Cosolvent-Based Protein Pharmacophore for Ligand Enrichment in Virtual Screening

Juan Pablo Arcon,[†,‡,⊥] Lucas A. Defelipe,[†,‡,⊥] Elias D. Lopez,[†,‡] Osvaldo Burastero,[†,‡] Carlos P. Modenutti,[†,‡] Xavier Barril,[§,∥] Marcelo A. Marti,*[,†,‡] and Adrian G. Turjanski*[,†,‡]

[†]Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, and [‡]IQUIBICEN/UBA-CONICET, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, Buenos Aires C1428EGA, Argentina
[§]Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluís Companys 23, Barcelona 08010, Spain
[∥]Faculty of Pharmacy and Institute of Biomedicine (IBUB), University of Barcelona, Av. Joan XXIII 27-31, Barcelona 08028, Spain

S *Supporting Information*

ABSTRACT: Virtual screening of large compound databases, looking for potential ligands of a target protein, is a major tool in computer-aided drug discovery. Throughout the years, different techniques such as similarity searching, pharmacophore matching, or molecular docking have been applied with the aim of finding hit compounds showing appreciable affinity. Molecular dynamics simulations in mixed solvents have been shown to identify hot spots relevant for protein−drug interaction, and implementations based on this knowledge were developed to improve pharmacophore matching of small molecules, binding free-energy estimations, and docking performance in terms of pose prediction. Here, we proved in a retrospective manner that cosolvent-derived pharmacophores from molecular dynamics (solvent sites) improve the performance of docking-based virtual screening campaigns. We applied a biased docking scheme based on solvent sites to nine relevant target proteins that have a set of known ligands or actives and compounds that are, presumably, nonbinders (decoys). Our results show improvement in virtual screening performance compared to traditional docking programs both at a global level, with up to 35% increase in areas under the receiver operating characteristic curve, and in early stages, with up to a 7-fold increase in enrichment factors at 1%. However, the improvement in pose prediction of actives was less profound. The presented application makes use of the AutoDock Bias method and is the only cosolvent-derived pharmacophore technique that employs its knowledge both in the ligand conformational search algorithm and the final affinity scoring for virtual screening purposes.
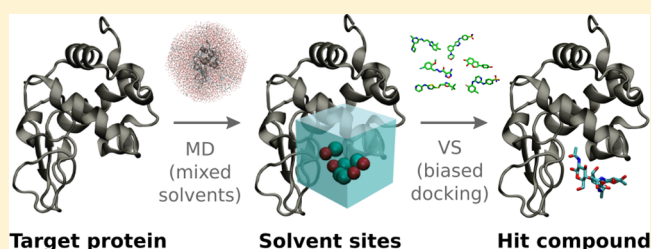
INTRODUCTION

Drug discovery is an expensive endeavor, costing between 0.16 and 1.8 billion dollars to introduce a new drug to the market.[1] The need for new compounds and more efficient tools to obtain them is illustrated by the outbreak of infectious diseases,[2] boosted by the emergence of drug resistance in microorganisms such as *Mycobacterium tuberculosis, Staphylococcus aureus,* or *Klebsiella pneumoniae* among others[3,4] and the need for better treatment of neglected diseases such as malaria, dengue, or sleeping sickness.[5] Furthermore, the discovery of new molecular targets to treat complex diseases such as cancer, autoimmune, and neurodegenerative diseases also requires the development of new drug candidates and pharmacological tools to elucidate their roles in physiological and pathological contexts.

Virtual screening (VS) of large compound databases, looking for potential ligands of a target protein, is a major tool in computer-aided drug discovery. Its main purpose is to find hit compounds with a certain affinity, usually around 1 μM, which can then be further optimized to become lead compounds and drug candidates faster, more efficiently, and at lower cost than

with experimental high-throughput screening campaigns.[6] Briefly, VS consists in rapidly estimating the binding free-energy of compounds to a given target by computational means and choosing a subset of these compounds with higher predicted affinities to be tested experimentally. A widespread VS technique starts from the target protein structure in the so-called docking methods and explicitly studies the interaction between the protein and the compounds to be evaluated.

Docking-based VS is far from perfect, and there are several issues that require improvement. Correct pose prediction is still a difficult task and accurate free-energy estimations are a major challenge, in part due to ignoring the flexibility of the target and desolvation effects.[7] Additional information about the target of interest, for example, ligand interactions revealed in a large number of cocrystal structures or similarity patterns between known ligands, is sometimes available, and its incorporation into VS campaigns can increase success rates.[8] The improved performance of knowledge-based docking is
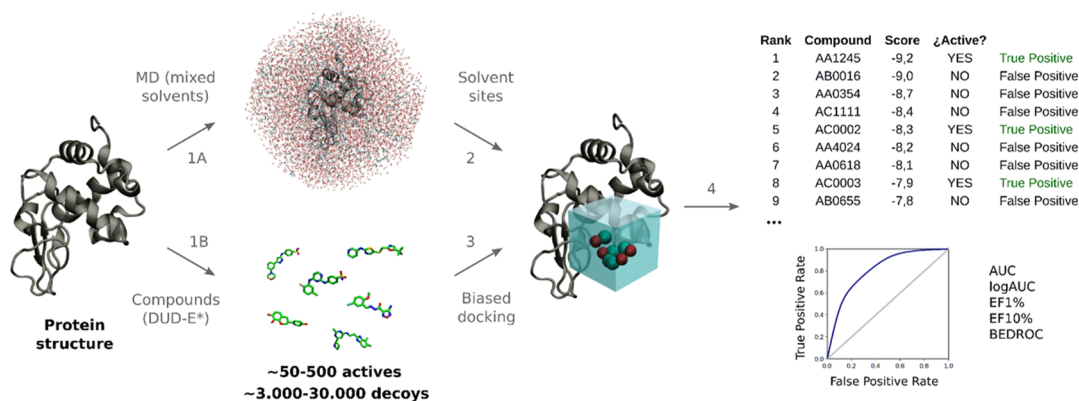
Figure 1. Biased docking scheme. (1A) The protein target is subjected to MD in mixed solvents and (2) solvent sites are determined to be used as docking biases. (1B) In parallel, active compounds against the target and their associated decoys are extracted from the DUD-E database[24] or generated accordingly*. (3) Actives and decoys are docked independently using the solvent sites biased method implemented in AutoDock Bias. (4) Compounds are ranked by their docking score, and ROC-based measures such as AUC, log AUC, EF$_{1\%}$, EF$_{10\%}$, and BEDROC are calculated to assess the suitability of the method for VS. *For targets not included in the DUD-E set, actives and decoys were retrieved as stated in Computational Methods.

highlighted by the different alternatives available in most common docking programs. Glide[9] and GOLD[10] allow hydrogen bonds and substructure-based constraints, whereas Glide also permits metal restraints to enforce coordination geometries. On the other hand, rDock[11] and MOE[12] are able to constrain generated poses to satisfy pharmacophores and thus bias the results toward important interactions and also perform template guided (or tethered) docking. Finally, DOCK6[13] has a conformational search option to bias the sampling toward poses in accordance with a defined set of known ligand structures. Besides docking methods, pharmacophore VS approaches evolved in parallel using either ligand or target information but ignoring explicit protein–ligand interaction.[14]

Molecular dynamics (MD) simulations of a protein in explicit water or mixed solvents that mimic specific types of ligand interactions have been increasingly used to analyze protein binding sites and to reveal protein–ligand interaction hot spots.[15−17] MD allows studying the structure and dynamics of water and cosolvent molecules which, as a consequence of the shape and charge distribution of the protein surface, are arranged inhomogeneously in the solvation shell. This gives rise to solvent sites, that is, confined space regions close to the protein surface where the probability of finding water or cosolvent molecules is significantly higher than that of the bulk. We have previously shown applications of solvent sites to binding mode and binding free-energy predictions,[18] especially using ethanol as a probe that is able to capture in a single molecule the hydrophobic, hydrogen bond donor and hydrogen bond acceptor moieties that are common in drug-like molecules.

On the basis of our work on ligand docking using information from water and solvent sites,[18−20] we have recently developed AutoDock Bias[21] for performing biased docking with AutoDock4[22] program. The versatile definition of the different types of biases in AutoDock Bias allows guided docking toward pharmacophoric interactions in a straightforward way for hydrogen bond and hydrophobic/aromatic interactions, and precise localization of any desired atom (e.g., metal) or group (e.g., substructure core of a congeneric ligand series or for fragment growth) in a defined region of the target protein, which may also be used as an anchor for covalent

docking studies. In the present work, we have retrospectively tested the performance of AutoDock Bias using the information from MD-derived solvent sites for its application in VS campaigns. As previously stated, we chose ethanol as a cosolvent, and once the ethanol sites were computed from the corresponding MD simulations (see Computational Methods), we implemented a biased docking scheme based on their location, binding free energy, and dispersion. We selected nine relevant target proteins that have a set of known ligands or *actives* and compounds that, presumably, do not bind to the protein (*decoys*). Our results show that supplementing the docking conformational search algorithm and scoring function for each target yielded improved VS performance both globally, as seen in receiver operating characteristic (ROC) curves and area under the ROC curve (AUC), and in early enrichment, shown by semilogarithmic ROC curves, log AUC, EF$_{1\%}$, EF$_{10\%}$, and Boltzmann-enhanced discrimination of ROC (BEDROC) metrics.

■ RESULTS AND DISCUSSION

The results are organized as follows. We first present the VS performance analysis of our biased docking method, AutoDock Bias,[21] using MD-derived solvent sites and compare it to conventional docking programs (AutoDock4[22] and Vina[23]) on nine different protein targets. Then, we move on to evaluate the relative scoring and pose prediction of known active molecules by the different docking methods to gain insight into the origin of observed improvement with AutoDock Bias. Finally, we outline the limitations of the proposed method and compare it with similar approaches.

For each of the nine proteins, we run a set of three 20 ns MD simulations in an aqueous solution of ethanol 20% v/v and compute the ethanol interaction sites (see Computational Methods). Interestingly, and confirming our previous study,[18] these MD-derived ethanol sites match the known pharmacophoric ligand interactions for these systems to a great extent. Details for each protein are included in the Supporting Information, and a graphical description is shown in Figure S1, where ethanol sites are superimposed on representative ligands. For their application in AutoDock Bias, both hydrophilic and hydrophobic ethanol sites were identified on each target binding/active site, ranging from one to five
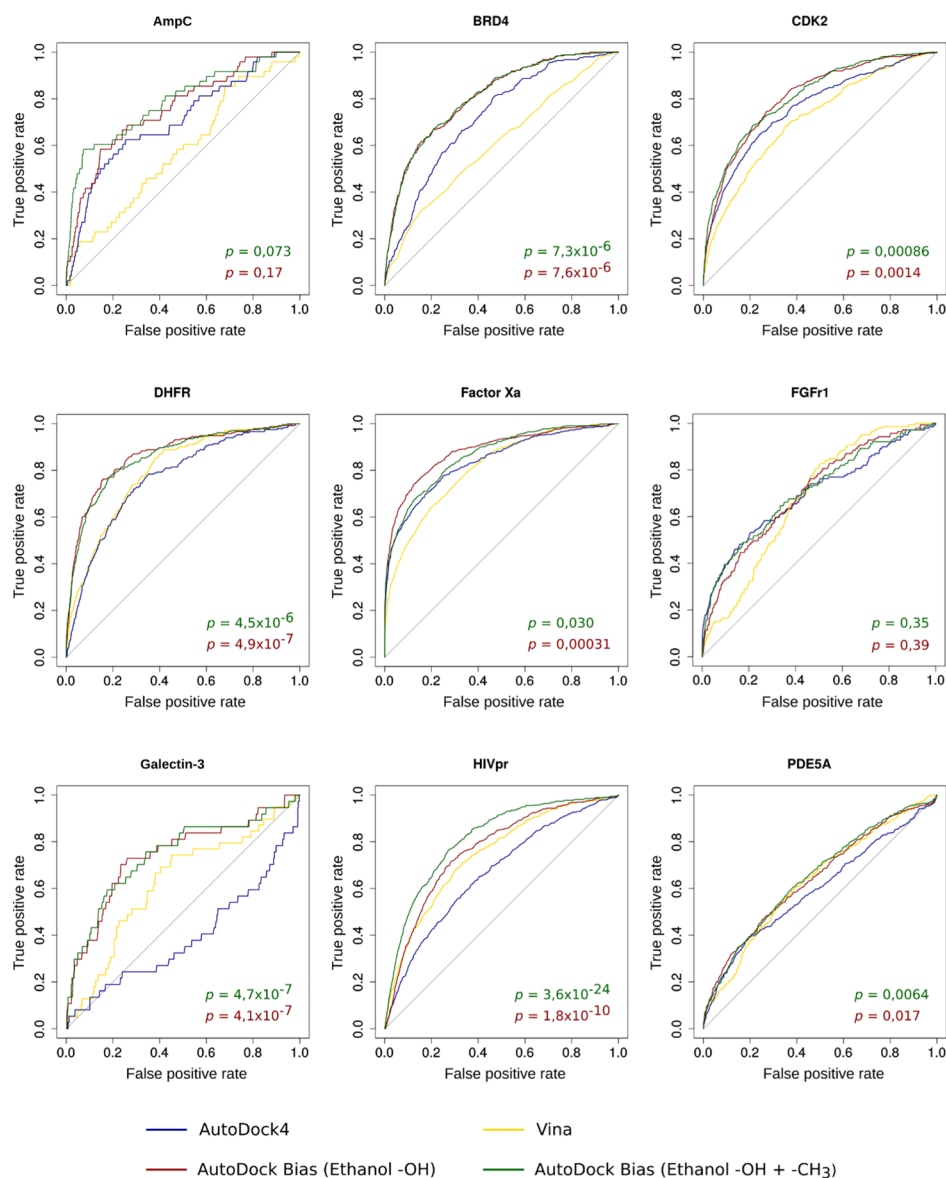
Figure 2. ROC curves for the docking of actives and decoys against nine protein targets. Results obtained with standard AutoDock4 (blue) and Vina (yellow) are compared with those obtained by AutoDock Bias using ethanol hydrophilic sites (red) and both ethanol hydrophilic and hydrophobic sites (green). The gray straight line is expected for a random selection of compounds from the database. The informed p-values correspond to the hypothesis test about the difference in AUC between the conventional method (AutoDock4) and AutoDock Bias (using ethanol hydrophilic sites in red and ethanol hydrophilic and hydrophobic sites in green).

interaction sites of each kind. Hydrophilic sites were used to bias hydrogen-bond interactions according to their donor and/or acceptor capacity, and hydrophobic sites were used to bias hydrophobic interactions involving aromatic rings (see Computational Methods for details).

We tested the high-throughput performance of AutoDock Bias in VS campaigns in a retrospective manner. Seven of the nine studied systems are relevant target proteins taken from the DUD-E database.[24] The two remaining systems, BRD4 and galectin-3, were both subjected to the DUD-E protocol to generate decoy molecules (http://dude.docking.org/generate)[24] and then added to the original set to enrich its diversity. All targets, therefore, have a set of (a) known true ligands called *actives* and (b) *decoys* or compounds that have similar physical and chemical properties as the actives but, presumably, do not bind to the target protein. After docking actives and decoys to their target, a ranking of the compounds

was built based on their score, and ROC-derived measures such as AUC, log AUC, $EF_{1\%}$, $EF_{10\%}$, and BEDROC were calculated as defined in the Supporting Methods. The whole process is outlined in Figure 1.

AutoDock Bias for VS. To quantify the ability of AutoDock Bias and the standard docking methods to rank active compounds in the first positions, that is, their VS performance, we calculated different metrics that evaluate both the overall performance (ROC curve and AUC) and the early ligand enrichment (semilogarithmic ROC curve, log AUC, enrichment factors, and BEDROC). The early ligand enrichment is of particular interest because, in practice, only a small subset of the original database, usually less than 10%, is experimentally tested. Therefore, it is desired that the first 1 or 10% of the ranking is significantly enriched in true ligands that can move on to the next stages of the drug-development process.

Table 1. Ligand Enrichment Measures[b]

protein	docking method	AUC	log AUC	EF$_{1\%}$ (%)	EF$_{10\%}$ (%)	BEDROC
AmpC	AutoDock4	0.71	0.12	2.1	39.6	0.19
	Vina	0.58	0.05	0.0	18.8	0.12
	AutoDock Bias (ethanol OH)	0.76	0.19	10.4	41.7	0.28
	AutoDock Bias (ethanol OH + CH$_3$)	0.79	**0.26**	14.6	58.3	0.41
BRD4	AutoDock4	0.72	0.13	6.5	25.4	0.18
	Vina	0.61	0.07	4.8	23.4	0.14
	AutoDock Bias (ethanol OH)	**0.81**	**0.23**	12.9	52.4	0.33
	AutoDock Bias (ethanol OH + CH$_3$)	**0.81**	**0.22**	12.5	51.6	0.32
CDK2	AutoDock4	0.76	0.20	13.9	43.0	0.30
	Vina	0.71	0.15	8.9	33.3	0.23
	AutoDock Bias (ethanol OH)	**0.81**	0.23	13.1	50.2	0.32
	AutoDock Bias (ethanol OH + CH$_3$)	**0.81**	**0.25**	17.5	51.1	0.36
DHFR	AutoDock4	0.76	0.15	3.9	39.4	0.22
	Vina	0.79	0.20	9.5	39.0	0.27
	AutoDock Bias (ethanol OH)	**0.86**	**0.30**	14.7	64.5	0.43
	AutoDock Bias (ethanol OH + CH$_3$)	**0.86**	**0.30**	19.5	64.1	0.42
FXa	AutoDock4	0.84	0.36	32.8	60.3	0.51
	Vina	0.81	0.27	21.4	49.5	0.40
	AutoDock Bias (ethanol OH)	**0.88**	**0.42**	39.3	70.4	0.58
	AutoDock Bias (ethanol OH + CH$_3$)	**0.86**	0.39	37.8	63.7	0.53
FGFr1	AutoDock4	0.70	0.19	16.5	39.6	0.29
	Vina	0.68	0.09	5.0	16.5	0.13
	AutoDock Bias (ethanol OH)	0.71	0.13	7.9	33.1	0.21
	AutoDock Bias (ethanol OH + CH$_3$)	0.71	0.17	10.8	39.6	0.27
Gal-3	AutoDock4	0.39	−0.01	2.7	8.1	0.08
	Vina	0.62	0.06	5.1	12.8	0.10
	AutoDock Bias (ethanol OH)	**0.74**	**0.18**	10.8	37.8	0.27
	AutoDock Bias (ethanol OH + CH$_3$)	**0.74**	**0.18**	13.5	37.8	0.29
HIVpr	AutoDock4	0.67	0.09	3.0	26.9	0.15
	Vina	0.74	0.15	5.6	36.6	0.22
	AutoDock Bias (ethanol OH)	**0.77**	**0.14**	2.8	36.6	0.19
	AutoDock Bias (ethanol OH + CH$_3$)	**0.81**	**0.20**	7.1	48.7	0.28
PDE5A	AutoDock4	0.60	0.08	4.5	23.6	0.15
	Vina	0.63	0.09	6.5	19.1	0.14
	AutoDock Bias (ethanol OH)	**0.64**	**0.11**	7.0	28.6	0.19
	AutoDock Bias (ethanol OH + CH$_3$)	**0.65**	**0.11**	7.5	24.6	0.17
mean (SD)[a]	AutoDock4	0.68 (0.13)	0.15 (0.10)	9.5 (10.1)	34.0 (14.9)	0.23 (0.13)
	Vina	0.69 (0.08)	0.13 (0.07)	7.4 (5.9)	27.7 (12.4)	0.19 (0.10)
	AutoDock Bias (ethanol OH)	0.78 (0.07)	0.21 (0.10)	13.2 (10.4)	46.1 (14.3)	0.31 (0.13)
	AutoDock Bias (ethanol OH + CH$_3$)	0.78 (0.07)	0.23 (0.08)	15.6 (9.3)	48.8 (13.0)	0.34 (0.10)

[a]Mean (standard deviation) for all systems. [b]Values highlighted in bold represent significant improvement in AUC and/or log AUC over standard docking procedures ($p < 0.05$, see text).

Figures 2 and S2 show the ROC and semilogarithmic ROC curves, respectively, for nine protein targets obtained with different docking methods, two standard docking programs (AutoDock4 in blue and Vina in yellow) and AutoDock Bias using ethanol sites. For AutoDock Bias, the performance using only hydrophilic ethanol sites and both hydrophilic and hydrophobic ethanol sites is shown in red and green, respectively. The expected behavior for a random extraction of compounds is also depicted in gray for comparative purposes. Table 1 shows the results obtained for ROC-derived metrics corresponding to the same set of proteins and docking methods.

The first general observation emerging from Figures 2 and S2 and Table 1 is that conventional methods, AutoDock4 and Vina, have very different performances depending on the tested system, showing poor metrics for several of them. This occurs both globally, as represented by ROC curves rising only

slightly above the line corresponding to a random extraction of compounds from the database and AUC values below 0.70, and also at an early stage, observed in semilogarithmic ROC curves not far from the random curve and log AUC values below 0.15. These arbitrary thresholds were derived from the mean values for each metric considering conventional methods. On three of the studied target systems, namely Factor Xa, DHFR, and CDK2, conventional methods have good (AUC > 0.70, log AUC > 0.15) to excellent (AUC > 0.80, log AUC > 0.25) performance, whereas on the other six, poor metrics are obtained for either one or both methods. Interestingly, both methods fail for the two systems that were not retrieved from the DUD-E database (BRD4 and galectin-3). The case of galectin-3 is particularly striking since AutoDock4 shows a worse overall ranking than what would be obtained by random selection, whereas the early enrichment measured by log AUC is practically zero. This is in agreement

with previous observations of poor results for carbohydrate docking.[19] Overall, the comparative performance between both conventional methods, especially at early enrichment, shows that for most systems, AutoDock4 gives substantially better results than Vina (AmpC, BRD4, CDK2, Factor Xa, FGFr1), whereas for some others, the trend is reversed (DHFR, galectin-3, HIV protease). Statistical differences between AutoDock4 and Vina performances are shown in Table S1, and a discussion in the context of a recent study[25] that concluded that Vina outperforms AutoDock4 using the CASF2013 data set is presented in the Supporting Discussion.

On the contrary, AutoDock Bias achieves good performance (AUC > 0.70, log AUC > 0.15) in five systems and excellent performance (AUC > 0.80, log AUC > 0.25) in other three systems. The only target where the performance of the biased methods is still poor, albeit better than the conventional ones, is phosphodiesterase 5A. Also, when using only ethanol hydrophilic (ethanol −OH) sites as a bias, a low early enrichment was achieved for FGFr1 and HIV protease. In relative terms, both biased methods have superior performance than conventional ones for all systems except FGFr1 at early enrichment. The conformation of the active site of this flexible kinase target[26] is fixed in the DUD-E structure to the so-called DFG-in state, which was the one sampled during the MD simulations used to determine the ethanol sites. It seems that the bias toward interactions in this particular conformation strongly penalized DFG-out ligands (actives) that might be present in the DUD-E dataset. This is expected since the DFG-in and DFG-out conformations are structurally and energetically dissimilar.[27] Finally, considering all targets, an increase in mean results is obtained for all metrics with the biased methods, although this has a limited value especially due to the difference in performance between systems. In this regard, the lower standard deviation values highlight that AutoDock Bias performance using both hydrophilic (−OH) and hydrophobic (−CH$_3$) ethanol sites is more consistent between diverse systems than conventional AutoDock4.

To statistically analyze the observed differences between conventional methods and AutoDock Bias, a one-tailed hypothesis test was performed using bootstrap (see Supporting Methods). The obtained p-values for the AUC are shown next to the ROC curves (Figure 2) and for the log AUC next to the semilogarithmic ROC curves (Figure S2) from which they are derived, for both biased methods. Taking the typical value of 0.05 as the significance level for the test, the AUC and log AUC values obtained using AutoDock Bias with both hydrophilic and hydrophobic ethanol sites are significantly higher than those of conventional AutoDock4 for seven out of the nine systems. The systems where no significant improvement (nor decline) was observed are FGFr1/AmpC for AUC and FGFr1/FXa for log AUC. A similar enhancement is achieved by the biased method with only ethanol −OH sites for the AUC values, but the improved performance is reduced to six out of nine targets for log AUC. The tests were performed against AutoDock4 because it is the conventional method that gave the best overall results and the one used as a baseline for our energy modifications. Beyond the statistical analysis of AUC and log AUC values, there are five targets (AmpC, BRD4, DHFR, galectin-3, and HIV protease) for which the improvement with AutoDock Bias using both types of ethanol sites is evident both qualitatively, by the separation between the corresponding ROC and semilogarithmic ROC curves, and quantitatively, especially in terms of very early

enrichments: EF$_{1\%}$ is 2−7 times higher than those obtained by the conventional methods. For example, AmpC has ca. 7-fold increase in its EF$_{1\%}$ for AutoDock Bias. This means that if a database is composed of 100 000 compounds out of which 100 are true ligands, we expect to experimentally obtain 2 hits if we select compounds based on standard docking software, but instead to get 15 hits if using AutoDock Bias, for every 999 confirmed nonbinders.

To summarize the overall performance of AutoDock Bias using hydrophilic and hydrophobic ethanol sites compared to that of conventional methods, regarding both global and early enrichment, Figure 3 shows their corresponding difference in
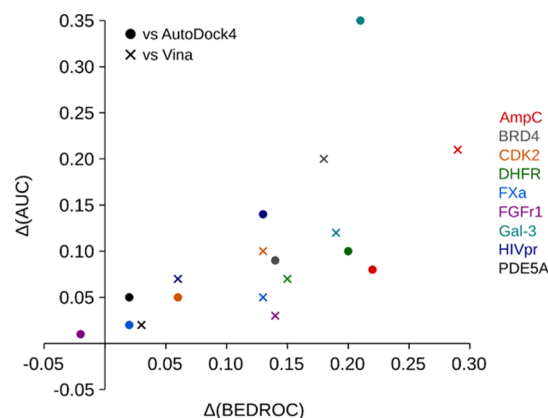

Figure 3. Δ(AUC) vs Δ(BEDROC) between AutoDock Bias using ethanol hydrophilic + hydrophobic sites and AutoDock4 (dots) or Vina (crosses). Results are color coded according to each system.

AUC and BEDROC for all systems. The Δ(AUC) versus Δ(BEDROC) plot displays improvement for AutoDock Bias in practically all of the systems, as revealed by the points in the (+,+) quadrant. Most systems have BEDROC differences greater than 0.10, which, according to the statistical interpretation of this parameter (see Supporting Methods), are values of substantial improvement. As previously analyzed, the only case in which AutoDock Bias does not attain better early parameters than AutoDock4 is FGFr1, yielding a value of BEDROC lower in 0.02 units and a drop of 5.7% in EF$_{1\%}$. However, the loss of enrichment is quickly reversed at EF$_{10\%}$ and thereafter, and improvement is observed with respect to Vina at all stages. Finally, the correlation for the increase in both global and early metrics seen in Figure 3 indicates that a general improvement is achieved with the biased method. However, the rise in global enrichment appears somewhat more moderate than the increase in early enrichment, suggesting that the bias causes a substantial increase in the score of some actives relative to decoys, but cannot fix grossly incorrect predictions, for example, due to inadequate protein conformation.

Scoring and Binding Mode of Actives. To analyze the effect of the biased conformational search and scoring on the docking of known binders, we compared it to conventional docking methods regarding (a) the relative ranking of active molecules and (b) the binding mode prediction of actives with a known bound crystal structure (see Computational Methods).

The rankings of active molecules obtained with standard AutoDock4 and AutoDock Bias using both hydrophilic and hydrophobic ethanol sites were compared by calculating the

Kendall tau coefficients (Table S2). Eight out of nine systems display a correlated order of active molecules in both rankings ($\tau \geq 0.5$). Therefore, when using AutoDock Bias, a general rise in the ranking of actives as a group with respect to decoys is accomplished, whereas the order between actives themselves is not significantly altered when compared to conventional docking results. This highlights that the biased docking method maintains the relative binding affinity predicted for actual ligands by AutoDock4, whose scoring function was calibrated with known binders, while better discriminating between binders and nonbinders by enforcing specific target important interactions.

We also compared actives docking pose with their crystallized structure with the aim of discerning if better ranking correlates with a more accurate pose prediction. Heavy atom root-mean-square deviations (rmsds) between the predicted pose for active molecules and the cocrystallized conformation of the same compound or a similar one (Tanimoto similarity index ≥ 0.85) are presented in Figure 4. In five of seven systems, there was an improvement in pose
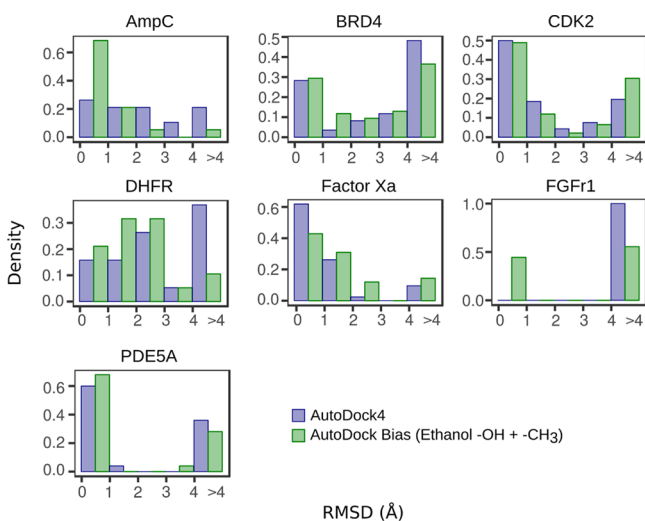

Figure 4. Histograms for heavy atom rmsd between predicted poses and crystallized conformations of ligands with Tanimoto similarity index equal to or greater than 0.85. Results for standard AutoDock4 are shown in blue and for AutoDock Bias with hydrophilic and hydrophobic ethanol sites in green.

prediction using AutoDock Bias: mild (5% more actives with rmsd < 2 Å) for PDE5A and DHFR, moderate (10%) for BRD4, and significant (45%) for FGFr1 and AmpC. However, there is slightly worse prediction for CDK2 (−5%) and a moderate one (−16%) for Factor Xa. Therefore, although improving pose prediction in the majority of tested proteins, AutoDock Bias performance remains system-dependent for significantly improving binding mode calculations.

It can be seen from Figure 4 that there are several systems with highly inaccurate binding mode predictions (rmsd > 4 Å). In BRD4, the reason seems to be that the main driving force for the binding is a single hydrogen bond interaction with Asn140, aided by variable nondirectional hydrophobic interactions. An additional difficulty for the docking is that there are known interactions bridged by water molecules. The poor pose prediction holds even with the biased method and is related to the lack of solvent sites—except the ones near the Asn140—that would aid in positioning the ligands correctly in

the whole binding site. On the contrary, the enhanced formation of the hydrogen bond between Asn140 and the ligands in the biased docking runs is enough to improve the VS performance, as revealed by significant increases in AUC and log AUC (Table 1). Unlike BRD4, DHFR constitutes a case where improvement in VS performance by the biased method is explained by better binding mode prediction. The conventional method has problems with several similar active ligands that repeatedly appear with their terminal fused aromatic rings inverted in the binding site, dramatically changing the interaction pattern (Figure S6A,B). The bias toward the hydrogen bond interaction between Glu30 and the diamino-pyrimidine ring (Figure S1D) corrects this issue for the biased method. Last, there are problematic ligands whose poses cannot be predicted by any of the methods and correspond to structures of considerable size with around 14 torsions and 50 atoms excluding nonpolar hydrogens (Figure S6C). Finally, for PDE5A and FGFr1, there is a pose prediction improvement with the biased method, but as previously discussed, the flexibility of both protein targets prevents from achieving better predictions since the protein conformations from the DUD-E set are substantially different from the holo structures found for certain types of ligands from the active set. Interestingly, the binding mode of ligands is certainly better captured by the biased method for FGFr1, a system whose improvement in enrichment parameters is quite poor (Table 1). We can rationalize this result taking into account that all ligands with rmsd below 2 Å are bound to the DFG-in protein conformation in the corresponding crystal structure, which matches the conformation selected from the DUD-E database for the docking experiments. This points out again that AutoDock Bias is probably incapable of assigning high scores to ligands that bind the DFG-out conformation and are present as actives in the test set.

Finally, it is interesting to dig deeper into the results for Factor Xa that show worse pose predictions with AutoDock Bias than with standard AutoDock4. This observation can be explained by the lack of directionality of the hydrophobic/aromatic pharmacophore derived from the MD, which is the binding driving force for this system. This feature is unique to Factor Xa since all other systems from the set have a robust polar pharmacophore which provides directionality for ligands (Figure S1). In this sense, it should be interesting to derive a better bias for the aromatic moieties that also includes directionality, for example, by using phenol as a cosolvent. In conclusion, the improvement observed in the classification between actives and decoys by the addition of the hydrophobic bias is relevant for VS but needs additional adjustment since ligand correct orientation is not well achieved when no hydrogen bonding is included. As a final remark, several FXa ligands bind benzamidine or other positively charged groups. Ethanol being a neutral molecule was able to identify this pharmacophoric hydrogen bond, but an ionic probe would probably assign better energy prioritization and predict correct poses for more ligands.

Discussion in Relation to Other Methods. Pharmacophores obtained from the protein structure, with no previous ligand-based knowledge, are versatile for the discovery of new drugs since they can be employed in novel targets for which limited information is available. Likewise, in the VS of large amounts of compounds, they increase the probability that the obtained hits have different chemical functionalities than the ligands already tested for the target of interest. Moreover, the

Table 2. Protein Set for VS Assessment

protein name	function	no. actives[a]	no. decoys[a]	PDB ID	resolution (Å)
AmpC β-lactamase (AmpC)	C–N hydrolase	48	2832	1l2s	1.94
bromodomain-containing protein 4[b] (BRD4)	bromodomain	248	15 009	4zc9	0.99
cyclin-dependent kinase 2 (CDK2)	kinase	474	27 830	1h00	1.60
dihydrofolate reductase (DHFR)	reductase	231	17 170	3nxo	1.35
Factor Xa (FXa)	protease	537	28 325	3kl6	1.45
fibroblast growth factor receptor 1 (FGFr1)	kinase	139	8651	3c4f	2.07
galectin-3[b] (Gal-3)	lectin	37	1694	1kjr	1.55
human immunodeficiency virus type 1 protease (HIVpr)	protease	536	35 674	1xl2	1.50
phosphodiesterase 5A (PDE5A)	esterase	398	27 521	1udt	2.30

[a]The number of actives and decoys correspond to unique compounds. Different species (tautomers, protomers, and conformers) were docked for each compound, thus increasing the total number of molecules. The final score for each compound corresponds to the most favorable score among all its species. [b]Systems not included in the DUD-E dataset.

protein-derived pharmacophores obtained in this study consider desolvation in an explicit way and protein flexibility to a certain degree because they were obtained from unrestrained MD simulations in mixed solvents.

To achieve more profound improvements with the developed method, we must answer two central questions. The first one is why some actives still do not appear in the upper part of the ranking. A careful analysis of ligand properties show that, in general, actives that are not retrieved at early or middle stages are relatively small molecules compared to those ranked higher (Figure S4). Although this is expected given that docking score scales with the size of compounds, AutoDock Bias seems less affected by the ligand size effect on the ranking, as revealed by the Spearman's correlation coefficients closer to 0, than the conventional method for seven of the nine systems (compare Figures S3 and S4). The second question is why some decoys remain in the upper portion of the ranking even when the biased method is applied. We found that many highly ranked decoys adopt excessively folded and seemingly entropically unfavorable conformations that allow them to coincide with several solvent sites and thus increase their estimated binding energy. This is possible because the scoring function has a poor balance between the interaction energy and conformational entropy. On the other hand, we also observed that there are cases in which the best-ranked decoys preferentially replace solvent sites that are not coincident with ligand-derived pharmacophoric sites when compared to best-ranked actives. In general, potential areas for future improvement include the use of additional solvent probes that capture other interaction types (e.g., ionic groups) and taking advantage of the information about displaceable/nondisplaceable water molecules that already comes out of the MD simulation in mixed solvents.

Finally, it is interesting to compare our method with other works using MD in mixed solvents in the context of VS. MacKerell's group pioneered the use of pharmacophore models obtained by cosolvent MD for screening a large number of compounds.[28,29] In their approach, ligand features are required to match all MD-derived pharmacophoric sites, and the score arises from the rmsd between these paired features, supplemented with free-energy grid calculations for the ligand pose(s). The method is useful to filter compounds that are not able to fulfill the pharmacophoric sites, but does not aid in the conformational search itself. Indeed, it relies on predefined ligand conformations that are not adjusted to the target structure. Nonetheless, since there is no rigid receptor, protein flexibility from MD simulations is totally accounted in

the pharmacophoric sites, and excluded space regions are determined from the absence of probe localization during the MD rather than from a given protein surface. On the other hand, Uehara and Tanaka used cosolvent MD to generate an ensemble of protein conformations for docking-based VS.[30] The ensemble selection relies on clustering the conformations according to binding site residues and ranking the clusters by predocking known actives, considering the ones achieving higher scores. Finally, ensemble docking is performed using standard Vina. This approach is totally complementary with our method since the final ensemble docking could be done with AutoDock Bias using solvent sites extracted from the same MD simulations required to obtain the target receptor ensemble. To our knowledge, the AutoDock Bias implementation presented here is the only method where cosolvent-derived information is used to modify the scoring function of a docking program, both for conformational search and final affinity estimation, with the aim of improving VS results.

■ CONCLUSIONS

In the present work, we demonstrate the potential of biased docking using pharmacophores derived from MD in mixed solvents (solvent sites) to increase the enrichment of true ligands in VS campaigns. Significantly higher AUCs and log AUCs were obtained with respect to conventional docking methods, indicating improvement both in the global ranking of compounds and, especially, in early stages. The latter important issue is additionally reflected in ligand enrichments at 1% which were multiplied by up to 7 times. The most successful variant is to perform MD simulations in the ethanol-water mixture, obtain both ethanol hydrophilic and hydrophobic interaction sites, and add this solvent information to the docking protocol of AutoDock Bias.

■ COMPUTATIONAL METHODS

Protein and Ligand Dataset. Nine protein targets were selected (Table 2) to test the performance of the biased docking method for VS. They belong to different families, have various functions, and represent a diverse set of targets relating its protein–ligand interactions (Figure S1). The protein structures for seven of the targets were obtained from the DUD-E database,[24] a repository of relevant target protein structures, each bearing a set of true ligands or *actives* with an experimentally measured affinity below 1 μM and nonbinders or *decoys*, which are compounds with similar physicochemical properties as the actives but a different topology based on two-dimensional (2D) fingerprints. In addition to these focused

decoys, molecules proved to be experimentally inactive (affinity greater than 30 μM) were also added to the database when that information was available. The DUD-E, together with its predecessor DUD,[31] has been proven useful for benchmarking VS procedures.[32] The two remaining systems not present in the benchmarking database are BRD4, a promising target for disrupting protein−protein interactions, and galectin-3, a docking-challenging carbohydrate binding protein with a binding site quite exposed to the solvent and known ligands with increased complexity due to high number of rotatable bonds. The structures of these two targets were obtained from the PDB in the holo form with a resolution lower than 2.5 Å and in a conformation that could accommodate the greater number of cocrystallized ligands. These resulted in PDB ID 4zc9 for BRD4 and 1kjr for galectin-3. Subsequently, a procedure analogous to that of the DUD-E was carried out. First, actives were recovered from BindingDB,[33] BindingMOAD,[34] and ChEMBL[35] databases, only considering compounds with experimentally measured K_i, K_d, or IC50 lower than 1 μM against the corresponding target. Then, a structural diversity filter was applied with DataWarrior program,[36] based on three-dimensional (3D) similarity according to protein-binding pharmacophoric characteristics (Flexophore similarity descriptor[37]). Finally, the decoys were obtained from these actives through the DUD-E web tool (http://dude.docking.org/generate),[24] which recovers 50 decoys extracted from ZINC[38] for each active protomer in the SMILES format. SMILES for actives and decoys of BRD4 and galectin-3 are available in the Supporting Information. In summary, the number of active ligands per system ranges from 37 to 537, while the number of decoys varies between 1694 and 35 674, so different sizes of databases were analyzed with the method (Table 2).

Workflow and ROC Metrics. To evaluate the performance of different docking methods in VS campaigns, we followed the process outlined in Figure 1 for each target. After retrieving actives and decoys, all compounds were docked against the protein structure stated in Table 2 and Supporting Methods, and a ranking was built according to their predicted affinity. The predominance of active ligands in the first positions of the ranking was quantified by the following metrics: ROC curve, AUC, semilogarithmic ROC curve, adjusted area under the semilogarithmic ROC curve (log AUC), enrichment factors at 1 and 10% ($EF_{1\%}$ and $EF_{10\%}$), and BEDROC.[39] The definition of each metric is detailed in the Supporting Methods. This process was repeated for each of the following docking methods: conventional AutoDock4, conventional Vina, AutoDock Bias using hydrophilic ethanol sites (ethanol −OH), and AutoDock Bias using both hydrophilic and hydrophobic ethanol sites (ethanol −OH + −CH$_3$). AutoDock4[22] and Vina[23] were chosen for comparison because they are the two most widely distributed freeware docking programs and because our biased method is based on the AutoDock4 scoring function.

The calculation of the different metrics from the rankings was made with R.[40] In particular, we use several functions from the "ROCR" package[41] to calculate ROC curves and AUCs, the "pracma" package to obtain the log AUC through integration using trapezoids, and the "pROC" package[42] to calculate the hypothesis test for the difference in AUC. The hypothesis test for the difference in log AUC was programmed in R following the guidelines included in the Supporting Methods.

MD in Mixed Solvents. The MD simulations in aqueous ethanol for AmpC, CDK2, DHFR, Factor Xa, FGF receptor 1, and PDE5A were previously performed and described.[18] BRD4, galectin-3, and HIV protease were simulated following the same protocol. Initial structures were downloaded from the PDB database[43] (PDB IDs 4zc9 for BRD4, 1kjr for galectin-3, and 1xl2 for HIVpr). All nonstructural ions and solvent and ligand molecules were removed. Missing side chains and hydrogen atoms were added using the LEaP module from Amber package (version 16).[44] Standard protonation state at physiological pH was assigned to all ionizable residues, except for HIV protease that was simulated with Asp124 protonated. The structures were then solvated with a truncated octahedral box of ethanol 20%V/V extending at least 10 Å from any protein atom using MDMix program.[15] TIP3P model was used for water molecules, and forcefield parameters for ethanol were assigned as previously reported.[18] Amber ff14SB forcefield was used for protein residues. All MD simulations were performed with the PMEMD implementation of SANDER for GPU from Amber 16.[45] Details on ethanol concentration and simulation protocol can be found in the Supporting Methods.

MD-Derived Solvent Sites. The MD-derived solvent sites (ethanol sites) were obtained as explained in our previous work.[18] Briefly, ethanol sites are defined as space regions adjacent to the protein surface where the probability of finding a probe atom of this cosolvent is higher than that in the bulk. We obtained both types of ethanol sites: hydrophilic sites in regions where the −OH resides for a significant amount of time during the MD simulation and hydrophobic sites in regions where the terminal −CH$_3$ had a considerable occupancy. The preference of the ethanol probe for being in the solvent site with respect to the bulk is given by the free energy of the site defined in eq 1

$$\Delta G = -RT \ln \frac{N_{\text{solvent site}}}{N_{\text{bulk}}} \tag{1}$$

where $N_{\text{solvent site}}$ is the number of times a probe atom is found inside the solvent site during the MD simulation and N_{bulk} is the number of times a probe atom is found in the bulk inside a sphere with the same volume of the solvent site (arbitrarily taken as 1 Å^3). Besides its free energy, a solvent site is also described by its coordinates, that is, the center of mass of the probe atoms defining the site and its radius or dispersion parameter, that is, the radius necessary to get a sphere containing 90% of the probe atoms defining the site.

Ethanol sites were obtained exclusively for the binding (or active) site of each target. For each protein, the binding pocket was defined using all protein−ligand cocrystal structures available in the PDB, having a resolution below 2.5 Å and a reported binding affinity (K_d, K_i or IC$_{50}$) below 1 mM. All protein residues that had at least one atom within 4 Å from any ligand atom in at least one structure were considered part of the pocket. To select the cosolvent pharmacophoric points (solvent sites) for the VS application, we first identified all possible solvent sites in the binding pocket. Then, we kept those with ΔG of binding above 0.95 kcal/mol ($\Delta G < -0.95$ kcal/mol, a threshold representing a probe-finding probability at least 5 times higher than the one in bulk, eq 1). Similar cutoffs have been used in previous studies for deriving pharmacophore profiles.[15,28] Finally, solvent sites clashing with the particular protein structure selected for docking, as a consequence of the receptor flexibility during the MD, were

removed. The distance cutoff between a solvent site and any protein heavy atom below which the solvent site was discarded was set at 2.1 Å for hydrophilic sites and 2.9 Å for hydrophobic sites (a 0.5 Å tolerance from the minimum values for hydrogen bonds and aliphatic/aromatic interactions, respectively, reviewed by Bissantz et al.[46]). As a final remark, the solvent sites overlapping the NADP cofactor from DHFR were also removed because the cofactor was retained for the docking as if it was part of the receptor.

Conventional and Biased Docking. *Protein, Ligand, and Grid Preparation for Docking.* Protein structures were obtained from the DUD-E database for seven systems (AmpC, CDK2, DHFR, Factor Xa, FGF receptor 1, HIV protease, and PDE5A) and directly from the PDB for BRD4 (4zc9) and galectin-3 (1kjr). They were checked and protonated when needed with Maestro[47] and then converted to PDBQT with AutoDockTools[22] (prepare_receptor4.py). Structural details can be found in the Supporting Methods. Each receptor was kept rigid for docking. Structural ligand files were obtained from the DUD-E database in the SDF format and converted to PDBQT format with Open Babel.[48] For the systems not present in the DUD-E database, the 3D structures of the compounds were generated with LigPrep[49] from the previously obtained SMILES (see "Protein and Ligand Dataset"), considering stereoisomers, tautomers, and ring conformers.

For building the grid that defined the docking site for each target, all protein−ligand crystallographic complexes available from the PDB, with resolution below 2.5 Å and a reported binding affinity (K_d, K_i or IC_{50}) below 1 mM, were used. For systems with less information, such as galectin-3, all complexes from the PDB were retrieved. The structures were aligned according to the binding site residues to the protein structure selected for docking (see above). The grid defining the binding pocket was centered in the geometric center of all aligned ligand atoms and extended to reach a rectangular prism with an edge distance equal to 2 times the maximum distance between any two ligand atoms in each direction (*x*,*y*,*z*). The spacing between grid points was set at 0.375 Å. All receptor structure files (PDBQT) are provided as Supporting Content, and docking grids position and dimensions are available in Table S3.

AutoDock4 and Vina. On the one hand, conventional docking experiments were performed with AutoDock4[22] program (version 4.2.6). Briefly, based on the protein structure, the energy maps were computed as usual inside the grid. One hundred independent docking runs were performed, and the resulting poses were clustered according to the ligand heavy atom rmsd using a cutoff of 2 Å, thus defining a population for each cluster. The Lamarckian genetic algorithm was applied with a maximum of 27 000 generations or 2 500 000 energy evaluations. On the other hand, when using Vina[23] for the docking runs, the parameters of the default search algorithm were used, and nine binding modes were generated for each ligand. The energy maps were calculated internally during the docking experiment using the same position and dimensions as for AutoDock4.

AutoDock Bias with Solvent Sites. The solvent site-biased docking method has been described in detail in our previous work.[18] Here, we only use solvent sites obtained from MD simulations in ethanol as a cosolvent because of the specificity and precision of the probe and the binding mode results previously shown.[18] To adapt the ethanol sites obtained from a flexible MD simulation to the rigid structure used in docking

experiments, a positional correction for the hydrophilic sites was made according to their hydrogen bond donor and/or acceptor capacity. The coordinates of the hydrophilic ethanol sites were translated to the ideal location of the hydrogen bond formed according to the conformation of the interactor group of the protein, following the guidelines of Boobbyer et al.[50] on which AutoDock4 scoring function is also based. Thus, acceptor sites remained located at ideal positions for heavy atoms and donor sites at those for H atoms. A single solvent site was chosen for each protein interactor atom, that is, not more than one solvent site of the same type can interact with the same protein atom. Then, AutoDock4 grid maps were modified such that O/N acceptor atoms from the ligand have an energy reward in the corrected location of acceptor ethanol sites while ligand H donor atoms have it in the corrected location of donor ethanol sites. In addition, aromatic ring centers from the ligand have an energy reward in the regions where hydrophobic ethanol sites were established (no positional correction was done because of lack of well-established directionality for these interactions). The rationale for guiding aromatic rings to hydrophobic sites derived from a nonaromatic probe comes from the finding made by Barril's group,[15] which showed that the terminal C of ethanol was able to reproduce ligand aromatic interactions with the same specificity as aliphatic ones. This was further supported by our previous work,[18] where we showed that ethanol has a sensitivity around 70% for ligand hydrophobic interactions and that, even though it does not have aromatic properties, has a slight preference for reproducing interactions of ligand aromatic rings compared to aliphatic rings or side chains.

The magnitude of the energy reward applied to the modified maps is directly related to the free energy of the ethanol sites according to eq 2

$$\Delta G_{bias} = \Delta G_{ori} + \sum_{i=1}^{N} \Delta G_i \times e^{-\frac{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}{r_i^2}} \tag{2}$$

where ΔG_{bias} is the resulting modified score, ΔG_{ori} is the original AutoDock4 score, ΔG_i is the free energy of the *i*-th solvent site, (*x*,*y*,*z*) are the grid point coordinates, (x_i,y_i,z_i), and r_i are the coordinates and radius of the *i*-th solvent site, and the sum extends over the total number N of ethanol sites. Modified OA/NA/HD energy maps built by AutoDock Bias are then used in the same way and in place of conventional AutoDock4 maps. For the aromatic bias, a new map was built using eq 2 with $\Delta G_{ori} = 0$ and affecting a dummy atom type located in the center of ligand aromatic rings. The rest of docking parameters were kept as with conventional AutoDock4 runs for sake of comparison.

Final Score. As it is common with AutoDock4, we used a combined scoring scheme. The consistency of a particular solution in repeated and independent docking calculations was evaluated by gathering the different predicted poses according to the rmsd of their atom coordinates, thus defining a population parameter that counts equivalent poses. Then, the scoring criterion used for each docked ligand consisted in retaining the pose with the lowest free energy of binding among those with a population equal or greater than 20% (if the population resulted in less than 20% for all poses, then the lowest global energy was taken). Finally, a ranking of compounds was generated according to the final score obtained in this way. For the same compound, only the protomer, conformer, tautomer, or stereoisomer with the most

favorable energy score was considered after applying the criterion just discussed. For Vina, the better ranked binding mode was merely taken with it associated score.

Binding Mode Prediction. Regarding pose prediction assessment, for each target, cocrystallized ligands with a Tanimoto similarity index equal to or above 0.85 against any of the actives of that target were retrieved from the PDB (see Supporting Methods for notes on the selected similarity threshold). The computation of similarity was done directly on the PDB server, using the extended connectivity fingerprint method from ChemAxon JChem. Target proteins were filtered by UniProt ID with sequence identity above 90% from the reference structure selected for docking. For each corresponding pair (docked active pose/crystallized ligand conformation), the protein structures were aligned according to the binding site, the maximum common substructure (MCS) between the ligands was found, and their heavy atom rmsd was computed using the RDKit.[51] When many MCSs alignments and/or crystallized ligands were available, the reported rmsd is the minimum of them. Crystallized ligands outside the docking grid were left out of the analysis. The procedure was performed for conventional and biased docking results and comprised 19 ligands for AmpC, 86 for BRD4, 92 for CDK2, 19 for DHFR, 84 for Factor Xa, 9 for FGFr1, and 25 for PDE5A. Galectin-3 lacks a proper number of cocrystallized ligands similar enough to the active molecules (it has only three crystallized ligands on the PDB which have binding data on ChEMBL, one of the databases used to build the active molecules set), and HIV protease structure from the DUD-E database differs enough from several other PDB entries, making the alignment unfeasible.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.9b00371.

> MD-derived solvent sites, semilogarithmic ROC curves for the docking of actives and decoys, number of heavy atoms in the active ligand vs. its position in the ranking, backbone rmsd for BRD4, Galectin-3, and HIV protease during the MD simulations, comparison of AUC and logAUC between AutoDock4 and Vina, Kendall tau correlation coefficient for rankings of actives, grid position and size defining the docking region, AutoDock4 and Vina performances, and supporting methods (PDF)

> SMILES for actives and decoys of BRD4 and Galectin-3 (ZIP)

> PDBQTs for target proteins (ZIP)

> The scripts for obtaining solvent sites are available upon request, and AutoDock Bias[21] implementation for biased docking is freely available and ready to use for any VS campaign as a part of AutoDockTools (http://ccsb.scripps.edu/mgltools/ or http://autodockbias.wordpress.com). A complete user guide including a tutorial that thoroughly explains how to apply biases coming from previous knowledge such as ligand-derived pharmacophore or from MD-derived solvent sites is delivered with the program.

AUTHOR INFORMATION

Corresponding Authors
*E-mail: marti.marcelo@gmail.com (M.A.M.).
*E-mail: adrian@qb.fcen.uba.ar (A.G.T.).

ORCID Ⓘ
Juan Pablo Arcon: 0000-0003-3350-1576
Lucas A. Defelipe: 0000-0001-7859-7300
Elias D. Lopez: 0000-0002-9956-0010
Carlos P. Modenutti: 0000-0001-5352-7234
Marcelo A. Marti: 0000-0002-7911-9340

Author Contributions
⊥J.P.A. and L.A.D. contributed equally to this work.

Notes
The authors declare no competing financial interest.

ABBREVIATIONS

AmpC, AmpC β-lactamase; AUC, area under the ROC curve; BEDROC, Boltzmann-enhanced discrimination of ROC; BRD4, bromodomain-containing protein 4; CDK2, cyclin-dependent kinase 2; DHFR, dihydrofolate reductase; DUD-E, directory of useful decoys, enhanced; EF1%, enrichment factor at 1%; EF10%, enrichment factor at 10%; FXa, Factor Xa; FGFr1, fibroblast growth factor receptor 1; Gal-3, galectin-3; HIVpr, human immunodeficiency virus type 1 protease; log AUC, adjusted area under the semilogarithmic ROC curve; MCS, maximum common substructure; MD, molecular dynamics; PDE5A, phosphodiesterase 5A; rmsd, root mean square deviation; ROC, receiver operating characteristic; VS, virtual screening

REFERENCES

(1) Morgan, S.; Grootendorst, P.; Lexchin, J.; Cunningham, C.; Greyson, D. The Cost of Drug Development: A Systematic Review. *Health Policy* **2011**, *100*, 4−17.

(2) Infectious Diseases Society of America. The 10 X '20 Initiative: Pursuing a Global Commitment to Develop 10 New Antibacterial Drugs by 2020. *Clin. Infect. Dis.* **2010**, *50*, 1081−1083.

(3) Manson, A. L.; Cohen, K. A.; Abeel, T.; Desjardins, C. A.; Armstrong, D. T., 3rd; Barry, C. E.; Brand, J.; Chapman, S. B.; Cho, S.-N.; Gabrielian, A.; Gomez, J.; Jodals, A. M.; Joloba, M.; Jureen, P.; Lee, J. S.; Malinga, L.; Maiga, M.; Nordenberg, D.; Noroc, E.; Romancenco, E.; Salazar, A.; Ssengooba, W.; Velayati, A. A.; Winglee, K.; Zalutskaya, A.; Via, L. E.; Cassell, G. H.; Dorman, S. E.; Ellner, J.; Farnia, P.; Galagan, J. E.; Rosenthal, A.; Crudu, V.; Homorodean, D.; Hsueh, P.-R.; Narayanan, S.; Pym, A. S.; Skrahina, A.; Swaminathan, S.; Van der Walt, M.; Alland, D.; Bishai, W. R.; Cohen, T.; Hoffner, S.; Birren, B. W.; Earl, A. M. Genomic Analysis of Globally Diverse Mycobacterium Tuberculosis Strains Provides Insights into the

Emergence and Spread of Multidrug Resistance. *Nat. Genet.* **2017**, *49*, 395−402.

(4) Holmes, A. H.; Moore, L. S. P.; Sundsfjord, A.; Steinbakk, M.; Regmi, S.; Karkey, A.; Guerin, P. J.; Piddock, L. J. V. Understanding the Mechanisms and Drivers of Antimicrobial Resistance. *Lancet* **2016**, *387*, 176−187.

(5) Hotez, P. J.; Pecoul, B.; Rijal, S.; Boehme, C.; Aksoy, S.; Malecela, M.; Tapia-Conyer, R.; Reeder, J. C. Eliminating the Neglected Tropical Diseases: Translational Science and New Technologies. *PLoS Neglected Trop. Dis.* **2016**, *10*, No. e0003895.

(6) Andricopulo, A.; Guido, R.; Oliva, G. Virtual Screening and Its Integration with Modern Drug Design Technologies. *Curr. Med. Chem.* **2008**, *15*, 37−46.

(7) Scior, T.; Bender, A.; Tresadern, G.; Medina-Franco, J. L.; Martínez-Mayorga, K.; Langer, T.; Cuanalo-Contreras, K.; Agrafiotis, D. K. Recognizing Pitfalls in Virtual Screening: A Critical Review. *J. Chem. Inf. Model.* **2012**, *52*, 867−881.

(8) Jansen, J. M.; Martin, E. J. Target-Biased Scoring Approaches and Expert Systems in Structure-Based Virtual Screening. *Curr. Opin. Chem. Biol.* **2004**, *8*, 359−364.

(9) 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. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739−1749.

(10) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking. *J. Mol. Biol.* **1997**, *267*, 727−748.

(11) Ruiz-Carmona, S.; Alvarez-Garcia, D.; Foloppe, N.; Garmendia-Doval, A. B.; Juhos, S.; Schmidtke, P.; Barril, X.; Hubbard, R. E.; Morley, S. D. rDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLoS Comput. Biol.* **2014**, *10*, No. e1003571.

(12) Corbeil, C. R.; Williams, C. I.; Labute, P. Variability in Docking Success Rates due to Dataset Preparation. *J. Comput.—Aided Mol. Des.* **2012**, *26*, 775−786.

(13) Allen, W. J.; Balius, T. E.; Mukherjee, S.; Brozell, S. R.; Moustakas, D. T.; Lang, P. T.; Case, D. A.; Kuntz, I. D.; Rizzo, R. C. DOCK 6: Impact of New Features and Current Docking Performance. *J. Comput. Chem.* **2015**, *36*, 1132−1156.

(14) Horvath, D. Pharmacophore-based virtual screening. *Methods in Molecular Biology*; Springer, 2011; Vol. 672, pp 261−98.

(15) Alvarez-Garcia, D.; Barril, X. Molecular Simulations with Solvent Competition Quantify Water Displaceability and Provide Accurate Interaction Maps of Protein Binding Sites. *J. Med. Chem.* **2014**, *57*, 8530−8539.

(16) Raman, E. P.; Yu, W.; Lakkaraju, S. K.; MacKerell, A. D., Jr. Inclusion of Multiple Fragment Types in the Site Identification by Ligand Competitive Saturation (SILCS) Approach. *J. Chem. Inf. Model.* **2013**, *53*, 3384−3398.

(17) Ghanakota, P.; Carlson, H. A. Driving Structure-Based Drug Discovery through Cosolvent Molecular Dynamics. *J. Med. Chem.* **2016**, *59*, 10383−10399.

(18) Arcon, J. P.; Defelipe, L. A.; Modenutti, C. P.; López, E. D.; Alvarez-Garcia, D.; Barril, X.; Turjanski, A. G.; Martí, M. A. Molecular Dynamics in Mixed Solvents Reveals Protein-Ligand Interactions, Improves Docking, and Allows Accurate Binding Free Energy Predictions. *J. Chem. Inf. Model.* **2017**, *57*, 846−863.

(19) Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Blanco, J. I.; Di Lella, S.; Martí, M. A. Solvent Structure Improves Docking Prediction in Lectin−carbohydrate Complexes. *Glycobiology* **2013**, *23*, 241−258.

(20) López, E. D.; Arcon, J. P.; Gauto, D. F.; Petruk, A. A.; Modenutti, C. P.; Dumas, V. G.; Marti, M. A.; Turjanski, A. G. WATCLUST: A Tool for Improving the Design of Drugs Based on Protein-Water Interactions. *Bioinformatics* **2015**, *31*, 3697−3699.

(21) Arcon, J. P.; Modenutti, C. P.; Avendano, D.; Lopez, E. D.; Defelipe, L. A.; Alessandra Ambrosio, F.; Turjanski, A. G.; Forli, S.; Marti, M. A. AutoDock Bias: Improving Binding Mode Prediction and

(22) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785−2791.

(23) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455−461.

(24) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J. Med. Chem.* **2012**, *55*, 6582−6594.

(25) Gaillard, T. Evaluation of AutoDock and AutoDock Vina on the CASF-2013 Benchmark. *J. Chem. Inf. Model.* **2018**, *58*, 1697−1706.

(26) Ravindranathan, K. P.; Mandiyan, V.; Ekkati, A. R.; Bae, J. H.; Schlessinger, J.; Jorgensen, W. L. Discovery of Novel Fibroblast Growth Factor Receptor 1 Kinase Inhibitors by Structure-Based Virtual Screening. *J. Med. Chem.* **2010**, *53*, 1662−1672.

(27) Klein, T.; Vajpai, N.; Phillips, J. J.; Davies, G.; Holdgate, G. A.; Phillips, C.; Tucker, J. A.; Norman, R. A.; Scott, A. D.; Higazi, D. R.; Lowe, D.; Thompson, G. S.; Breeze, A. L. Structural and Dynamic Insights into the Energetics of Activation Loop Rearrangement in FGFR1 Kinase. *Nat. Commun.* **2015**, *6*, 7877.

(28) Yu, W.; Lakkaraju, S. K.; Raman, E. P.; MacKerell, A. D. Site-Identification by Ligand Competitive Saturation (SILCS) Assisted Pharmacophore Modeling. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 491−507.

(29) Yu, W.; Lakkaraju, S. K.; Raman, E. P.; Fang, L.; MacKerell, A. D., Jr. Pharmacophore Modeling Using Site-Identification by Ligand Competitive Saturation (SILCS) with Multiple Probe Molecules. *J. Chem. Inf. Model.* **2015**, *55*, 407−420.

(30) Uehara, S.; Tanaka, S. Cosolvent-Based Molecular Dynamics for Ensemble Docking: Practical Method for Generating Druggable Protein Conformations. *J. Chem. Inf. Model.* **2017**, *57*, 742−756.

(31) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49*, 6789−6801.

(32) Lagarde, N.; Zagury, J.-F.; Montes, M. Benchmarking Data Sets for the Evaluation of Virtual Ligand Screening Methods: Review and Perspectives. *J. Chem. Inf. Model.* **2015**, *55*, 1297−1307.

(33) Gilson, M. K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; Chong, J. BindingDB in 2015: A Public Database for Medicinal Chemistry, Computational Chemistry and Systems Pharmacology. *Nucleic Acids Res.* **2016**, *44*, D1045−D1053.

(34) Ahmed, A.; Smith, R. D.; Clark, J. J.; Dunbar, J. B., Jr; Carlson, H. A. Recent Improvements to Binding MOAD: A Resource for Protein-Ligand Binding Affinities and Structures. *Nucleic Acids Res.* **2015**, *43*, D465−D469.

(35) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2012**, *40*, D1100−D1107.

(36) Sander, T.; Freyss, J.; von Korff, M.; Rufener, C. DataWarrior: An Open-Source Program for Chemistry Aware Data Visualization and Analysis. *J. Chem. Inf. Model.* **2015**, *55*, 460−473.

(37) von Korff, M.; Freyss, J.; Sander, T. Flexophore, a New Versatile 3D Pharmacophore Descriptor That Considers Molecular Flexibility. *J. Chem. Inf. Model.* **2008**, *48*, 797−810.

(38) Irwin, J. J.; Shoichet, B. K. ZINC − A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45*, 177−182.

(39) Truchon, J.-F.; Bayly, C. I. Evaluating Virtual Screening Methods: Good and Bad Metrics for the "Early Recognition" Problem. *J. Chem. Inf. Model.* **2007**, *47*, 488−508.

(40) R Core Team. *R: A Language and Environment for Statistical Computing*, R version 3.4.1; R Foundation for Statistical Computing: Vienna, Austria, 2017; https://www.R-project.org/ (2017-06-30).

Virtual Screening Using Known Protein-Ligand Interactions. *Bioinformatics* **2019**, DOI: 10.1093/bioinformatics/btz152.

(41) Sing, T.; Sander, O.; Beerenwinkel, N.; Lengauer, T. ROCR: Visualizing Classifier Performance in R. *Bioinformatics* **2005**, *21*, 3940−3941.

(42) Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M. pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinf.* **2011**, *12*, 77.

(43) Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(44) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, S.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman, P. A. *AMBER 2016*; University of California: San Francisco, 2016.

(45) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.

(46) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53*, 5061−5084.

(47) *Schrödinger Release 2017-1*; Maestro, Schrödinger, LLC: New York, NY.

(48) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminf.* **2011**, *3*, 33.

(49) Schrödinger, L. *LigPrep*, Version 2.5, New York, NY, 2011.

(50) Boobbyer, D. N. A.; Goodford, P. J.; McWhinnie, P. M.; Wade, R. C. New Hydrogen-Bond Potentials for Use in Determining Energetically Favorable Binding Sites on Molecules of Known Structure. *J. Med. Chem.* **1989**, *32*, 1083−1094.

(51) Landrum, G. *RDKit*; Open-Source Cheminformatics Software, 2016..