J Comput Aided Mol Des (2008) 22:161–168
DOI 10.1007/s10822-007-9165-4

Can we use docking and scoring for hit-to-lead optimization?

Istvan J. Enyedy · William J. Egan

Received: 11 November 2007 / Accepted: 18 December 2007 / Published online: 9 January 2008
© Springer Science+Business Media B.V. 2008

Abstract Docking and scoring is currently one of the tools used for hit finding and hit-to-lead optimization when structural information about the target is known. Docking scores have been found useful for optimizing ligand binding to reproduce experimentally observed binding modes. The question is, can docking and scoring be used reliably for hit-to-lead optimization? To illustrate the challenges of scoring for hit-to-lead optimization, the relationship of docking scores with experimentally determined IC_{50} values measured in-house were tested. The influences of the particular target, crystal structure, and the precision of the scoring function on the ability to differentiate between actives and inactives were analyzed by calculating the area under the curve of receiver operator characteristic curves for docking scores. It was found that for the test sets considered, MW and sometimes ClogP were as useful as GlideScores and no significant difference was observed between SP and XP scores for differentiating between actives and inactives. Interpretation by an expert is still required to successfully utilize docking and scoring in hit-to-lead optimization.

Keywords Docking and scoring · Glide · Hit optimization · ROC analysis · Kinase · KDR · CDK2 · ABL

I. J. Enyedy (✉) · W. J. Egan
Global Discovery Chemistry, Computer-Aided Drug Discovery, Novartis Institutes for BioMedical Research, 250 Massachusetts Avenue, Cambridge, MA 02139, USA
e-mail: istvan.enyedy@novartis.com

Introduction

Hit finding and hit optimization are the two primary applications of molecular modeling in drug discovery. In silico hit finding (virtual screening) can be undertaken by searching corporate compound collections or databases of commercially available or virtual compounds. Possible search queries include 1D fingerprints (e.g., Tanimoto similarity), 2D substructure, topology, or QSAR models, and 3D models (e.g., pharmacophore or structure-based docking). The aim of in silico hit finding is to generate a subset of the original compound collection enriched in actives (the hit list). The hit rate represents the percentage of true positives in the hit list. The idea is to tune the search criteria to minimize both the size of the hit list (percentage of false positives) and the percentage of false negatives.

Docking software and scoring functions are usually optimized to reproduce the observed binding geometries of ligands in crystal structures found in commercial or public repositories. Structure-based docking methods automatically sample ligand conformations and protein–ligand interactions with a specified region of the protein surface. Docking results are reported as the lowest energy/highest scoring pose(s) for each ligand, typically one pose for a large compound collection. Docking and scoring methods have been evaluated in a number of studies [1–20]. Docking scores have been successfully used for identifying active compounds through virtual screening [21–23] by filtering out those that do not fit into the binding site. Differentiating activity among a set of actual hits requires more accurate estimation of protein–ligand interactions than what is used for virtual screening.

In silico hit optimization is the process by which the activity of validated hits are improved to a level required for human testing. In silico hit optimization is aimed at

differentiation of target activity among a set of known active compounds in one or more specific series of interest. At this point in a drug discovery project, the dataset contains a small number of compounds drawn from a few, or even just one, chemical series. Hit optimization involves expert analysis of ligand docking poses aimed at qualitative rank ordering of compounds proposed for synthesis. Individual docking score terms may provide insights about key interactions needed for activity. Induced fit of the protein may be used, along with customized modifications to the docking and scoring parameters. The aim is to select/prioritize the best possible molecules for synthesis. High false positive and false negative rates are not acceptable in this case. There is a widely, although not universally, held belief among modelers that docking and scoring software is not good enough to rank order compounds vs. activity [18, 24, 25]. This is understandable since a difference of 1.36 kcal/mol in binding free energy leads to a 10-fold difference in Ki. However, there are reports about good correlation between docking score and experimental binding energies [26].

In this paper, we examine the relationship of docking scores with experimental IC_{50} data. Previously published studies [26, 27] have used a very small number of compounds. Small datasets do not contain the diversity and information necessary to characterize computational methods. Thus, it is necessary to test how docking software performs by using larger datasets of compounds having experimental data obtained in the same laboratory using the same method. Three kinases were chosen as targets: kinase insert domain protein receptor (KDR), cyclin dependent kinase 2 (CDK2), and Abelson tyrosine kinase (C-ABL). We considered more than 4,300 in-house compounds with measured IC_{50}s for each kinase. The influences of the particular target, differences due to multiple crystal structures, and the precision of the scoring function were tested. This situation reproduces the challenges a modeler would face in hit optimization phase of a kinase drug discovery project.

Methods

Structures of kinases were obtained from the Protein Databank [28]. The structures used were 1YWN [29], 2P2H [30], 2P2I [30] for KDR, 1E9H [31] and 1QMZ [32] for CDK2, and 1IEP[33] for ABL.

Glide 4.5 [27, 34, 35] was used for docking and scoring. All structures were prepared for docking using the "Protein Preparation Wizard" in Maestro 8.0.11 [36]. Water molecules in the crystal structures were deleted and termini were capped by adding ACE and NMA residues. Grids were defined by centering them around the ligand in the crystal structure using the default box size. One grid was built using the default setting in Glide: scaling of Van der Waals radii of protein atoms partial atomic charge of less then 0.25 by 1.0. Two additional grids were built with Van der Waals scaling of 0.8 and 0.5 to investigate the influence of softer protein surface on docking and scoring. Grids for KDR structures included constraint to the backbone NH of Cys919 in the hinge region.

Docking was done in standard precision (SP) and extra precision (XP) modes [27]. Flips of 5-and 6-member rings were allowed and non-planar conformation of amide bonds was penalized. Van der Waals radii of ligand atoms with partial atomic charge less than 0.15 was scaled by 0.8 (default) or 1.0. After docking in SP mode postdocking minimization was done on five ligand poses using 100 iterations and a dielectric constant of 2.0.

3D coordinates of all compounds used for docking were generated using Ligprep [37] from Schrodinger. The data set included 4,904 compounds tested in the KDR assay, 5,149 compounds tested in the CDK2 assay, and 4,346 compounds tested in the C-ABL assay.

Docking poses were further minimized using eMBrAcE module from Macromodel 9.5.213[38] in energy difference mode for estimating the effect of solvation and protein flexibility on ranking. Protein atoms that were 8 Å around the ligand were considered flexible. Minimization was done using GB/SA water solvation model [39], 5,000 PRCG steps and OPLS2005 force field. The energy difference was calculated using the following equation:

$$\Delta E = E_{complex} - E_{ligand} - E_{protein}$$

Triballeau et al. [40] used receiver operator characteristic (ROC) curves for comparing docking protocols. We implemented ROC analysis using Matlab 7.4 R2007a (MathWorks, Natick, MA, USA) and computed the area under the curve (AUC) using a trapezoidal integration.

Results and discussion

There are several filters implemented in Glide that discard ligand conformations or ligands that do not fit into the defined binding site [34]. The role of these filters is to reduce the time required for docking. In our tests the usefulness of Glide filters varied with the crystal structure used for docking (Table 1). When docking into KDR structures, more compounds were discarded when using the 2P2I structure than when using 2P2H or 1YWN structure. The number of compounds discarded when docking into the KDR structures was much higher when using a constraint to Cys919 backbone NH in the hinge region of KDR than when docking without a constraint. Constraints to backbone atoms of residues from the hinge are often

Table 1 Compounds discarded during docking by the Glide filters

Target	Struc	SP0.8 (%)	XP0.8 (%)	SP0.8C (%)	XP0.8C (%)	SP1.0 (%)	XP1.0 (%)	SP1.0C (%)	XP1.0C (%)
KDR 4904	1YWN	0.4	0.4	4.0	31.8	0.4	0.4	5.8	41.0
	2P2H	0.4	0.4	5.5	44.8	0.4	0.4	7.8	53.9
	2P2I	11.0	3.3	56.4	77.8	20.1	5.5	72.3	94.8

The numbers after the docking precision represent scaling factors used for ligand atoms with partial charge less than 0.15. A "C" after the number means that constraint to the "hinge" was used

considered as a way of better separating actives from inactives. When using the structure 2P2I with constraint to the hinge and with no scaling of Van der Waals radii of non-polar protein and ligand atoms, XP1.0C, 94.8% of the compounds were discarded. This resulted in a large number of actives also being discarded suggesting that docking with constraints must be done with caution. Reducing the scaling factor applied to non-polar protein atoms from the default 1.0 to 0.8 or 0.5 eliminated the above problem. The KDR structures 1YWN and 2P2I have an "out" conformation of the DFG motif while 2P2H has a "DFG in" conformation [41]. This shows the importance of considering multiple structures for docking especially for kinases that are very flexible and can have multiple binding site conformations depending on the compound bound.

The Glide XP scoring function was developed for semiquantitatively ranking compound binding to a target protein [27]. Thus we were interested in the efficiency of GlideScores in rank ordering compounds that were tested in the same assay. This should eliminate errors that come from various setups that different groups use. Glide SP score is also of interest since docking in SP mode is much faster than in XP mode and we were interested if the extra time spent on docking compounds leads to much better ranking. There should be some correlation between the docking score and the experimental IC_{50} for rank ordering compounds. Glide SP scores calculated from docking into the KDR structure 1YWN showed no correlation with IC_{50} ($r^2 = 0.04$) data for KDR (Fig. 1) and there was no absolute score that could be used for differentiating between actives and inactives. This was not surprising since this score was not intended for rank ordering compounds. However, the plot suggests that there is an enrichment of actives as the score decreases, which agrees with the common belief about the usefulness of the Glide SP score for virtual screening. The same results were observed for all the other docking runs.

We next tested the performance of Glide XP score that was developed for rank ordering active compounds [27]. Docking in extra precision mode takes about 10 times longer than in standard precision mode [19]. All our tests showed no correlation between the XP score and the experimental IC_{50}. Figure 2 plots XP score vs. IC_{50} for the

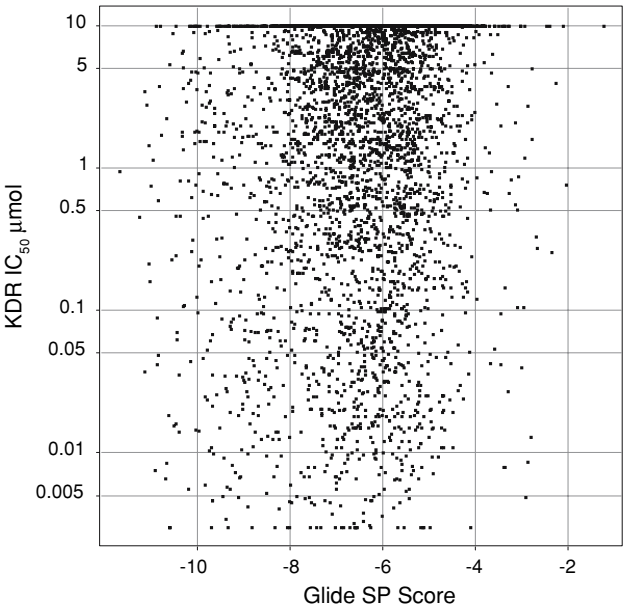

Fig. 1 Correlation of Glide SP score with IC_{50} when using default setup for protein grid and compound docking. The structure used was 1YWN for KDR. Compounds with reported $IC_{50} > 10$ μM were plotted as 10 μM

KDR structure 1YWN. This correlation is also quite small ($r^2 = 0.05$) and no absolute XP score could be used either to separate actives from inactives or to rank order compounds. This showed that the precision of calculating docking scores was not enough for hit optimization when small improvements in potency are desired. This was in good agreement with other studies [18, 25, 42, 43] suggesting that more developments in scoring are needed.

Bytheway and Cochran [26] noted that Emodel correlated better with binding affinity than GlideScore. Thus we tested the efficiency of Emodel on all our docking runs. The experimental IC_{50} values and corresponding Emodel scores calculated from Glide XP scores in docking into the KDR structure 1YWN are shown in Fig. 3. The correlation between Emodel and IC_{50} values was slightly better ($r^2 = 0.12$). For a large set of compounds, the Emodel score also performed poorly for rank ordering compounds and could not be used for differentiating between actives and inactives. The same was noticed for all the other structures.

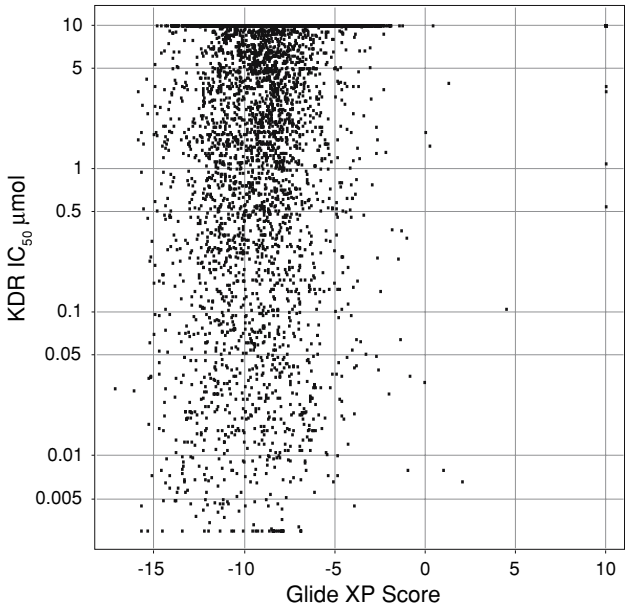

Fig. 2 Correlation of Glide XP score with IC_{50} when using default setup for protein grid and compound docking. The structure used was 1YWN for KDR. Compounds with reported $IC_{50} > 10$ µM were plotted as 10 µM

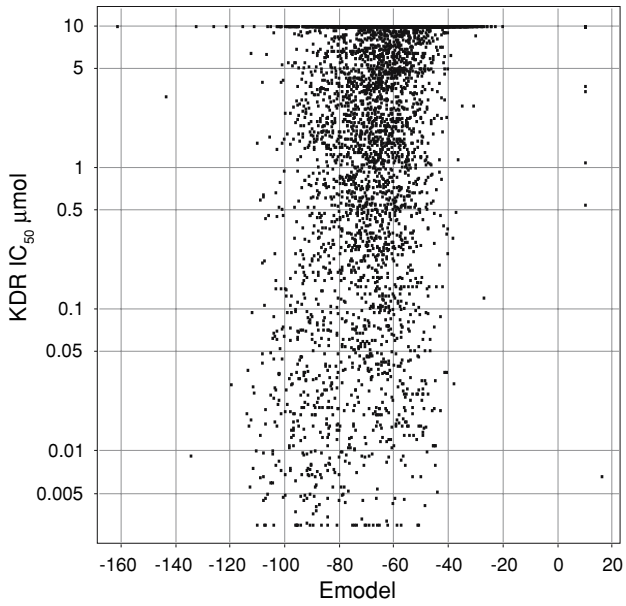

Fig. 3 Correlation of Emodel score with IC_{50} when using default setup for protein grid and compound docking. The structure used was 1YWN for KDR. Compounds with reported $IC_{50} > 10$ µM were plotted as 10 µM

We also tried rescoring docking poses using eMBrAcE in energy difference mode. It was possible that factoring in protein flexibility and desolvation could lead to better correlation with IC_{50} values. Unfortunately, Fig. 4 shows that the binding energy calculated with eMBrAcE also

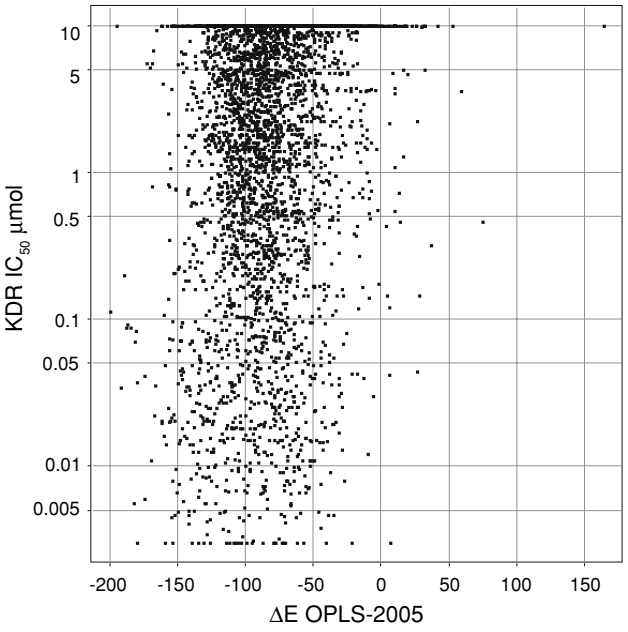

Fig. 4 Correlation of Embrace energy difference score with IC_{50} when using default setup for protein grid and compound docking. The structure used was 1YWN for KDR. Compounds with reported $IC_{50} > 10$ µM were plotted as 10 µM

performs poorly at rank ordering compounds and does not correlate with IC_{50} ($r^2 = 0.03$).

Figures 1–4 showed that docking scores are useful for enrichment of actives in the top scoring compounds. We considered ROC analysis for comparing the performance of scoring functions tested since it considers both the sensitivity and specificity of the score considered [44]. ROC AUC values range from 0.5 (random) to 1.0 (perfect separation of actives versus inactives). ROC analyses were performed on the docking results to assess the extent to which GlideScores can differentiate between actives and inactives. Standard errors for the ROC AUCs were computed using the method of Hanley and McNeil [45]. Standard errors were generally small, in the range 0.007–0.01. ROC analysis requires setting criteria for considering compounds inactive and active. Inactive compounds in our test set were considered those that had reported IC_{50} values >10 µM. Two cutoffs were considered for defining actives, the first set included those with IC_{50} values <10 µM and the second set included those with IC_{50} values <10 nM resulting in more than 1,000-fold difference between actives and inactives. Ligprep generates all enantiomers for compounds with chiral centers and no specified chiralities thus each compound can have multiple scores if more than one stereomer was built. Only the top-scoring configuration of each compound was considered in the analysis below.

Table 2 shows a summary of ROC AUC analysis of the docking results for the chosen kinase structures. When using a 10 μM cutoff, GlideScores show little improvement over random selection for all structures tested and XP scores were not found to be better than SP scores in separating actives from inactives. This was surprising since the XP methodology was optimized to produce large improvements in database enrichment when compared to other scoring functions within Glide [27]. These results did not justify the extra computer time necessary to optimize compounds in the binding site in extra precision mode.

Emodel was found to better differentiate actives from inactives than Glide SP or XP scores. ROC AUCs further suggested that molecular weight, MW, performed similar to Emodel in differentiating actives from inactives, and ClogP performed similar to Glide SP and XP in this regard. This might be due to the fact that larger compounds tend to be more active because these can form more interactions with the target protein, and kinase inhibitor activity tends to increase with logP.

Performance of the docking methods was also tested using a cutoff of 10 nM to classify compounds as active and >10 μM for inactives, Fig. 5. This did not lead to significant changes in ROC AUCs when using Glide SP or XP scoring, Table 2. However, this did lead to better prediction of activity when using MW and ClogP or when using Emodel for the KDR 2P2I crystal structure with XP scoring. Using Glide XP scores still did not improve docking results versus Glide SP scores.

Docking and scoring was tested for three KDR and two CDK2 crystal structures. As we noted, the standard error of the ROC AUCs was approximately 0.01 which means that differences of 0.04 are significant. As shown in Table 2, there was little variation in GlideScore and Emodel ROC AUCs using the 10 μM cutoff for actives in KDR docking.

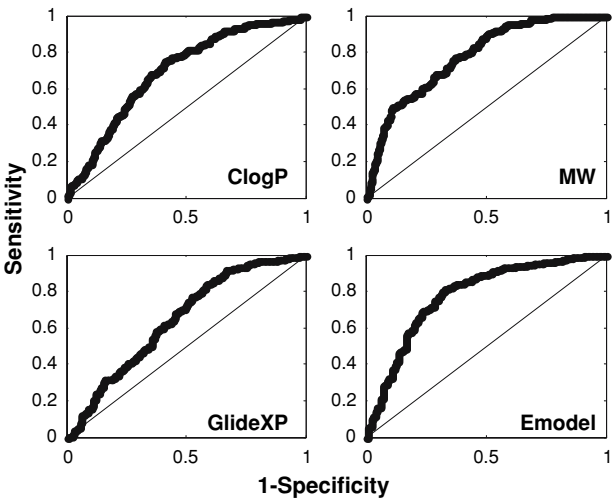

Fig. 5 ROC curves obtained for docking compounds tested in the KDR assay in the 1YWN structure

However, considering a 10 nM cutoff for actives resulted in higher Glide SP score ROC AUCs for KDR structures 1YWN and 2P2I than 2P2H (0.62 and 0.66 vs. 0.59 and 0.54). Docking in 2P2H had a ROC AUC of 0.53 for XP score and 0.81 for the corresponding Emodel. Using Emodel led to higher ROC AUCs for XP scores than for SP scores. Structural changes did not affect ROC AUCs for Emodel for KDR (0.78 vs. 0.81 vs. 0.82, see Table 2). In contrast, there were large differences in ROC AUCs when docking into CDK2 1QMZ versus 1E9H structures (both have "DFG in" conformation) using GlideScore and Emodel. In this case, Emodel did not have higher ROC AUCs than GlideScore. Thus, the structure of the target could influence docking performance. This should depend on the binding site size, geometry, and flexibility and the structure of the compounds docked.

Table 2 ROC AUC values for docking compounds into the corresponding kinases using 10 μM and 10 nM cutoff respectively

ROC AUC	GlideScore		Emodel		MW		ClogP	
	10 μM	10 nM	10 μM	10 nM	10 μM	10 nM	10 μM	10 nM
KDR_1YWN(SP0.8)	0.59	0.62	0.67	0.71	0.68	0.78	0.54	0.70
KDR_1YWN(XP0.8)	0.63	0.65	0.70	0.78	0.68	0.78	0.54	0.70
KDR_2P2H(SP0.8)	0.58	0.54	0.69	0.70	0.68	0.78	0.54	0.70
KDR_2P2H(XP0.8)	0.56	0.53	0.72	0.81	0.68	0.78	0.54	0.70
KDR_2P2I(SP0.8)	0.61	0.66	0.65	0.74	0.70	0.80	0.57	0.71
KDR_2P2I(XP0.8)	0.54	0.66	0.69	0.82	0.69	0.79	0.55	0.71
CDK2_1QMZ(SP0.8)	0.58	0.53	0.53	0.50	0.57	0.63	0.63	0.51
CDK2_1QMZ(XP0.8)	0.56	0.59	0.51	0.55	0.57	0.63	0.63	0.51
CDK2_1E9H(SP0.8)	0.69	0.61	0.63	0.65	0.57	0.63	0.62	0.51
CDK2_1E9H(XP0.8)	0.66	0.58	0.54	0.62	0.57	0.63	0.63	0.51
CABL_1IEP(SP0.8)	0.59	0.60	0.66	0.71	0.68	0.78	0.56	0.74
CABL_1IEP(XP0.8)	0.56	0.62	0.71	0.79	0.68	0.78	0.56	0.75

Most of the kinase inhibitors form hydrogen bonds to the "hinge" main-chain NH group. These hydrogen bonds were shown to be crucial to the activity of the inhibitors [6, 46]. Thus we tried using constraints to the "hinge" NH for improving the docking outcome in the case of KDR (Table 3). No improvement in ROC AUCs was observed for the three KDR structures used. This might suggest that the scoring functions are already optimized enough that constraints are not necessary for distinguishing actives from inactives.

All the above docking runs were done using the default 1.0 scaling factor for Van der Waals radii of protein atoms with partial atomic charge of less than 0.25 and a scaling factor of 0.8 for Van der Waals radii of ligand atoms with partial atomic charge of less than 0.15. Proteins are flexible and making their surface softer can slightly compensate for this flexibility. This was tested using the 2P2I structure of KDR by reducing the scaling factors for non-polar protein atoms to 0.8 and 0.5, and increasing the scaling factor for non-polar ligand atoms to 1.0 (Table 4). The latter change should lower the conformational energy of bound ligands. Using a scaling factor of 1.0 for both protein and ligand was worse than using a scaling factor of 1.0 for the protein and 0.8 for ligands when testing Glide and Emodel scores. A scaling factor of 0.8 for protein atoms and 1.0 for ligands led to slightly worse results when using a 10 μM cutoff but better results when using a 10 nM cutoff in IC_{50}. A scaling factor of 0.8 for protein atoms also reduced the number of compounds discarded by Glide filters and increased chances of retrieving more actives. Further reducing the scaling factor for protein atoms to 0.5 had almost no effect on Glide SP score but had a large effect on Glide XP score and Emodel.

Conclusions

Docking scores are useful for optimizing ligand geometry in the binding site to reproduce experimentally observed binding modes. However, our tests showed that docking scores could not rank order compounds and even could not always differentiate individual active molecules from inactive ones. Docking scores were found useful for enrichment of actives that was in good agreement with their use for screening large compound collections. For the test sets considered, MW consistently outperformed GlideScore with ROC AUCs of 0.78 for KDR and C-ABL, 0.63 for CDK2. ClogP outperformed GlideScores for KDR (ROC AUC 0.70) and C-ABL (ROC AUC 0.74) while it was comparable or worse than SP or XP score for CDK2 (ROC AUC 0.51). No significant difference was observed between SP and XP scores for differentiating between actives and inactives; XP scoring sometimes had lower ROC AUC than the corresponding SP score. It is recommended to use SP scores for virtual screening since the extra time required for optimization of compounds using the XP method does not lead to better ROC AUCs. The docking setup and the structure of the target also influenced the ROC AUCs. It is also recommended to scale the Van der Waals radii of non-polar protein atoms when constraints are used in docking because without scaling, actives might be discarded by Glide filters.

These results suggest that (1) modeling expertise is required to translate docking poses into a good understanding of what drives affinity and how a molecule should be optimized, (2) more improvements are needed in scoring algorithms by eliminating approximations in calculating components of the score, (3) a scoring function should

Table 3 Results obtained for docking compounds with constraint to the hinge into the KDR structures

ROC AUC	GlideScore		Emodel		MW		ClogP	
	10 μM	10 nM	10 μM	10 nM	10 μM	10 nM	10 μM	10 nM
1YWN(SP0.8)	0.64	0.62	0.75	0.71	0.79	0.78	0.63	0.70
1YWN(SP0.8C)	0.62	0.56	0.70	0.69	0.69	0.79	0.55	0.71
1YWN(XP0.8)	0.66	0.65	0.79	0.78	0.78	0.78	0.63	0.70
1YWN(XP0.8C)	0.64	0.63	0.72	0.80	0.71	0.79	0.58	0.70
2P2H(SP0.8)	0.60	0.54	0.75	0.70	0.76	0.78	0.62	0.70
2P2H(SP0.8C)	0.58	0.58	0.68	0.61	0.68	0.79	0.55	0.71
2P2H(XP0.8)	0.56	0.53	0.82	0.81	0.79	0.78	0.63	0.70
2P2H(XP0.8C)	0.54	0.57	0.72	0.76	0.70	0.78	0.56	0.72
2P2I(SP0.8)	0.68	0.66	0.73	0.74	0.80	0.80	0.65	0.71
2P2I(SP0.8C)	0.68	0.76	0.65	0.79	0.70	0.83	0.60	0.74
2P2I(XP0.8)	0.61	0.66	0.77	0.82	0.78	0.79	0.64	0.71
2P2I(XP0.8C)	0.61	0.67	0.69	0.84	0.69	0.82	0.60	0.74

Table 4 Dependence of the docking outcome using scaling factors of 1.0, 0.8, and 0.5 for Van der Waals radii of protein atoms with partial charge of less than 0.25 for KDR 2P2I structure KDR 2P2I

ROC AUC	GlideScore		Emodel		MW		ClogP	
	10 μM	10 nM	10 μM	10 nM	10 μM	10 nM	10 μM	10 nM
G1.0 SP1.0	0.59	0.62	0.60	0.70	0.70	0.80	0.58	0.71
G1.0 XP1.0	0.55	0.61	0.65	0.78	0.69	0.79	0.55	0.71
G1.0 SP0.8	0.68	0.66	0.73	0.74	0.80	0.80	0.65	0.71
G1.0 XP0.8	0.61	0.66	0.77	0.82	0.78	0.79	0.64	0.71
G0.8 SP1.0	0.61	0.71	0.67	0.76	0.70	0.79	0.56	0.71
G0.8 XP1.0	0.60	0.75	0.71	0.84	0.69	0.79	0.55	0.71
G0.5 SP1.0	0.60	0.72	0.59	0.72	0.69	0.79	0.55	0.70
G0.5 XP1.0	0.50	0.52	0.66	0.73	0.68	0.79	0.54	0.70

minimally perform better than MW and logP to be considered a viable method for rank ordering compounds, (4) better models for estimating vibrational and conformational entropy, desolvation, polarization effects are needed to improve the quality of scoring functions [47], and (5) further improvements could be obtained via local models where scoring function parameters are adjusted to make them target and compound set specific [48]. In the literature, it has been suggested that another possibility of improving the performance of scoring functions might be achieved by considering scoring of ensembles of the same compound instead of just one orientation for better rank ordering compounds [49, 50]. Molecular mechanics Poisson-Bolzmann surface area (MM-PBSA) and molecular mechanics generalized Born surface area (MM-GBSA) methods have been used to estimate binding free energy in a reasonable time [51]. There are several publications showing successful rank ordering of compounds using these approaches [52–55]. However, Page and Bates concluded that the assumptions and approximations made for calculating the binding free energy using MM-PBSA are too coarse for reliable predictions of binding energy on a diverse set of compounds and targets [47]. As computer speed continues to increase, free energy perturbation and thermodynamic integration methods should become feasible for routine use in hit optimization and improve the binding affinity predictions [43].

Acknowledgment The authors thank Robert Pearlstein, Richard Lewis, and Hanneke Jansen for their reviews of the manuscript.

References

1. Gschwend DA, Good AC, Kuntz ID (1996) J Mol Recog 9:175
2. Perola E, Walters WP, Charifson PS (2004) Proteins 56:235
3. Muegge I, Rarey M (2001) In: Boyd DB, Lipkowitz KB (eds) Reviews in computational chemistry, vol 17. Wiley-VCH, New York, Ch 1, p 1
4. Muegge I, Enyedy I (2003) In: Abraham D (ed) Burger's medicinal chemistry 5th edn. vol 1. Wiley, New York, p 243
5. Muegge I, Enyedy I (2004) In: Bultinck P, De Winter H, Langenaeker W, Tollenaere J (eds) Computational medicinal chemistry for drug discovery. Marcel Dekker Inc, New York, p 405
6. Muegge I, Enyedy IJ (2004) Curr Med Chem 11:693
7. Bissantz C, Folkers G, Rognan D (2000) J Med Chem 43:4759
8. Kroemer RT (2003) Biochem Soc Trans 31:980
9. Cotesta S, Giordanetto F, Trosset JY, Crivori P, Kroemer RT, Stouten PFW, Vulpetti A (2005) Proteins 60:629
10. Stouten PFW, Kroemer RT (2006) Comprehensive Medicinal Chemistry II 4:255
11. Kroemer RT (2007) Curr Protein Pept Sci 8:312
12. Wang R, Lu Y, Wang S (2003) J Med Chem 46:2287
13. Smith R, Hubbard RE, Gschwend DA, Leach AR, Good AC (2003) J Mol Graphics Model 22:41
14. Ferrara P, Gohlke H, Price DJ, Klebe G, Brooks CL III (2004) J Med Chem 47:3032
15. Kontoyianni M, McClellan LM, Sokol GS (2004) J Med Chem 47:558
16. Kontoyianni M, Sokol GS, McClellan LM (2005) J Comput Chem 26:11
17. Cummings MD, DesJarlais RL, Gibbs AC, Mohan V, Jaeger EP (2005) J Med Chem 48:962
18. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS (2006) J Med Chem 49:5912
19. Zhou Z, Felts AK, Friesner RA, Levy RM (2007) J Chem Inf Model 47:1599
20. McGaughey GB, Sheridan RP, Bayly CI, Culberson JC, Kreatsoulas C, Lindsley S, Maiorov V, Truchon JF, Cornell WD (2007) J Chem Inf Model 47:1504
21. Vangrevelinghe E, Zimmermann K, Schoepfer J, Portmann R, Fabbro D, Furet P (2003) J Med Chem 46:2656
22. Enyedy IJ, Ling Y, Nacro K, Tomita Y, Wu X, Cao Y, Guo R, Li B, Zhu X, Huang Y, Long Y-Q, Roller P, Yang D, Wang S (2001) J Med Chem 44:4313
23. Enyedy IJ, Lee S-L, Kuo AH, Dickson RB, Lin C-Y, Wang S (2001) J Med Chem 44:1349
24. Tirado-Rives J, Jorgensen WL (2006) J Med Chem 49:5880
25. Leach AR, Shoichet BK, Peishoff CE (2006) J Med Chem 49:5851
26. Bytheway I, Cochran S (2004) J Med Chem 47:1683
27. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) J Med Chem 49:6177
28. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindialov IN, Bourne PE (2000) Nucleic Acids Res 28:235

29. Miyazaki Y, Matsunaga S, Tang J, Maeda Y, Nakano M, Philippe RJ, Shibahara M, Liu W, Sato H, Wang L, Nolte RT (2005) Bioorg Med Chem Lett 15:2203
30. Hodous BL, Geuns-Meyer SD, Hughes PE, Albrecht BK, Bellon S, Bready J, Caenepeel S, Cee VJ, Chaffee SC, Coxon A, Emery M, Fretland J, Gallant P, Gu Y, Hoffman D, Johnson RE, Kendall R, Kim JL, Long AM, Morrison M, Olivieri PR, Patel VF, Polverino A, Rose P, Tempest P, Wang L, Whittington DA, Zhao H (2007) J Med Chem 50:611
31. Davies TG, Tunnah P, Meijer L, Marko D, Eisenbrand G, Endicott JA, Noble MEM (2001) Structure 9:389
32. Brown NR, Noble ME, Endicott JA, Johnson LN (1999) Nat Cell Biol 1:438
33. Nagar B, Bornmann WG, Pellicena P, Schindler T, Veach DR, Miller WT, Clarkson B, Kuriyan J (2002) Cancer Res 62:4236
34. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) J Med Chem 47:1739
35. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) J Med Chem 47:1750
36. Maestro, version 8.0.110, (2007) Schrodinger LLC, New York
37. LigPrep, version 1.2, (2007) Schrodinger LLC, New York
38. MacroModel, version 9.5.213, (2007) Schrodinger LLC, New York
39. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) J Am Chem Soc 112:6127
40. Triballeau N, Acher F, Brabet I, Pin JP, Bertrand HO (2005) J Med Chem 48:2534
41. Kornev AP, Haste NM, Taylor SS, Ten Eyck LF (2006) Proc Natl Acad Sci USA 103:17783
42. Tirado-Rives J, Jorgensen WL (2006) J Med Chem 49:5880
43. Pearlman DA, Charifson PS (2001) J Med Chem 44:3417
44. Triballeau N, Acher F, Brabet I, Pin JP, Bertrand HO (2005) J Med Chem 48:2534
45. Hanley JA, McNeil BJ (1982) Radiology 143:29
46. Verma S, Nagarathnam D, Shao J, Zhang L, Zhao J, Wang Y, Li T, Mull E, Enyedy I, Wang C, Zhu Q, Altieri M, Jordan J, Dang TTA, Reddy S (2005) Bioorg Med Chem Lett 15:1973
47. Page CS, Bates PA (2006) J Comput Chem 27:1990
48. Coupez B, Lewis RA (2006) Curr Med Chem 13:2995
49. Sheu SH, Kaya T, Waxman DJ, Vajda S (2005) Biochemistry 44:1193
50. Prasad JC, Goldstone JV, Camacho CJ, Vajda S, Stegeman JJ (2007) Biochemistry 46:2640
51. Wang J, Hou T, Xu X (2006) Curr Comput Aided Drug Des 2:287
52. Xiong Y, Li Y, He H, Zhan CG (2007) Bioorg Med Chem Lett 17:5186
53. Ferrari AM, Degliesposti G, Sgobba M, Rastelli G (2007) Bioorg Med Chem 15:7865
54. Zhou Z, Bates M, Madura JD (2006) Proteins 65:580
55. Lyne PD, Lamb ML, Saeh JC (2006) J Med Chem 49:4805