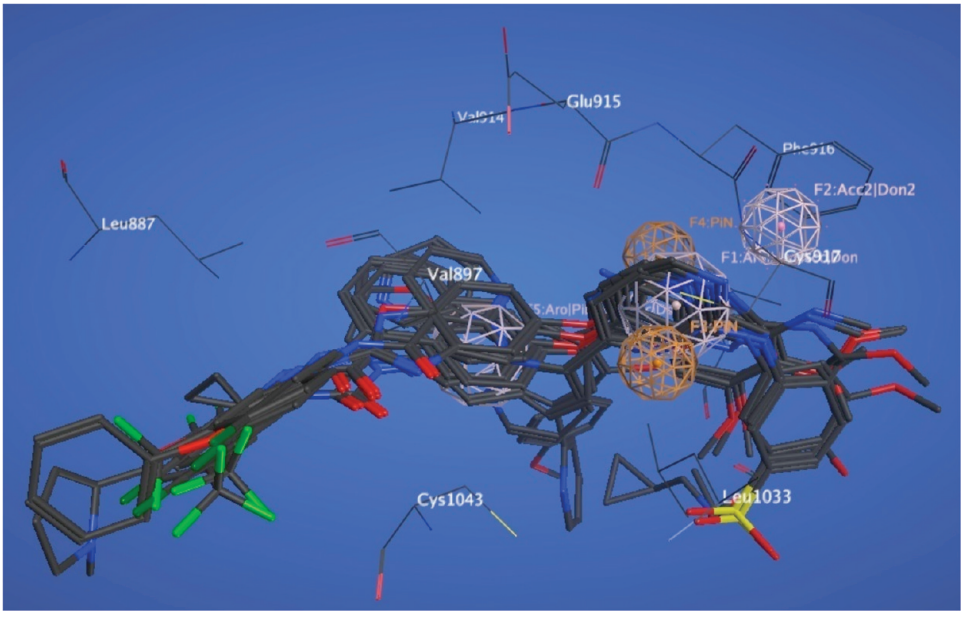


Figure 6. Pharmacophore model resulting from VEGFR-2 PDB complexes alignment and Pharmacophore Consensus analysis. Key interactions (features) defining the pharmacophore are colored in orange (F3 and F4) and pink (F1, F2, and F5). Each feature indicates one or more interactions at the same time. Feature F2 is projected out of feature F1 to facilitate visualization.

Radii of the features were optimized, by running several docking calculations and postfiltering docked poses, changing the pharmacophore radii values, until the maximum ratio of ligands/decoys was reached. The best results were achieved with two main variants of the same pharmacophore query (Figure 6). First, when the radii of the features F4 and F5 were 1.5 Å and 2 Å, respectively, 69.3% of the ligands satisfied the pharmacophore query, but also 35.2% of the decoys satisfied the pharmacophore filter, with an overall enrichment factor of 1.83. Second, if the radii of the features F4 and F5 were 1.0 Å and 1.5 Å, respectively, there were 59.0% of docked and postfiltered ligands, and 20.8% of false positives with an enrichment factor of 2.69 (see Figure 7).

Figure 8 and Figure 9 show the resulting enrichment plot and ROC plot, respectively, when the pharmacophoric filter is applied to docked poses and hits are ranked by their Alpha HB scoring value. Early recovery values are also shown in both cases. It can be observed that the early performance considerably increases using the pharmacophore postfilter (% true positives at 1% of false positives: 0.0% using only Alpha HB scoring function, and % true positives at 1% of false positives: 10.26% with pharmacophore postfilter).

Similarity search postprocessing was also applied to the pharmacophore postfiltered results. Figure 10 shows the resulting ROC plot when pharmacophore postfiltered docked poses

dx.doi.org/10.1021/ci1002763 |*J. Chem. Inf. Model.* 2011, 51, 777–787

were ranked by MACCS similarity to the congnate ligand using the Tanimoto index. When the three steps (docking, pharmacophore postfilter and similarity search postprocessing) are used, the greatest improvement observed is in the early recovery performance (% true positives at 1% of false positives: 0.0% using only Alpha HB scoring function versus % true positives at 1% of false positives: 16.67% with pharmacophore postfilter and similarity postprocess, % true positives at 5% of false positives: 2.56% using only Alpha HB scoring function versus % true positives at 5% of false positives: 30.77% with pharmacophore postfilter and similarity postprocess, and % true positives at 10% of false positives: 6.41% using only Alpha HB scoring function versus % true positives at 10% of false positives: 41.03% with pharmacophore postfilter and similarity postprocess).

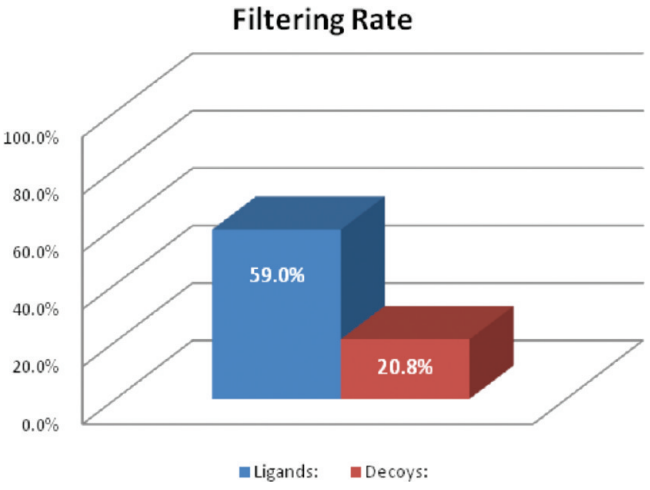
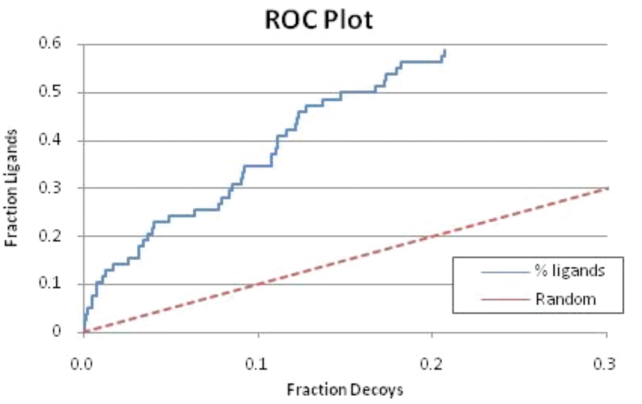


Figure 7. Filtering rate bar chart. Percentage of ligands and decoys recovered when docked poses are postfiltered with a pharmacophore.

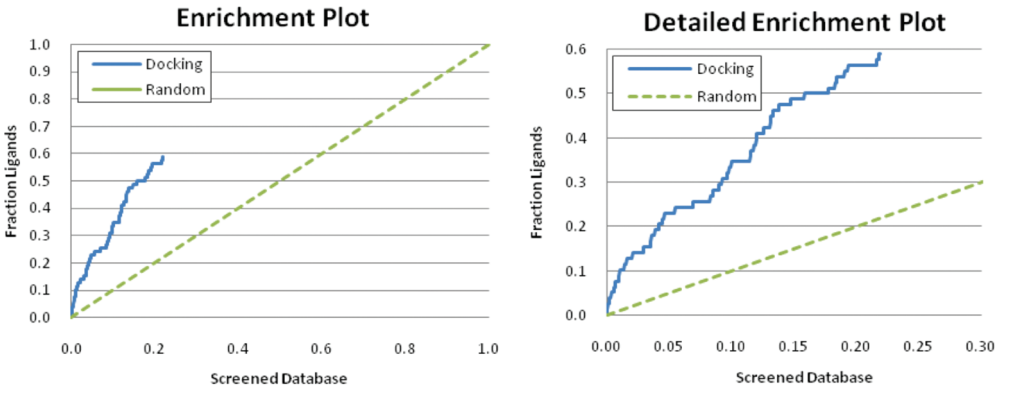
DISCUSSION

Docking-based VS aims to find the lowest energy binding modes for a set of ligands inside the receptor pocket site. In addition, scoring functions are expected to distinguish active



% True Pos at 1% of False Pos	% True Pos at 5% of False Pos	% True Pos at 10% of False Pos.
10.26%	24.36%	34.62%

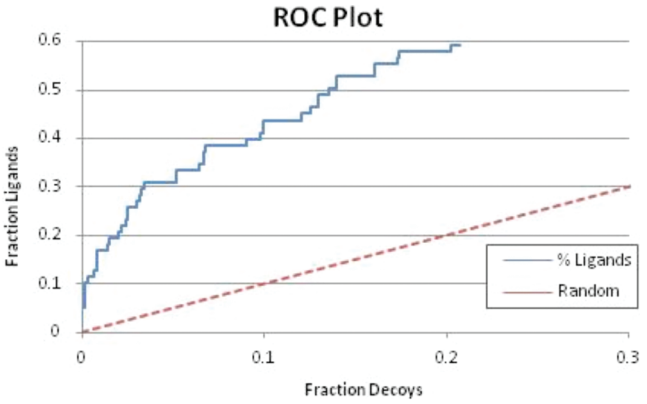
Figure 9. ROC plot obtained when the pharmacophore postfilter is applied to the docked poses ranked by Alpha HB scoring function. The curve only shows ligands and decoys whose docked poses match pharmacophoric requirements. No value is returned for docked poses that do not match the pharmacophore constrains; therefore, they cannot be plotted. In the table, the ROC plot shows early recognition values.



% Database Screened	True Positives (Ligands)	False Positives (Decoys)	EF (Enrichment Factor)	% True Positives at 1% False Positives	% True Positives at 5% False Positives	% True Positives at 10% False Positives
1%	8	20	**9.37**	10.26%	24.36%	34.62%
5%	18	112	**4.54**			
10%	**27**	**231**	**3.43**			
20%	44	469	**2.81**			
30%	46	515	**2.69**			
50%	46	515	**2.69**			

Figure 8. Enrichment plot obtained when the pharmacophore postfilter is applied to the docked poses ranked by Alpha HB scoring function. The curve only shows ligands and decoys whose docked poses match pharmacophoric requirements (22% of the total database). No value is returned for docked poses that do not match the pharmacophore constrains; therefore, they cannot be plotted. On the left, detailed enrichment plot for the first 30% of screened database. In the table, the plots show the EF at the first percentages of database screened.

dx.doi.org/10.1021/ci1002763 |J. Chem. Inf. Model. 2011, 51, 777–787



% True Pos at 1% of False Pos	% True Pos at 5% of False Pos	% True Pos at 10% of False Pos.
16.67%	30.77%	41.03%

Figure 10. ROC plot obtained when postfiltered docked poses are ranked by MACCS similarity to the congnate ligand (Tanimoto index). The curve only shows ligands and decoys whose docked poses match pharmacophoric requirements. In the table, the ROC plot shows early recognition values.

molecules from inactives based on protein−ligand binding interactions. In many cases, docking techniques can predict correct ligand binding modes, but screening between high activity binders and decoys using scoring functions do not always provide good results. In those cases, it is highly convenient to use pharmacophoric filters to set the key interactions needed for a high activity binder and, hence, to discriminate whether a molecule is an active or a decoy, improving in this way ligands from decoy selection. Thus, docked poses well fitted in the binding site which also satisfy the pharmacophoric features ought to be actives. The most difficult aspect is to build a high selective pharmacophore model for filtering the database and discriminate all decoys from the actives. Moreover, if ligand information is added to docking results, such as a similarity search step, then VS is expected to perform better. In this work we combined structure-based and ligand-based information through three steps for VEGFR-2 retrospective VS. It has been shown that when docked poses were ranked by their Alpha HB binding energy values (best MOE scoring function found for this kind of protein−ligand interactions), selection between ligands and decoys was worse than random (AUC = 0.33), and early performance was also very low (true positives at 1% of false positives: 0.0%, see Figure 2). However when poses were reranked by MACCS molecular similarity (Tanimoto index) instead of binding energy, ligand selection improves significantly (AUC = 0.70) and early performance also did (% true positives at 1% of false positives: 14.10%, see Figures 2 and 3). When the pharmacophore postfilter was applied to the docked poses, 59.0% of the ligands and 20.8% of decoys were selected. In this case, the early performance becomes more effective (% true positives at 1% of false positives: 0.0% using only Alpha HB scoring function, and % true positives at 1% of false positives: 10.26% with pharmacophore postfilter, see Figures 2 and 9). ROC curves for pharmacophore postfiltered poses improved selection when the similarity search was also applied as a third step because similar structures are assumed to have similar biological activity. However, the highest improvement was observed in the early performance values also applying the three steps (% true positives at 1% of false positives: 0.0% using only Alpha HB scoring function versus % true positives at 1% of false positives: 16.67% with pharmacophore postfilter and similarity postprocess, see Figures 2 and 10).

Previous VS approaches applied to the VEGFR-2 DUD test set show that it is difficult to achieve high performance when working with this target.[17,18,48] Poor overall AUC has been retrieved in many cases: Kirchmair et al. have obtained AUC values from 0.43 to 0.48, using different scoring functions. Tawa et al. achieved AUC values from 0.599 to 0.614 using queries computed by ROCS. Enyedi et al.,[32] working with pdb 1YWN and testing an in-house database with different scoring functions, achieved an AUC of 0.78 but using an IC_{50} value cutoff. Moreover, there are not much early performance data published, and in all these cases very low values were achieved (EF values from 0 to 1.4 at 1%, EF values from 0 to 0.5 at 5%, and EF values from 0.4 to 1.1 at 10% in Kirchmair et al.). Here we show that by applying a three-step VS protocol, i.e., the conventional docking step, a pharmacophore postfilter step, and a similarity search postprocess, we get better early performance than that retrieved in previous work using only the most common scoring functions.

Furthermore, the VS protocol applied is able to better distinguish between correct and incorrect binding modes rather than using only a scoring function, which is logical given that we are using protein−ligand information in the second step. First-ranking binding modes retrieved by the protocol used here keep the main expected interactions for active ligands. Hence, this procedure ensures that selected ligands will have the same binding mode (key interactions) and presumably the same biological response than the query used. It is increasingly the case that pharmaceutical companies have multiple ligands for a given target and these bind in different binding modes to the same pocket. Using this procedure, we ensure that ligands selected bind in the same mode as the query used to select them.

CONCLUSION

Successful docking-based VS approaches must find good binding modes for active ligands and be able to discriminate actives from decoys using a scoring function. Previous VS approaches applied to VEGFR-2 inhibitors show that it is difficult to achieve high performance when working with the VEGFR-2 target. We carried out retrospective VS on the VEGFR-2 DUD data set with the aim to improve ligands from decoy selection. We first validated pose prediction in the VEGFR-2 target using the 1YWN complex (1.71 Å resolution) and the most accurate scoring function that matched the cognate ligand binding mode obtained from docking with the cocrystallized pose. This evaluation was performed by measuring the best rmsd value and best rmsd mean value for the top 10 rmsd poses using several MOE scoring and placement functions. Discrimination of ligands from decoys was then evaluated first through docking into the VEGFR-2 binding pocket, using the selected Alpha HB and Alpha Triangle scoring and placement functions. Docking results were postfiltered according to defined pharmacophore protein−ligand interactions, and the selected poses were again ranked according to their molecular similarity (MACCS fingerprint) to the 1YWN cognate ligand. Results show that by applying this three-step protocol, the overall performance (both AUC and EF) improves, mainly increasing the early performance. Moreover, performance metrics usually retrieve worse

dx.doi.org/10.1021/ci1002763 |J. Chem. Inf. Model. 2011, 51, 777–787

values when an energy scoring function is used for ranking compounds rather than when a similarity search index is used. On the other hand, this VS protocol is better suited for finding correct ligand binding modes than using only a scoring function. First-ranking binding modes retrieved keep the main key protein−ligand interactions needed for active ligands. Hence, this procedure can be useful for prospective VS of novel VEGFR-2 inhibitors, mainly 1YWN VEGFR-2 cognate ligand analogues, with both chemical structure and binding mode similar to that of the 1YWN cognate ligand, and is an ongoing study.

■ ASSOCIATED CONTENT

Supporting Information. Performance of cross-docking validation using all different MOE scoring and placement functions. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +34-93-267.20.00. Fax: +34-93-205.62.66. E-mail: violeta.perez@iqs.url.edu.

■ ACKNOWLEDGMENT

We are grateful to Chemical Computing Group, Inc., for the support provided. V.I.P.-N. thanks the Generalitat de Catalunya − DURSI for a grant within the Formació de Personal Investigador (2008FI) program.

■ REFERENCES

(1) Pradeep, C. R.; Sunila, E. S.; Kuttan, G. Expression of vascular endothelial growth factor (VEGF) and VEGF receptors in tumor angiogenesis and malignancies. *Integr. Cancer Ther.* **2005**, *4*, 315–321.

(2) Cee, V. J.; Albrecht, B. K.; Geuns-Meyer, S.; Hughes, P.; Bellon, S.; Bready, J.; Caenepeel, S.; Chaffee, S. C.; Coxon, A.; Emery, M. Alkynylpyrimidine amide derivatives as potent, selective, and orally active inhibitors of Tie-2 kinase. *J. Med. Chem.* **2007**, *4*, 627–640.

(3) Hasegawa, M.; Nishigaki, N.; Washio, Y.; Kano, K.; Harris, P. A.; Sato, H.; Mori, I.; West, R. I.; Shibahara, M.; Toyoda, H. Discovery of novel benzimidazoles as potent inhibitors of TIE-2 and VEGFR-2 tyrosine kinase receptors. *J. Med. Chem.* **2007**, *18*, 4453–4470.

(4) Pytel, D.; Sliwinski, T.; Poplawski, T.; Ferriola, D.; Majsterek, I. Tyrosine kinase blockers: new hope for successful cancer therapy. *Anti-Cancer Agent Me.* **2009**, *1*, 66–76.

(5) Kiselyov, A. S.; Semenova, M.; Semenov, V. V. 1, 2, 3-Triazol-4-yl) benzenamines: Synthesis and activity against VEGF receptors 1 and 2. *Bioorg. Med. Chem. Lett.* **2009**, *5*, 1344–1348.

(6) Peifer, C.; Selig, R.; Kinkel, K.; Ott, D.; Totzke, F.; Schachtele, C.; Heidenreich, R.; Rocken, M.; Schollmeyer, D.; Laufer, S. Design, synthesis, and biological evaluation of novel 3-aryl-4-(1*H*-indole-3-yl)-1,5-dihydro-2*H*-pyrrol-2-ones as vascular endothelial growth factor receptor (VEGF-R) inhibitors. *J. Med. Chem.* **2008**, *13*, 3814–3824.

(7) Mannion, M.; Raeppel, S.; Claridge, S.; Zhou, N.; Saavedra, O.; Isakovic, L.; Zhan, L.; Gaudette, F.; Raeppel, F.; Déziel, R. N-(4-(6, 7-Disubstituted-quinolin-4-yloxy)-3-fluorophenyl)-2-oxo-3-phenylimidazolidine-1-carboxamides: A novel series of dual c-Met/VEGFR2 receptor tyrosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *23*, 6552–6556.

(8) Ruel, R.; Thibeault, C.; L'Heureux, A.; Martel, A.; Cai, Z. W.; Wei, D.; Qian, L.; Barrish, J. C.; Mathur, A.; D'Arienzo, C. Discovery and preclinical studies of 5-isopropyl-6-(5-methyl-1, 3, 4-oxadiazol-2-yl)-N-(2-methyl-1H-pyrrolo [2, 3-b] pyridin-5-yl) pyrrolo [2, 1-f][1, 2, 4] triazin-4-amine (BMS-645737), an in vivo active potent VEGFR-2 inhibitor. *Bioorg. Med. Chem. Lett.* **2008**, *9*, 2985–2989.

(9) Polverino, A.; Coxon, A.; Starnes, C.; Diaz, Z.; DeMelfi, T.; Wang, L.; Bready, J.; Estrada, J.; Cattley, R.; Kaufman, S. AMG 706, an oral, multikinase inhibitor that selectively targets vascular endothelial growth factor, platelet-derived growth factor, and kit receptors, potently inhibits angiogenesis and induces regression in tumor xenografts. *Cancer Res.* **2006**, *17*, 8715–8721.

(10) Pujadas, G.; Vaque, M.; Ardevol, A.; Blade, C.; Salvado, M.; Blay, M.; Fernandez-Larrea, J.; Arola, L. Protein-ligand Docking: A Review of Recent Advances and Future Perspectives. *Curr. Pharm. Anal.* **2008**, *1*, 1–19.

(11) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2009**, *6*, 1455–1474.

(12) Cheng, T.; Li, X.; Li, Y.; Liu, Z.; Wang, R. Comparative Assessment of Scoring Functions on a Diverse Test Set. *J. Chem. Inf. Model.* **2009**, *4*, 1079–1093.

(13) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *23*, 6789–6801.

(14) Rohrer, S. G.; Baumann, K. Maximum Unbiased Validation (MUV) Data Sets for Virtual Screening Based on PubChem Bioactivity Data. *J. Chem. Inf. Model.* **2009**, *2*, 169–184.

(15) Graves, A. P.; Brenk, R.; Shoichet, B. K. Decoys for docking. *J. Med. Chem.* **2005**, *11*, 3714–3728.

(16) *MOE (Molecular Operating Environment)*, 2008.10 release; Chemical Computing Group, Inc.: Montreal, Canada, 2004.

(17) Kirchmair, J.; Distinto, S.; Markt, P.; Schuster, D.; Spitzer, G. M.; Liedl, K. R.; Wolber, G. How To Optimize Shape-Based Virtual Screening: Choosing the Right Query and Including Chemical Information. *J. Chem. Inf. Model.* **2009**, *3*, 678–692.

(18) Tawa, G. J.; Baber, J. C.; Humblet, C. Computation of 3D queries for ROCS based virtual screens. *J. Comput.-Aided Mol. Des.* **2009**, *12*, 853–868.

(19) Cowan-Jacob, S. W. Structural biology of protein tyrosine kinases. *Cell. Mol. Life Sci.* **2006**, *22*, 2608–2625.

(20) Berman, H.; Battistuz, T.; Bhat, T.; Bluhm, W.; Bourne, P.; Burkhardt, K.; Feng, Z.; Gilliland, G.; Iype, L.; Jain, S. The protein data bank. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *6*, 899–907.

(21) Harris, P.; Cheung, M.; Hunter, R., 3rd; Brown, M.; Veal, J.; Nolte, R.; Wang, L.; Liu, W.; Crosby, R.; Johnson, J. Discovery and evaluation of 2-anilino-5-aryloxazoles as a novel class of VEGFR2 kinase inhibitors. *J. Med. Chem.* **2005**, *5*, 1610–1619.

(22) 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**, *3*, 213–228.

(23) Tuccinardi, T.; Botta, M.; Giordano, A.; Martinelli, A. Protein Kinases: Docking and Homology Modeling Reliability. *J. Chem. Inf. Model* **2010**, *8*, 1432–1441.

(24) Jahn, A.; Hinselmann, G.; Fechner, N.; Zell, A. Optimal assignment methods for ligand-based virtual screening. *J. Cheminf.* **2009**, *1*, 14–37.

(25) Peach, M. L.; Nicklaus, M. C. Combining docking with pharmacophore filtering for improved virtual screening. *J. Cheminf.* **2009**, *1*, 6–21.

(26) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T.; Murray, C. W.; Taylor, R. D.; Watson, P. Virtual screening using protein-ligand docking: avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *3*, 793–806.

(27) Good, A. C.; Oprea, T. I. Optimization of CAMD techniques 3. Virtual screening enrichment studies: a help or hindrance in tool selection?. *J. Comput.-Aided Mol. Des.* **2008**, *3*, 169–178.

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

(29) Mackey, M. D.; Melville, J. L. Better than Random? The Chemotype Enrichment Problem. *J. Chem. Inf. Model.* **2009**, *5*, 1154–1162.
(30) Triballeau, N.; Acher, F.; Brabet, I.; Pin, J. P.; Bertrand, H. O. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J. Med. Chem.* **2005**, *7*, 2534–2547.
(31) Liebeschuetz, J. W. Evaluating docking programs: keeping the playing field level. *J. Comput.-Aided Mol. Des.* **2008**, *3*, 229–238.
(32) Enyedy, I. J.; Egan, W. J. Can we use docking and scoring for hit-to-lead optimization?. *J. Comput.-Aided Mol. Des.* **2008**, *3*, 161–168.
(33) Miyazaki, Y.; Matsunaga, S.; Tang, J.; Maeda, Y.; Nakano, M.; Philippe, R. J.; Shibahara, M.; Liu, W.; Sato, H.; Wang, L. Novel 4-amino-furo [2, 3-d] pyrimidines as Tie-2 and VEGFR2 dual inhibitors. *Bioorg. Med. Chem. Lett.* **2005**, *9*, 2203–2207.
(34) Bostrom, J.; Hogner, A.; Schmitt, S. Do structurally similar ligands bind in a similar fashion?. *J. Med. Chem.* **2006**, *23*, 6716–6725.
(35) Wolber, G.; Seidel, T.; Bendix, F.; Langer, T. Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discovery Today* **2008**, *13*, 23–29.
(36) Johnson, M. A.; Maggiora, G. M. *Concepts and applications of molecular similarity*; Wiley: New York: 1990; pp 1−13.
(37) Eckert, H.; Bajorath, J. Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches. *Drug Discovery Today* **2007**, *5*−*6*, 225–233.
(38) Kortagere, S.; Krasowski, M. D.; Ekins, S. The importance of discerning shape in molecular pharmacology. *Trends Pharmacol. Sci.* **2009**, *3*, 138–147.
(39) Ghose, A. K.; Herbertz, T.; Salvino, J. M.; Mallamo, J. P. Knowledge-based chemoinformatic approaches to drug discovery. *Drug Discovery Today* **2006**, *23*−*24*, 1107–1114.
(40) Neaz, M. M.; Pasha, F. A.; Muddassar, M.; Lee, S. H.; Sim, T.; Hah, J. M.; Cho, S. J. Pharmacophore based 3D-QSAR study of VEGFR-2 inhibitors. *Med. Chem. Res.* **2009**, *2*, 127–142.
(41) Yu, H.; Wang, Z.; Zhang, L.; Zhang, J.; Huang, Q. Pharmacophore modeling and in silico screening for new KDR kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *8*, 2126–2133.
(42) Sammond, D. M.; Nailor, K. E.; Veal, J. M.; Nolte, R. T.; Wang, L.; Knick, V. B.; Rudolph, S. K.; Truesdale, A. T.; Nartey, E. N.; Stafford, J. A. Discovery of a novel and potent series of dianilinopyrimidineurea and urea isostere inhibitors of VEGFR2 tyrosine kinase. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3519–3523.
(43) Gracias, V.; Ji, Z.; Akritopoulou-Zanze, I.; Abad-Zapatero, C.; Huth, J. R.; Song, D.; Hajduk, P. J.; Johnson, E. F.; Glaser, K. B.; Marcotte, P. A. Scaffold oriented synthesis. Part 2: Design, synthesis and biological evaluation of pyrimido-diazepines as receptor tyrosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *8*, 2691–2695.
(44) Sridhar, J.; Akula, N.; Sivanesan, D.; Narasimhan, M.; Rathinavelu, A.; Pattabiraman, N. Identification of novel angiogenesis inhibitors. *Bioorg. Med. Chem. Lett.* **2005**, *18*, 4125–4129.
(45) Harris, P. A.; Boloor, A.; Cheung, M.; Kumar, R.; Crosby, R. M.; Davis-Ward, R. G.; Epperly, A. H.; Hinkle, K. W.; Hunter, R. N., III; Johnson, J. H. Discovery of 5-[[4-[(2,3-Dimethyl-2*H*-indazol-6-yl)methylamino]-2-pyrimidinyl]amino]-2-methylbenzenesulfonamide (Pazopanib), a Novel and Potent Vascular Endothelial Growth Factor Receptor Inhibitor. *J. Med. Chem.* **2008**, *15*, 4632–4640.
(46) Potashman, M. H.; Bready, J.; Coxon, A.; DeMelfi, T. M., Jr.; DiPietro, L.; Doerr, N.; Elbaum, D.; Estrada, J.; Gallant, P.; Germain, J. Design, synthesis, and evaluation of orally active benzimidazoles and benzoxazoles as vascular endothelial growth factor-2 receptor tyrosine kinase inhibitors. *J. Med. Chem.* **2007**, *18*, 4351–4373.
(47) Dakshanamurthy, S.; Kim, M.; Brown, M. L.; Byers, S. W. In-silico fragment-based identification of novel angiogenesis inhibitors. *Bioorg. Med. Chem. Lett.* **2007**, *16*, 4551–4556.
(48) Zhou, Z.; Felts, A. K.; Friesner, R. A.; Levy, R. M. Comparative performance of several flexible docking programs and scoring functions: enrichment studies for a diverse set of pharmaceutically relevant targets. *J. Chem. Inf. Model.* **2007**, *4*, 1599–1608.