Identification of Compounds with Nanomolar Binding Affinity for Checkpoint Kinase-1 Using Knowledge-Based Virtual Screening

Paul D. Lyne,*,† Peter W. Kenny,‡ David A. Cosgrove,‡ Chun Deng,† Sonya Zabludoff,† John J. Wendoloski,† and Susan Ashwell†

Cancer Discovery, AstraZeneca R&D Boston, Waltham, Massachusetts 02451, and Cancer Discovery, AstraZeneca, Alderley Park, Macclesfield, Cheshire, SK10 4TG, U.K.

Received October 7, 2003

A virtual screen of a subsection of the AstraZeneca compound collection was performed for checkpoint kinase-1 (Chk-1 kinase) using a knowledge-based strategy. This involved initial filtering of the compound collection by application of generic physical properties followed by removal of compounds with undesirable chemical functionality. Subsequently, a 3-D pharmacophore screen for compounds with kinase binding motifs was applied. A database of approximately 200K compounds remained for docking into the active site of Chk-1 kinase, using the FlexX-Pharm program. For each compound that docked successfully into the binding site, up to 100 poses were saved. These poses were then postfiltered using a customized consensus scoring scheme for a kinase, followed by visual inspection of a selection of the docked compounds. This resulted in 103 compounds being ordered for testing in the project assay, and 36 of these (corresponding to four chemical classes) were found to inhibit the enzyme in a dose−response fashion with IC_{50} values ranging from 110 nM to 68 μM.

Introduction

Many of the proteins involved in the regulation of the cell cycle are interesting targets for the development of therapeutics for cancer.[1] In particular, DNA damage checkpoint kinases have recently emerged as attractive targets for cancer therapy.[2] The checkpoint kinase-1 (Chk-1 kinase) prevents cells with damaged DNA from entering mitosis by arresting them at the G2/M checkpoint. DNA damage activates the ATM/ATR pathway, allowing the cells to arrest and repair their DNA. ATR kinase activates Chk-1 kinase, which then phosphorylates the phosphatase Cdc25C, which in turn inactivates cdc2. Inactivation of cdc2 leads to arrest at the G2/M checkpoint.[3] It has therefore been suggested that inhibition of Chk-1 kinase would lead to abrogation of the G2/M checkpoint, allowing cells to enter a lethal mitosis with damaged DNA. Recently, the Chk-1 inhibitors UCN-01 and SB-218078, staurosporine analogues, have been reported to abrogate the G2/M arrest induced by DNA-damaging agents in vivo and thus to enhance the cytotoxicities of the DNA-damaging agents.[4,5] Therefore, Chk-1 kinase inhibitors are potential sensitizing agents to be used in combination with standard therapies that induce DNA damage. In this paper a virtual screening strategy that was used to identify inhibitors of the Chk-1 kinase is described.

Virtual screening is a technology that is gaining increased use in drug discovery.[6−9] It is seen as a complementary approach to experimental screening (high-throughput screening) and, when coupled with structural biology, promises to increase the number and enhance the success of projects in the lead identification stage of the discovery process. Virtual screening involves analyzing large collections of compounds with computational methods with a view to identifying a small subset of compounds either for testing in an assay or for synthesis and eventual biological testing. Structure-based virtual screening involves explicit molecular docking of each ligand into the binding site of the target, producing a predicted binding mode for each database compound together with a measure of the quality of the fit of the compound in the target binding site. This information is then used to rank the compounds with a view to selecting and testing experimentally a small subset for biological activity.

A number of studies aimed at validating approaches to virtual screening have been published. These have focused on an evaluation of the docking methodologies employed in virtual screening experiments with respect to reproduction of known crystal structure binding modes or to enrichment of hit lists from databases of decoys.[10−13] In addition, there have been some examples published of successful applications of structure-based virtual screening in a drug discovery project.[14−17]

There are many challenges associated with the technology employed in virtual screening. The process of docking large databases containing on the order of 10^5−10^6 compounds dictates that a number of approximations in describing the nature of the interactions between a protein and a ligand need to be made in order for the virtual screen to be completed in a timely manner. Approximations in the scoring functions employed,[7,18] the lack of a proper treatment of solvation, the neglect of protein flexibility, poor assessment of the protonation states of active site residues or ligands, and occasionally (admittedly, rarely) incorrect or ambiguous X-ray assignment of protein side chains (such as rotamer positions or the orientations of an amide side chain of Gln or Asn) are all potential sources of error in the

* To whom correspondence should be addressed. E-mail: paul.lyne@astrazeneca.com. Phone: +1 (781) 839 4808. Fax: +1 (781) 839 4357.
† AstraZeneca R&D Boston.
‡ AstraZeneca, Alderley Park.

10.1021/jm030504i CCC: $27.50
Published on Web 03/13/2004

docking process and contribute to the identification of false positives and false negatives in virtual screening. Virtual screening, however, can avail itself of additional knowledge about a specific receptor and use this information to guide the docking and selection process.

Protein kinases are currently a source of many targets of interest to the pharmaceutical industry[19,20] and are attractive targets for virtual screening because of the wealth of structural and binding information available for protein−inhibitor complexes.[21] In this paper a strategy for a virtual screen of Chk-1 kinase is described that exploits the structural and chemical information available for kinase inhibitors. The strategy was to dock only those compounds in the corporate database that have a minimal kinase binding motif and to guide the docking to specific hot spots common to all kinase ATP binding sites. Subsequently, the docked compounds were rescored using a scoring scheme customized for a kinase, and finally compounds were selected for screening on the basis of the rescored list and visual inspection of the docked poses. The approach proved to be very successful with 36 of 103 compounds tested found to inhibit the enzyme in a dose−response fashion with IC_{50} values ranging from 110 nM to 68 μM (a 35% hit rate). The hits found by the virtual screen corresponded to four chemical classes, with each class deemed attractive for further investigation by the medicinal chemistry team.

Methods

Database Preparation. The original source of all the compounds considered was in SMILES format.[22] Initially, the database was prefiltered to remove compounds with molecular weight greater than 600 Da or with greater than 10 rotatable bonds. Protonation and tautomeric states were generated with Leatherface, an in-house molecular editor coded using the Daylight programming tool kits. Leatherface modifies molecular connection tables according to rules specified in SMARTS notation.[23] Leatherface is also capable of enumerating forms that are appropriate for representing relatively unbiased equilibria. A 3-D version of the database was then generated using Corina,[24,25] with explicit enumeration of stereocenters (generating a maximum of eight stereoisomers per molecule). A conformational version of the database was then generated using the program Omega,[26] with default parameters. A primary virtual screen was applied to the database using Plurality, an in-house 3-D pharmacophore screening program.

Plurality allows the specification of a 3-D pharmacophore query in terms of combinations of SMARTS targets such as hydrogen-bond donors and acceptors, ring centroids, etc. A multiconformational database, such as those produced by Omega, may be searched to find compounds matching the pharmacophore, with the option of the hits being overlaid onto the query pharmacophore. It may be run in parallel over a large computer network with very short search times (for example, a simple four-point pharmacophore search on 150 million conformations takes approximately 15 min on a 50 processor Linux farm).

Plurality was used in this instance to filter out compounds that did not have an appropriate kinase binding motif. This motif was defined as a hydrogen bond donor and acceptor pair with a distance range of 1.35−2.40 Å. It is well-known that there are examples of kinase inhibitors that can bind to the adenine binding subsite by forming one, two, or three hydrogen bonds. A two-point pharmacophore was employed here because it was felt that a one-point pharmacophore (requiring molecules to have a hydrogen bond acceptor) was far too lenient and that a three-point pharmacophore (donor−acceptor−donor) was too restrictive. This left approximately 200K compounds (single conformers per molecule) to dock into the ATP binding site of Chk-1 kinase.

Protein Preparation. The crystal structure of the Chk-1 kinase domain was used. The protonation states of residues in the binding site were adjusted to the dominant ionic forms at pH 7.4. The crystallographic waters were removed, and the bound ligand was used to define the active site for the docking run, with residues within 6.5 Å of the bound inhibitor included in the active site definition. The bound inhibitor was not included in the docking run.

Docking Protocol. A guided docking strategy was employed using the FlexX-Pharm[27] program to exploit the large volume of internally and publicly available data on kinase inhibition. FlexX-Pharm employs a docking algorithm that takes account of ligand flexibility but keeps the protein rigid. On the basis of several X-ray structures of protein kinases, it is known that kinases adopt a common fold and that the structure of the ATP binding pocket is highly conserved. The ATP binding pocket of kinases may be considered to comprise five specific subsites (see Figure 1).[28,29] One subsite, the adenine binding region, recognizes the purine group of ATP through the formation of hydrogen bonds to the N1 and N6 atoms of adenine. In almost all known complexes of kinases bound to ATP competitive inhibitors, the inhibitors mimic the purine of ATP by binding to the adenine binding subsite of the kinase. For the purposes of the virtual screen, this was considered an essential interaction, and the docking of the compounds in the database was biased to ensure that these hydrogen-bonding interactions were satisfied by specifying that an interaction be made with the backbone NH of Cys87 and the backbone carbonyl of Glu85. Other subsites found in the kinase active site include the phosphate binding site, the sugar binding site, the selectivity pocket, and the solvent channel. A study of protein structures of kinase−inhibitor complexes reveals that usually one of these subsites interacts with the inhibitor in addition to the interactions with the adenine binding region. To account for this, the docking of compounds was guided by having optional interactions with the side chains of Lys33, Tyr86, Glu91, Asp94, or Ser147. In each case, a maximum of 100 poses were saved for each docked compound, although typically many fewer were saved because biased sampling was used in the docking process.

Postprocessing and Compound Selection. The poses for each docked compound were rescored according to a consensus scoring scheme derived from an enrichment study for Cdk-2. A popular strategy for postprocessing the results of a virtual screen is to use the concept of consensus scoring.[30] In this approach, the poses generated from a docking program are then rescored using multiple scoring functions. Only com-

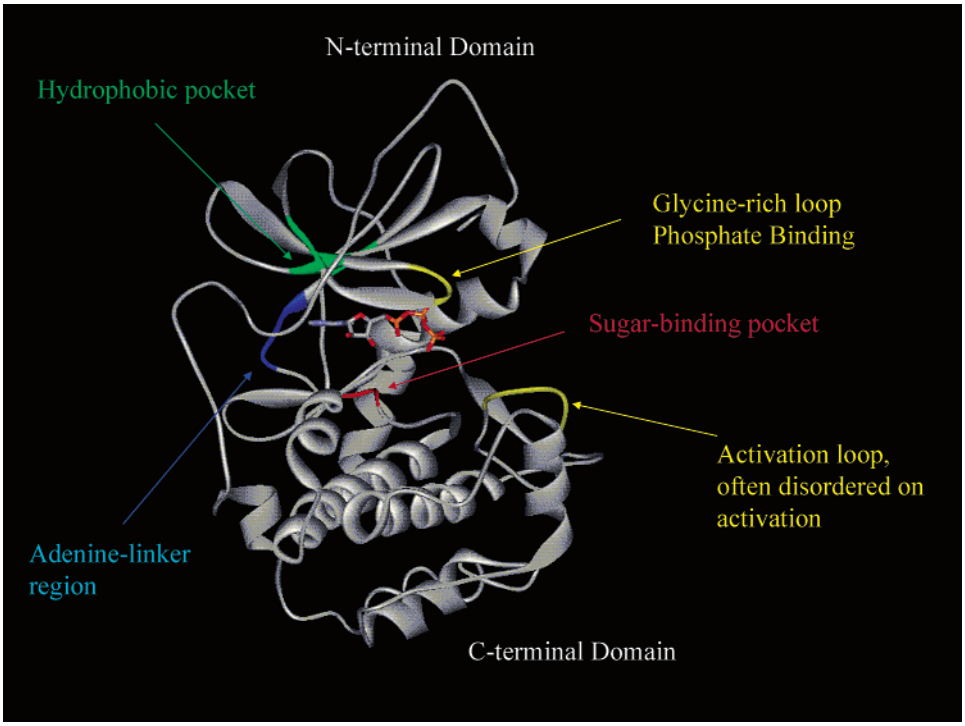

Figure 1. Structure of a kinase domain highlighting the five subsites present in the ATP binding site.

pounds that score well with each scoring function (the consensus) are considered further for biological testing. Retrospective virtual screening studies have shown that this approach can improve the enrichment of true hits.[10,11,30] The exact nature of the consensus scoring scheme used is dependent on the target being considered.[7] For the purposes of the virtual screen against Chk-1 in this study, Cdk-2 was chosen as a surrogate for Chk-1 in determining the consensus scoring scheme because at the time there was a much richer set of activity data for the Cdk-2 project available. A database of 8000 compounds containing 100 known Cdk-2 inhibitors were docked into the active site of Cdk-2 with FlexX. For the purposes of rescoring, up to 300 poses per molecule were saved. Typically docking methods identify the correct experimental pose as the top pose on average 50% of the time. This is based on published docking evaluations by several groups using many different docking methods against a wide range of protein structures.[12,13,31] There is no systematic way of telling, for any specific situation, if the correct experimental pose will be found in the top 10, top 50, or top 100 of the poses generated by a docking program. To offset this uncertainty, it was decided to keep 300 docked poses for each molecule in the hope that the correct pose would exist among that set.[11] The saved 300 poses per docked compound were rescored using the Cscore module of Tripos. The Cscore module includes versions of the FlexX,[31] Dock (electrostatic),[32] Gold,[12] PMF,[33] and ChemScore[34] scoring functions. All possible combinations of these scoring functions were used to determine enrichment curves for Cdk-2.

For any combination of scoring functions, the consensus score of a pose, Z, was defined as

$$Z = \sum_i z_i$$

where z_i is the normal deviate score z for a pose with a specific scoring function i. z is defined as

$$z = \frac{x - \bar{x}}{\sigma}$$

x is the raw score (for example, the Chemscore or Goldscore), $\bar{x}$ is the average raw score for all the poses, and σ is the standard deviation of the raw scores for all the poses. It is assumed that for each individual scoring function the raw scores for all the poses are approximated by a normal distribution. The enrichment curves found for all the consensus schemes are shown in Figure 2. For the case of Cdk-2, it was determined that the consensus scoring combination of the PMF and FlexX scoring functions gave the best enrichment curves. The compounds docked into Chk-1 kinase were rescored with the PMF and FlexX consensus score, and the distribution of scores is shown in Figure 3. Also indicated on the distribution are the scores of known micromolar inhibitors of Chk-1 that were used to guide the cutoff employed for selecting compounds for visual inspection. On the basis of the observed distribution and the locations of the known compounds, a cutoff of 2.5 was used, giving approximately 250 compounds for visual inspection. The visual inspection step identified compounds that exhibited unfavorable interactions with the binding site or compounds that adopted unrealistic conformations when docked into the active site. Removal of those compounds resulted in a final list of 103 compounds being selected for testing.

Chk-1 Kinase Assay. Chosen compounds were profiled, and IC_{50} values were determined in a scintillation proximity assay using purified human Chk-1 and a biotinylated cdc25c peptide substrate. In brief, N-terminal GST-Chk-1 protein was expressed in baculoviral infected insect cells and purified using glutathione sepharose resin. Purified protein and peptide in HEPES-based buffer were added to a 384-well plate containing

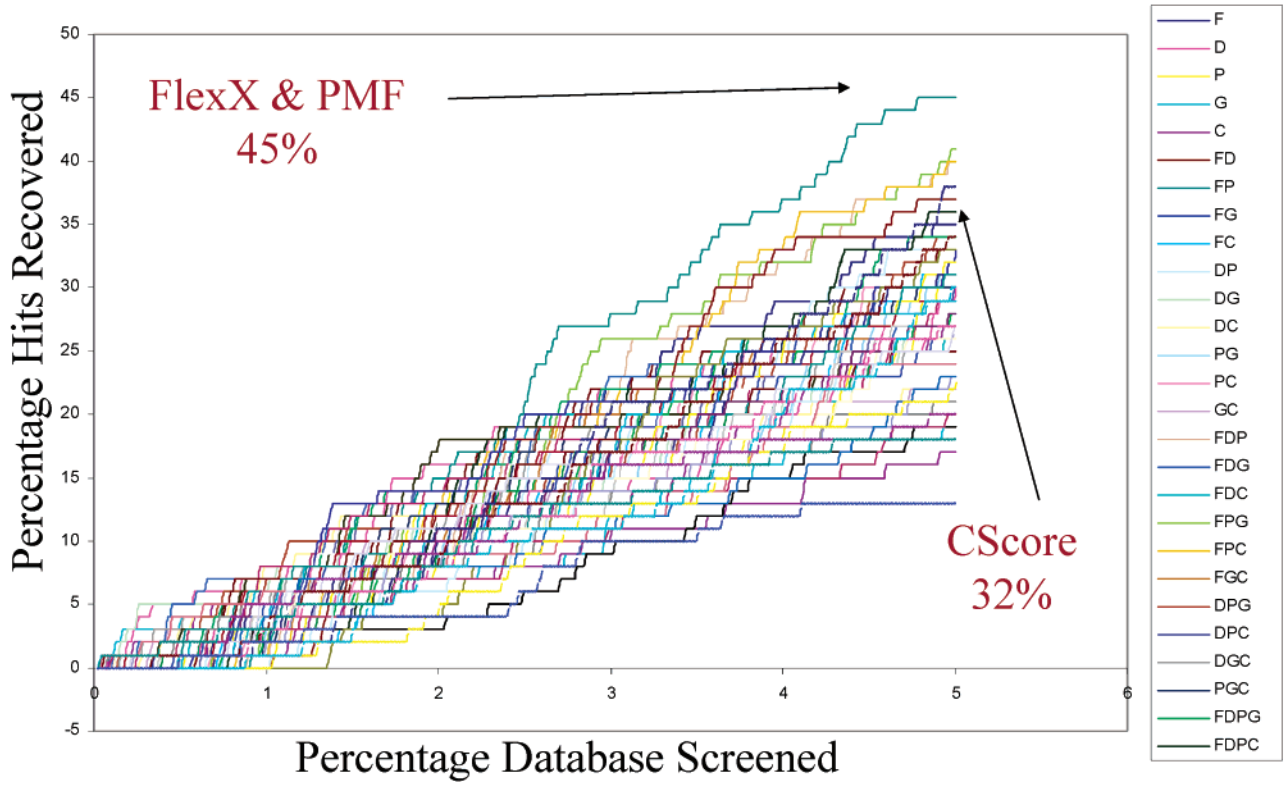

Figure 2. Enrichment curves for Cdk-2 using a number of consensus scoring schemes: P = PMF; D = Dockscore; C = Chemscore; F = FlexXscore; G = Goldscore.

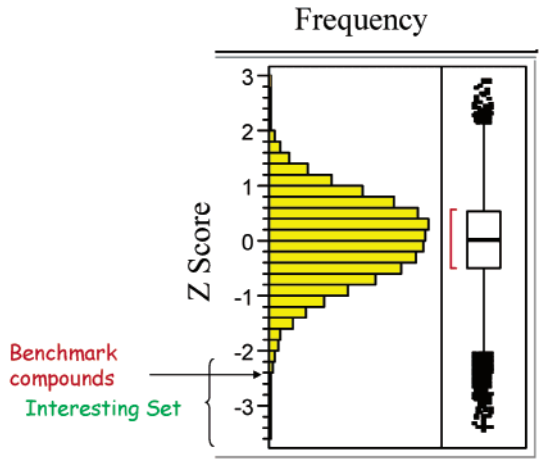

Figure 3. Frequency distribution of consensus *Z* scores for compounds docked into Chk-1 kinase. The arrows indicate the scores of known Chk-1 inhibitors ($Z \cong -2.2$), and the solid line indicates the cutoff ($Z = -2.0$) chosen for visual inspection.

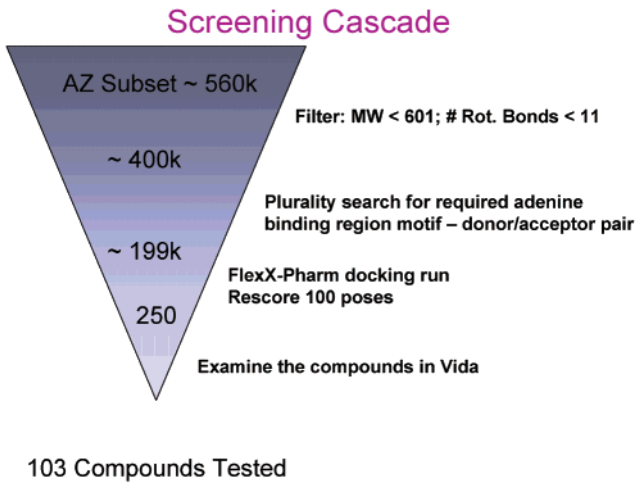

Figure 4. Cartoon illustrating the tiered approach taken to the virtual screen of a subset of the AZ corporate database for Chk-1 kinase inhibitory activity.

diluted compounds (final eight-point concentration range from 100 μM to 10 nM). Reactions were initiated by adding a mix of nonradiolabeled ATP and [^{33}P]ATP. Plates were then incubated for 2 h at room temperature, and reactions were terminated by addition of EDTA, SPA beads, and CsCl. Plates were then read in a Topcount NXT microplate scintillation counter. IC_{50} values were determined using XLFit within Activity Base.

Results

A tiered approach for virtual screening (Figure 4) was employed with the aim of enriching subsets with kinase-like inhibitors initially and with specific Chk-1 kinase inhibitors at later stages of the process. The approach capitalized on a rich source of structural information, available publicly and internally, on kinase−inhibitor complexes and on a scoring scheme that has been demonstrated to give superior enrichments for a kinase (Cdk-2) in pilot virtual screening studies. The virtual screen successfully enriched with hits the small subset of the corporate database that was selected for testing in the project assay, with 36 hits out of 103 screened ranging in potency (IC_{50} data) from 110 nM to 68 μM. In general, in drug discovery projects attrition rates in the lead optimization process due to poor cellular efficacy, poor DMPK profiles, or poor physical properties of the leads are particular challenges that project teams face. As a consequence, it is desirable to have several

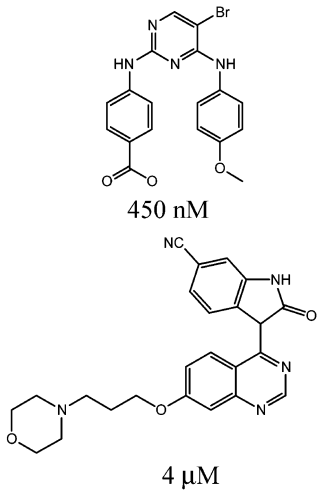

Figure 5. Examples of the hits found by the virtual screen. The activities reported are dose−response data, measured in triplicate.

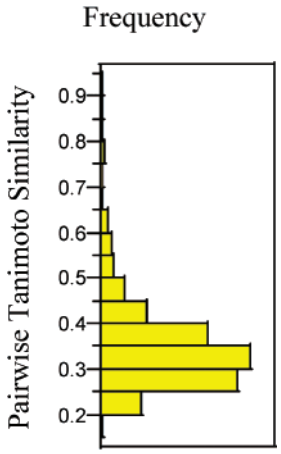

Figure 6. Frequency distribution of the pairwise Tanimoto similarities of the 36 active hits using Daylight fingerprints.

chemical series represented in the hits that are carried forward by the project team. The hits found by the virtual screen correspond to four chemical classes (some examples are shown in Figure 5), each deemed by the medicinal chemistry team to have desirable leadlike properties for further synthetic investigation as lead series, and they have been shown to be ATP-competitive. A frequency distribution of the pairwise Tanimoto similarities of the hits is shown in Figure 6. The mean pairwise Tanimoto similarity for the active compounds is 0.35 with a standard deviation of 0.12.

Discussion

Molecular docking programs face several challenges in trying to predict correctly the binding mode of a protein−inhibitor complex and to predict the relative affinities of a set of compounds for a specific target. The docking problem comprises sampling of the binding site, sampling of the conformational space of the ligand, and a prediction of the interaction energy of every pose found for a given ligand in the binding site. The form of the scoring function employed is necessarily approximate, and as a consequence, docking programs are often not very accurate at reproducing the correct binding modes or predicting the relative affinities of inhibitors for a specific target.[7,18,35] In general, it has been observed in several in-house projects that many of the standard docking methods perform particularly poorly for kinase targets. Usually the top pose identified does not involve interactions with the adenine binding region but is often drawn toward the phosphate or sugar binding pockets. The ATP binding site of kinases is highly conserved with specific residues in the binding site being conserved absolutely across the whole gene family. The binding site is predominantly hydrophobic in nature, but it does have charged or highly polar regions. One charged region that is invariable across the whole gene family is a lysine (Lys33 in Chk-1) that binds the γ-phosphate of ATP. The sugar binding region is polar in nature and often is lined by the acidic residues Asp or Glu. These highly charged regions in the binding site of kinases are often the source of error in docking methods that are guided by optimizing polar interactions between a ligand and the protein or that have simplistic treatments of electrostatic interactions in their scoring functions. This results in poses biased toward this region of the binding site, without making interactions to the adenine-binding region.

Almost all known kinase−inhibitor crystal structures demonstrate that the adenine-binding region is a key molecular recognition motif in the binding sites of protein kinases. As a consequence, it is reasonable to impose a constraint on the docking procedure in the virtual screening to diminish the deficiencies in the scoring function employed. In general, applying constraints to a specific subsite of a protein binding site might result in certain compounds being misclassified as not being inhibitors, but in the specific case of applying a constraint to the adenine-binding region of kinase targets, this risk is small.

Similarly, since the bias of the virtual screen is to identify inhibitors of Chk-1 kinase that bind in the ATP pocket while making interactions to the adenine-binding region, there is no need to dock the full corporate collection because not every compound will have the necessary chemical functionality to interact with the adenine-binding region. By employment of a modest pharmacophore filter for a kinase binding motif, the size of the database was reduced by half to the chemistry space known to be relevant to kinase inhibition. This rational approach saved time in the docking and enabled more poses to be saved for the postscoring stage of the screen.

Taking the poses generated from the docking stage of a virtual screen and rescoring them according to a consensus scoring scheme is a common tactic employed in virtual screening.[10,30] The philosophy behind this approach is to use several different types of scoring function to arrive at a consensus of which compounds bind well to the protein, in an attempt to overcome the well-known deficiencies associated with individual scoring functions.[18] Typically, the scoring functions employed in the consensus scoring scheme are representatives of the three main classes of scoring function, namely, empirical-based scoring functions (Goldscore,[12] Chemscore,[34] FlexX,[31] Ludi,[37] PLP[38]), force-field-based scoring functions (Dock Energy Score),[32] or knowledge-based scoring functions (PMF,[33] Drugscore,[39] Bleep[40]). Usually the consensus uses five or six scoring functions

Table 1. Distribution of Selected Pose (Based on Postscoring Analysis) of the 36 Active Compounds in the Ranked Poses Generated by FlexX-Pharm[a]

postion in docked list	no. of actives
top 3	5
top 20	22
top 50	32
top 100	36

[a] Only five of the actives identified were ranked in the top three poses by FlexX-Pharm alone.

together, as implemented in the Cscore module provided by Tripos, for example. However, it is more prudent to identify which combination of scoring functions used as a consensus performs best for the specific system at hand because it is conceivable that the standard combination of several scoring functions in the consensus may not give the best enrichment. As described in the Methods, the optimal consensus scoring scheme was identified for a related cell cycle kinase, Cdk-2, and this was then applied successfully to rescoring the docking poses generated for Chk-1.

The data presented in Table 1 illustrate that the identification of true-positive hit compounds was dependent on retaining a large number of docked poses and rescoring these on the basis of the consensus scoring scheme outlined above. For the 36 active compounds found by the virtual screening, only five were found in the top three poses generated by FlexX-Pharm alone (without postscoring analysis). These are the results for this specific study, and it is not anticipated that the exact distribution of poses and active hits found here is what would be found in general with virtual screening. However, this and related studies[11] do suggest that the strategy employed here may be useful in other virtual screening campaigns.

In conclusion, a successful virtual screen of the Chk-1 kinase was performed, capitalizing on in-house experience and publicly available information in the field of protein kinase inhibition with ATP competitive compounds. A hit rate of 36% was achieved by screening 103 compounds. The hits corresponded to four chemical series with correct modes of action (ATP competitive), and in some instances, the binding mode has been confirmed by X-ray crystallography. The knowledge-based virtual screening strategy employed here could in principle be applied to the identification of ATP competitive inhibitors of other kinases.

Acknowledgment. The authors thank Michelle Lamb, Tim Perkins, and Dave Timms for helpful comments on the manuscript.

References

(1) Hanahan, D.; Weinberg, R. A. The Hallmarks of Cancer. *Cell* **2000**, *100*, 57−70.
(2) Zhou, B.-B. S.; Elledge, S. J. The DNA Damage Response: Putting Checkpoints in Perspective. *Nature* **2000**, *408*, 433−439.
(3) Sanchez, Y.; Wong, C.; Thoma, R. S.; Richman, R.; Wu, Z.; et al. Conservation of the DNA Checkpoint Pathway in Mammals: Linkage of DNA Damage to Cdk Regulation Through Cdc25. *Science* **1997**, *277*, 1497−1501.
(4) Graves, P. R.; Yu, L.; Schwarz, J. K.; Gales, J.; Sausville, E. A.; et al. The Chk1 protein kinase and the Cdc25C regulatory pathways are targets of the abticancer agent UCN-01. *J. Biol. Chem.* **2000**, *275*, 5600−5605.
(5) Jackson, J. R.; Gilmartin, A.; Imburgia, C.; Winkler, J. D.; Marshall, L. A.; et al. An indolocarbazole inhibitor of human checkpoint kinase (Chk1) abrogates cell cycle arrest caused by DNA damage. *Cancer Res.* **2000**, *60*, 566−572.
(6) Bajorath, J. Integration of virtual and high throuput screening. *Nat. Rev. Drug Discovery* **2002**, *1*, 882−894.
(7) Lyne, P. D. Structure-based virtual screening: an overview. *Drug Discovery Today* **2002**, *7*, 1047−1055.
(8) Schneider, G.; Bohm, H. J. Virtual screening and fast automated docking methods. *Drug Discovery Today* **2002**, *7*, 64−70.
(9) Abagyan, R.; Totrov, M. High-throughput docking for lead generation. *Curr. Opin. Chem. Biol.* **2001**, *5*, 375−382.
(10) Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43*, 4759−4767.
(11) Stahl, M.; Rarey, M. Detailed analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.
(12) 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.
(13) Kramer, B.; Rarey, M.; Lengauer, T. Evaluation of the FlexX incremental construction algorithm for protein ligand docking. *Proteins: Struct. Funct., Genet.* **1999**, *37*, 228−241.
(14) Doman, T. N.; McGovern, S. L.; Witherbee, B. J.; Kasten, T. P.; Kurumbail, R.; et al. Molecular docking and high throughput screening for novel inhibitors of protein tyrosine phosphatase 1B. *J. Med. Chem.* **2002**, *45*, 2213−2221.
(15) Gruneberg, S.; Stubbs, M. T.; Klebe, G. Successful virtual screening for novel inhibitors of human carbonic anhydrase: Strategy and experimental confirmation. *J. Med. Chem.* **2002**, *45*, 3588−3602.
(16) Perola, E.; Xu, K.; Kollmeyer, T. M.; Kaufmann, S. H.; Prendergast, F. G.; et al. Successful virtual screening of a chemical database for farnesyltransferase inhibitor leads. *J. Med. Chem.* **2000**, *43*, 401−408.
(17) Filikov, A. V.; Mohan, V.; Vickers, T. A.; Griffey, R. H.; Cook, P. D.; et al. Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 593−610.
(18) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. Principles of docking: an overview of search algorithms and a guide to scoring functions. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 409−443.
(19) Cohen, P. Protein kinases, the major drug targets of the twenty-first century? *Nat. Rev. Drug Discovery* **2002**, *1*, 309−315.
(20) Blume-Jensen, P.; Hunter, T. Oncogenic kinase signalling. *Nature* **2001**, *411*, 335−365.
(21) Hubbard, S. R. Protein tyrosine kinases: autoregulation and small molecule inhibition. *Curr. Opin. Struct. Biol.* **2002**, *12*, 735−741.
(22) *SMILES*; Daylight Chemical Information Systems Inc.: Santa Fe, NM, Vol. 471.
(23) *SMARTS*; Daylight Chemical Information Systems: Santa Fe, NM, Vol. 471.
(24) Gasteiger, J.; Rudolph, C.; Sadowski, J. Automatic generation of 3-D atomic coordinates for organic molecules. *Tetrahedron Comput. Methodol.* **1990**, *3*, 537−547.
(25) Networks, M. *Corina*; Gasteiger Research: Erlangen, Germany.
(26) *Omega*; Openeye Science Software: Sante Fe, NM, Vol. 1.0.
(27) Hindle, S. A.; Rarey, M.; Buning, C.; Lengauer, T. Flexible docking under pharmacophore type constraints. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 129−149.
(28) Traxler, P.; Furet, P. Strategies toward the design of novel and selective tyrosine kinase inhibitors. *Pharmacol. Ther.* **1999**, *82*, 195−206.
(29) Toledo, L. M.; Lydon, N. B.; Elbaum, D. The structure-based design of ATP-site directed protein kinase inhibitors. *Curr. Med. Chem.* **1999**, *6*, 775−805.
(30) Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* **1999**, *42*, 5100−5109.
(31) 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.
(32) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 411−428.
(33) Muegge, I. A knowledge-based scoring function for protein−ligand interactions: Probing the reference state. *Perspect. Drug Discovery Des.* **2000**, *20*, 99−114.
(34) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions. 1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.
(35) Pearlman, D. A.; Charifson, P. S. Are free energy calculations useful in practice? A comparison with rapid scoring functions

for the p38 MAP kinase protein system. *J. Med. Chem.* **2001**, *44*, 3417−3423.
(36) Hubbard, S. R.; Till, J. H. Protein tyrosine kinase structure and function. *Ann. Rev. Biochem.* **2000**, *69*, 373−398.
(37) Boehm, H.-J.; Boehringer, M.; Bur, D.; Gmuender, H.; Huber, W.; et al. Novel inhibitors of DNA Gyrase: 3-D structure based biased needle screening. Hit validation by biophysical methods, and 3-D guided optimization. A promising alternative to random screening. *J. Med. Chem.* **2000**, *43*, 2664−2674.
(38) Gehlhaar, D. K.; Verkhirker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. B.; Fogel, L. J.; Freer, S. T. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming *Chem. Biol.* **1995**, *2*, 317−324.
(39) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based scoring function to predict protein−ligand interactions. *J. Mol. Biol.* **2000**, *295*, 337−356.
(40) Mitchell, J. B. O.; Laskowski, R. A.; Alex, A.; Forster, M. J.; Thornton, J. M. Bleep−potential of mean force describing protein−ligand interactions: II. Calculation of binding energies and comparison with experimental data. *J. Comput. Chem.* **1999**, *20*, 1177−1185.

JM030504I