5578 J. Med. Chem. 2009, 52, 5578–5581
DOI: 10.1021/jm901045w

Journal of Medicinal Chemistry Letter

Pursuing Aldose Reductase Inhibitors through in Situ Cross-Docking and Similarity-Based Virtual Screening

Sandro Cosconati,[†] Luciana Marinelli,*[,†] Concettina La Motta,[‡] Stefania Sartini,[‡] Federico Da Settimo,[‡] Arthur J. Olson,[#] and Ettore Novellino*[,†]

[†]*Dipartimento di Chimica Farmaceutica e Tossicologica, Università di Napoli "Federico II", Via D. Montesano 49, 80131 Napoli, Italy,* [‡]*Dipartimento di Scienze Farmaceutiche, Università di Pisa, Via Bonanno 6, 56126 Pisa, Italy, and* [#]*Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037*

Received July 14, 2009

Abstract: Aldose reductase (ALR2) is a critical enzyme in the development of the major complications of diabetes mellitus. Herein, new molecular entities active against ALR2 were discovered through an integrated receptor- and ligand-based virtual screening campaign. Twelve candidates were found to inhibit this enzyme in the micromolar range including two ligands having an IC_{50} below 3 μM. Six new compounds, structurally unrelated to the known ARIs, have been identified, opening up opportunity for lead optimization.

Diabetes mellitus (types 1 and 2) is a pandemic metabolic disorder characterized by high levels of blood glucose, resulting from defects in insulin production, insulin action, or both. In different tissues, the prolonged exposure to hyperglycemia causes biochemical and functional alterations which eventually lead to the complications of this disease (blindness, renal failure, neuropathy, limb amputation, myocardial infarction, and stroke). These alterations are linked to an increased flux of glucose through the polyol pathway[1] in which aldose reductase (ALR2[a]) is the enzyme that catalyzes the reduction of glucose to sorbitol. Under hyperglycemic conditions, this pathway is overactivated resulting in high concentrations of intracellular sorbitol which lead to increased cellular osmolarity and oxidative stress.[2] Recent studies clearly demonstrated a correlation between overexpression of the human ALR2 gene and the likelihood of the development of complications among diabetic patients.[3,4] Thus, inhibition of ALR2 represents a good opportunity for drug intervention aimed at preventing the onset, progression, and severity of diabetic complications. Despite the big strides in the identifications of ALR2 inhibitors (ARI), most of them failed in clinical trials because of poor pharmacokinetic properties and unexpected side effects.[5] Thus, there is still a great interest in the identification of novel ALR2 inhibitors. Herein, we report the results of a virtual screening (VS) campaign in which an integrated ligand- and receptor-based approach was used. From the structural point of view, ALR2 represents one of the most striking examples of protein induced fit upon ligand binding. Deep inspection of the X-ray ARI/ALR2 complex has clearly shown that the enzyme can adopt at least three main different binding site conformations depending on the bound ligand.[6] Thus, sorbinil (**1**), IDD594 (**2**), and tolrestat (**3**) conformations exist (see Supporting Information). In particular, the ARL2 binding site can be divided into two different regions having diverse plasticity properties. The first, which could be considered the most rigid one, is named "anion binding pocket" and is made up of the catalytic site and the flanking cofactor. Conversely, the second, known as specificity pocket, displays a high degree of flexibility, especially regarding V297-L300, W219, C303, and Y309 residues. Thus, given the intrinsic flexibility displayed by ALR2 for the VS calculations, an "in situ cross-docking" (ISCD) approach was adopted to simultaneously address multiple target conformations.[7] The underlying idea behind the use of ISCD method was to take advantage of the existence of different protein conformations that can possibly host structurally heterogeneous compounds, thus increasing the overall success rate. The ISCD approach was performed using the AutoDock4 (AD4) software as the search engine.[8] In this methodology, first reported by Sotriffer and co-workers,[7] the ligand is simultaneously docked into different protein conformations in a single docking run. Basically, the ISCD takes advantage of grid-based docking software in which a single large grid encompassing all protein conformations can be calculated. Interestingly, in 2006, the reliability of this methodology has been assessed using different conformations of ALR2. Specifically, the porcine **1** and **3** and the human **2** conformations of ALR2 were used for this purpose.[9] Through the ISCD approach and AutoDock as the docking program, all the inhibitors, for which the preferential ALR2 conformation and the binding mode were known, were correctly docked. However, although the ISCD approach has already been validated to simultaneously address multiple targets[7] and different conformations of the same target,[9] to the best of our knowledge, this is the first time that this computational technique is used in a real VS campaign. Different from Sotriffer and co-workers, we used the three most divergent X-ray structures of human ALR2 (**1** conformation with code 2PDK,[10] **2** conformation with PDB code 1US0,[10] and **3** conformation with code 2FZD.[10] Figure 1 in Supporting Information depicts the "multiconformational" ALR2 receptor used for docking calculations along with the box enclosing the searched area.

The calculated grids (one for each atom type) were then used to perform a preliminary docking simulation on a set of 141 active ligands against the receptor used in the ISCD. Structures and inhibition data for known inhibitors were downloaded from the BindingDB database.[11] These simulations were useful to probe the performances of the in situ cross-docking using the above-mentioned ALR2 conformations. ISCD results were first analyzed on the basis of the predicted binding free energies (ΔG_{AD4}). These values ranged from −11.22 to −4.24 kcal/mol with an average −8.60 kcal/mol binding free energy. Interestingly, the docked conformations of the positive controls determined by AD4 with the lowest ΔG_{AD4} were in good agreement with the experimental data. Since the ISCD approach allows us to simultaneously dock ligands in different protein conformations, we also

*To whom correspondence should be addressed. For L.M.: phone, +39 (081) 678-619; fax, +39 (081) 678-629; e-mail, lmarinel@unina.it. For E.N.: phone, +39 (081) 678-643; fax, +39 (081) 678-629; e-mail, ettore.novellino@unina.it.

[a] Abbreviations: ALR2, aldose reductase; VS, virtual screening; ARI, aldose reductase inhibitors; ISCD, in situ cross-docking; AD4, Autodock 4.0.

 Published on Web 08/31/2009

considered at which ALR2 conformations the known inhibitors were predicted to preferentially bind and compared these results with the experimental binding modes, when present. Indeed, as already suggested by Sotriffer et al., ISCD was successful in selecting the appropriate binding site conformation from the three different ALR2 conformations. On the basis of these encouraging results, VS calculations employing the ISCD method were performed on the commercially available Maybridge HitFinder database (14 400 compounds). VS results were then sorted on the basis of the predicted ΔG_{AD4} values which ranged from −13.87 to −3.49 kcal/mol. Solutions with a predicted ΔG_{AD4} value higher than the average ΔG_{AD4} calculated for the known active compounds (−8.00 kcal/mol) were discarded. On the basis of this criterion, 7468 compounds were retained, representing almost the 52% of the database. Therefore, a second screening step, using a different methodology, was required in order to identify a reasonable number of potential hits.

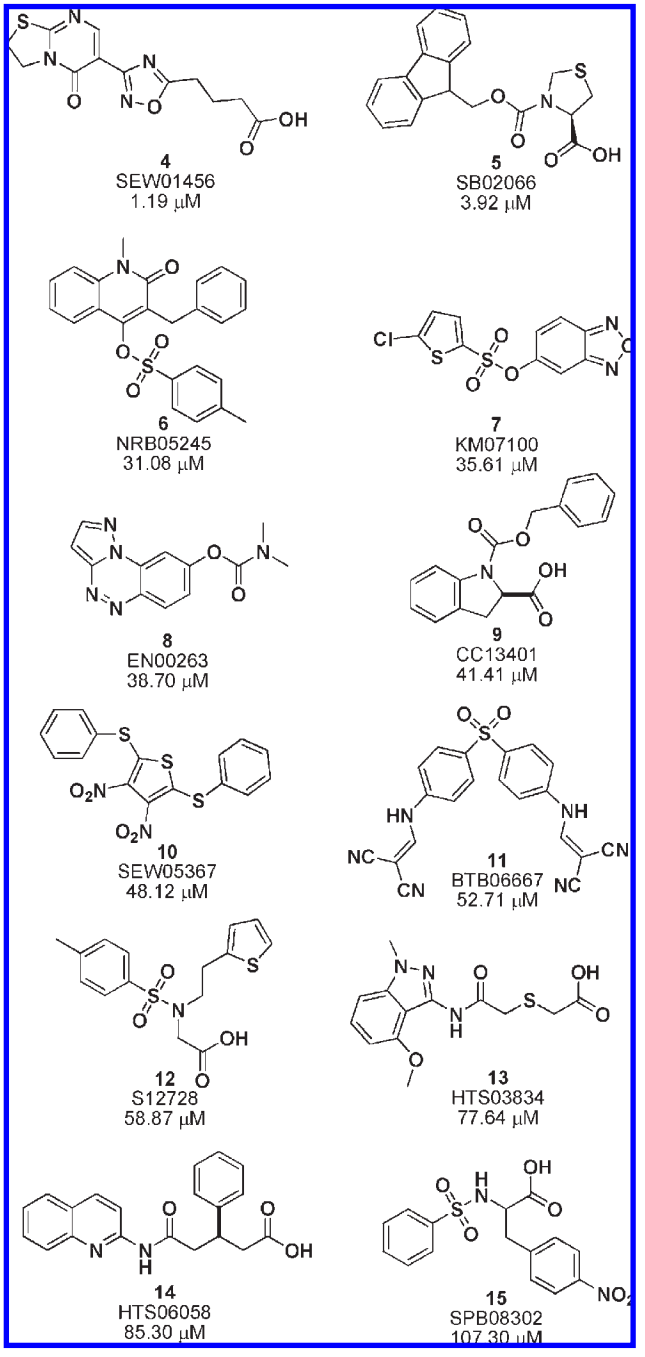
Very recently, Bajorath and co-workers[12] demonstrated that the combination of docking-based VS and similarity search generally improved the compound recall and that in the ALR2 case the MolPrint2D[13] software worked particularly well. Therefore, in our studies the above-mentioned program was used to perform a similarity search on the MayBridge database. In our VS, the so-called atom environment fingerprints (MOLPRINT 2D descriptors) were calculated using the set of 141 known active ligands, which were previously used as positive controls in the ISCD VS, and a library of inactive molecules. For the latter library we decided to use the DUD database which is a set of decoys, with 2950 ligands for 40 different targets including ALR2.[14] The created atom environment fingerprints were then used to rank the Maybridge database according to the similarity to the active and inactive models. MolPrint2D scores ranged from −116.23 to 77.10 with the one having the highest score being the most similar to the active compounds.

Thus, only compounds having a ΔG_{AD4} value lower than −8.00 kcal/mol and a MolPrint2D score higher than 10 (106 ligands) were selected for further analysis. Noteworthy, this MolPrint2D cutoff (> 10) allowed us to also retain compounds that are only vaguely similar to the known ARIs. Finally, as a last criterion of selection, we introduced the visual inspection of the putative best ranking ligand/receptor complexes. In this regard, we decided to discard all the molecules for which AD4 did not predict favorable interactions (i.e., H-bonds) with at least one of the anion binding site residues such as H110, W111, and Y48. Another selection criterion resided in the presence of a $\pi-\pi$ interaction with W111 flanking the selectivity pocket. As a result, 57 candidates were selected and evaluated for their efficacy against ALR2. All the selected hits were tested for their ability to inhibit ALR2 purified from rat lenses. Enzyme inhibition was studied in vitro through a spectrophotometric assay, by monitoring at 340 nm the decrease in absorbance resulting from the oxidation of NADPH to $NADP^+$ catalyzed by the enzyme. The change in pyridine coenzyme concentration/min was determined in a phosphate buffered reaction mixture containing the enzyme extract in the presence of D,L-glyceraldehyde, as the substrate, and NADPH, as the cofactor. The inhibitory activity of the tested hits was assayed by adding them to this mixture, at a routine concentration of 100 μM. Those compounds found to be active were tested at additional concentrations between 10 and 0.1 μM. Tolrestat, a well-known ALR2 inhibitor, was used in the same assay conditions as those for the reference compound. All the products were soluble in the reaction mixture, with the exception of four which, giving rise to a visible precipitate, were discarded. Therefore, only 53 compounds could be experimentally tested for their inhibition potency against ALR2 activity. Twelve out of the 53 soluble compounds were shown to be ARIs, resulting in a VS success rate of 22%.

Chart 1 lists the structures of the active inhibitors and their IC_{50} values, which were deduced from the linear regression analysis of the log dose response curves (see Supporting Information). To exclude any possible nonspecific/promiscuous inhibition of ALR2, we deepened our hit validation, repeating all the assays pertaining the active compounds in the presence of 0.01% Triton X-100, as suggested by Shoichet et al.[15] As none of the observed inhibitory activity was affected by the addition of the nonionic detergent, we can assume the 12 active compounds as effective and viable leads deserving further development.

The binding free energy values (ΔG_{AD4}) calculated for the ligands/ALR2 complexes and the associated MOLPRINT 2D scores are reported in Supporting Information. Interestingly, the differences of the computed ΔG_{AD4} for each ligand in each ALR2 conformation are always below 2 kcal/mol (see Supporting Information), which is the average error obtained in AD4 calibration experiments. Therefore, at least for the 12 active ligands a preferential binding to a specific ALR2 conformation could not be postulated. Nevertheless, for the 12 ARIs identified in this study the **2** conformation is able to produce slightly lower binding free energies. Therefore, the predicted binding poses of most active and structurally diverse ligands in such a protein conformation were further analyzed to explain at the molecular level the reasons behind their inhibitory activity against ALR2 and to give some hints for future lead optimization. Among the newly discovered ARIs, several of them are characterized by the presence of a carboxylic acid moiety (SEW01456, **4**; SB02066, **5**; CC13401, **9**; S12728, **12**; HTS03834, **13**; HTS06058, **14**) with the most active compounds being **4** and **5** having an IC_{50} of 1.19 and 3.92 μM, respectively. Docking results achieved for **4** (Figure 1a) revealed that the ligand carboxylate is well inserted in the anion binding pocket H-bonding with Y48, H110, and W111 side chains and engaging an electrostatic interaction with the nicotinamide moiety of the $NADP^+$ cofactor. On the other hand, the presence of the oxadiazole ring well orients the polycyclic dihydrothiazolopyrimidinone ring in the specificity pocket making favorable hydrophobic contacts with W79, W111, and L300 with the N1 atom of the same ring being at H-bonding distance with Y309 side chain. Interestingly, an analogue of **4** had already been found to be active against ALR2 in another successful VS study reported by Klebe and co-workers.[16] Almost the same ligand/receptor interaction pattern was also found for **5** (Figure 1b) in which the thiazolidine ring optimally orients the attached carboxylate in the anion binding pocket and the fluorene aromatic ring in the specificity pocket where the polycyclic aromatic ring establishes well-oriented charge transfer interactions with W111 and Y309 side chains. Indeed, **4** and **5** do display good inhibitory activity against ALR2 in the low micromolar range. On the other hand, the other carboxylic acid derivatives (**9**, **12**, **13**, and **14**) are still able to establish similar contacts with the enzyme, which would be in contrast with their experimental IC_{50} values. Actually, **9** and **14** were tested as racemates; therefore, it is likely that only one enantiomer is responsible for the inhibitory activity against ALR2. This is further supported by docking calculations showing that only the (*S*)-enantiomer would be able to simultaneously place the

Chart 1. Structures, Labels, and Activities (IC_{50}[a]) of the New ARIs Found in VS Experiments



[a] IC_{50} (95% CL) values represent the concentration required to produce 50% enzyme inhibition.

carboxylate group in the anion site and the aromatic moieties in the selectivity cleft. Therefore, we believe that the lower activity recorded for these ligands might be ascribed to the presence in the racemic mixture of an inactive isomer. We are currently testing this hypothesis by synthesizing the above-described ligands as pure enantiomers and testing their activity against ALR2. For **12** and **13**, we postulate that one of the reasons for the observed low activity could be ascribed to the presence of electron-donating groups (methoxy and methyl, respectively) on the W111 interacting moieties, which negatively affect the established $\pi-\pi$ stacking interactions.

Compounds **6** (NRB05245), **11** (BTB06667), and **15** (SPB-08302) are particular intriguing, as they are structurally

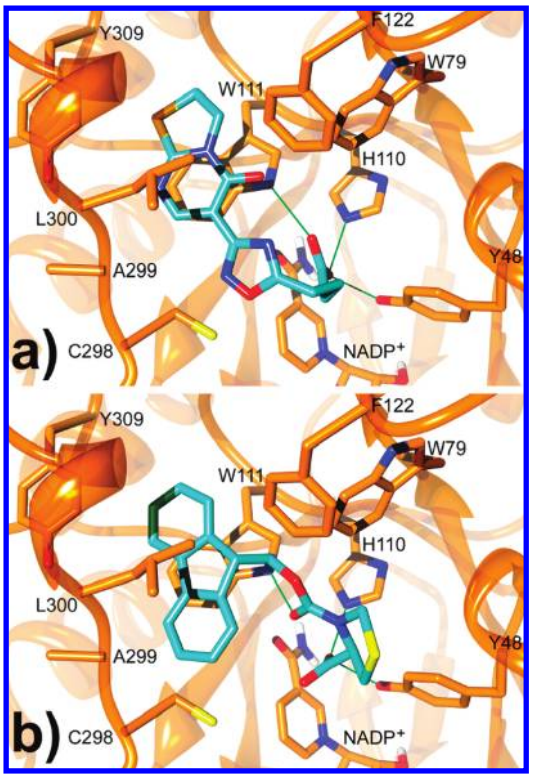


Figure 1. Docked conformations of **4** (a) and **5** (b) in ALR2 structure. Hydrogens are omitted for clarity. Ligand carbon atoms are displayed in cyan, and key binding site residues are displayed as orange sticks. Hydrogen bonds are represented by green lines.

unrelated to the inhibitors known so far. These ligands insert their SO_2 group in the anion binding pocket, while the aromatic moiety is placed in the specificity pocket. In all the docked conformations of **15** only the SO_2 group is able to positively H-bond with W111 and C298. Unfortunately, the presence of the bulky phenyl ring attached to the SO_2 group sterically hampers the direct formation of H-bonds with H110 and Y48 in the anion binding site. In such a binding conformation, the carboxylate group points toward a rather lipophilic enzyme region establishing unfavorable contacts. Therefore, the presence of ineffective ligands/receptor contacts might explain the activity of **15** which is the least potent ARI found in this study (IC_{50} = 107.30 μM). Conversely, **6** is the most potent non carboxylic acid compound presented herein (IC_{50} = 31.08 μM). As happened for **15**, the presence of a tolyl substituent directly attached to the SO_2 group could indeed determine less productive interactions with the anion binding site (Figure 2a). Conversely, the quinolinone ring optimally orients the pendent benzyl group in the specificity pocket sandwiched between W111 and L300 side chains. We believe that removal of the tolyl substituent and substitution of the benzyl ring with electron-withdrawing groups would further enhance the activity of such a lead. Different from the other SO_2 containing leads, AD4 places the benzoxadiazole ring of **7** (KM07100) in the anion pocket H-bonding with W111, H110, and Y48 side chains (Figure 2b). The ligand aromatic ring (5-Cl-thiophene) is lodged in the specificity pocket forming a $\pi-\pi$ stacking interaction with W111 and hydrophobic contacts with L300 and Y309. The extension of aromaticity of the specificity pocket interacting group might result in more potent ARIs.

Regarding **8** (EN00263), AD4 calculated a well clustered binding pose in which the dimethylcarbamate moiety H-bonds

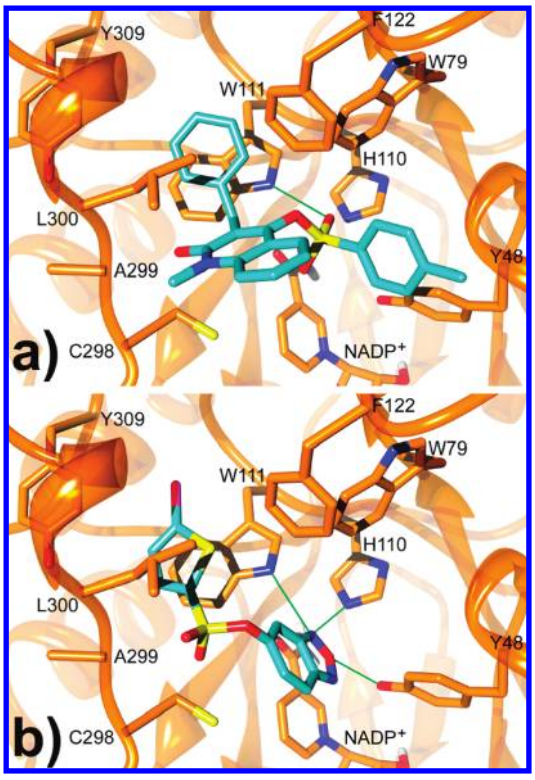


Figure 2. Docked conformations of **6** (a) and **7** (b) in ALR2 structure. Hydrogens are omitted for clarity. Ligand carbon atoms are displayed in cyan, and key binding site residues are displayed as orange sticks. Hydrogen bonds are represented by green lines.

with its carbonyl oxygen to the W111 side chain while the benzopyrazolotriazine nucleus fits in the specificity pocket engaging π-stacking and hydrophobic interactions with W111, F115, Y309, and L300 (see Supporting Information). Thus, the inhibitory potency (IC_{50} = 38.70 μM) of this compound has to be ascribed to the tight interactions of the polycyclic aromatic ring with the specificity pocket rather than to the less productive contacts of the carbamate moiety with the anion binding site. Therefore, a substitution of the latter group with negatively charged groups (i.e., carboxylate) would definitively lead to better inhibitory potencies. As for **10** (SEW05367), while the dinitrothiophene central core is inserted in the anion binding site, forming with both its NO_2 groups H-bonds with Y48, H110, and C298 side chains, one of the pendent thiophenyl substituents is inserted in the specificity cleft engaging hydrophobic contact and π-stacking interactions with W79, W111, F122, and L300 (see Supporting Information).

In conclusion, virtual screening of the publicly available Maybridge HitFinder database was performed to seek new ALR2 inhibitors. The ISCD methodology was employed to simultaneously calculate the binding mode of the inspected ligands in three different ALR2 conformations to take advantage of the ALR2 demonstrated capacity to flexibly adapt to different inhibitors and substrates. The major improvement of the ISCD method, with respect to other docking methodologies that include protein flexibility during the calculation,[17] is that differences among the enzyme structures need not to be confined to small side chain reorientations. The top ranked compounds coming from the ISCD-based VS were then subjected to a similarity search to prioritize the visual inspection of the ligand/enzyme complexes. Finally, 53 compounds were assayed in vitro for inhibition of ALR2. Notably, 12 candidates were found to inhibit this enzyme in the micromolar range including two ligands having an IC_{50} below 3 μM. Most important, six new chemical entities that are structurally unrelated to the known ARIs have been identified. The structural diversity of these inhibitors has provided valuable alternative series for ongoing medicinal chemistry optimization.

Supporting Information Available: Additional figures, charts, and table; molecular modeling methods; and experimental procedure for enzymatic inhibition assays. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) King, H.; Aubert, R. E.; Herman, W. H. Global burden of diabetes, 1995−2025. Prevalence, numeric estimates, and projections. *Diabetes Care* **1998**, *21*, 1414–1431.
(2) *Textbook of Diabetes*; Pickup, J. C., Williams, G., Eds.; Blackwell Science Ltd.: Oxford, U.K., 1997; Vol. 1, Chapter 42.
(3) Ho, E. C. M.; Lam, K. S. L.; Chung, S. S. M. Aldose reductase deficient mice are alleviated from depletion of GSH in the peripheral nerve and MNCV deficit associated with diabetes. *Diabetes* **2001**, *50*, 59.
(4) Chung, S. S.; Chung, S. K. Genetic analysis of aldose reductase in diabetic complications. *Curr. Med. Chem.* **2003**, *10*, 1375–1387.
(5) Costantino, L.; Rastelli, G.; Vianello, P.; Cignarella, G.; Barlocco, D. Diabetes complications and their potential prevention: aldose reductase inhibition and other approaches. *Med. Res. Rev.* **1999**, *19*, 3–23.
(6) Zentgraf, M.; Steuber, H.; Koch, C.; La Motta, C.; Sartini, S.; Sotriffer, C. A.; Klebe, G. How reliable are current docking approaches for structure-based drug design? Lessons from aldose reductase. *Angew. Chem., Int. Ed.* **2007**, *46*, 3575–3578.
(7) Sotriffer, C. A.; Dramburg, I. "In situ cross-docking" to simultaneously address multiple targets. *J. Med. Chem.* **2005**, *48*, 3122–3125.
(8) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.* **2007**, *28*, 1145–1152.
(9) Zentgraf, M.; Fokkens, J.; Sotriffer, C. A. Addressing protein flexibility and ligand selectivity by "in situ cross-docking". *ChemMedChem* **2006**, *1*, 1355–13359.
(10) (a) Howard, E. I.; Sanishvili, R.; Cachau, R. E.; Mitschler, A.; Chevrier, B.; Barth, P.; Lamour, V.; Van Zandt, M.; Sibley, E.; Bon, C.; Moras, D.; Schneider, T. R.; Joachimiak, A.; Podjarny, A. Ultrahigh resolution drug design I: details of interactions in human aldose reductase−inhibitor complex at 0.66 Å. *Proteins* **2004**, *55*, 792–804. (b) Steuber, H.; Zentgraf, M.; Gerlach, C.; Sotriffer, C. A.; Heine, A.; Klebe, G. Expect the unexpected or caveat for drug designers: multiple structure determinations using aldose reductase crystals treated under varying soaking and co-crystallisation conditions. *J. Mol. Biol.* **2006**, *363*, 174–187. (c) Steuber, H.; Heine, A.; Podjarny, A.; Klebe, G. Merging the binding sites of aldose and aldehyde reductase for detection of inhibitor selectivity-determining features. *J. Mol. Biol.* **2008**, *379*, 991–1016.
(11) Liu, T.; Lin, Y.; Wen, X.; Jorrisen, R. N.; Gilson, M. K. BindingDB: a Web-accessible database of experimentally determined protein−ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198–D201.
(12) Tan, L.; Geppert, H.; Sisay, M. T.; Gutschow, M.; Bajorath, J. Integrating structure- and ligand-based virtual screening: comparison of individual, parallel, and fused molecular docking and similarity search calculations on multiple targets. *ChemMedChem* **2008**, *3*, 1566–1571.
(13) Bender, A.; Mussa, H. Y.; Glen, R. C.; Reiling, S. Molecular similarity searching using atom environments, information-based feature selection, and a naive Bayesian classifier. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 170–178.MOLPRINT 2D. http://www.molprint.com.
(14) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49*, 6789–6801.
(15) (a) 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. (b) McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K. A specific mechanism of non-specific inhibition. *J. Med. Chem.* **2003**, *46*, 4265–4272.
(16) Kraemer, O.; Hazemann, I.; Podjarny, A. D.; Klebe, G. Virtual screening for inhibitors of human aldose reductase. *Proteins* **2004**, *55*, 814–823.
(17) Osterberg, F.; Morris, G. M.; Sanner, M. F.; Olson, A. J.; Goodsell, D. S. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 34–40.