



Modified AutoDock for accurate docking of protein kinase inhibitors

Oleksandr V. Buzko*, Anthony C. Bishop[†] & Kevan M. Shokat

Department of Cellular and Molecular Pharmacology, University of California, San Francisco, 513 Parnassus Ave, San Francisco, CA 94143-0450, U.S.A. [†]Current address: Skaggs Institute for Chemical Biology, Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, U.S.A.

Received 23 November 2001; Accepted 12 April 2002

Key words: approximation of solvation, AutoDock, force field, hydrogen bonding, inhibitor design, scoring function

Summary

Protein kinases are an important class of enzymes controlling virtually all cellular signaling pathways. Consequently, selective inhibitors of protein kinases have attracted significant interest as potential new drugs for many diseases. Computational methods, including molecular docking, have increasingly been used in the inhibitor design process [1]. We have considered several docking packages in order to strengthen our kinase inhibitor work with computational capabilities. In our experience, AutoDock offered a reasonable combination of accuracy and speed, as opposed to methods that specialize either in fast database searches or detailed and computationally intensive calculations.

However, AutoDock did not perform well in cases where extensive hydrophobic contacts were involved, such as docking of SB203580 to its target protein kinase p38. Another shortcoming was a hydrogen bonding energy function, which underestimated the attraction component and, thus, did not allow for sufficiently accurate modeling of the key hydrogen bonds in the kinase-inhibitor complexes.

We have modified the parameter set used to model hydrogen bonds, which increased the accuracy of AutoDock and appeared to be generally applicable to many kinase-inhibitor pairs without customization. Binding to largely hydrophobic sites, such as the active site of p38, was significantly improved by introducing a correction factor selectively affecting only carbon and hydrogen energy grids, thus, providing an effective, although approximate, treatment of solvation.

Introduction

Protein kinases represent one of the largest groups of proteins in the human genome with over 500 kinases identified to date [2–5]. These enzymes play a major role in eukaryotic signal transduction, via regulation of the phosphorylation states and, thus, cellular functions of substrate proteins. The use of ATP as the common phosphodonor substrate imposes similar structural requirements on all kinase active sites thus resulting in conserved ATP binding sites across the entire superfamily. Design of potent and selective kinase inhibitors has been challenging because the majority

of inhibitors target the conserved ATP binding pocket [6, 7].

Recently, our laboratory has circumvented this problem by engineering a single kinase to contain an active site pocket not present in any wild-type protein kinase [8, 9]. The engineered kinase can be specifically inhibited by molecules designed to complement the newly introduced active site pocket. This chemical-genetic approach affords highly specific kinase-inhibitor pairs and has been successfully used to study a wide range of protein kinases [10–14].

In connection with our efforts to design highly selective inhibitors of mutant protein kinases, we have explored computational approaches for predicting the geometry of kinase-inhibitor complexes, as well as

*To whom correspondence should be addressed. E-mail: buzko@cmp.ucsf.edu

evaluating the relative affinity of different compounds. We report here the optimization of the docking suite AutoDock for modeling of ADP and four classes of kinase inhibitors bound to a number of protein kinases from different subfamilies.

Selection of a docking method

A number of docking algorithms have been made publicly available over the past several years, including those based on shape matching (DOCK [15, 16]), simulated annealing (AutoDock [17, 18]), genetic algorithm (GOLD [19, 20]), tabu search [21] and combinations of the above [19]. We were interested in testing a limited number of known kinase inhibitor scaffolds, thus, a fast docking search of a structure database was not one of the requirements. Such fast searches are performed by DOCK developed by Kuntz and coworkers [16]. Unfortunately, the speed of such analysis does not allow high accuracy, which is required for modeling closely related inhibitors. GOLD, on the other hand, relies heavily on networks of hydrogen bonds within the complex, which, effectively, eliminates hydrophobic sites from being evaluated with this software [20].

Simulated annealing offers modeling of the physical process of ligand binding guided by an energy gradient along the path of approach. As implemented in AutoDock 2.4 [18], this algorithm translates the ligand from an arbitrary point in space into the protein binding site through a series of steps of translation and rotation. The ligand is treated as a flexible entity, while the protein target remains rigid. Rigidity of the protein is clearly a limitation of the method. Since protein conformation can change upon ligand binding, a crystal structure of enzyme-inhibitor complex, rather than that of the free protein, is used for docking putative inhibitors to enzyme mutants.

Scoring is performed through the use of energy grids that are pre-calculated for the entire area of interest, usually centered on the binding site. Each grid point stores an energy of interaction of the corresponding atom type with the rest of the protein. Calculation of grids is performed only once for each protein target, which eliminates the need to compute interaction energies at every step of the simulation. Several docking runs are performed independently of each other with different initial positions, in order to arrive at the most statistically significant result. Evaluation of the binding energy is based on a set of interactions, including

van der Waals dispersion forces, electrostatics and hydrogen bonding. We used the scoring function built into AutoDock 2.4 to study several kinase-inhibitor complexes and, in most cases, it offered an adequate description of binding.

Results and discussion

Docking ADP to c-Src

The Src family of protein tyrosine kinases has been used by our laboratory as a design template [9, 11, 12, 22, 23]. The structures of two Src family members, c-Src and HCK, with ADP and inhibitors bound have been solved by X-ray diffraction [24, 25].

The first question we asked was whether the standard AutoDock package was capable of accurately reproducing crystallographically determined substrate and inhibitor conformations. As an initial test, we simulated binding of ADP to c-Src. Each experiment consisted of 15 to 20 independent runs with starting positions randomly selected by the software. Initial annealing temperature was set at 610 RT, temperature reduction factor 0.95 per cycle, with a total of 120 cooling cycles and a maximum of 12 000 accepted and 12 000 rejected steps. Each cooling cycle started with the minimum state found during the previous cycle.

AutoDock generated docked structures of the complex wherein ADP occupied a position similar to that derived from crystallographic data (Figure 1A). The extent of similarity was quantified using the value of root-mean square deviation (RMSD) of the docked structure from the experimental position. The value of RMSD ranged from 1.5 to 2 Å. Most of this error was due to a shift of the purine ring and displacement of the diphosphate moiety. A hydrogen bond between the N⁶ amino hydrogen of ADP and a protein backbone carbonyl oxygen (E339) was not reproduced (Figure 1A). This particular hydrogen bond is found in all nucleotide triphosphate-kinase co-crystal structures [24, 26, 27].

In an effort to correct these inconsistencies, we introduced several changes to the hydrogen bonding potential and added a 'polar hydrogen' atom type. The corresponding potential function (1) [28] has been augmented by coefficients that define a deeper potential energy well and a slightly greater equilibrium distance between the donor and the acceptor (2.1 Å and 7 kcal/mol vs. the original settings of 1.9 Å and 5 kcal/mol for the case of OH as the donor). The new

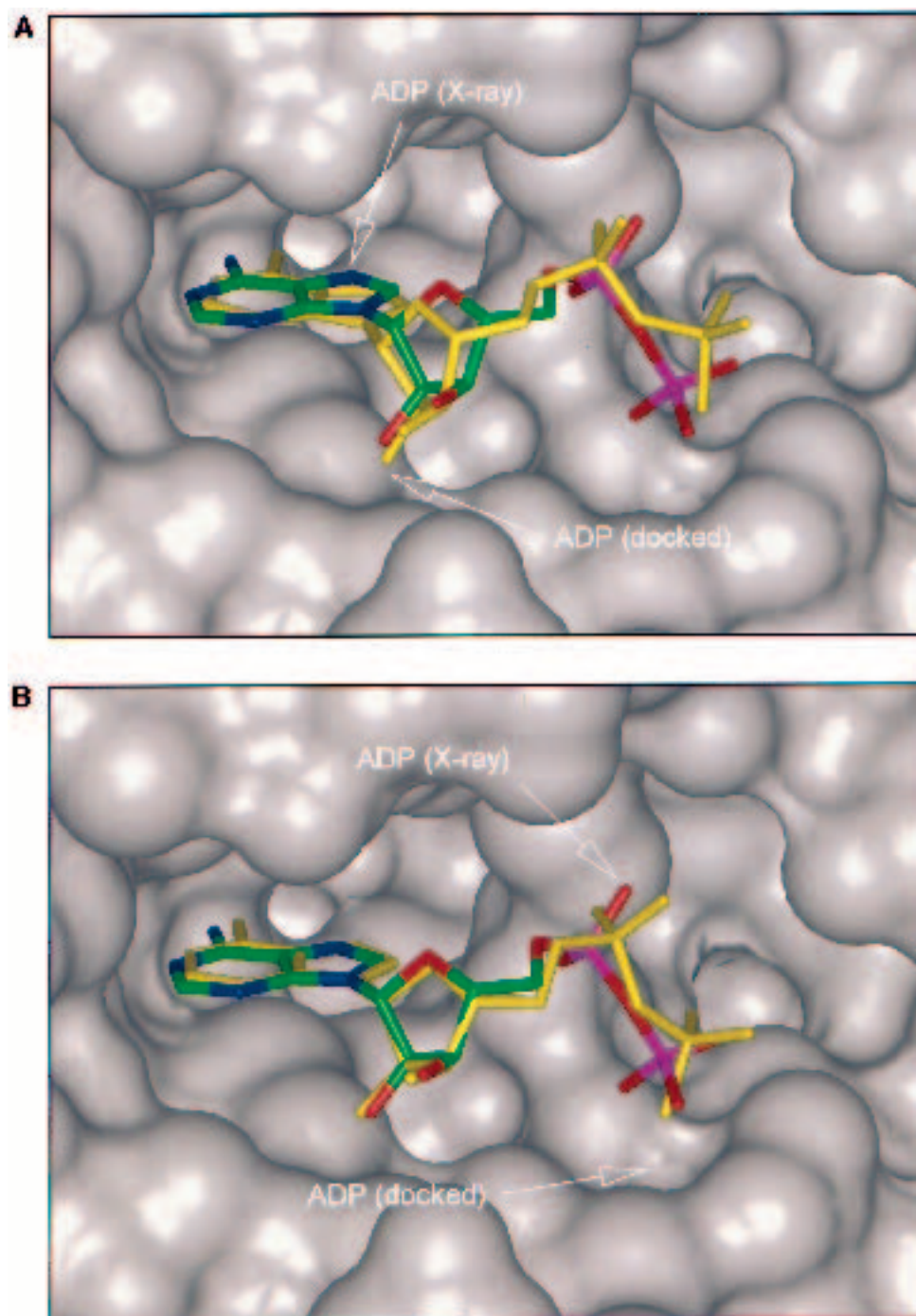


Figure 1. (A) ADP docked to c-Src using AutoDock 2.4 with the structure of ADP from the X-ray structure overlaid for reference (2SRC) [58]. The docked molecule is colored yellow, ADP derived from the crystal structure is colored by atom. Some residues were removed for clarity. (B) ADP docked to c-Src with the additional polar hydrogen atom type introduced into AutoDock. Reference molecule of ADP derived from the X-ray data is colored by atom type, the docked molecule is shown yellow. Some residues were removed for clarity.

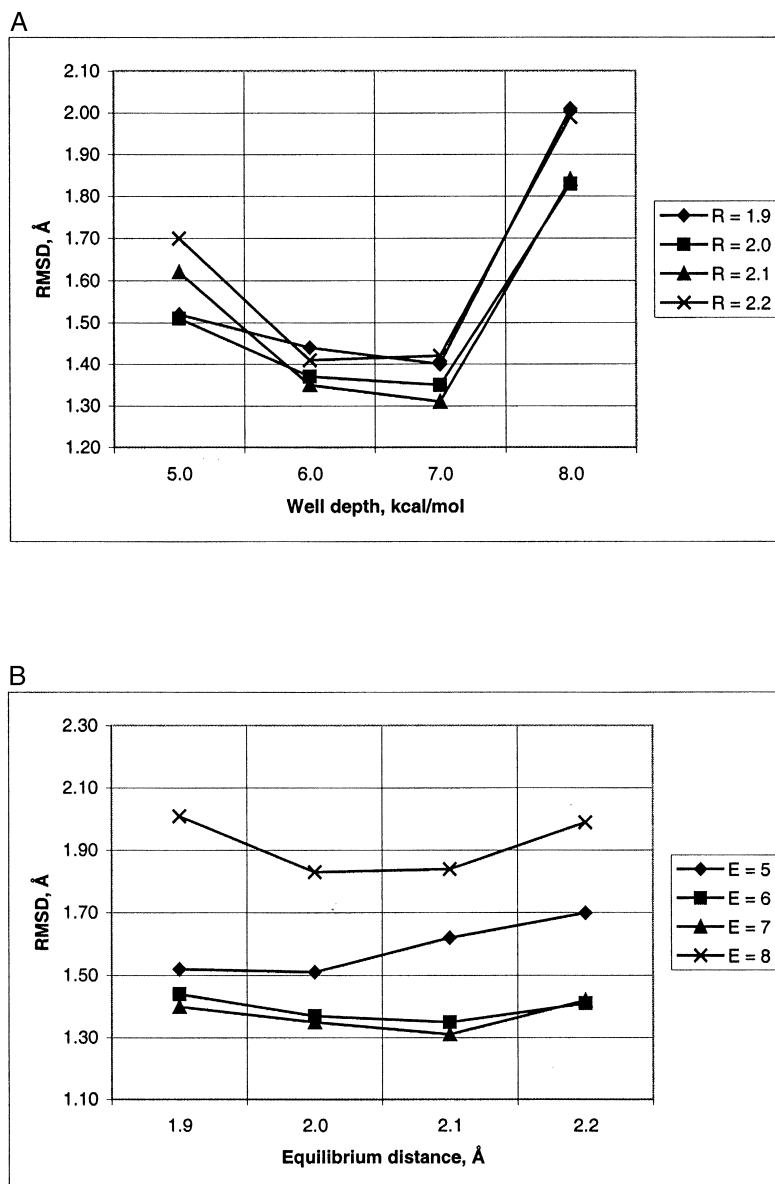


Figure 2. Dependence of RMSD in docking of ATP to wild type c-Src (2SRC) on potential energy well depth and equilibrium distance in the hydrogen bonding potential function. Each data point represents a docking run with the specified values of parameters R_0 and E_0 . Minimum RMSD for this experiment have been achieved at $R_0 = 2.1$ Å and $E_0 = 7$ kcal/mol.

values of well depth and equilibrium distance were determined by docking of ADP with different values of both parameters. One of them was varied while the other kept constant. Subsequently, we selected the values that provided the smallest RMSD (Figure 2A, 2B).

$$E_{HB} = \Sigma(C_{12}/R_{ij}^{12} - C_{10}/R_{ij}^{10}), \quad (1)$$

where $C_{12} = 12/(10-12) \epsilon r^{12}$ and $C_{10} = 10(12-10) \epsilon r^{10}$.

With the introduction of the additional atom type, aliphatic and aromatic hydrogen atoms could be treated explicitly without merging them with carbons. This change, along with a “stronger” hydrogen bonding potential, allowed for the production of a structure of the c-Src-ADP complex with an RMSD as low as 1.3 Å (Figure 1B). Both critical hydrogen bonds

(E339-N⁶ of ADP and M341-N¹ of ADP) have been accurately reproduced, and position of the purine ring system matched its crystallographic position.

In order to assess the generality of the hydrogen bond parameters utilized in docking ADP to c-Src, we used them for modeling of an ADP conformation in the binding site of enterococcal kanamycin kinase APH(3′)-IIIa. This kinase is only distantly related to c-Src sharing 6% sequence identity (14% sequence homology) in its catalytic domain. These kinases are sufficiently different to provide a good test of generality of the modified parameter set. A structure of APH(3′)-IIIa with bound ADP has been solved crystallographically [29] and was used as a reference. Docking of ADP to APH(3′)-IIIa was performed in 15 independent runs and resulted in a conformation of the nucleotide diphosphate that matched that of the crystallographically determined structure exceptionally well (RMSD 1.22 Å in the lowest-energy conformation). The final structure of the complex is shown in Figure 3B.

We also used the standard AutoDock 2.4 force field to model the complex of APH(3′)-IIIa and ADP. Conditions of docking, such as starting position of the ligand, initial annealing temperature and number of temperature reduction cycles, were identical to those set in the experiment with a modified scoring function. The resulting minimum-energy conformation of ADP had an RMSD of over 5.3 Å relative to the X-ray structure (Figure 3A). Most of the error originated from the misplaced adenine ring, which failed to establish the two required hydrogen bonds. Positions of the triphosphate moiety and the ribose ring were reproduced correctly, but overall accuracy was not satisfactory. The most likely cause of such a performance difference is insufficiently strong hydrogen bonding attraction, which has been augmented by the modified version.

Docking of derivatives of PP1 to a c-Src mutant

The ability to reproduce the crystallographic orientation of a ligand bound to a protein active site is only one aspect of inhibitor design. The energetics of the complex should ideally be scored as well. Unfortunately, AutoDock 2.4 does not provide binding energy of the protein-ligand complex, i.e., total free energy change upon the complex formation expressed in kcal/mol. Despite significant effort and some success in this area [30, 31], accurate prediction of the binding energy still remains a daunting challenge.

Predictions are complicated by a multitude of interactions involved in small molecule-protein complex formation, as well as by the lack of a reliable and generally applicable set of scoring functions [32–34]. Molecular mechanics force fields [35–38] have been developed with limited training sets, and hence cannot be applied successfully to all protein-ligand interactions. Alternatively, models based on a more firm theoretical foundation (such as those using quantum mechanics) are still at an early stage of development, although these methods are starting to show promise as a valuable addition to conventionally used MM methods [39–42]. Unfortunately, such models require considerable computational resources.

Nonetheless, analysis of our AutoDock 2.4 output has indicated that useful quantitative information, such as relative binding affinity, could be extracted from the total energy of the complex. The latter is calculated in the process of simulation and is used by AutoDock to rank the binding orientations. We used the scoring function modified as described above to dock a panel of pyrazolopyrimidine 1 (PP1, Figure 4A) inhibitors to an engineered c-Src kinase. The enzyme active site was modified by introduction of a single mutation (T338G) designed to accommodate the naphthyl ring of the compound **1** (Figure 4A) [23]. The panel of candidate inhibitors consisted of PP1 derivatives containing differing substituents at the position 3 of the pyrazolopyrimidine ring (Figure 4A) [9, 11].

The simulation produced structures of the protein-inhibitor complex with all candidate inhibitors located in the ATP binding site with the R-group in the pocket created by the T338G mutation. This result was expected from analysis of the X-ray co-crystal structure of the close c-Src relative, lymphocyte-specific kinase lck, and PP1 (PDB ID: 1QPE) [43]. Clearly, the mutation alone is unable to create space sufficient to accommodate ligands with larger R groups, such as compounds **3**, **4**, or **9** (Figure 4A). However, according to the crystal structure of the Lck-PP1 complex, removal of T338 opens approach to a large, but otherwise inaccessible, pocket between the two lobes of the catalytic domain. This mutation allows positioning of the bulky R groups in the pocket while maintaining the key hydrogen bonds with backbone residues of the kinase. The binding mode of 1-naphthyl PP1 (compound **1**, Figure 4A) with respect to the position of PP1 is shown in Figure 5 and agrees well with the predicted binding orientation [23]. Earlier modeling of some of the complexes using the standard AutoDock gave very similar results for PP1 and 1-naphthyl PP1

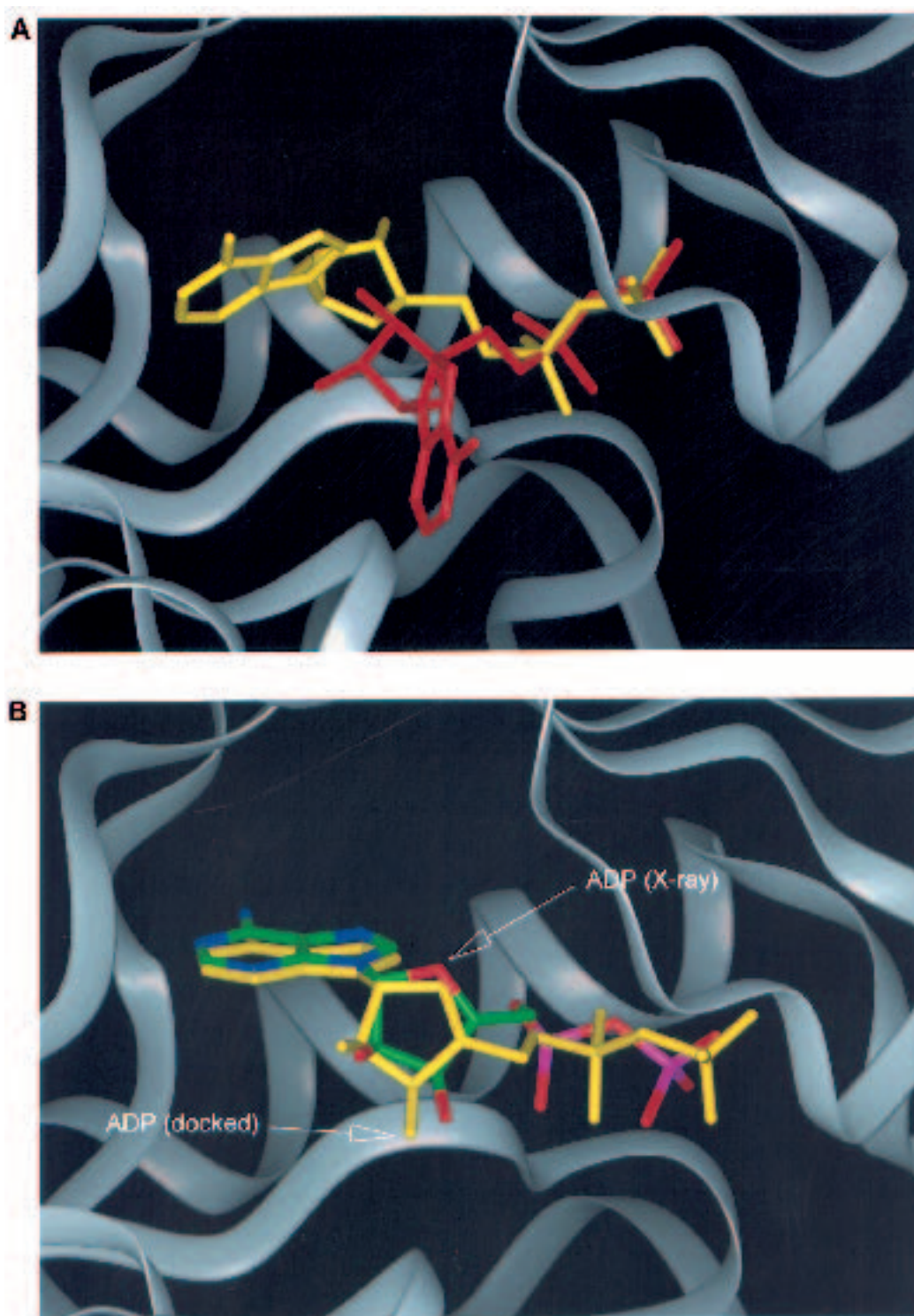


Figure 3. (A) ADP docked to APH(3')-IIIa using standard AutoDock. Two lowest-energy resulting positions are colored red and yellow. Reference structure of ADP as determined by X-ray crystallography is shown in Figure 3B. (B) ADP docked to APH(3')-IIIa with the set of parameters identical to that used for modeling of *c-Src*-ADP complex. Docked ADP is shown yellow, and the reference structure of ADP from the X-ray structure [29] is colored by atom type.

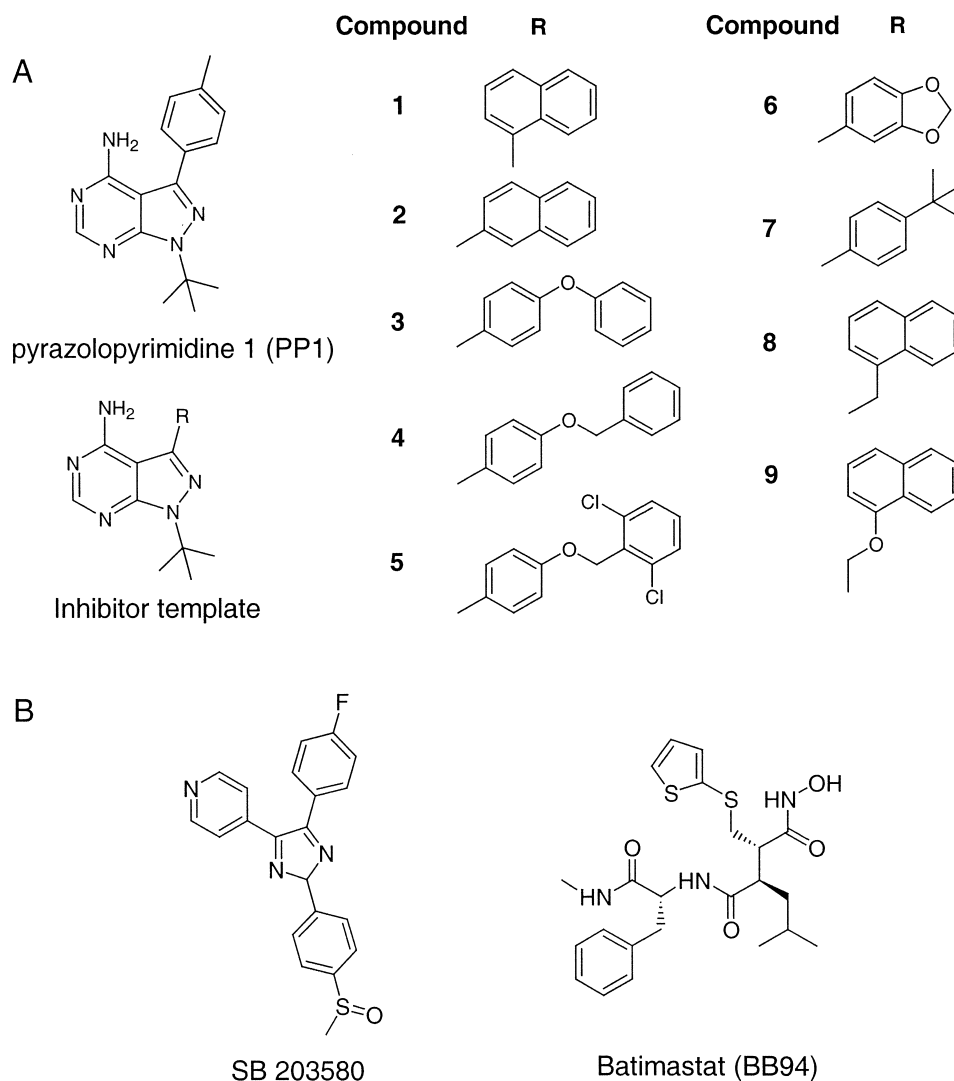


Figure 4. (A) The family of pyrazolopyrimidine 1 (PP1) inhibitors used in the docking studies. (B) Chemical structures of SB203580 and batimastat.

(compound **1**, Figure 4A), compared to the modified version. However, in the cases of bulky aromatic substituents (compounds **3**, **4**, **5** and **9**, Figure 4A) the ligands did not retain one of the key hydrogen bonds with the protein backbone, which were correctly reproduced when using the updated hydrogen bonding function.

Correlation between total binding energies calculated by AutoDock and experimentally determined IC_{50} values is shown in Figure 6. The subset of six compounds (**1** through **6**, Figure 4A) was used as a training set to derive a relationship between experimental values of IC_{50} and total energy values provided by AutoDock. This linear dependency was

used to make predictions of inhibition in the remaining cases, producing remarkably good correlation with experimental values of IC_{50} . The corresponding correlation coefficient for the complete set of compounds is 0.8766.

Significant correlation between calculated binding energies and measured IC_{50} allows for discrimination between candidate compounds whose measured affinity differs by one or two orders of magnitude. This approach may not be sufficiently sensitive for a comparison between compounds with more similar affinities, however, it may prove useful at early stages of screening to identify and discard weak binders.

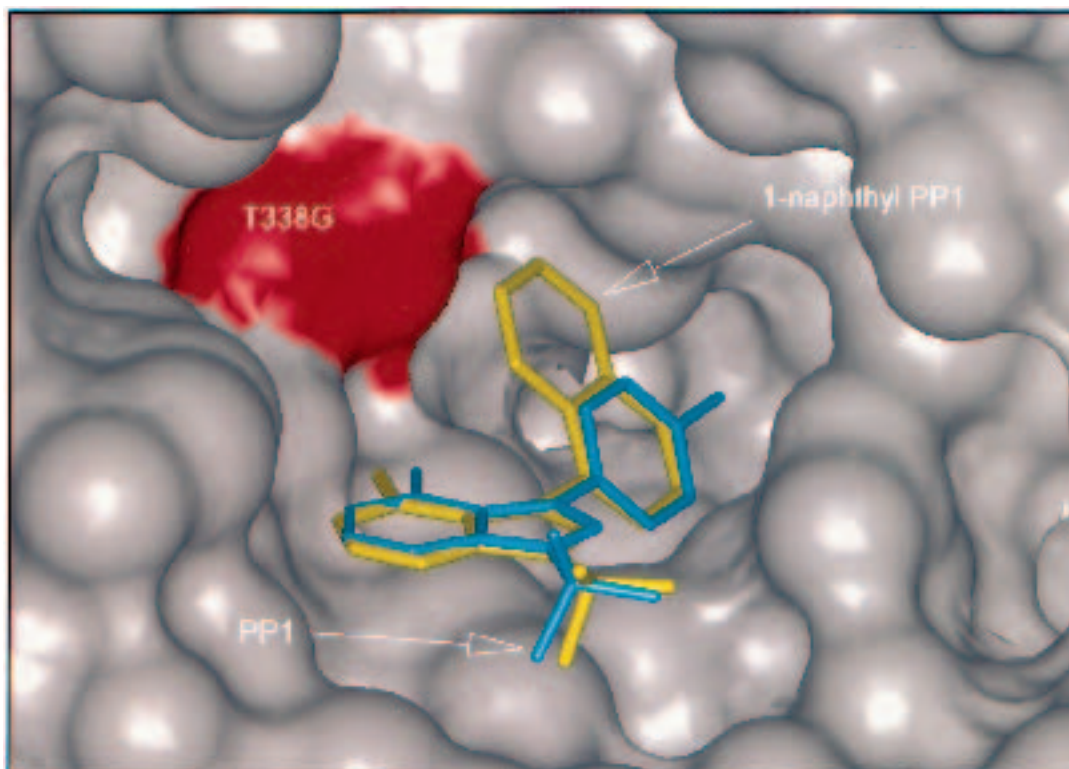


Figure 5. 1-Naphthyl PP1 (shown in yellow) docked to T338G c-Src mutant and superimposed with a docked structure of PP1 (colored blue). Some residues are removed for clarity.

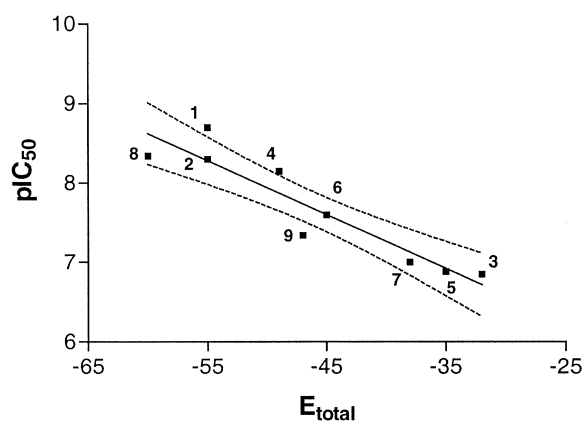


Figure 6. Relationship between binding energy in the protein-ligand complex as given by AutoDock and experimentally determined extent of inhibition (pIC_{50}). The 95% confidence interval is shown as dotted lines. Compound numbers are given at each data point.

Prediction of binding energies as described above can only be made within a series of compounds that belong to the same structural class. Ligands with widely differing surface areas or number of hydrogen

bonding contacts will yield very different total binding energies, while their actual binding affinities may be quite similar. Such discrepancies are likely to arise due to unbalanced treatment of the interactions involved in protein-inhibitor complex formation.

Approximation of solvation in complexes of p38 MAP kinase

A wide variety of heterocycle scaffolds have been identified as kinase inhibitors [8]. Thus, any useful docking algorithm must be sufficiently versatile to allow accurate treatment of a range of small molecule inhibitors. As a test of our modified version of AutoDock, we turned to the p38 MAP kinase. Mitogen activated protein kinase p38 is an important link in the transfer of extracellular signals to the nucleus and has been implicated in cytokine and stress responses, as well as antiapoptotic activity [5, 44, 45]. As opposed to c-Src, a protein tyrosine kinase that features a closed and well defined ATP binding site, p38 is a serine/threonine protein kinase, which has an open binding cleft with relatively few potential van der

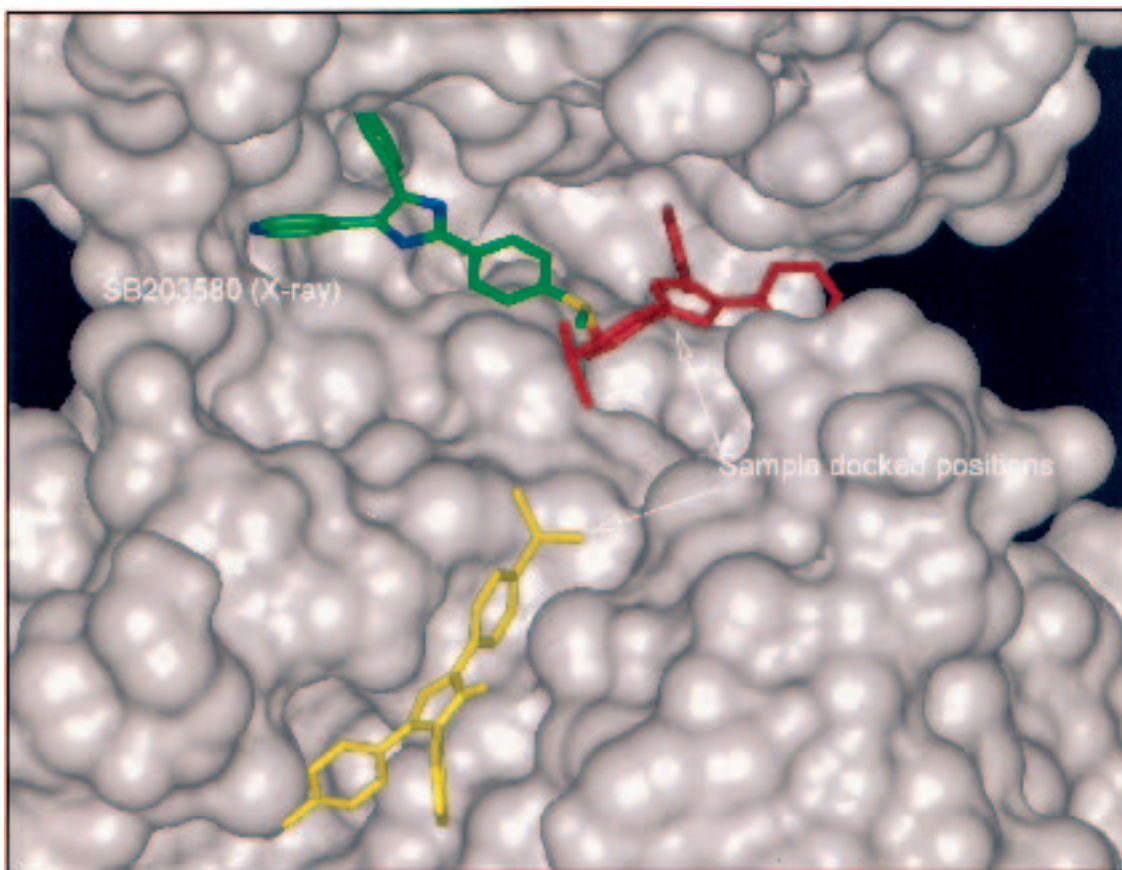


Figure 7. Structure of SB203580 derived from the crystallographic data (colored by atom type) and conformations resulting from docking using standard AutoDock 2.4 with no treatment of solvation effects (colored red and yellow).

Waals or hydrogen bonding contacts (11AN in PDB) [46]. In the crystal structure of the complex of p38 and p38-specific inhibitor SB203580 (Figures 4B and 7), only one hydrogen bond is present – between the backbone amide of M109 and the pyridine moiety of the inhibitor. The lack of strong van der Waals interactions and hydrogen bonding in this complex severely limited utility of AutoDock 2.4. Multiple docking runs with different run options (grid spacing, starting position of the inhibitor, use of the additional polar hydrogen atom type, etc.) resulted in orientations of the ligand completely different from that experimentally observed (Figure 7). Visual inspection of the structure of the complex suggested that binding of SB203580 to p38 is directed mostly by hydrophobic contacts made by the 4-fluorophenyl group of the inhibitor in the pocket formed by residues 75, 84–86, 104 and 105. The ligand is oriented by the hydrogen bond with the amide hydrogen of M109. An adequate treatment of solvation was, however, beyond the capabilities of

AutoDock. A newer version, AutoDock 3.0 [47], does contain a desolvation term in its free energy function, which is calculated only for aliphatic and aromatic carbons regardless of their environment. Such generalized treatment, unfortunately, does not accurately treat local differences within the binding site and tends to only increase affinity between the components of the complex, rather than discriminate between subsites (data not shown).

The problem of modeling hydrophobic interactions/solvation effects appears to be very general, as many other protein-ligand systems require explicit treatment of these effects in order to produce meaningful results [30, 48, 49]. Several approaches to this problem have been suggested to date, one of them being a solvent-accessible surface area (SASA) model proposed by Eisenberg and McLachlan [50]. The SASA methodology is based on the assumption that energy of a hydrophobic interaction is directly proportional to the area of the surface exposed to the

solvent. Total free energy change upon solvation is calculated using (2):

$$\Delta G_{\text{solv}} = \sum \Delta \sigma_i A_i, \quad (2)$$

where A_i is solvent-accessible surface of the atom and $\Delta \sigma_{i_g}$ its atomic solvation parameter. The latter corresponds to the free energy change associated with exposure of a unit surface area of the atom to the solvent. Values of atomic solvation parameters $\Delta \sigma_i$ are calculated from a training set of small molecules whose solvation energy has been determined experimentally [50]. Coefficients are then fitted to provide consistency between experimental and calculated solvation energy.

Another approach to treating hydrophobic interactions involves calculation of the solvent-excluded volume for each component atom [51–53], and calculation of the hydrophobic effect by summation over all component atoms without restriction to the surface residues. The value of ΔG is calculated as follows [53]:

$$\Delta G_{\text{solv}}(i) = \text{SolPar}(i) \cdot (\text{Occ}_{\text{max}}(i) - \text{Occ}(i)), \quad (3)$$

where the $\Delta G_{\text{solv}}(i)$ is atomic free energy of solvation, $\text{SolPar}(i)$ and Occ_{max} are the atomic solvation parameter and the maximum occupancy of the atom i , respectively. The occupancy of the atom is calculated from the expression

$$\text{Occ}(i) = \sum \text{Vol}(j) \cdot \exp(-r_{ij}^2/2\sigma^2), \quad (4)$$

where $\text{Vol}(j)$ is the fragmental volume of neighbor atom j , r_{ij} is the distance between atoms i and j , and the exponent is the envelope function [53].

Both of these methods provide adequate treatment of small molecules, however, analysis of proteins is complicated by various possible ionization states of the residues. Correct parameterization of atom types in proteins is difficult due to significant differences in electrostatic properties of free amino acids and amino acid residues of proteins. A number of other, mathematically more rigorous, methods are also available [54–56], however, they require significant computational resources and still do not provide sufficiently accurate description of desolvation effects.

Given the shortcomings of the existing methods, we set out to find a simple and efficient alternative. In this context, manipulation of the content of scoring grids appeared to be the most clear and straightforward way to introduce treatment of solvation effects. Indeed, as we analyzed scoring grids generated by *autogrid*, we found differences between subsites featuring mostly hydrophobic interactions and those having

a strong electrostatic component. Generally, a more hydrophobic area is characterized by strong attraction on the carbon and hydrogen grids and relatively low values of electrostatic energy. Conversely, more polar subsites showed considerably higher values of electrostatic interaction energy. At this point, we made the assumption that stronger electrostatics facilitate solvation of the given subsite, causing a less favorable energetic outcome when water is displaced by the ligand. The opposite would be true for more hydrophobic subsites. Therefore, we predicted an increase the attraction in hydrophobic subsites would compensate for the lack of an explicit hydrophobic function.

In order to implement a search for potential hydrophobic sites, we modified the AutoDock module responsible for calculating scoring grids (*autogrid*). The search is carried out by testing values of electrostatic energy and C and H attraction at the current grid point. When the condition of high hydrophobicity is ‘true’, the carbon attraction component is increased by a factor δ :

$$\Delta E_i'' = \Delta E_i' \cdot \delta, \quad (5)$$

where $\Delta E_i'$ is the original value of interaction energy at the grid point and $\Delta E_i''$ is the corrected energy of interaction for atom i . The correction factor δ is determined by values of ΔE_C and ΔE_e as follows:

$$\delta = \begin{cases} 1.7 & \text{if } \Delta E_C < 0.5 \cup \Delta E_H < 0.5 \cup -0.5 < \Delta E_e < 1.2 \\ 1 & \text{otherwise} \end{cases} \quad (6)$$

where ΔE_C , ΔE_H and ΔE_e are values of interaction energy in the carbon, hydrogen and electrostatic grids, respectively. The thresholds have been determined in a series of simulations, where the finally accepted values were found to provide the best description of a hydrophobic site.

The optimal value of δ has been determined empirically using p38-SB203580 complex as a model system. A set of scoring grids with different values of δ was calculated, and SB203580 was subsequently docked to each of them with the expanded atom type set (*vide supra*). Results of each simulation were evaluated by the RMSD of the docked inhibitor with respect to the crystallographically determined position (Figure 8).

Remarkably, all of the values used in the training set produced better RMSD values than the original simulation procedure, which indicated that our approach, despite its simplicity, had introduced changes sufficient to handle the mostly hydrophobic environment in the active site of p38. Final simulation with the optimized δ (1.7) in 8 runs out of 12 produced

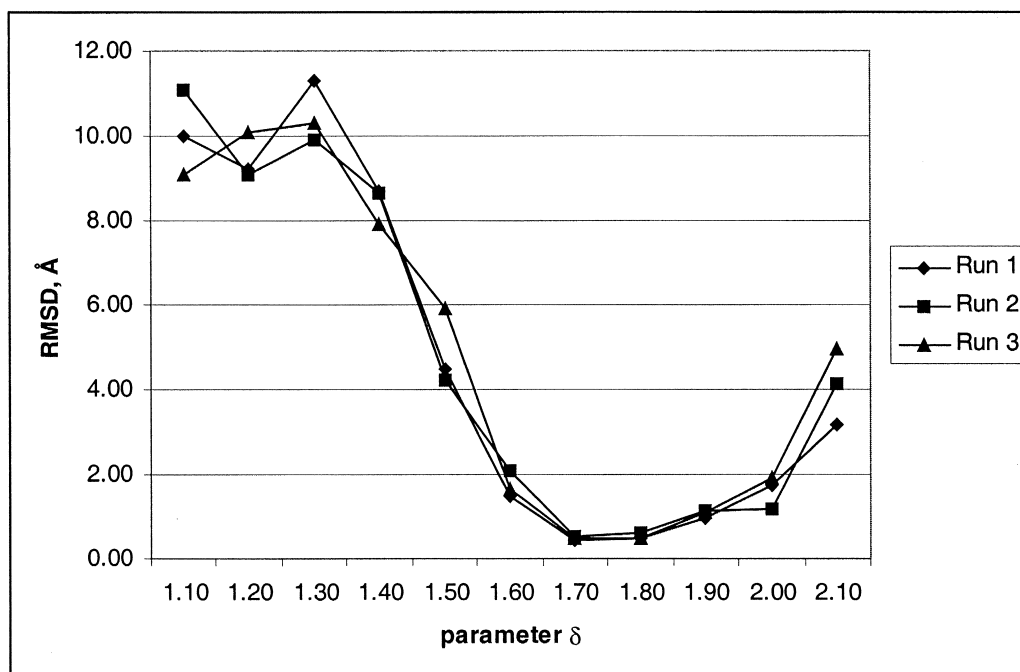


Figure 8. Relationship between value of δ coefficient and RMSD of the resulting orientation of SB203580 with reference to the crystallographically determined position in the active site of p38.

ligand orientations in excellent agreement with the experimental X-ray structure with RMSD ranging from 0.84 to 1.42 Å (Figure 9). All important features of the complex were correctly reproduced, including the position of the 4-fluorophenyl group in the hydrophobic pocket and the key hydrogen bond between the backbone amide of M109 and the pyridine group of SB203580. The only deviation from the crystallographic position was the overall tilt ($\sim 10^\circ$) of the ligand in the binding site. This might have been a result of overestimated attraction between SB203580 and surface residues 35, 38 and 53 due to the absence of explicit treatment of solvent. Molecules of water could potentially occupy the space between the ligand and these residues and, thus, keep the molecule in a more balanced position. It has to be noted, however, that the p38-SB203580 crystal structure did not show water molecules coordinated in this site [46].

Docking to other protein-ligand complexes

Having introduced an adequate hydrophobic correction in the p38 simulation, we set out to discover if the modified program would maintain or enhance performance of AutoDock in other docking examples. These studies were carried out to test the balance of hydrophobic and van der Waals forces modeled by

the δ parameter. Since manual modification of the δ parameter for each protein of interest would be unacceptable, the ability of the program to accommodate other cases is essential to the broad applicability of the modified algorithm.

In order to fully test the new capabilities of AutoDock, we selected a complex of human MMP-12 (macrophage elastase) with a broad-specificity inhibitor batimastat (Figure 4B). According to the crystal structure solved at 1.1 Å (1JK3 in PDB) [57], the ligand is positioned in a shallow binding site virtually on the surface of the protein. It participates in several hydrogen bonds with the protein backbone and makes a hydrophobic contact with its isobutyl group (Figure 10B). The hydroxamic acid moiety is contacting Zn, which is also coordinated by three histidine residues (113, 117 and 123).

We modeled Zn by replacing it in the original protein with a polar hydrogen and assigning it a +2 charge. The standard and modified AutoDock were run with identical parameters: initial temperature and RT reduction (at 615 and 0.95, respectively), coordinates of the starting position, number of cooling cycles (20), number of accepted and rejected steps (12000 in each case).

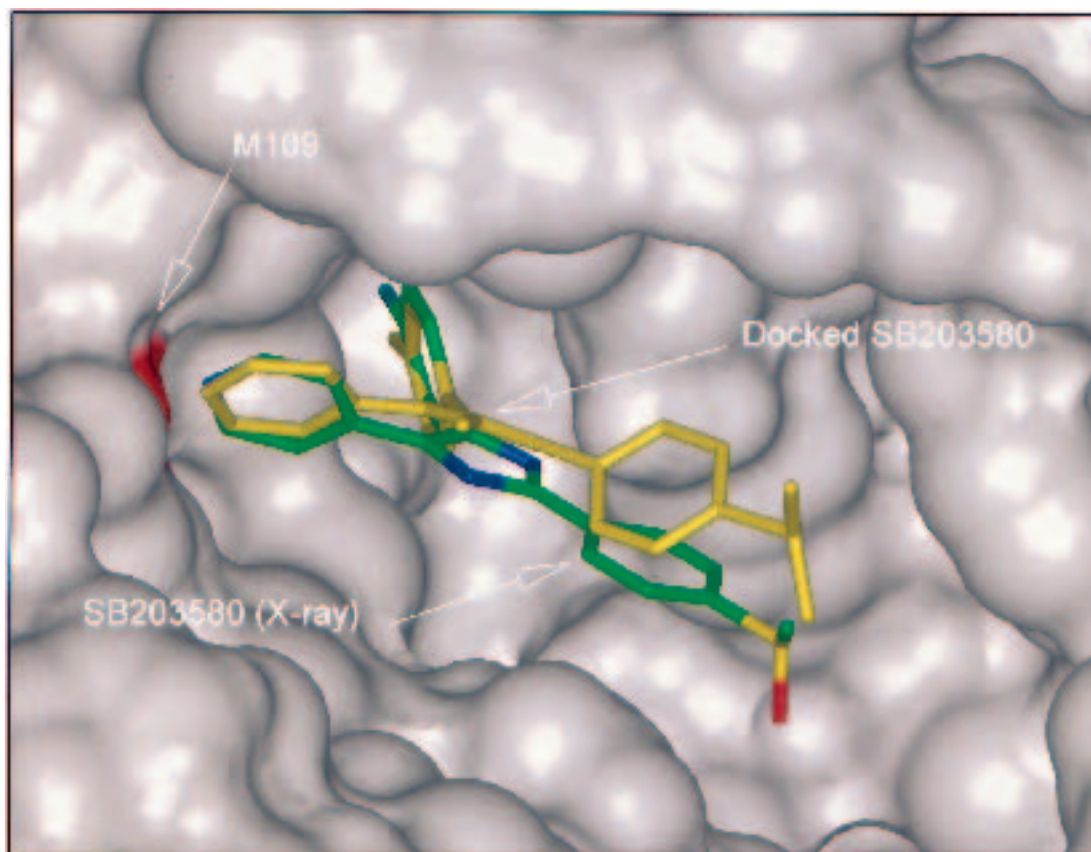


Figure 9. Overlaid structures of SB203580 derived from the crystallographic data (colored by atom type) and docked using the modified AutoDock with approximation of solvation effects (shown yellow). The root mean-square deviation averages 0.44 Å over the 8 best docking runs.

The standard AutoDock produced a number of ligand conformations, where the molecule was placed in the target pocket. However, none of them correctly reproduced the experimentally determined position. In particular, the bidentate coordination of the hydroxamic moiety with the Zn was replaced by a single contact with a carbonyl group of the inhibitor. In another case, only the carbonyl oxygen of hydroxamic acid contacted Zn. In addition, the overall position of the ligand reflected absence of a hydrophobic attraction component, as it either did not bind in the pocket normally occupied by the isobutyl group or adopted a conformation different from the crystallographically determined (Figure 10A). The RMSD value in this series ranged from 5.1 to 9.4 Å.

The modified version of the suite, in contrast, generated several lowest energy models with the ligand placed in the correct position. Although the local deviations from the experimentally determined structure amount to a total RMSD of 2.87 Å for the lowest en-

ergy conformation, the overall position and the key interactions are reproduced correctly (Figure 10B). The hydroxamic acid group is rotated by approximately 30 deg, however it did not change the coordination with the Zn ion, and neither did it interfere with either of the histidine residues participating in the complex. In addition, positions of the aromatic groups were reproduced correctly, as well as binding of the isobutyl group.

Thus, the modified version of AutoDock was capable of correctly modeling the overall position of the ligand and specific interaction within the complex, including the key hydrogen bonds and hydrophobic contacts. Considering that neither the target nor the ligand are related to the p38-SB203580 system or protein kinases in general, this result appears to be a reliable demonstration of a wider applicability of modified AutoDock than protein kinase complexes.

We also used the modified AutoDock to simulate systems more amenable to basic docking, such

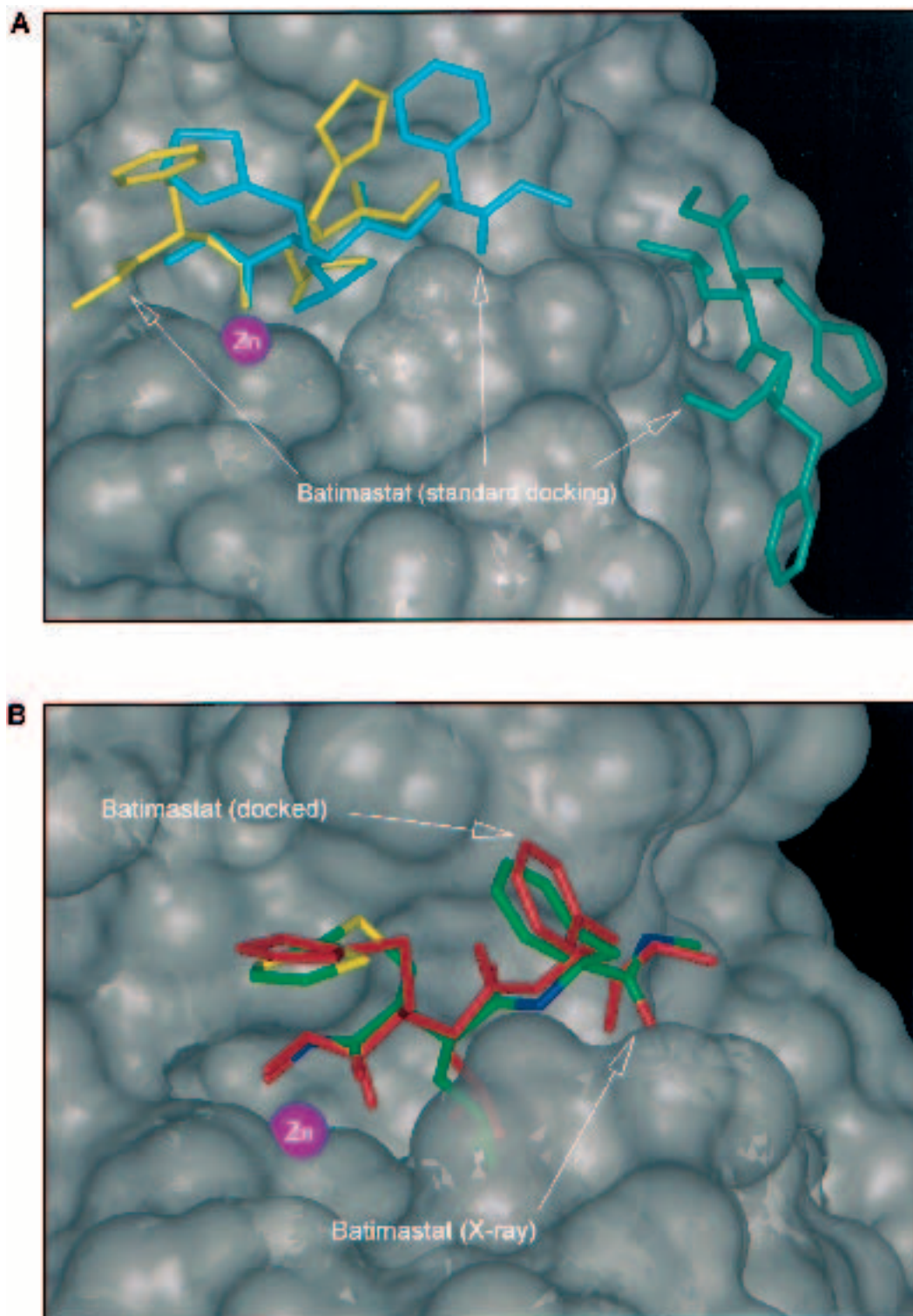


Figure 10. (A) Representative positions of batimastat complexed to MMP-12 as predicted by standard AutoDock 2.4. (B) Complex of human macrophage elastase (MMP-12) and batimastat (BB94) modeled with modified AutoDock. The crystallographic conformation is colored by atom, the docked ligand is colored red.

as protein kinase HCK with quercetin and cyclin-dependent kinase CDK with its picomolar inhibitor staurosporine. In these cases, the modified AutoDock reproduced the experimentally determined orientations as well or slightly better than the standard version with RMSD 0.4 and 0.5 Å for the above complexes. This indicates that the new version is capable of handling 'conventional' cases as well or better than the AutoDock 2.4 and, in addition, it can be an essential tool in modeling more complex docking situations where simulation of strong hydrogen bonding and explicit treatment of hydrophobics are required.

Conclusions

Protein kinases are one of the largest families of enzymes in nature controlling key aspects of almost every known cell signaling pathway. Consequently, they are very popular drug targets, thus, making fast and accurate methods for evaluating potential kinase ligands (substrates or inhibitors) especially important. Computational approaches have been used extensively in this area as a modeling and predictive tool. Among these, docking methods have shown much promise in modeling of enzyme-inhibitor interactions. AutoDock provided highly accurate modeling of interactions within kinase-inhibitor complexes in the studied cases. However, in situations where binding is governed largely by extensive hydrophobic contacts, additional elements are required to produce an adequate model.

In order to improve performance, we have modified the docking suite AutoDock 2.4 to include additional atom types. In addition, we have introduced an empirical function for modeling solvation effects. Despite its simplicity, the modification is sufficiently effective to produce correct binding modes of SB203580 in the binding site and to account for the major hydrophobic component. The described method is a rather simplified approximation of the hydrophobic effect, and a model with a charge-dependent correction factor may prove to be more accurate. However, parameters derived for p38 are applicable to other protein systems, as confirmed by further tests.

Acknowledgements

This work was supported by grants from the NIH (ROI AI44009, ROI CA70331). We would like to thank Dr

A. M. Berghuis for the coordinates of the APH(3')-IIIa-ADP complex. We would like to thank Dr Yi Liu for helpful comments on the manuscript.

References

1. Muegge, I. and Rarey, M., *Rev. Comp. Chem.*, 17 (2001) 1.
2. Baselga, J. and Mendelsohn, J., *J. Mammary Gland Biol. Neoplasia*, 2 (1997) 165.
3. Pessin, J.E. and Saltiel, A.R., *J. Clin. Invest.*, 106 (2000) 165.
4. Reuter, C.W., Morgan, M.A. and Bergmann, L., *Blood*, 96 (2000) 1655.
5. Okamoto, S., Krainc, D., Sherman, K. and Lipton, S.A., *Proc. Natl. Acad. Sci. USA*, 97 (2000) 7561.
6. Davies, S.P., Reddy, H., Caivano, M. and Cohen, P., *Biochem. J.*, 351 (2000) 95.
7. Dumas, J., Hatoum-Mokdad, H., Sibley, R., Riedl, B., Scott, W.J., Monahan, M.K., Lowinger, T.B., Brennan, C., Natero, R., Turner, T., Johnson, J.S., Schoenleber, R., Bhargava, A., Wilhelm, S.M., Housley, T.J., Ranges, G.E. and Shrikhande, A., *Bioorg. Med. Chem. Lett.*, 10 (2000) 2051.
8. Bishop, A.C., Shah, K., Liu, Y., Witucki, L., Kung, C. and Shokat, K.M., *Curr. Biol.*, 8 (1998) 257.
9. Liu, Y., Bishop, A., Witucki, L., Kraybill, B., Shimizu, E., Tsien, J., Ubersax, J., Blethrow, J., Morgan, D.O. and Shokat, K.M., *Chem. Biol.*, 6 (1999) 671.
10. Bishop, A., Buzko, O., Heyeck-Dumas, S., Jung, I., Kraybill, B., Liu, Y., Shah, K., Ulrich, S., Witucki, L., Yang, F., Zhang, C. and Shokat, K.M., *Annu. Rev. Biophys. Biomol. Struct.*, 29 (2000) 577.
11. Bishop, A.C. and Shokat, K.M., *Pharmacol. Ther.*, 82 (1999) 337.
12. Bishop, A.C., Ubersax, J.A., Petsch, D.T., Matheos, D.P., Gray, N.S., Blethrow, J., Shimizu, E., Tsien, J.Z., Schultz, P.G., Rose, M.D., Wood, J.L., Morgan, D.O. and Shokat, K.M., *Nature*, 407 (2000) 395.
13. Gillespie, P.G., Gillespie, S.K., Mercer, J.A., Shah, K. and Shokat, K.M., *J. Biol. Chem.*, 274 (1999) 31373.
14. Weiss, E.L., Bishop, A.C., Shokat, K.M. and Drubin, D.G., *Nat. Cell Biol.*, 2 (2000) 677.
15. DesJarlais, R.L., Sheridan, R.P., Dixon, J.S., Kuntz, I.D. and Venkataraghavan, R., *J. Med. Chem.*, 29 (1986) 2149.
16. Kuntz, I.D., Blaney, J.M., Oatley, S.J., Langridge, R. and Ferrin, T.E., *J. Mol. Biol.* 161 (1982) 269.
17. Goodsell, D.S., Morris, G.M. and Olson, A.J., *J. Mol. Recognit.*, 9 (1996) 1.
18. Morris, G.M., Goodsell, D.S., Huey, R. and Olson, A.J., *J. Comput. Aided Mol. Des.*, 10 (1996) 293.
19. Jones, G., Willett, P. and Glen, R.C., *J. Mol. Biol.*, 245 (1995) 43.
20. Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R., *J. Mol. Biol.* 267 (1997) 727.
21. Baxter, C.A., Murray, C.W., Waszkowycz, B., Li, J., Sykes, R.A., Bone, R.G., Perkins, T.D. and Wylie, W., *J. Chem. Inf. Comput. Sci.*, 40 (2000) 254.
22. Liu, Y., Shah, K., Yang, F., Witucki, L. and Shokat, K.M., *Chem. Biol.*, 5 (1998) 91.
23. Bishop, A.C., Kung, C.-Y., Shah, K., Witucki, L., Shokat, K.M. and Liu, Y., *J. Am. Chem. Soc.*, 121 (1999) 627.
24. Xu, W., Harrison, S.C. and Eck, M.J., *Nature* 385 (1997) 595.
25. Schindler, T., Sicheri, F., Pico, A., Gazit, A., Levitzki, A. and Kuriyan, J., *Mol. Cell* 3 (1999) 639.

26. Sicheri, F., Moarefi, I. and Kuriyan, J., *Nature*, 385 (1997) 602.
27. Schulze-Gahmen, U., De Bondt, H.L. and Kim, S.H., *J. Med. Chem.*, 39 (1996) 4540.
28. Kollman, P., *Chem. Rev.*, 93 (1993) 2395.
29. Hon, W.C., McKay, G.A., Thompson, P.R., Sweet, R.M., Yang, D.S., Wright, G.D. and Berghuis, A.M., *Cell*, 89 (1997) 887.
30. Klebe, G. and Bohm, H.J., *J. Recept. Signal Transduct. Res.*, 17 (1997) 459.
31. Ajay, M.A. Murcko, *J. Med. Chem.*, 38 (1995) 4953.
32. Holloway, M.K., *Perspect. Drug Disc. Res.*, 9/10/11 (1998) 63.
33. Oprea, T.I. and Marshall, G.R., *Perspect. Drug Disc. Des.*, 9/10/11 (1998) 35.
34. Tame, J.R.H., *J. Comp. Aided Mol. Des.*, 13 (1998) 99.
35. Tokarski, J.S., Hopfinger, A.J., *J. Chem. Inf. Comput. Sci.*, 37 (1997) 792.
36. Weiner, S.J., Kollman, P.A., Nguyen, D.T., Case, D.A., *J. Comput. Chem.*, 7 (1986) 230.
37. Allinger, N.L., Yuh, Y.H. and Lii, J.-H., *J. Am. Chem. Soc.*, 111 (1989) 8551.
38. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W. and Kollman, P.A., *J. Am. Chem. Soc.*, 117 (1995) 5179.
39. Gao, J. and Xia, X., *Science*, 258 (1992) 631.
40. Han, W.G., Tajkhorshid, E. and Suhai, S., *J. Biomol. Struct. Dyn.*, 16 (1999) 1019.
41. Friesner, R.A. and Beachy, M.D., *Curr. Opin. Struct. Biol.*, 8 (1998) 257.
42. Perakyla, M. and Pakkanen, T.A., *Proteins*, 20 (1994) 367.
43. Zhu, X., Kim, J.L., Newcomb, J.R., Rose, P.E., Stover, D.R., Toledo, L.M., Zhao, H. and Morgenstern, K.A., *Structure Fold. Des.*, 7 (1999) 651.
44. Somwar, R., Perreault, M., Kapur, S., Taha, C., Sweeney, G., Ramlal, T., Kim, D.Y., Keen, J., Cote, C.H., Klip, A. and Marette, A., *Diabetes*, 49 (2000) 1794.
45. Hutter, D., Chen, P., Barnes, J. and Liu, Y., *Biochem. J.*, 352 Pt 1 (2000) 155.
46. Wilson, K.P., Fitzgibbon, M.J., Caron, P.R., Griffith, J.P., Chen, W., McCaffrey, P.G., Chambers, S.P. and Su, M.S., *J. Biol. Chem.*, 271 (1996) 27696.
47. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J., *J. Comp. Chem.*, 19 (1998) 1639.
48. Shoichet, B.K., Leach, A.R. and Kuntz, I.D., *Proteins*, 34 (1999) 4.
49. Richards, W.G., King, P.M. and Reynolds, C.A., *Protein Eng.*, 2 (1989) 319.
50. Eisenberg, D. and McLachlan, A.D., *Nature*, 319 (1986) 199.
51. Kang, Y.K., Nemethy, G. and Scheraga, H.A., *J. Phys. Chem.*, 91 (1987) 4105.
52. Motoc, I. and Marshall, G.R., *Chem. Phys. Lett.*, 116 (1985) 415.
53. Stouten, P.F.W., Frommel, C., Nakamura, H. and Sander, C., *Mol. Simulation*, 10 (1993) 97.
54. Bashford, D. and Case, D.A., *Annu. Rev. Phys. Chem.*, 51 (2000) 129.
55. Dennis, S., Camacho, C.J. and Vajda, S., *Proteins*, 38 (2000) 176.
56. Lee, M.R., Duan, Y. and Kollman, P.A., *Proteins*, 39 (2000) 309.
57. Lang, R., Kocourek, A., Braun, M., Tschesche, H., Huber, R. and Bode, Maskos, K., *J. Mol. Biol.*, 312 (2001) 731.
58. Xu, W., Doshi, A., Lei, M., Eck, M.J. and Harrison, S.C., *Mol. Cell.*, 3 (1999) 629.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.