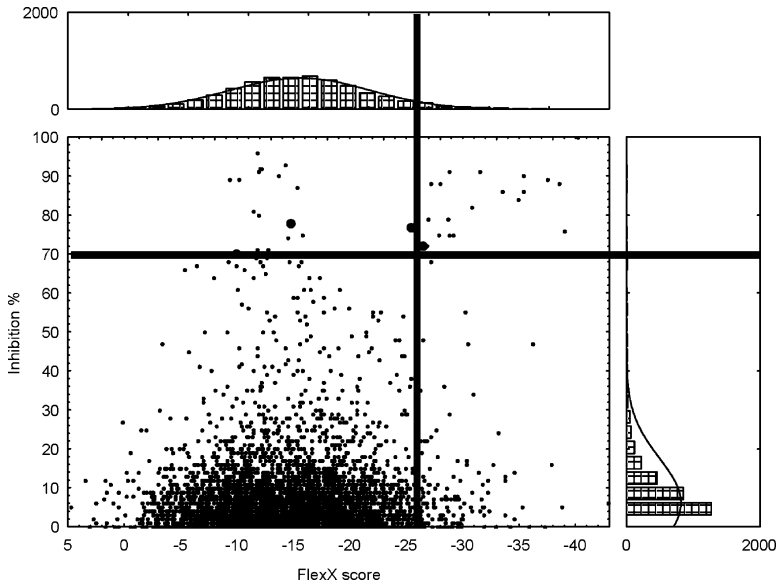
Journal of Medicinal Chemistry

Article

Comparative Virtual and Experimental High-Throughput Screening for Glycogen Synthase Kinase-3β Inhibitors

Tmea Polgr, Andrea Baki, Gyrgyi I. Szendrei, and Gyrgy M. Keseru

J. Med. Chem., **2005**, 48 (25), 7946-7959• DOI: 10.1021/jm050504d • Publication Date (Web): 12 November 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

Journal of Medicinal Chemistry is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Comparative Virtual and Experimental High-Throughput Screening for Glycogen Synthase Kinase-3β Inhibitors

Tímea Polgár,[†] Andrea Baki,[†] Györgyi I. Szendrei,[‡] and György M. Keserű*,[†]

CADD&HTS, and Synthetic Chemistry II, Gedeon Richter Ltd., P.O. Box 27, H-1475 Budapest, Hungary

Received May 30, 2005

Glycogen synthase kinase-3β (GSK-3β) is a serine/threonine kinase that has recently emerged as a key target for neurodegenerative diseases and diabetes. As an initial step of our lead discovery program, we developed a virtual screen to discriminate known GSK-3β inhibitors and inactive compounds using FlexX, FlexX-Pharm, and FlexE. The maximal enrichment factor (EF = 28) suggests that our protocol identifies potential GSK-3β inhibitors effectively from large compound collections. The effectiveness of our screening protocol was further investigated by comparative experimental and virtual high-throughput screens (HTSs) performed for the same subset of our corporate library. Enrichment factors, the significantly higher hit rate of virtual screening (12.9%) than that of the HTS (0.55%), and also the comparison of active clusters suggest that our virtual screening protocol is an effective tool in GSK-3β-based library focusing. Head-to-head comparison of true/false positives and negatives revealed the two approaches to be complementary rather than competitive.

Introduction

Protein kinases phosphorylate several cellular proteins providing control mechanisms for various signaling pathways. Most protein kinases are subjected to positive and negative regulation by phosphorylation of serine, threonine, and tyrosine residues at specific sites. Glycogen synthase kinase-3β (GSK-3β), one of the major isoforms of GSK-3, is a regulatory serine/threonine kinase that phosphorylates glycogen synthase.[1] GSK-3 isoforms are related by a high degree of similarity in the catalytic domain and differ at the N and C termini.[2] To date several crystal structures of GSK-3β have been published.[3−9] The overall structures of the enzyme complexed with inhibitors are similar to that of the apoenzyme, and these X-ray structures provided information, for the first time, about how the enzyme interacts with selective and nonselective inhibitor scaffolds.[10]

The development of selective protein kinase inhibitors is currently of considerable interest as a drug discovery target. A number of drug discovery programs have yielded small-molecule ATP-competitive inhibitors that are presently in various stages of development.[11] Because each member of the kinase family binds the same cofactor, the dogma had been that the ATP-binding cleft of the enzyme would not prove to be a good target for drug discovery. However, within the last several years many selective ATP-competitive protein kinase inhibitors have been identified. Given the association of abnormal GSK-3 activity with various human diseases, it is emerging as a promising therapeutic target,[12] and GSK-3 inhibitors might prove useful in the treatment of conditions associated with elevated enzyme activity such as type 2 diabetes or Alzheimer's disease. Several new GSK-3 inhibitors have recently been identified, all of which are ATP competitive. Since its discovery, GSK-3 has provided insights into a number of biological processes, several of which have proven to be new signaling paradigms. It seems likely that this will continue, and for example, by site-directed mutagenesis studies with selective inhibitors, there is hope that the true cellular roles for this multifaceted kinase will be understood and hence its utility as a therapeutic target will be realized.[13,14]

In this paper, we report the development of an effective virtual screening protocol that was validated in a real screening situation. The enrichment study was first performed to search for GSK-3β inhibitors using the publicly available structural information. The proper preprocessing of the input database and the application of pharmacophore filters[15] improved the enrichment factors given by FlexX.[16] Docking into protein ensembles[17] by FlexE and the restriction of the ligand flexibility, however, did not improve the enrichment factors further. Our results support the observation of Lyne and co-workers,[18] suggesting that for structurally similar active sites of different protein targets very similar scoring combinations can be effective. In fact, we show here that the PMF and FlexX scoring scheme applied by Lyne for CDK-2 outperforms other combinations when screening against GSK-3β. Combining results obtained on CDK-2,[18] CHK-1,[18] GSK-3β and p38,[19,20] we argue that a pharmacophore-driven virtual screening with FlexX and PMF scoring functions might be a generally applicable protocol for serine/threonine kinases that share similar active sites. Second, our screening protocol was experimentally validated for the first time by screening the *same* sublibrary of our corporate collection in silico and in vitro. Similar enrichment factors observed for both of the in silico and the in vitro screen suggest that our virtual screening protocol truly identifies active ligands from large librar-

* Corresponding author. Phone: +36-1-431-4605. Fax: +36-1-432-6002. E-mail: gy.keseru@richter.hu.
[†] CADD&HTS.
[‡] Synthetic Chemistry II.

10.1021/jm050504d CCC: $30.25
Published on Web 11/12/2005

ies in real screening situations. In fact, we found that the hit rate in virtual screening (12.9%) is significantly higher than that obtained by the high-throughput screen (HTS) (0.55%). Direct comparison of complete hit lists and clusters of hits revealed that the overlap between the results of the two approaches is acceptable. The large number of false positives and false negatives in virtual screening, however, suggest this method to be a complementary approach to HTS that could be useful in the design of screening libraries. On the basis of these results, we concluded that our virtual screening protocol might be an effective prescreening filter in the GSK-3β lead discovery but could not replace experimental screening.

Results

Developing a Virtual Screening Protocol. Comparison of X-ray Structures. Comparison of the X-ray structures has revealed that overall the folds of both the apo enzyme and its ligand-bound form are very similar. GSK-3β complexed with AMP-PNP (adenyl imidodiphosphate) provides a good starting point for understanding how ATP-mimetic inhibitors interact with GSK-3β. The main polar interactions are formed between Asp133, Val135, Gln185, and the adenine group of the substrate. Structural comparison of GSK-3β complexes (PDB codes 1Q4L, 1Q3W, 1Q41, 1UV5, 1R0E, 1Q5K, 1Q3D, 1PYX) strengthens the importance of these polar interactions. In each of the complexes, two conserved structural water molecules could also be identified. In 1Q4L, crystallized with ligand I-5, Gln185 adopts a conformation that is significantly different from all the other structures; the carboxylic oxygen of this residue substitutes the O atom of one of the conserved water molecules. Apart from 1Q4L, the conformation of Gln185 divides ligand-bound GSK-3β X-ray structures into two major classes: 1Q3W, 1Q41, 1UV5, and 1R0E and 1Q5K, 1Q3D, and 1PYX (see Figure 1a,b). Since the conformation of the active site has significant impact on the enrichment, we selected 1Q4L and one representative from each class (1UV5 and 1Q3D, respectively) for virtual screening.

Virtual Screening Protocol. Virtual screening on GSK-3β structures complexed with staurosporine (1Q3D), 6-bromoindirubin-3′-oxime (1UV5), and I-5 (1Q4L) was performed by FlexX, FlexX-Pharm, and FlexE (Figures 2−5).

To show the effectiveness of our screening protocol, we compared the enrichment factors calculated in docking runs with various settings. Enrichment factors assess the quality of the rankings:

$$\mathrm{EF}(\%) = \frac{(N_{\mathrm{active}(\%)}/N_{(\%)})}{(N_{\mathrm{active}}/N_{\mathrm{all}})}$$

where EF(%) is given as the percentage of the ranked database, $N_{\mathrm{active}(\%)}$ is the number of active compounds in a selected subset of the ranked database, $N_{(\%)}$ is the number of compounds in the subset, and N_{active} and N_{all} are the number of active molecules and the number of compounds in the screening database. Enrichment factors were not scaled; that is, absolute values depend on the ratio of active and inactive molecules and should be compared to the maximum (ideal) achievable EFs. In our screening protocol, 1% of the database was selected for calculating the enrichment factors as a fixed threshold that enables the comparison of enrichment factors when studying the effectiveness of the screening protocol.

FlexX provided enrichment factors (Table 1) that were comparable with those reported by Lyne and co-workers for the structurally similar serine/threonine kinase, CDK2.[18] Using five different scoring functions, we first extracted the best poses from the 30 saved ones generated for all of the ligands. Then ligands with the lowest scores were ranked with the help of the same five scoring functions; thereby the pose extraction and ranking was separated resulting in 25 enrichment factors for each docking run. For the sake of clarity, Figures 2−5 show only 15 enrichment factors obtained by the three best performing scores in pose extraction. For 1UV5 and 1Q4L, ranking by Gold score gave the highest enrichment factor, while for 1Q3D, FlexX score best ranked the poses. Although ranking by different scores showed significant effect on enrichment factors, the scoring function used for pose selection has only limited impact. When compounds were docked with FlexX a maximal enrichment factor of 18 was achieved for 1Q4L and the best enrichment factor of 14 was calculated for both 1UV5 and 1Q3D. Introduction of pharmacophore constraints by FlexX-Pharm increased enrichment factors up to 28 (1UV5), 21 (1Q4L) and 19 (1Q3D), which is better than that obtained for CDK2.[18] For the three structures, FlexX and PMF scores best ranked the poses. Enrichment factors obtained under pharmacophoric constraints were found to be more sensitive to pose selection. In pose selection, the FlexX score for 1Q4L provided the best enrichment factor; for 1UV5, the PMF score best selected the poses; and for 1Q3D, PMF score provided the highest enrichment factor. Apart from the FlexX and PMF scores, sometimes the Dock score gave reasonable enrichment factors as well. When the selected poses were ranked, FlexX and PMF scores provided the best enrichment factors. For 1UV5, with pose selection by PMF and ranking by FlexX an enrichment factor of 28 was achieved, which is by far the best compared to those for 1Q4L or 1Q3D structures.

Because FlexX-Pharm could provide conformations with impossible torsion energy values, we introduced torsion energy constraints by setting the maximal torsion energy term to 12 kJ/mol from a 20 kJ/mol default value to eliminate these artifacts. With this setup, PMF and FlexX scores gave outstanding enrichment factors, but no significant improvement was observed compared to FlexX-Pharm.

Presently, FlexE cannot be combined with pharmacophore constraints, so a virtual screen by FlexE for these three structures was performed without any constraints. Enrichment factors are comparable to those given by FlexX. PMF and Gold scores ranked the poses well, while ChemScore, FlexX, and Dock scores best selected the poses. The best enrichment factor was 14, which is lower than the best one given by FlexX alone (EF = 18). Therefore the reduction in maximal torsion energy term and the usage of ensemble proteins were not used for our in-house library screening.

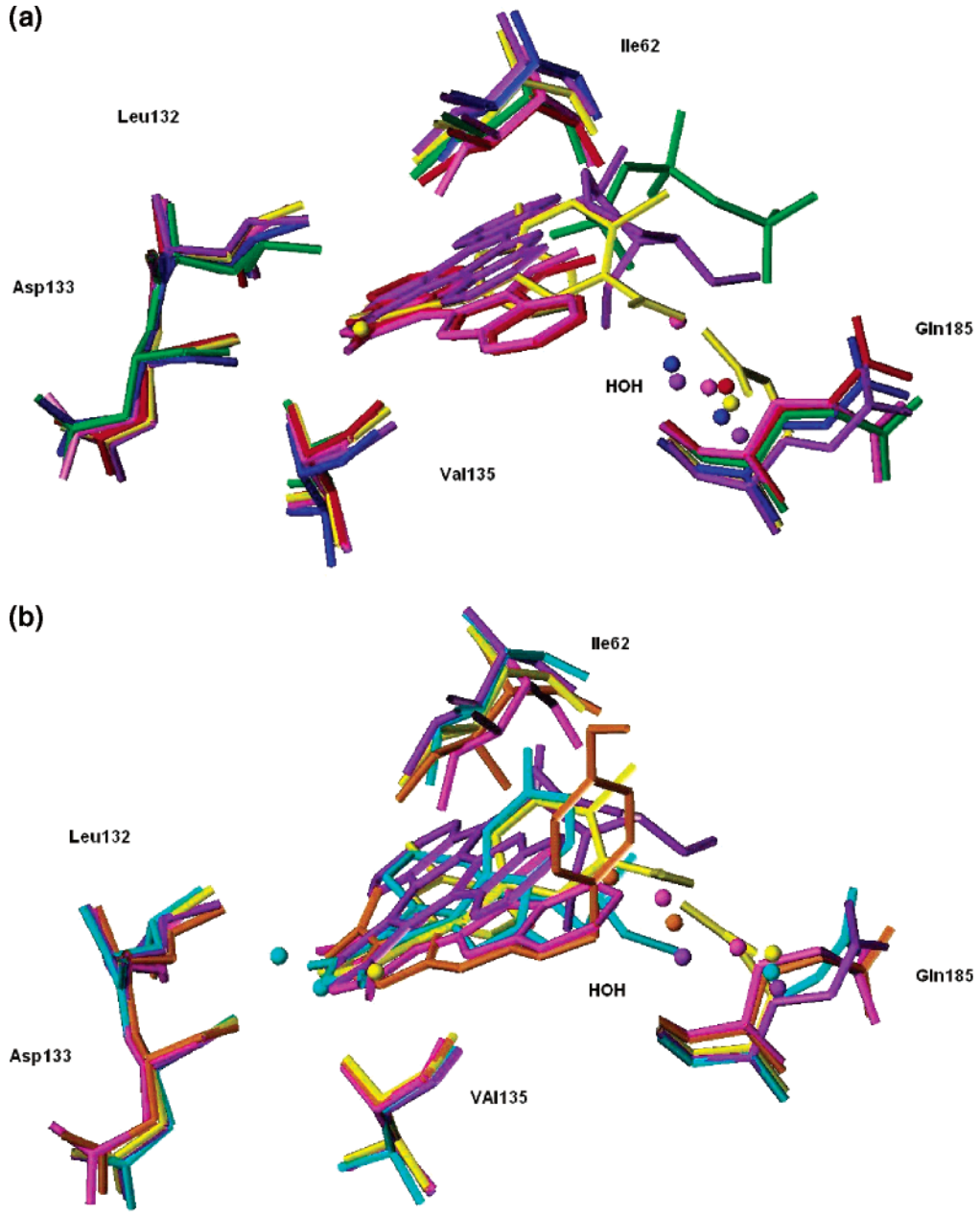


Figure 1. Active sites of (a) six X-ray structures and (b) X-ray structures used for virtual screening and the two recently released X-ray structures: 1Q3W blue; 1PYX green; 1UV5 magenta; 1Q41 red; 1Q3D violet; 1Q4L yellow; 1Q5K orange; 1R0E cyan. Only those residues that are given in the pharmacophore constraints are visible. Conserved water molecules are also assigned (HOH).

When FlexX or FlexX-Pharm are run on an Intel Xeon MP CPU 2.50 GHz processor, the average run time for docking one ligand is about 50 s, while on average the run time of FlexE is about factor of two lower than the accumulated time needed by FlexX or FlexX-Pharm for placing a ligand sequentially into all ensemble structures.

High-Throughput Screening Protocol. The Kinase-Glo luminescent kinase assay[22] developed by Koresawa and Okabe[21] is a homogeneous high-throughput screening method for measuring kinase activity by quantifying the amount of ATP remaining in the solution following a kinase reaction. This assay can be performed with any kinase and substrate combination and does not require radiolabeled components. The Kinase-Glo assay was adapted and optimized for screening against GSK-3β. To get the best performance in selecting between active and inactive compounds, the optimization of kinase reaction conditions was performed regarding both Promega's protocol[22] and the results of Koresawa et al.

Determining Optimal ATP Concentration. A direct linear relationship exists between the luminescence measured and the amount of ATP in the assay buffer from 0.003 to 3 μM ($R^2 = 0.999$) (Figure S1, see Supporting Information). The measured signal was stable and decreased only ~15% in 100 min at room temperature. The 1% DMSO concentration used for dissolving the test compounds did not influence the ATP standard curve (data not shown). The optimal ATP concentration for the kinase reaction was determined by varying the ATP concentration in an excess of kinase substrate and using as much kinase as reasonable (estimated to be 20 ng of enzyme per assay when the

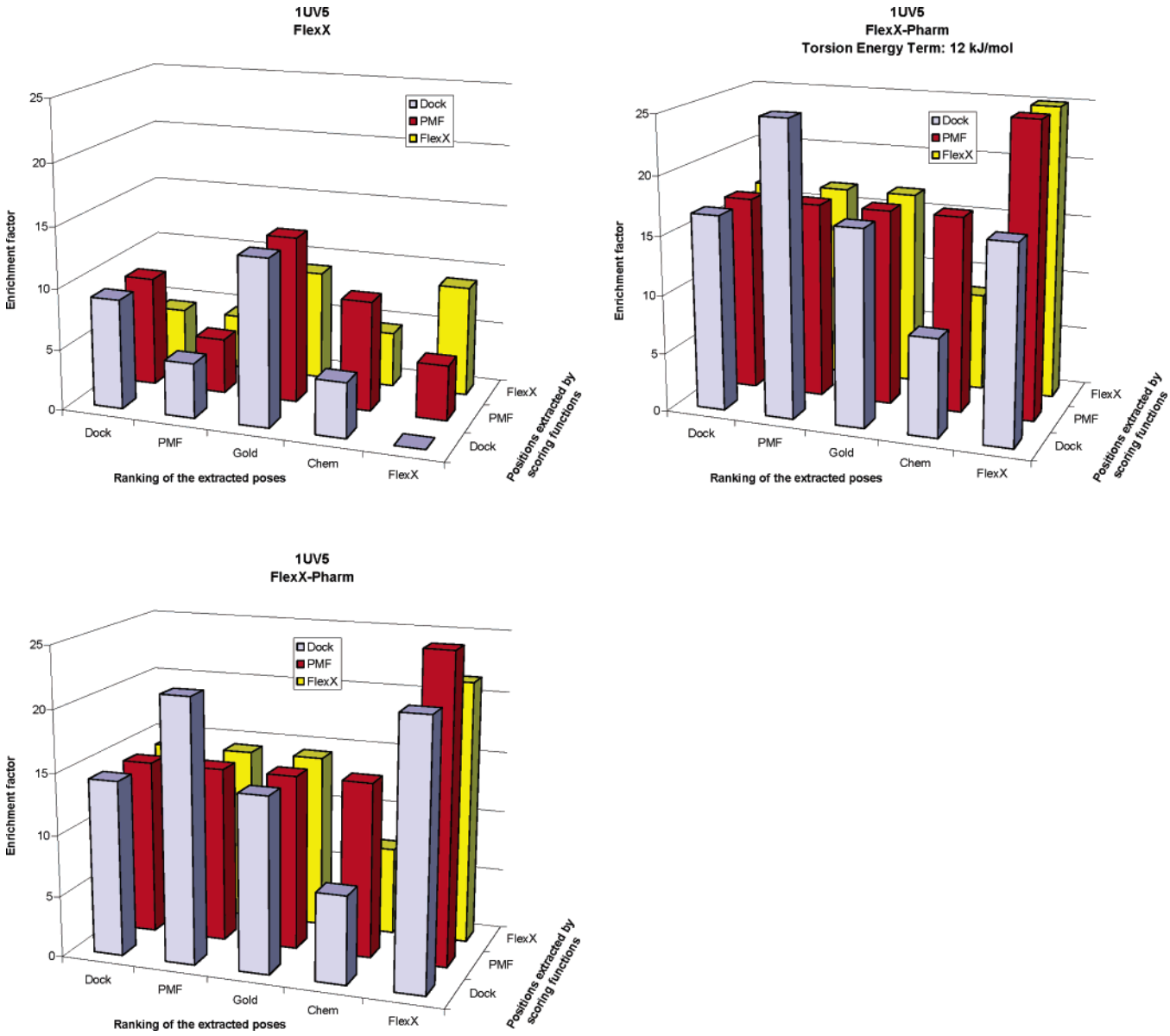


Figure 2. Enrichment factors observed at 1% of the ranked database for 1UV5 structure. Docking algorithms FlexX and FlexX-Pharm were used. The torsion energy term was also reduced to 12 kJ/mol where indicated; otherwise, the default value was set (20 kJ/mol).

enzyme activity was previously tested, data not shown) to achieve the maximal signal/noise ratio. The same ATP titration without GSK-3β was also performed as a control (Figure S2). The kinase was allowed to consume as much ATP as possible (~70%), enabling our assay to identify potent GSK-3β inhibitors. Although high substrate conversion may not be suitable for kinetic investigations, for HTS purposes, the robust signal and sensitivity are far more important requirements.[23] This substrate conversion might increase the number of false positive molecules while not causing the loss of the true positives. One micromolar ATP concentration was selected for high-throughput screening, as a reaction window large enough to select potent inhibitors from inactive compounds.

Determining Optimal Substrate Concentration. Serial dilutions of the prephosphorylated polypeptide kinase substrate were prepared: 1, 5, 25, 50, 100, and 200 μM containing 1 μM ATP and 20 ng of enzyme (Figure S3). Twenty-five micromolar substrate concentration was chosen as the optimum that provided a large change in luminescence (a high ΔRLU value) when kinase reaction wells were compared to wells containing no enzyme.

Determining Optimal GSK-3β Concentration. The ATP consumption was compared at different enzyme concentrations (Figure S4). The reaction rate was in the linear range up to 20 ng of GSK-3β. Linearity of the kinase reaction was preserved within 30 min of incubation at 30 °C (data not shown).

The optimized parameters for the GSK-3β assay are 25 μM prephosphorylated polypeptide substrate, 1 μM ATP, and 20 ng of enzyme in a final volume of 40 μL at 30 °C. The reaction was stopped after 30 min.

The optimized assay was first validated by measuring three known GSK-3β inhibitors as reference compounds, TDZD-8,[24] SB-41528635,[25] and AR-A014418.[26] The GSK-3β inhibitory activities were measured using Kinase-Glo and a radioisotope assay provided by Upstate.[27]

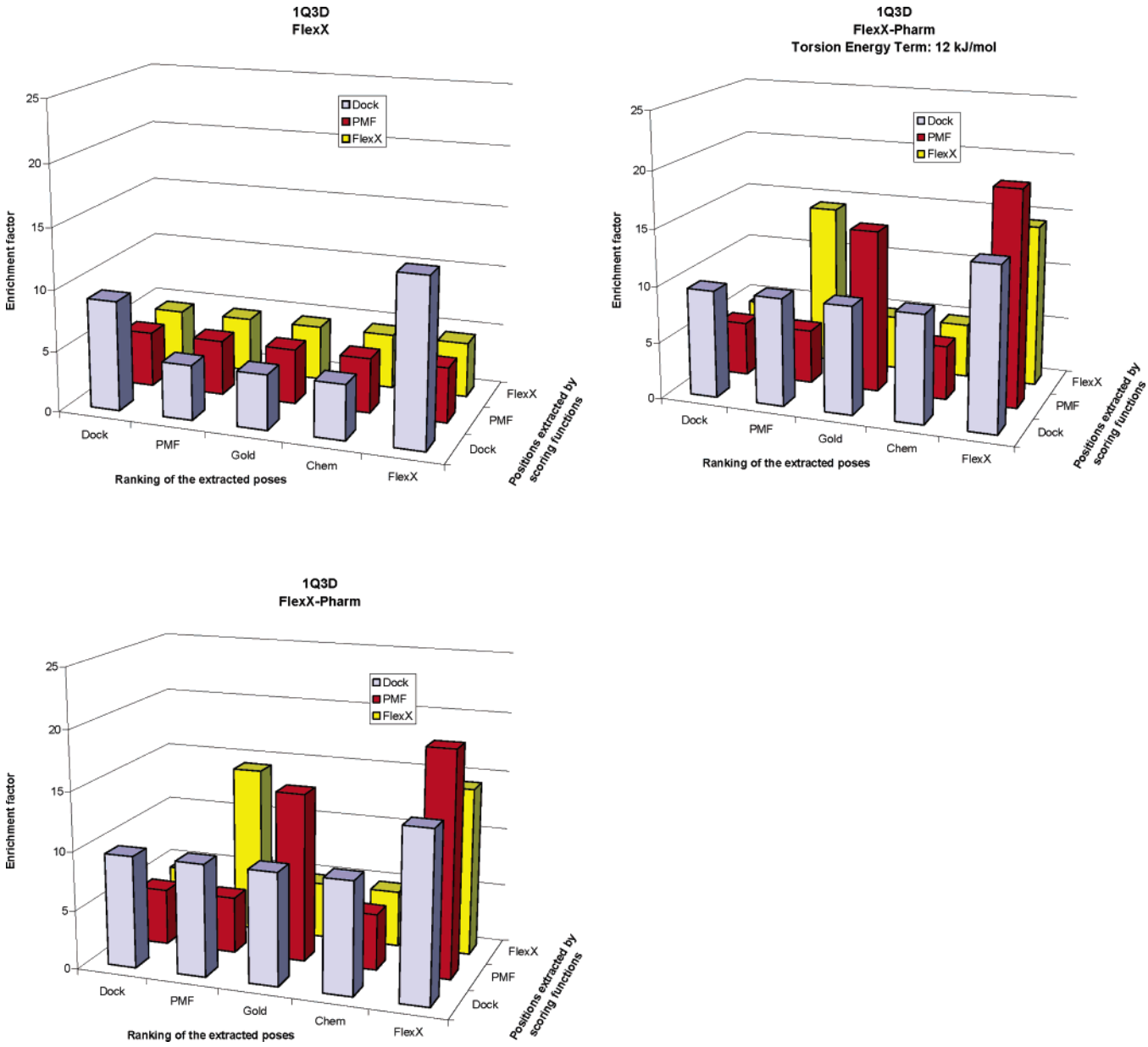


Figure 3. Enrichment factors observed at 1% of the ranked database for 1Q3D structure. Docking algorithms FlexX and FlexX-Pharm were used. The torsion energy term was also reduced to 12 kJ/mol where indicated; otherwise the default value was set (20 kJ/mol).

Measured values show good correlation and agree well with those described in the literature (Table 2).

High-Throughput Screening of the Corporate Sublibrary. The screening campaign involved testing 16 299 molecules, a diverse subset of our corporate database. The campaign was realized on AssayStation, an in-house robotic system developed for screening by homogeneous assays.[28] Compounds were plated to 96-well black test plates. Each plate contained 80 test compounds, negative controls (100 nM and 5 μM SB415826 for 50% and 100% inhibition, respectively), and positive controls (without inhibitor). Quality of measurements was enumerated by calculating the Z' factor for each plate:

$$Z \text{ factor} = 1 - ((3\mathrm{SD}(\mathrm{neg}) + 3\mathrm{SD}(\mathrm{pos}))/(\mathrm{avg}(\mathrm{neg}) - \mathrm{avg}(\mathrm{pos})))$$

where SD(neg) and SD(pos) stand for the standard deviation of data obtained for the negative and positive controls, respectively, and avg(neg) and avg(pos) are averages obtained for negative and positive controls, respectively.[29]

Z' factor varied between 0.6 and 0.9 for assay plates indicating that there is a large separation of data points between the baseline and the positive signal.

For each tested compound, we calculated the inhibition % as follows:

$$\mathrm{inhibition}\% = 100.0(\mathrm{raw\_data\_of\_compound} - \mathrm{avg}(\mathrm{pos}))/(\mathrm{avg}(\mathrm{neg}) - \mathrm{avg}(\mathrm{pos}))$$

Compounds having inhibition larger than 70% in a screen concentration of 10 μM were identified as hits and were then subjected to validation measurements in triplicate. Table 3 shows the summary of our results. Validated hits consist of 90 molecules that were clustered using chemical similarity principles. Molecules

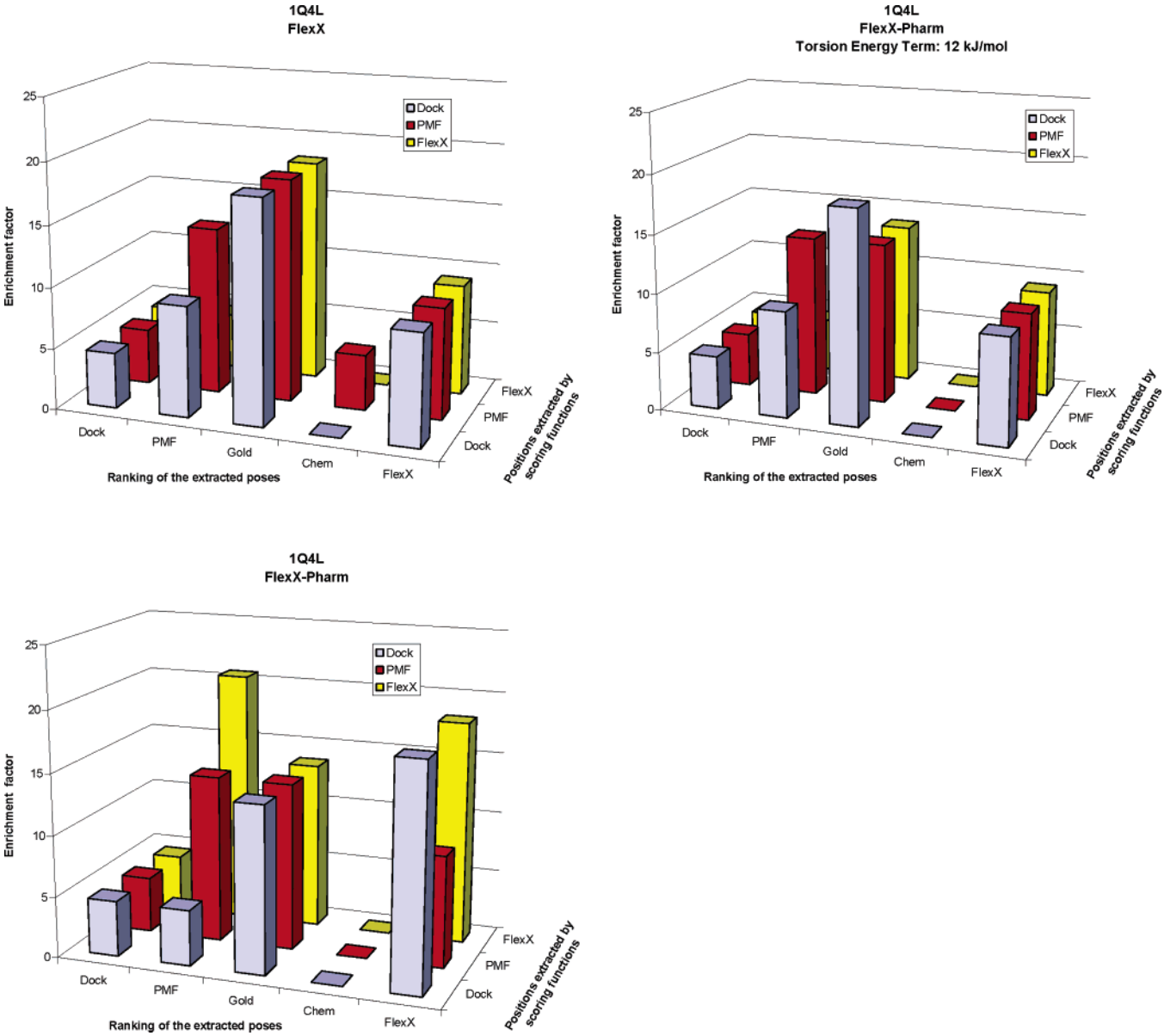


Figure 4. Enrichment factors observed at 1% of the ranked database for 1Q4L structure. Docking algorithms FlexX and FlexX-Pharm were used. The torsion energy term was also reduced to 12 kJ/mol where indicated; otherwise the default value was set (20 kJ/mol).

were represented by their 2D UNITY fingerprints, and the similarity of compounds was calculated as Tanimoto distances between 2D fingerprints. The resulting clusters of the validated hits (Table 4) allowed us to identify six compound series as promising GSK-3β inhibitors.

To exclude the promiscuous compounds,[30,31] the 90 validated hits were tested in three different assays including an aspartic protease, a glutamate receptor, and a peptidergic G-protein-coupled receptor (GPCR) test (brief descriptions of these assays are available in Supporting Information). None of the compounds showed activity against more than one target out of these three (Table S1), which suggests our validated hits to be free from frequent hitters.

Virtual Screening of the Corporate Sublibrary. The virtual screen of the sublibrary was performed against the 1UV5 X-ray structure by FlexX-Pharm, which provided the best results during the development phase (Table 5) without restricting the torsion energy term. Hits were evaluated using FlexX and PMF scores for both extracting poses and ranking compounds. Enrichment factors were calculated on the basis of the validated HTS hits (see Table 5) at 1%, 2%, 5%, and 10% of the ranked database (Table S2) indicating that all active compounds in the top 10% were already found at 1% of the ranked database.

One percent of the top ranking molecules (162 compounds) were then clustered as described for the HTS dataset, and chemical classes were compared to those of the HTS hits (Table 4). The evaluation of hits identified by virtual screening resulted in four chemical classes similar to those identified by HTS. Direct comparison of the chemical space covered by the two approaches was made by calculating the Tanimoto distance matrix among all of the validated hits obtained. Multidimensional scaling was used to illustrate the coverage of the chemistry space in two dimensions (Figure 6). This comparison indicated that the majority of the GSK-3β-related chemistry space identified by the HTS has also been explored by virtual screening.

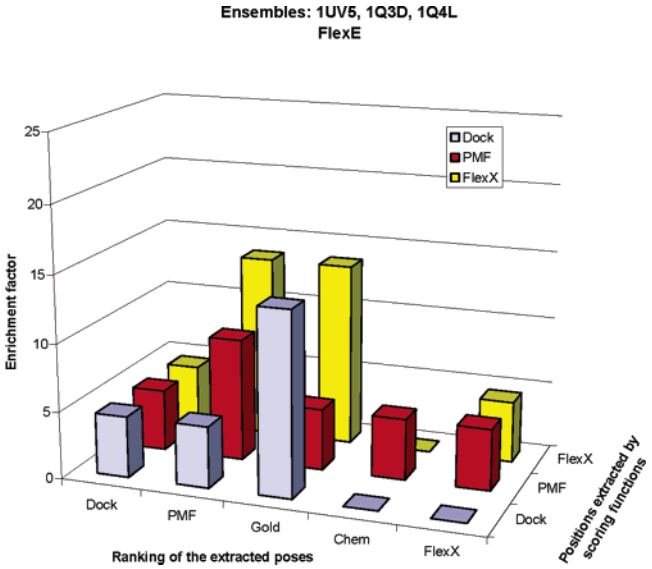


Figure 5. Enrichment factors observed at 1% of the ranked database for the ensemble. Docking algorithm FlexE was used.

Table 1. Best Enrichment Factors at 1% of Ranked Database for GSK-3β Structures[a]

structure	FlexX	FlexX-Pharm	FP-torsion
1Q3D	14	19	15
1UV5	14	28	25
1Q4L	18	21	18
ensemble	14		

[a] FP-torsion: FlexX-Pharm with torsion energy constraints.

Table 2. Comparison of IC_{50} Values of Reference Compounds

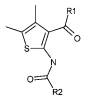
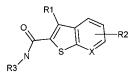
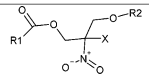
	IC_{50} value literature data	IC_{50} value in-house measured	
ref compd	radioactive assay (μM)	radioactive assay (μM)	Kinase-Glo assay (μM)
TDZD-8	2.000[a]	0.990	1.300
SB-415286	0.077	0.132	0.103
AR-A014418	0.104	N/D	0.187

[a] Measured with rabbit enzyme.

Table 3. A Short Summary of the Validated HTS Hits

	compds tested	hits having inhibition% >70%	hit rate	chemical classes
HTS	16 299	90	0.0055	6

Table 4. Chemical Classes of Hits Found by High-Throughput Screen and Virtual Screen[a]

	Chemical class	HTS	VS	VS+SS
Cluster 1		8 (83)	2 (81)	34 (19)
Cluster 2		6 (83)	2 (90)	4 (87)
Cluster 3		6 (75)	2 (83)	2 (83)
Cluster 4		3 (90)	1 (90)	2 (93)
Cluster 5		4 (90)	0	0
Cluster 6		4 (100)	0	0

[a] HTS = the number of hits found by HTS; VS = the number of hits found by VS; VS+SS = the number of hits found by VS combined with similarity seaching (similarity index > 0.85). The average inhibitory activities are shown in parentheses.

Table 5. Enrichment Factors at 1% of Ranked Corporate Database for GSK-3β (1UV5)[a]

	selecting position	
ranking	FlexX score	PMF score
FlexX score	23 (21)	23 (14)
PMF score	19 (28)	12 (14)

[a] Virtual screen was performed by FlexX-Pharm in consistency with parameters applied during the enrichment study. Enrichment factors in parentheses are observed during the enrichment study for 1UV5 when using FlexX-Pharm.

Discussion

Developing a Virtual Screening Protocol. Structural comparison of active sites in available GSK-3β structures revealed that inhibitor-bound forms show significant conformational differences. The Gln185 side chain adopts different rotamer conformations; it is in gauche conformation in 1UV5, 1Q41, 1R0E, and 1Q3W, has gauche′ conformation in 1Q3D, 1PYX, and 1Q5K, and shows a unique anti conformation in 1Q4L. Three different X-ray structures, one from each conformational group, were therefore selected for virtual screening (Figure 1a,b).

Since polar interactions are dominant within the GSK-3β active site, scores that consider these interactions are expected to give adequate results. Poses generated by FlexX were best ranked by Gold and FlexX scores. Gold[32,33] score is a GOLD-like function that focuses on H-bonding interactions, so this performs well in situations where significant polar interactions exist.[19,20] FlexX score performs well when the ligand has many site points (heteroatoms and ring centroids),[16] which is typically the case for cocrystallized GSK-3β inhibitors having at least four site points.

Introduction of pharmacophore constraints by FlexX-Pharm improved the enrichment factors significantly, which could be attributed to filtering out the inactive molecules and decreasing the number of false positives. This time, FlexX and PMF scoring combination did well for all of the structures. Lyne and co-workers used FlexX and PMF scoring combinations for CHK-1 and CDK-2 kinases.[18] Their studies suggested that proteins with structurally similar or identical active sites can be screened by the same scoring functions. Because CDK-2, CHK-1, and GSK-3 share a structurally identical

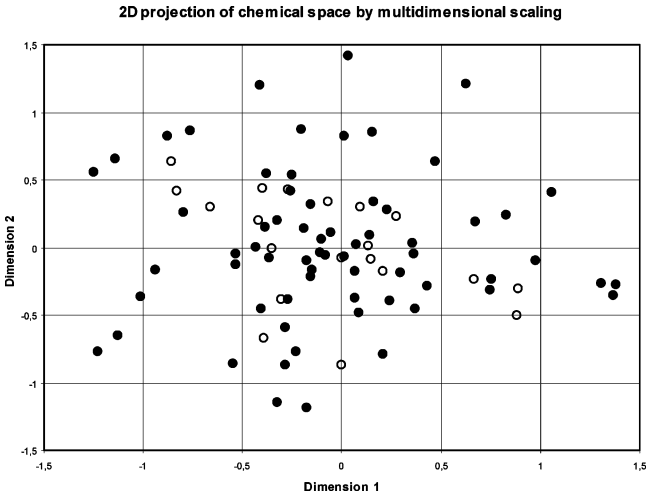


Figure 6. 2D projection of chemical space by multidimensional scaling based on Tanimoto distances calculated for hits obtained by virtual and experimental screening: (●) false negative molecules, 69; (○) true positive compounds, 21.

active site, the scoring combination suggested by Lyne was expected to be outstanding. In our case, independent from the conformation of the active site, the combination of PMF and FlexX scores for ranking and selecting poses gave the best enrichment factors, which gives further support to the application of similar scoring combinations for closely related active sites. Reducing the possibility of improper torsions by decreasing the maximal torsion energy term showed no effect on enrichment.

The conformational diversity of Gln185 within the GSK-3β active site prompted us to test FlexE, a docking tool developed for virtual screening of active sites represented by conformational ensembles. FlexE provided enrichment factors comparable to those given by FlexX. FlexE could not exceed FlexX-Pharm and did not even improve the enrichment given by FlexX. FlexE seems to provide the combination of enrichment factors of FlexX but cannot eliminate the false hits of FlexX and in consequence could not increase the enrichment. Comparing results obtained by FlexX and FlexX-Pharm, we assume that pharmacophore constraints could give significant improvements when combined with FlexE. In its present form, FlexE reduced the average run time for docking one ligand by almost 50%, which indicates that FlexE is a useful tool for docking to conformationally diverse active sites in a shorter time compared to independent FlexX runs.

Developing a High-Throughput Screening Protocol. Kinase-Glo is a homogeneous, nonradioactive, luminescent kinase assay for measuring the kinase activity by quantifying the ATP that is not consummated during the kinase reaction. This method was found to be highly sensitive, robust, and reproducible in the case of GSK-3β, producing acceptable Z' factors of 0.6−0.9 in the screening campaign. A potential disadvantage of the assay is that compounds with luciferase inhibitory activity or absorbing the luminescence emission are detected as false negatives. In fact, we could not detect such a compound during the validation of primary hits. IC_{50} values measured by this assay were comparable to those obtained by the conventional radioactive ^{32}P-γ-ATP-based assay and also to literature data. The robustness of the assay allowed us to adapt the process to our in-house HTS system, AssayStation.

Evaluating the High-Throughput and Virtual Screening Results. Screening of the 16 299-membered diverse sublibrary of our corporate collection afforded 117 primary hits demonstrating enzyme inhibition larger than 70% at 10 μM concentration. These hits were validated by measuring inhibition at the screening concentration in triplicate, which resulted in 90 validated hits. Validated hits were then clustered to six different chemical classes (see Table 4). This dataset obtained by experimental screening was used for the performance evaluation of our virtual screening protocol.

Structure-based virtual screening has long been applied in lead discovery and has demonstrated its effectiveness in several applications.[34] Validation of the methodology applied for virtual screening in these cases involves nothing else but picking several 10s to 100s of compounds (virtual hits) and checking their activity in a biological test. In such an evaluation, only compounds with high scores are considered that experimentally turn out to be true or false positives. However, false negatives, that is, compounds with low scores but high experimental activity are never explored and reported. Validation, by definition, needs exactly two procedures run in parallel: (i) experimental screening of a compound library consisting N compounds and (ii) virtual screening the very same N-membered library. Although there are some studies that reported a comparison of experimental and virtual screening, to the best of our knowledge, such a real validation has never been published.

Doman et al.[35] compared the performance of DOCK[36] and experimental screening against protein tyrosine phosphatase 1B by docking 235 000 commercially available compounds and screening a library of 400 000 compounds from the corporate collection. The hit rate of virtual screening (34.8%) was significantly higher than that obtained experimentally (0.021%). The unusually high enrichment factor of 1700 was rationalized by differences in drug likeness of the libraries and altered assay conditions used for evaluating hits identified by virtual screening. Unusually high enrichment achieved by virtual screening indicates that different libraries used for experimental and virtual screening not only prevent direct comparison but also have significant impact on the difference in hit rates achieved by the approaches.

In a combined virtual and experimental screening against peptide deformylase,[37] Howard and co-workers utilized distinct libraries (10 000 in-house compounds for HTS and 528 439 commercially available compounds for virtual screening). As in the previous case, libraries used for HTS and virtual screening were different not only in vendors and numbers but also in physicochemical character, drug likeness, etc., which prevented the direct comparison again. Although both HTS and virtual screening identified chemically similar inhibitors, hit rates and enrichment factors were not reported and compared.

Jenkins et al. reported a case study on small-molecule angiogenin inhibitors;[38] 18 111 compounds were experimentally tested in an ANG-catalyzed oligonucleotide

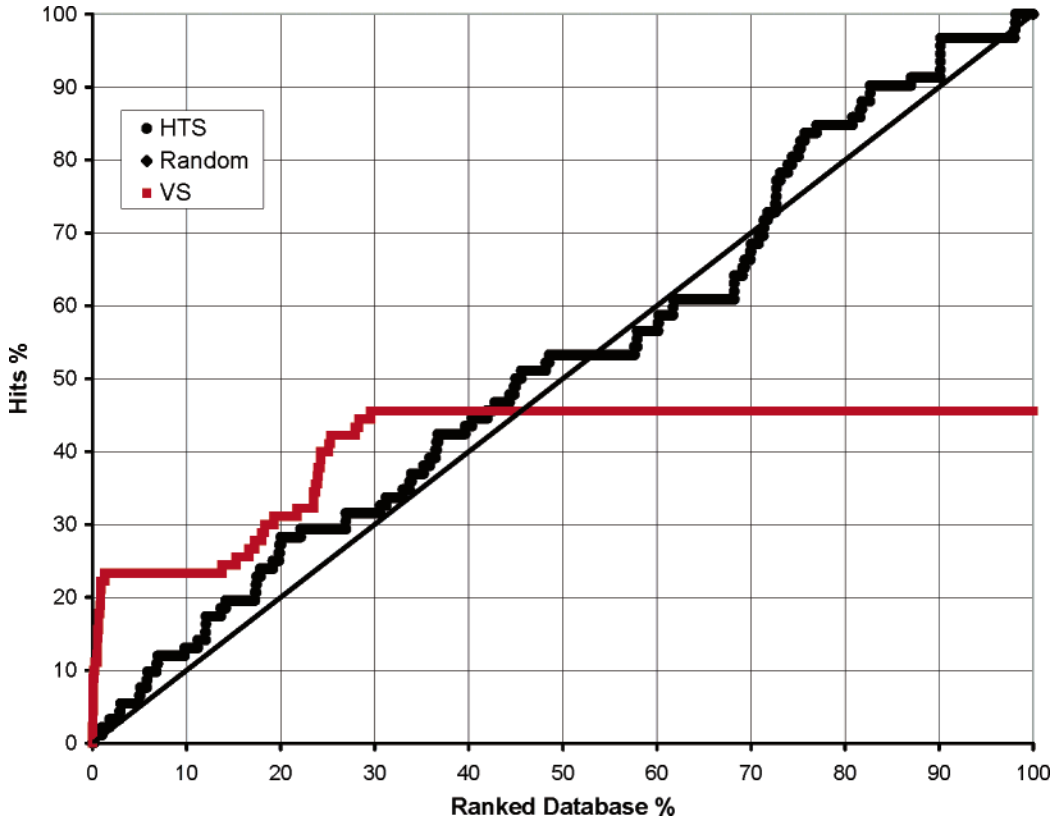


Figure 7. Enrichment plot of VS and HTS data. Red curve is the enrichment plot of virtual screening calculated as the ratio of the hits found by the virtual screen vs the ratio of the ranked database. The linear thin line indicates the random distribution of active molecules. The thick black curve is the enrichment plot of high-throughput screening calculated as the ratio of the hits found by the high-throughput screen vs the ratio of the ranked database; here the database is ranked by time. Deviation of the HTS curve from random is the effect of plate design.

cleavage assay, which resulted in 12 validated hits (hit rate 0.066%). The same library was subjected to virtual screening in which GOLD yielded the best hit rate of 1.4%; the enrichment factor was around 20. Although datasets in this case allowed direct comparison, the authors demonstrated the ability of virtual screening to enrich HTS hit lists (true and false positives) and did not analyzed hits from different approaches comparatively.

The most direct comparison between HTS and virtual screening up to now has been published by Paiva et al. reporting the discovery of dihydrodipicolinate inhibitors.[39] The authors used the very same source, the Merck chemical collection, to select compounds for screening that ensured the physicochemical character and drug likeness to be more balanced. Experimental and virtual screening were, however, performed on sublibraries different in nature and size. HTS and virtual screening gave hit rates of 0.2% and 6%, respectively, which resulted in an enrichment factor of about 30. In the last two cases, when both approaches utilized the same source for compounds, the enrichment achieved by virtual screening was 20−30-fold, that is, within the range reported for high-throughput docking studies when compared to random. This finding again indicates the importance of the source of compounds used for comparative studies.

Here we report for the first time a direct comparison of experimental and virtual screening utilizing the protocol developed for GSK-3β on a single sublibrary of our corporate collection. Comparative evaluation of screening results involved the following criteria to be investigated: (i) enrichment achieved, (ii) true positives vs false positives and true positives vs false negatives, and (iii) clusters of actives identified.

The effectiveness of our protocol was first checked by calculating the enrichment factors, as shown in Table 5. The performance of the virtual screening protocol was very similar to that observed in the dedicated enrichment study performed during its development. This result suggests that adequate virtual screening protocols can identify active molecules in both enrichment studies (artificial libraries mixed from active compounds and hopefully inactive molecules) and real screening situations.

Enrichment plots are useful when comparing the performance of different virtual screening protocols in enrichment studies using artificial libraries. In this case, the performance of the protocol is compared to a random situation. The availability of HTS data allowed us to calculate the enrichment plot for the real screening library (Figure 7) where time ranking was considered.

Figure 7 shows that enrichment factors in real screening might be somewhat different from those in the random case depending on the plate design of the screening library. For between 20% and 30% of tested compounds, the enrichment factors calculated for the HTS are better than random and more interestingly are comparable to those achieved by virtual screening when the top 20%−30% of the ranked database is considered. However, when less than 15% of the tested/ranked database is considered, the enrichment achieved by

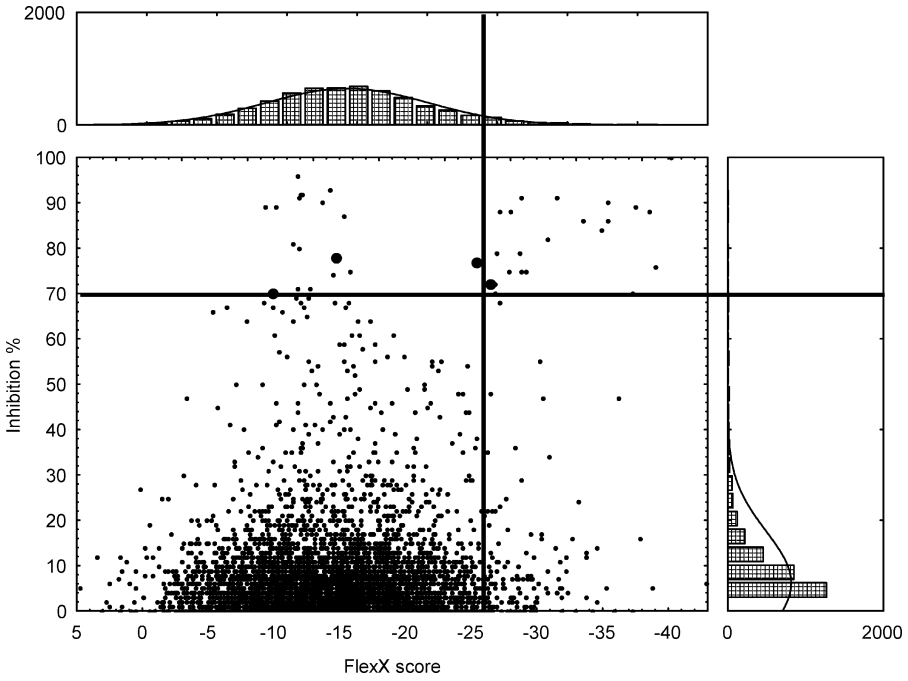


Figure 8. Inhibition % vs FlexX score. Filled black circles show HTS hits that were not validated. The 1% level of the ranked database is at FlexX score = −26.4. From the 90 validated HTS hits, only 41 could be docked using FlexX-Pharm; 49 did not fulfill the pharmacophore constraints; therefore, these cannot be plotted here.

virtual screening is significantly higher than that which occurred randomly in HTS. Comparison of experimental and virtual screening data in the top 5% of the ranked database is the basis of the efficacy of our virtual screening protocol. Figure 7 demonstrates that true hits were most frequently identified in the top 1−2% of the ranked database. This performance resembles that of previously reported successful virtual screens that calculated enrichment factors at 1−5% of the ranked database. It should be noted, however, that the enrichment factor achieved by our virtual screen reaches the maximum of 45 at around 40% of the ranked database. After this point, both the experimental and the truly random situation outperform the virtual screen since the latter could only identify 45% of the experimental hits. Since our investigation is the first comparative study using identical databases this behavior of virtual screening has never been reported previously.

Our next criterion involved the comparison of true and false positives and negatives identified by the two approaches. Clustering the first 1% of the ranked database obtained by virtual screening (162 compounds) resulted in four clusters out of six identified by HTS. Activity data obtained by HTS were compared directly to scores calculated during virtual screening. This comparison was possible for compounds that were successfully docked and scored by FlexX-Pharm (5904 datapoints). We used 70% as a cutoff for actives in experimental screening, while hits of virtual screening were identified as the top 1% of the ranked database. Compounds without any docking solution and score were considered as inactives in virtual screening. Figure 8 depicts the result of this analysis.

Considering primary data, we found that virtual screening identified 21 validated true positives and 141 false positives. The calculated hit rate (12.9%) is higher than that of the experimental screening (0.55%). The higher hit rate of virtual screening goes at the expense of the relatively high rate of false positives. The 23-fold enrichment factor achieved by virtual screening suggests that our protocol is effective. In other words, considering the top 1% of the ranked database, virtual screening selects 162 compounds for biological testing and out of them 21 are active, which parallels structure-based virtual screens reported in the literature.[34]

It was interesting to see, however, that docking scores showed no correlation with GSK-3β inhibitory activities as measured by HTS. Although single-point HTS results are far from accurate and there are indications in the literature that even the rank correlation between hit lists of HTS and virtual screening could be limited, there is at least a hope that biological activity and docking score run parallel. In our case, we were surprised when we recognized that not only the ranking but also the distributions of activity measures are different. Distribution of HTS data shows a well-known picture with maximum centered in the low activity range around 5%, which can be considered as the noise of the assay. In contrast, distribution of docking scores follows a quasi-normal distribution centered at ∼50% of the ranked database, which indicates significant uncertainty in activity predictions.

One of the major consequences of this uncertainty is the high number of false negatives (69 compounds in total) indicating that most of the primary hits were lost by virtual screening. Our GSK-3β assay was performed using the recombinant protein, and the inhibitory action of the compounds was measured by quantifying the remaining ATP after the kinase reaction, which allows the identification of both ATP-competitive and non-competitive inhibitors. The virtual screening protocol utilized the 3D structure of the very same protein, and we only considered the ATP binding site. Consequently, a chance for binding to alternate binding sites might rationalize the larger diversity of HTS hits.

The inability of scoring functions can also be considered as a potential source of false negatives. Although the combined scoring function was able to give reasonable enrichment factors in both mimicked and real screening situations, we speculate that the inaccuracy of the docking scores might be partially responsible for this limitation of virtual screening.

In addition to scoring problems, we suspect pharmacophore constraints as being responsible for the limited performance. To clarify this issue, false negatives were further analyzed. Considering that FlexX-Pharm could only dock 5904 compounds out of the 16 299, one can conclude that false negatives might be (i) compounds docked successfully but misscored and (ii) compounds docked unsuccessfully. Careful analysis of the false negatives revealed that FlexX-Pharm could not find any docking solution for 49 compounds. The remaining 20 compounds out of the 69 false negatives were misscored. Further analysis of misdocked compounds suggested that the lack of the docking solution could be caused by the docking algorithm itself or the pharmacophore constraints applied. The misdocked compounds were therefore redocked by FlexX without any pharmacophore constraint. FlexX docked all but one compound successfully revealing that the majority of misdocked false negatives are due to pharmacophore constraints.

The enrichment factor achieved during the development of our virtual screening protocol was found to be 14 as calculated at 1% of the ranked database, suggesting that the distribution of the further 86% of actives is not known within the 99% of the screening library. In other words, one can expect a false negative rate as high as 86% when comparing results to those of experimental screening. In fact our false negative rate was found to be ~77%, rather close to that suggested by the enrichment factor calculated. Since the success criterion of virtual screening is rather to find active compounds than exploring all possible ligands in the library, the huge number of false negatives might be acceptable. Since this acceptance is closely related to the number of clusters successfully identified by virtual screening, we compared the clusters of actives identified by the two approaches.

Comparison of the chemical space covered by HTS and virtual screening is shown in Figure 6, which reveals these spaces to be comparable. In fact, virtual screening successfully identified four of the six HTS clusters, and we only found two active clusters that were lost (Table 4). This situation is unexpected when using different libraries, and it was suggested to be the main reason virtual and experimental screenings are complementary rather than competitive. Hit lists of experimental and virtual screening were markedly dissimilar for PTP-1B. Virtual and experimental hits identified for dihydrodipicolinate,[39] angiogenin,[38] and peptide deformylase[37] showed some partial overlap, but different libraries used prevented direct comparisons. In our case, using the same library for both virtual and experimental screening, this overlap is significant and indicates the success of our virtual screening protocol. It should be noted, however, that clusters were only partially populated by virtual screening (Table 4). Virtual screening identified 25−33% of actives in clusters 1−4 and failed to pick up any hit from clusters 5 and 6. Limited population of virtual screening hits in active HTS clusters suggests again that scoring functions and pharmacophore constraints could exclude some part of the active chemistry space identified by HTS. Disfavored side chain orientations within the active site, however, could also limit the identification of compounds with identical cores but different substituents located within a single cluster. One can therefore conclude that the limited flexibility of side chains would also be a factor responsible for the lost hits. These limitations could be overcome by similarity searches coupled to virtual screening. As seen in Table 4, similarity search using a Tanimoto distance ratio of 0.85[40] around virtual hits retrieves more actives in clusters 2 and 4. Although similarity search identified a large amount of compounds around the virtual screening hits of cluster 1, their average activity was found to be low because of the limited active content of this cluster.

Conclusion

We developed a successful virtual screening protocol for GSK-3β inhibitors using the FlexX suite. The effects of pharmacophore and torsional constraints, as well as protein flexibility, were investigated. Pharmacophore constraints are an effective way of filtering out many unwanted structures in a database of virtual compounds. They offer great help in enrichment studies by overcoming the uncertainties of current scoring functions due to their strictness and filtering power. On one hand, however, they could also prevent one from finding truly novel ligands, which could be a rationale for the larger diversity of HTS hits when compared to that of virtual screening. On the other hand, they showed less impact on false positives and especially on false negatives.

Reasonable enrichment factors have been achieved for both crystal structures of GSK-3β using constrained (FlexX-Pharm) and unconstrained (FlexX) docking algorithms. Increased performance of FlexX-Pharm over FlexX was also shown for GSK-3β. The same combination of scoring schemes observed for CDK-2, CHK-1, and GSK-3β suggest that for similar active sites unique methods can be applied, but differences in enrichment factors could be rationalized by qualitatively different input libraries. It might show that score selection is dependent on the active site structure itself, while the enrichment is dependent on the docking library. In addition to this effect, it should be noted that the inactive and active ratio strongly influences the enrichment.

Although this study gives further support to the hypothesis that similar virtual screening protocols can be applied for structurally similar proteins, we showed that considering different protein conformations by FlexE is far from adequate for taking protein flexibility into account.

Unfortunately, FlexE does not include a solution for the problem of ligand-induced fit, so it is essential to build in protein flexibility by including docking calculations when investigating selectivity issues. At the moment, the "flexibility" solution provided by FlexE does not seem to reflect true flexibility, but FlexE could be a useful tool for merging different crystal structures and docking into them simultaneously.

We conclude that pharmacophore constraints can be applied successfully for GSK-3β kinase, and the scoring combination suggested by Lyne et al. was found to be effective in our virtual screening study. We provided here some further support for applying general/similar scoring schemes for various proteins sharing structurally similar active sites.

The effectiveness of our screening protocol has been partially demonstrated by comparing the results of a virtual screen to those obtained by experimental screening. Virtual screening picked up four out of six series of compounds identified by HTS, but the large number of false positives and the high rate of false negatives indicate significant limitations. Although FlexX-Pharm with the combined FlexX/PMF scoring functions was able to give reasonable enrichment factors in both artificial and real screening situations, we showed that the inaccuracy of the docking scores, the application of pharmacophore constraints, and the neglected flexibility of the active site might be responsible for these limitations.

On the basis of the success criteria used in the screening literature, our virtual screening protocol was successful in producing remarkable enrichment and identifying the majority of active clusters at much lower cost and time relative to HTS. These results suggest that this protocol can be useful for prefiltering our in-house library and can complement experimental screening when investigating large, commercially available virtual libraries for GSK-3β inhibitors.

We feel that this work, being the first truly comparative study, might be interesting for the broader community of lead discovery teams when planning experimental and virtual screening activities.

Materials and Methods

Virtual Screening. Preparation of the Ligand Database. Our screening library involves a subset of World Drug Index (WDI) as inactive molecules that were specifically designated to reduce artificial enrichment.[39] The WDI was first filtered to eliminate compounds having molecular weight lower than 200 and greater than 800, log *P* larger than 7, and rotatable bonds more than 15. Inorganic compounds were also filtered. Out of the ~40 000 remaining molecules, 5354 were chosen on the basis of a diversity selection (2D UNITY Fingerprint); Tanimoto index was lower than 0.68. The GSK-3β ligand set (25 known inhibitors) was compiled manually from Prous Integrity Drugs & Biologics database[40] and other public sources.[41−49] Our main selection criterion was to cover as many different compound classes as possible to establish structural diversity among the inhibitors. The input library comprised the 5354 inactive and 25 active compounds (active molecules content ~0.5% such as in a typical real high-throughput screen).

The preprocessing of a subset of our in-house library was the same as mentioned above. This subset is comprised of 16 299 molecules in total.

Preparation of Protein Coordinates and Definition of Active Sites for Docking with FlexX, FlexX-Pharm, and FlexE. Crystal structures of GSK-3β were used to screen a benchmark library of known inhibitors. X-ray structures 1Q3W[3] (alsterpaullone, 2.3 Å), 1Q41[3] (indirubin-3′-monoxime, 2.1 Å), 1Q4L[3] (inhibitor I-5, 2.77 Å), 1PYX[3] (AMP−PNP, 2.4 Å), 1UV5[5] (6-bromoindirubin-3′-oxime, 2.8 Å), 1R0E[4] (release date 12 Oct-2004, 3-indolyl-4-arylmaleimide, 2.25 Å), 1Q5K[4] (release date 10-Aug-2004, *N*-(4-methoxybenzyl)-*N*′-(5-nitro-1,3-thiazol-2-yl) urea, 1.94 Å), and 1Q3D[3] (staurosporine, 2.2 Å) complexed with inhibitors were considered. Crystal structures 1UV5, 1Q3D, and 1Q4L were used for virtual screening as a visual selection based on the conformation of Gln185 (Figure 1a, violet−1Q3D, magenta−1UV5, yellow−1Q4L). Recently released structures such as 1R0E and 1Q5K have only been used for structural analysis (Figure 1b). The "A" chains were used for virtual screening in each case. Actives sites were defined as all the residues within 6.5 Å of the bound ligand. Two conserved water molecules in case of 1UV5 and 1Q3D and one water molecule in 1Q4L were also added to the active sites.

For screening the corporate sublibrary, we used the 1UV5 X-ray structure as described above.

FlexX 1.13.2, FlexX-Pharm, and FlexE Docking. Standard parameters were used as implemented in SYBYL 6.9.2 package.[50] Formal charges and the particle concept options were always checked. Thirty poses were saved in mol2 files for further analysis. All stored poses were rescored using the CScore module of SYBYL 6.9.2 comprising the following functions: Dock, Chem, FlexX, PMF, and Gold scoring function. The torsional energy term, beside the default value (20 kJ/mol), was set to 12 kJ/mol to reduce the number of impossible torsions.

Pharmacophore constraints assigned by comparative analysis of ligand-bound X-ray structures were given as essential interaction constraints for Val135 and optional interaction constraints for Leu132, Asp133, Tyr134, and Gln185.

Structures 1UV5, 1Q3D, and 1Q4L were aligned by homology for docking with FlexE. Structure 1UV5 was treated as a reference structure for the alignment. One water molecule was added to 1Q4L, and heteroatom files were generated for 1UV5 and 1Q3D containing two structural water molecules. Positions were scored externally using the CScore module of the SYBYL 6.9.2 package.

High-Throughput Screen. Reagents. Human recombinant glycogen synthase kinase-3β was purchased from Upstate (14-306). The prephosphorylated polypeptide substrate was synthesized by American Peptide Incorporation. Kinase-Glo luminescent kinase assay was obtained from Promega (V6711). The ATP (A-7699) and all other reagents were from SIGMA Chemical Co. The water used for buffer preparation was Millipore grade pure. Assay buffer contained 50 mM Hepes, pH = 7.5, 1 mM EDTA, 1 mM EGTA, and 15 mM magnesium acetate.

GSK-3β Assay Protocol. Kinase-Glo assays were performed in assay buffer in black 96-well-plates. In a typical assay, 10 μL (10 μM) of test compound (dissolved in DMSO) and 10 μL (20 ng) of enzyme were added to each well, followed by 20 μL of assay buffer containing 25 μM subtrate and 1 μM ATP. The final DMSO concentration in the reaction mixture does not exceed 1%. After incubation at 30 °C for 30 min, 40 μL of Kinase-Glo reagent was added to stop the reaction. The luminescene was recorded after 10 min using FLUOstar-POLARstar Optima (BMG Labtech) multimode reader. For background corrections, luminescence was measured in wells containing the reaction mixture without enzyme. The activity is expressed as a difference of the consumed ATP and total ATP. The inhibitory activities are expressed in percentage of maximal activities (in the absence of inhibitor). IC_{50} values were calculated using Origin software.

Acknowledgment. The authors thank for Ms. Éva Horváth and Ms. Krisztina Recska for technical assistance.

Supporting Information Available: Brief description of the assays used for checking promiscuous inhibitors; figures demonstrating the relationship between luminescent signal measured and the ATP concentration in our GSK-3β assay, showing the ATP-luminescence standard curve and calibration curves used for determining the optimal substrate concentration and the optimal GSK-3β concentration; tables summarizing the measured inhibition % of GSK-3β hits in three different assay systems and enrichment factors calcu-

lated at 1%, 2%, 5%, and 10% of the ranked database for enrichment studies and virtual screening of the corporate sublibrary. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Meijer, L.; Flajolet, M.; Greengard, P. Pharmacological inhibitors of glycogen synthase kinase 3. *Trends Pharmacol. Sci*. **2004**, *25*, 471−480.

(2) Woodgett, J. R. Molecular cloning and expression of glycogen synthase kinase-3/factor A. *EMBO J.* **1990**, *9*, 2431−2438.

(3) Bertrand, J. A.; Thieffine, S.; Vulpetti, A.; Cristiani, C.; Valsasina, B.; Knapp, S.; Kalisz, H. M.; Flocco, M. Structural characterization of the gsk-3beta active site using selective and nonselective ATP-mimetic inhibitors. *J. Mol. Biol.* **2003**, *333*, 393−407.

(4) Bhat, R.; Xue, Y.; Berg, S.; Hellberg, S.; Ormo, M.; Nilsson, Y.; Radesater, A. C.; Jerning, E.; Markgren, P. O.; Borgegard, T.; Nylof, M.; Gimenez-Cassina, A.; Hernandez, F.; Lucas, J. J.; Diaz-Nido, J.; Avila, J. Structural insights and biological effects of glycogen synthase kinase 3-specific inhibitor AR-A014418. *J. Biol. Chem.* **2003**, *278*, 45937−45945.

(5) Meijer, L.; Skaltsounis, A.-L.; Magiatis, P.; Polychronopoulous, P.; Knockaert, M.; Leost, M.; Ryan, X. P.; Vonica, C. A.; Brivanlou, A.; Dajani, R.; Crovace, C.; Tarricone, C.; Musacchio, A.; Roe, S. M.; Pearl, L. H.; Greengard P. Gsk-3-Selective Inhibitors derived from tyrian purple indirubins. *Chem. Biol*. **2003**, *10*, 1255−1266.

(6) Bax, B.; Carter, P. S.; Lewis, C.; Guy, A. R.; Bridges, A.; Tanner, R.; Pettman, G.; Mannix, C.; Culbert, A. A.; Brown, M. J.; Smith, D. G.; Reith, A. D. The structure of phosphorylated GSK-3beta complexed with a peptide, FRATtide, that inhibits beta-catenin phosphorylation. *Structure* **2001**, *9*, 1143−1152.

(7) Dajani, R.; Fraser, E.; Roe, S. M.; Young, N.; Good, V.; Dale, T. C.; Pearl, L. H. Crystal structure of glycogen synthase kinase 3 beta: structural basis for phosphate-primed substrate specificity and autoinhibition. *Cell* **2001**, *105*, 721−732.

(8) ter Haar, E.; Coll, J. T.; Austen, D. A.; Hsiao, H. M.; Swenson, L.; Jain, J. Structure of GSK-3beta reveals a primed phosphorylation mechanism. *Nat. Struct. Biol.* **2001**, *8*, 593−596.

(9) Dajani, R.; Fraser, E.; Roe, S. M.; Yeo, M.; Good, V.; Thompson, V.; Dale, T. C.; Pearl, L. H. Structural basis for recruitment of glycogen synthase kinase 3beta to the axin-apc scaffold complex. *EMBO J.* **2003**, *22*, 494−501.

(10) Krupa, A.; Preethi, G.; Srinivasan, N. Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in ser/thr and tyr kinases. *J. Mol. Biol.* **2004**, *339*, 1025−1039.

(11) Toledo, L.; Lydon, N. B.; Elbaum, D. The structure-based design of ATP-site directed protein kinase inhibitors. *Curr. Med. Chem.* **1999**, *6*, 775−805.

(12) Eldar-Finkelman, H. Glycogen synthase 3: an emerging therapeutic target. *Trends Mol. Med.* **2002**, *8*, 126−132.

(13) Lucas, J. J.; Hernandez, F.; Gomez-Ramos, P.; Moran, M.; Hen, R.; Avila, J. Decreased nuclear β-catenin, tau hyperphosphopylation and neurodegeneration in GSK-3β conditional transgenic mice. *EMBO J.* **2000**, *20*, 27−39.

(14) Doble, B. W.; Woodgett, J. R. GSK-3 tricks of the trade for multitasking kinase. *J. Cell. Sci.* **2003**, *116*, 1175−1186.

(15) Hindle, S. A.; Rarey, M.; Buning, C.; Lengaue, T. Flexible docking under pharmacophore type constraints. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 129−149.

(16) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.

(17) Claussen, H.; Buning, C.; Rarey, M.; Lengauer, T. FlexE: Efficient molecular docking considering protein structure variations. *J. Mol. Biol.* **2001**, *308*, 377−395.

(18) Lyne, P. D.; Kenny, P. W.; Cosgrove, D. A.; Deng, C.; Ashwell, S.; Zabludoff, S.; Wendoloski, J. J. Identification of compounds with nanomolar binding affinity for checkpoint kinase-1 using knowledge-based virtual screening. *J. Med. Chem.* **2004**, *47*, 1962−1968.

(19) Shultz-Gasch, T.; Stahl, M. Binding site characteristics in structure-based virtual screening: evaluation of current docking tools. *J. Mol. Model.* **2003**, *9*, 47−57.

(20) Stahl, M.; Rarey, M. Detailed Analysis of Scoring Functions for Virtual Screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.

(21) Koresawa, M.; Okabe, T. High-Throughput Screening with Quantitation of ATP Consumption: A Universal Non-Radioisotope, Homogeneous Assay for Protein Kinase. *Assay Drug Dev. Technol.* **2004**, *2*, 153−160.

(22) *Kinase-Glo Luminescent Kinase Assay*; Promega Corporation: Madison, WI, 2002.

(23) Wu, G.; Yuan, Y.; Hodge, C. N. Determining Appropriate Substrate Conversion for Enzymatic Assay in High-Troughput Screening. *J. Biomol. Screening* **2003**, *8*, 694−700.

(24) Martinez, A.; Alonso, M.; Castro, A.; Perez, C.; Moreno, F. J. First non-ATP competitive glycogen synthase kinase 3β (GSK-3β) inhibitors: thiadiazolidinones (TDZD) as potential drugs for the treatment of Alzheimer's disease. *J. Med. Chem.* **2002**, *45*, 1292−1299.

(25) Coghlan, M. P.; Culbert, A. A.; Cross, D. A. E.; Corcoran, S. L.; Yates, J. W.; Pearce, N. J.; Rausch, O. L.; Murpy, G. J.; Carter, P. S.; Cox, L. R.; Mills, D.; Brown, M. J.; Haigh, D.; Ward, R. W.; Smith, D. G.; Murray, K. J.; Reith, A. D.; Holder, J. C. Selective small molecule inhibitors of glycogen synthase kinase-3 modulate glycogen metabolism and gene transcription. *Chem. Biol.* **2000**, *7*, 793−803.

(26) Bhat, R.; Xue, Y.; Berg, S.; Hellberg, S.; Ormö, M.; Nilsson, Y.; Radesater, A. C. Jerning, E.; Markgren, P. O.; Borgegard, J. J.; Diaz-Nido, J.; Avila, J. Structural insights and biological effects of glycogen synthase kinase 3-specific inhibition: AR-A014418. *J. Biol. Chem.* **2003**, *278*, 45937−45945.

(27) Upstate cell signaling solutions, a Serologicals Company, 48 Barn Road, Lake Placid, NY 12946.

(28) Bánki, Zs.; Báthor, M.; Molnár, L.; Bielik, A.; Keserű G. M. Scheduling a flexible, open-architecture robotic workstation under LabWindows. *JALA* **2005**, *10*, 149−154.

(29) Zhang, J.-H.; Chung, T. D. Y.; Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screening* **1999**, *4*, 67−73.

(30) McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K. A specific mechanism of nonspecific inhibition. *J. Med. Chem*. **2003**, *46*, 4265−4272.

(31) McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem*. **2002**, *45*, 1712−1722.

(32) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43−53.

(33) 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.

(34) Barril, X.; Hubbard, R. E.; Morley, S. D. Virtual screening in structure-based drug design. *Mini-Rev. Med. Chem*. **2004**, *4*, 779−791.

(35) Doman, T. N.; McGovern, S. L.; Witherbee, B. J.; Kasten, T. P.; Kurumbail, R.; Stallings, W. C.; Connolly, D. T.; Shoichet, B. K. Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J. Med. Chem.* **2002**, *45*, 2213−2221.

(36) Oshiro, C. M.; Kuntz, I. D.; Dixon J. S. Flexible ligand docking using a genetic algorithm. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 113−130.

(37) Howard, M. H.; Cenizal, T.; Gutteridge, S.; Hanna, W. S.; Tao, Y.; Totrov, M.; Wittenbach, V. A.; Zheng, Y. J. A novel class of inhibitors of peptide deformylase discovered through high-throughput screening and virtual ligand screening. *J. Med. Chem.* **2004**, *47*, 6669−6672.

(38) Jenkins, J. L.; Kao, R. Y.; Shapiro, R. Virtual screening to enrich hit lists from high-throughput screening: a case study on small-molecule inhibitors of angiogenin. *Proteins* **2003**, *50*, 81−93.

(39) Paiva, A. M.; Vanderwall, D. E.; Blanchard, J. S.; Kozarich, J. W.; Williamson, J. M.; Kelly, T. M. Inhibitors of dihydrodipicolinate reductase, a key enzyme of the diaminopimelate pathway of *Mycobacterium tuberculosis*. *Biochim. Biophys. Acta* **2001**, *1545*, 67−77.

(40) Kitchen, D. B.; Stahura, F. L.; Bajorath, J. Computational techniques for diversity analysis and compound classification. *Mini-Rev. Med. Chem.* **2004**, *4*, 1029−1039.

(41) Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T.; Murray, C. W.; Taylor, R. D.; Watson, P. J. Virtual screening using protein−ligand docking: avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 793−816.

(42) *Prous Integrity Database*; Prous Science: Barcelona, Spain, 1995−2004; www.prous.com.

(43) Alonso, M.; Martinez, A. GSK-3 Inhibitors: Discoveries and Developments. *Curr. Med. Chem.* **2004**, *11*, 755−763.

(44) Peat, A. J.; Boucheron, J. A.; Dickerson, S. H.; Garrido, D.; Mills, W.; Peckham, J.; Preugschat, F.; Smalley, T.; Schweiker, S. L.; Wilson, J. R.; Wang, T. Y.; Zhou, H. Q.; Thomson, S. A. Novel pyrazolopyrimidine derivatives as GSK-3 inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2121−2125.

(45) Witherington, J.; Bordas, V.; Gaiba, A.; Garton, N. S.; Naylor, A.; Rawlings, A. D.; Slingsby, B. P.; Smith, D. G.; Takle, A. K.; Ward, R. W. 6-Aryl-pyrazolo[3,4-b]pyridines: potent inhibitors of glycogen synthase kinase-3 (GSK-3). *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1577−1580.

(46) Kunick, C.; Lauenroth, K.; Wieking, K.; Xie, X.; Schultz, C.; Gussio, R.; Zaharevitz, D.; Leost, M.;, Meijer, L.; Weber, A.; Jørgensen, F. S.; Lemcke, T. Evaluation and comparison of 3D-QSAR COMSIA models for CDK1, CDK5, and GSK-3 inhibition by paullones. *J. Med. Chem.* **2004**, *47*, 22−36.
(47) Polychronopoulos, P.; Magiatis, P.; Skaltsounis, A.-L.; Myrianthopoulos, V.; Mikros, E.; Tarricone, A.; Musacchio, A.; Roe, S. M.; Pearl, L.; Leost, M.; Greengard, P.; Meijer, L. Structural basis for the synthesisi of indirubins as potent and selective inhibitors of glycogen synthase kinase-3 and cyclin dependent kinase. *J. Med. Chem.* **2004**, *47*, 935−946.
(48) Leclerc, S.; Garnier, M.; Hoessel, R.; Marko, D.; Bibb, J. A.; Snyder, G. L.; Greengard, P.; Biernat, J.; Wu, Y.-Z.; Mandelkow, E.-M.; Eisenbrand, G.; Meijer, L. Indirubins inhibit glycogen synthase kinase-3β and CDK5/P25, two protein kinases involved in abnormal tau phosphorylation in Alzheimer's disease. *J. Biol. Chem.* **2001**, *276*, 251−260.
(49) Mettey, Y.; Gompel, M.; Thomas, V.; Garnier, M.; Leost, M.; Caballos-Picot, I.; Noble, M.; Endicott, J.; Vierfond J.-M.; Meijer, L. Aloisines, a new family of CDK/GSK-3 inhibitors. SAR study, crystal structure in complex with CDK2, enzyme selectivity, and cellular effects. *J. Med. Chem.* **2003**, *46*, 222−236.
(50) *SYBYL 6.9.2*; Tripos Inc.: St. Louis, MO, 2004.

JM050504D