J Comput Aided Mol Des (2008) 22:897–906
DOI 10.1007/s10822-008-9224-5

Structure-based design of Aurora A & B inhibitors

Anders Poulsen · Anthony William · Angeline Lee · Stéphanie Blanchard · Eeling Teo · Weiping Deng · Noah Tu · Evelyn Tan · Eric Sun · Kay Lin Goh · Wai Chung Ong · Chee Pang Ng · Kee Chuan Goh · Zahid Bonday

Received: 27 March 2008 / Accepted: 8 June 2008 / Published online: 24 June 2008

Abstract The Aurora family of serine/threonine kinases are mitotic regulators involved in centrosome duplication, formation of the bipolar mitotic spindle and the alignment of the chromosomes along the spindle. These proteins are frequently overexpressed in tumor cells as compared to normal cells and are therefore potential therapeutic oncology targets. An Aurora A high throughput screen revealed a promising sub-micromolar indazole-benzimidazole lead. Modification of the benzimidazole portion of the lead to a C2 linker with a phenyl ring was proposed to achieve novelty. Docking revealed that a conjugated linker was optimal and the resulting compounds were equipotent with the lead. Further structure-guided optimization of substituents on the 5 & 6 position of the indazole led to single digit nanomolar potency. The homology between the Aurora A & Aurora B kinase domains is 71% but their binding sites only differ at residues 212 & 217 (Aurora A numbering). However interactions with only the latter residue may be used for obtaining selectivity. An analysis of published Aurora A and Aurora B X-ray structures reveals subtle differences in the shape of the binding sites. This was exploited by introduction of appropriately sized substituents in the 4 & 6 position of the indazole leading to Aurora B selective inhibitors. Finally we calculate the conformational energy penalty of the putative bioactive conformation of our inhibitors and show that this property correlates well with the Aurora A binding affinity.

A. Poulsen (✉) · A. William · A. Lee · S. Blanchard · E. Teo · W. Deng · N. Tu · E. Tan · E. Sun · K. L. Goh · W. C. Ong · C. P. Ng · K. C. Goh · Z. Bonday
S*BIO Pte Ltd., 1 Science Park Road, #05-09 The Capricorn, Singapore Science Park II, 117 528 Singapore, Singapore
e-mail: anders_poulsen@sbio.com

Keywords Aurora kinase · Inhibitor · Structure-based design · Docking · Subtype selectivity · Bioactive conformation · Conformational energy penalty

Introduction

Aurora kinases are a highly conserved family of mitotic serine-threonine kinases that play important roles in mitotic spindle formation, centrosome maturation and segregation of chromosomes into daughter cells. These kinases have been shown to be over-expressed in several solid tumors including breast, ovary, prostate, pancreas and colorectal cancer [1, 2].

There are three mammalian aurora kinase genes, encoding Aurora A, B and C. Intense investigation has focused mainly on Aurora A and B which are expressed ubiquitously in proliferating cells and are closely associated with carcinogenesis [3]. Aurora A localizes to the duplicated centrosomes and has been implicated in centrosome maturation and separation [4]. Aurora B is a chromosomal passenger protein and localizes to the centromeric regions of chromosomes in the early stages of mitosis [5]. During mitosis, Aurora B is required for phosphorylation of histone H3 on serine 10, and this might be important for chromosome condensation [6]. Further, Aurora B is also required for correct chromosome alignment and segregation and is one of the upstream kinases needed for spindle-checkpoint function and cytokinesis [7]. Not much is known about Aurora C, a chromosomal passenger protein that binds directly to INCENP and survivin in vitro [8, 9].

Aurora kinases comprise mainly two domains: a regulatory domain in the NH2 terminus and a catalytic domain in the COOH terminus [10]. The regulatory domain is

diverse, whereas the catalytic domain shares >70% homology among Aurora A, Aurora B, and Aurora C. Phosphorylation at the threonine site within the catalytic domain is required for the kinase activity. Crystal structures of the Aurora A and Aurora B catalytic domain have been resolved and have been shown to contain two lobes in their catalytic domain joined by a hinge, like other serine-threonine kinases [11, 12]. The N lobe is responsible for positioning the ATP phosphate group through an αC helix, whereas the activation loop within the C lobe is able to harbor substrates. Only when the two lobes arrive at a certain conformation can the Aurora kinases fulfil their kinase functions.

The first evidence to implicate this family of kinases in tumorigenesis was the discovery of human Aurora A as the product of gene BTAK (breast tumor amplified kinase, also named STK15) on chromosome 20q13, a region that is commonly amplified in primary breast tumors and colorectal cancers [13]. Subsequent studies identified other tumor types, including pancreatic, ovarian and gastric tumors, in which Aurora kinases were amplified or overexpressed [14, 15]. Aurora kinases therefore emerged as candidate targets for anti-tumor therapy. Hence, we embarked on a drug discovery program in search of inhibitors of Aurora kinases. An in vitro enzyme assay was set up to screen for small molecules that inhibit the enzymatic activity of recombinant human Aurora A kinase. A high-throughput screen using this assay was performed against a library of ~206,600 compounds and identified new chemotypes for optimization [16]. Subsequent derivatives were also tested in similar in vitro enzyme assays to test for selectivity versus Aurora B kinase. The initial project objective was to prepare Aurora A inhibitors with or without selectivity over Aurora B. Later it was realized that Aurora A and B inhibitors have distinct modes of action and possibly distinct toxicities. Hence the project objectives subsequently shifted to making selective Aurora A as well as selective Aurora B inhibitors.

Experimental and computational methods

Conformational search, force fields and solvation model

The molecules were built using Maestro 8.0.308 [17] or converted to 3D structures from the 2D structure using LigPrep version 2.1.207 [17]. Basic amines were protonated as in aqueous solution at physiological pH. The conformational space was searched using the Monte Carlo (MCMM) method [18] as implemented in MacroModel version 9.5.207 [17]. All heavy atoms and hydrogens on heteroatoms were included in the test for duplicate conformations. All rotatable single bonds were included in the conformational search and all aliphatic rings were ring-opened and quaternary atoms were allowed to invert. Each search was continued until the global energy minima were found at least ten times. The energy minimizations were carried out using the truncated Newton conjugate gradient algorithm (TNCG) and the MMFFs force field [19] as implemented in MacroModel. Default parameters were used. The conformational searches were done for aqueous solution using the generalized Born/solvent accessible surface (GB/SA) continuum solvation model [20, 21].

Docking

The Aurora A (entry 1MQ4 [12]) and Aurora B (entry 2BFY [22]) X-ray structures were accessed from the protein data bank (PDB). The protein structures were prepared using the protein preparation wizard in Maestro with standard settings. Grids were generated using Glide version 4.5.208 [17] following the standard procedure recommended by Schrödinger. A hydrogen bond constraint was included in the grid files. The backbone NH of the residue Ala213 (Aurora A) and Ala173 (Aurora B) must form a hydrogen bond to the ligand. This is the hydrogen bond donor in the hinge region of the kinase to which the 1 position of the purine of ATP forms a hydrogen bond. The compounds were docked using Glide with standard settings in both standard and extra precision mode. The docked poses discussed in this paper were not necessarily the highest scoring, but were selected as the highest scoring pose with a reasonable conformation and binding mode as judged by the modeler.

Compound selection

Compounds were selected for synthesis by two criteria. Either for Structure Activity Relationship (SAR) reasons by the synthetic chemists or they were selected by the modeler. Compounds selected by the modeler were not necessarily the highest scoring, but were selected if they docked with a reasonable conformation and binding mode, i.e. the poses had a low conformational energy and good lipophilic and electrostatic interactions with the protein. Compounds designed by the synthetic chemists had to follow some general guidelines determined by the modeler. These were: Ethenyl and ethynyl linkers are preferred. No substituents in the 2 and 6 position of the styrene or equivalent position of other ring systems. Solubility tags may be attached to the 3 or 4 position of the styrene. The 3 and 5 positions of the styrene may not both be substituted. No substituents on the 1 or 7 position of the indazole. Only flat (coplanar) preferably lipophilic substituents allowed in the 4 position of the indazole. Larger lipophilic preferably non-flat substituents allowed in the 5 position. Only methyl or chloride allowed in the 6 position of the indazole. This

was later changed to substituents not larger than methylpiperazine when the first Aurora B X-ray structure was released from PDB. Thus structure-based design was used for setting general guidelines for the synthesized compounds as well as selecting specific substituents. A few compounds were synthesized against the advice of the modeler to challenge the modeling of which compound **4** is an example.

Calculation of the conformational energy penalty

The docked conformation was minimized with MacroModel using flat bottom Cartesian constraint with a half width of 1.0 Å and the default restraining force constant of 100 kJ/mol. This allows the docked conformations to relax (adjust) to the MMFFs force field. Without the relaxation the energy calculated by MMFFs would be meaningless. This relaxation does not change the conformation as RMS between the docked and relaxed structures are <0.1 Å. The conformational energy penalty of the docked conformations was calculated by subtracting the internal (steric) energy of the preferred conformation in aqueous solution (i.e. the energy of the global energy minimum in solution excluding the hydration energy) from the calculated energy of the docked conformation. Since the conformational ensemble was represented by only the global energy minima, entropy effects have not been taken into account. For flexible molecules this leads to an underestimation of the energy penalty. A limit of 3 kcal/mol (12.6 kJ/mol) for acceptable energy penalties was imposed as recommended by Boström et al. [23].

Calculation of physical properties

The 2D structures were converted to 3D using LigPrep with the -qik option. This option neutralizes the output structures as appropriate for QikProp input. Physical properties were calculated using QikProp version 3.0.207 [17] with standard settings.

DFT calculations

The program Jaguar version 7.0.207 [17] was used for B3LYP DFT energy minimization. The chosen basis set was 6-311G**++. The PBF [24] solvation model was used with water as solvent. Accuracy level was set to "Accurate" and grid density was set to "Fine", otherwise default parameters were used.

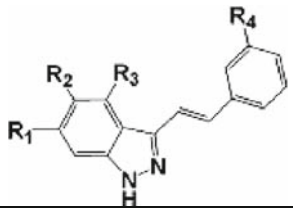
Aurora A assay

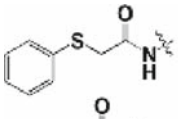
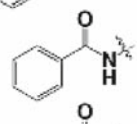
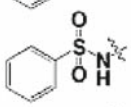
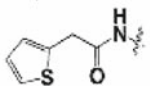
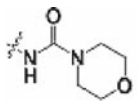
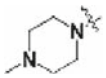
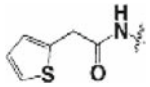
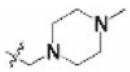
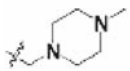
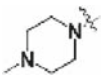
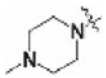
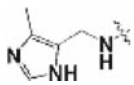
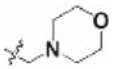
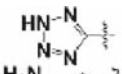
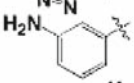
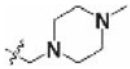
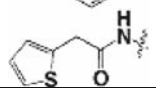
The recombinant Aurora A kinase is a hexahistidine-tagged full-length protein expressed in *Hi-5* insect cells using the expression plasmid pFBHTb (Invitrogen, USA). The activity assay uses the PKLight^{TM} HTS Protein Kinase Assay Kit from Cambrex Corporation (New Jersey, USA). This assay measures the consumption of ATP and is based on the bioluminescent detection of the ATP remaining in the wells after the kinase reaction. The reaction mixture consisted of 10–15 nM of recombinant kinase and 5.0 μM of ATP (its Km value) in 40 μL assay buffer (25 mM Hepes pH 7.5, 25 mM KCl, 5 mM $MgCl_2$, 5 mM β-glycerophosphate, 1 mM dithiothreitol). The affinity-purified kinase undergoes robust autophosphorylation and no exogenous substrate is added. The compounds were tested at 8 concentrations prepared from 4-fold serial dilution starting at 10 μM. The reaction was incubated at room temperature for 2 h. 20 μL of PKLight^{TM} ATP detection reagent was added and the reaction was incubated for another 10 min. Luminescence signals were detected on a multi-label plate reader (Victor^2 V 1420, Perkin-Elmer). IC_{50} values were generated from the raw data using the Prism 4.0 software (GraphPad Software Pte Ltd). All IC_{50} values are the mean from at least 2 independent experiments. The historical coefficient of variance (CV) is less than 55%, with the lowest and highest IC_{50} values differing by less than 3-fold except for compounds **1**, **9**, **12** and **14** (Table 1) .

Aurora B assay

The recombinant Aurora B kinase is a GST-tagged full-length protein expressed in *Hi-5* cells using the expression plasmid pFB-GST (Invitrogen, USA). The activity assay is based on the principles of time-resolved fluorescence (TRF). The reaction mixture consisted of 20–40 nM of enzyme, 0.5 μM of biotinylated substrate (Cell Signaling Technology, Inc (CST) Cat#1300) and 40 μM of ATP (its Km value) in 50 μL assay buffer (CST Kinase Buffer Cat# 9802). The compounds were tested at 8 concentrations prepared from 4-fold serial dilution starting at 10 μM. The reaction was incubated at room temperature for 1 hr. 25 μL of STOP buffer (60 mM Hepes pH 7.5, 60 mM EDTA) was added and 30 μL quenched mixture was transferred to Neutravidin coated plates (Cat#15402B, Pierce). After incubation at room temperature (RT) for 30 min, the plates were washed with 4 × 80 μL Tris-buffered saline with 0.5% Tween-20 (TBST). About 50 μL of antibody mixture (CST antibody Cat# 5070 and PerkinElmer antibody Cat# AD0124) was added and the reaction incubated for 30 min, followed by washing with 4 × 80 μL TBST. About 50 μL of Dissociation-enhanced lanthanide fluorescence immunoassay (DELFIA) Enhancement Solution (PerkinElmer Cat#4001-0010) was added and the reaction incubated for 5 min. TRF signals were detected on a multi-label plate reader (Victor^2 V 1420, PerkinElmer). IC_{50} values were generated from the raw data using the Prism 4.0 software

Table 1 SAR table



Cpd.	AurA IC_{50} (μM)	CV %	AurB IC_{50} (μM)	CV %	R1	R2	R3	R4
1	0.21	108.5	0.14	10.1		See Figure 1		
2	0.37	42.5	1.7	12.9		See Figure 1		
3	>10		N.A.			See Figure 1		
4	0.31	7.0	N.A.			See Figure 1		
5	0.12	6.2	0.34	8.3		See Figure 1		
6	1.2	11.8	0.22	12.9		See Figure 1		
7	0.015	12.0	N.A.			See Figure 1		
8	0.017	43.1	0.078	20.1		See Figure 1		
9	0.18	88.0	N.A.		H	NH2	H	H
10	0.030	35.9	N.A.		Me		H	H
11	0.15	52.5	N.A.		Me		H	H
12	0.074	162.9	N.A.		Me		H	H
13	0.009	50.4	0.036	69.2	Me		H	
14	1.7	66.1	0.18	7.8		NH2	H	H
15	0.020	14.2	0.029	4.9	Cl		H	
16	3.2	2.2	0.30	2.3	H	NH2	Me	
17	0.75	14.3	0.28	12.9	H	NO2		H
18	>10		0.50	10.0			H	H
19	0.26	13.3	0.065	3.2	H	CN	Me	H
20	0.13	5.7	0.036	19.9	H	CN	Me	
21	0.16	4.6	0.54	2.6	H		Me	H
22	0.34	6.2	0.031	0.0	H		Me	
23	0.018	12.8	0.047	9.0	Me		H	H

Activities are in micro molar (uM). The Markush structure in this header defines the substiturens R1-R4. A detailed description of the synthesis can be found in Blanchard et al. [16]

Cpd.: Compound; CV: Historical Coefficient of Variance; N.A.: Not available

(GraphPad Software Pte Ltd). All IC_{50} values are the mean from at least 2 independent experiments. The historical coefficient of variance is less than 25%, with the exception of compound **13** for which it is 69%. The lowest and highest IC_{50} values differ by less than 1.5-fold except for compound **13** where the difference is 3.5-fold (Table 1).

Results and discussion

Selection of linker

A high throughput screen (HTS) revealed indazole-benzimidazole **1** (Fig. 1 and Table 1) to be a potent Aurora A inhibitor. Due to its relatively small size (354 g/mol) and good ligand efficiency (0.34 kcal/mol/heavy atom) a medicinal chemistry optimization program was activated. A literature search revealed that a number of closely related compounds were already published [25]. The docking of compound **1** into Aurora A is shown in Fig. 2. Compound **1** forms 3 hydrogen bonds to the backbone of the hinge region of Aurora A, one hydrogen bond to the carbonyl oxygen of residue Glu211 and a hydrogen bond to both the carbonyl oxygen and the NH of residue Ala213. These are highlighted in yellow in Fig. 2. The preferred bioisostere of an imidazole is an amide and such a replacement would conserve the hydrogen bond formed to the carbonyl of Ala213. Compounds of this type were already reported in the literature but further analysis of the structure suggested novelty could be obtained if the imidazole portion of the benzimidazle was replaced with a C2 linker other than amide and if the indazole had 3 substituents [16]. A closer inspection of the carbonyl-Ala213 hydrogen bond revealed that the geometry was poor for optimal hydrogen bonding ($\theta = 162°$ and $\Phi = 58°$) and therefore probably did not contribute much to the binding affinity. The imidazole part of compound **1** is sandwiched between Leu139 in the top of the binding pocket and

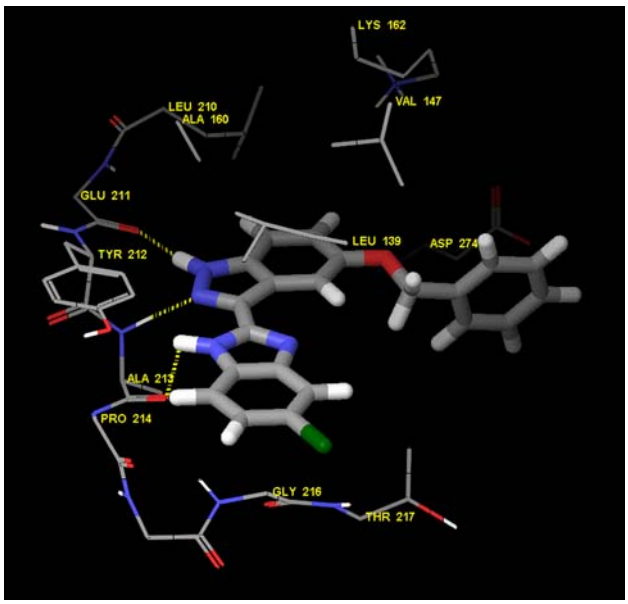


Fig. 2 Tube representation of Compound **1** docked into Aurora A. Hydrogen bonds are displayed with dotted yellow lines. The residues in the binding site are displayed in thin lines with the alpha carbons labeled with residue type and number

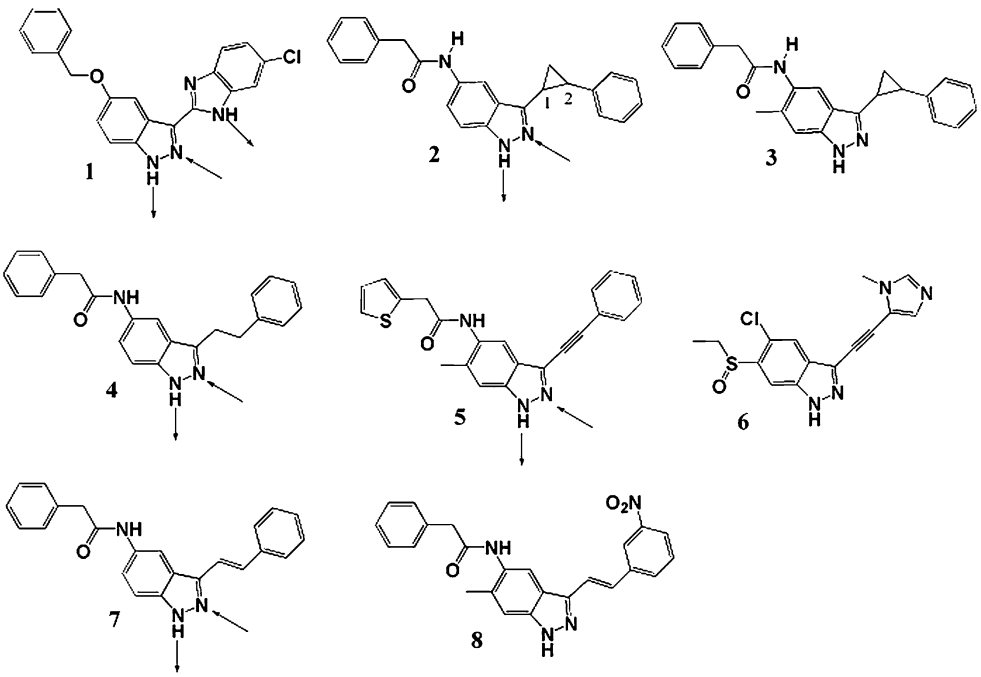


Fig. 1 The lead structure identified during the Aurora A high throughput screen is compound **1**. The hydrogen bonds formed to the protein backbone of the hinge region are shown as arrows. The proposed imidazole replacements have one hydrogen bond donor less as shown in Compounds **2**, **4**, **5** and **7**

Leu263 in the bottom of the binding pocket as seen in Fig. 2. It was hypothesized that a replacement of the imidazole for a hydrophobic linker without hydrogen bonding capability would retain activity. Four different linkers were proposed and compounds incorporating these were synthesized; cyclopropyl linker (compounds **2** and **3**), ethyl linker (compound **4**), ethyne linker (compounds **5** and **6**) and ethenyl linker (compounds **7**–**22**).

The first linker proposed was the cyclopropyl linker, which has 2 chiral centers, hence 4 isomers are possible: 2 cis isomers (1R2S & 1S2R) and 2 trans isomers (1S2S & 1R2R). The 2 cis isomers could not be docked into the Aurora A binding site by any of the docking algorithms used. Of the 2 trans isomers the 1R2R docked well using Glide SP, but had a rather high conformational energy penalty (compounds **2** and **3**, Table 2). Interestingly the 1S2S enantiomer docked better than the 1R2R using Glide XP and the conformational energy penalty was low for compound **2** but high for **3**. Although the cyclopropyl compounds had two drawbacks (2 chiral centers and a putative bioactive conformation with rather high conformational energy for the trisubstituted indazole **3**) we concluded that these compounds were worthwhile synthesising. Unfortunately the Aurora A activity of compound **2** was mediocre and compound **3** was not active at the highest concentration tested (Table 1).

Table 2 Docking scores and conformational energy penalties for the poses docked into Aurora A

Compound	Glide XP docking score	Glide XP conformational energy (kJ/mol)	Glide SP docking score	Glide SP conformational energy (kJ/mol)
1	−13.0	−0.2	−9.3	7.4
2[a]	−14.1	4.3	−7.9	27.3
3[a]	−13.4	18.8	−9.4	37.0
4	−13.3	22.5	−8.9	20.9
5	−14.3	0.4	−9.5	10.8
6	−13.6	18.8	−9.2	13.1
7	−13.5	2.7	−9.4	8.3
8	−13.8	6.4	−9.5	7.8
9	−12.3	3.0	−8.9	7.9
10	−12.9	9.3	−9.1	9.5
11	−12.4	10.0	−9.3	10.0
12	−11.0	10.7	−8.3	10.8
13	−13.4	8.6	−9.1	9.4
14	−9.4	74.8	−6.1	61.8
15	−12.2	10.8	−9.6	2.6
16	−13.5	34.7	−8.7	13.2
17	−10.2	95.0	−7.2	44.5
18	−12.3	107.6	−9.2	86.7
19	−13.8	3.0	−9.4	6.6
20	−14.3	11.7	−9.3	10.6
21	−13.8	9.1	−9.4	9.9
22	−15.0	13.0	−9.5	22.8

These are not necessarily the best scoring poses. Compound 23 was included in this paper after the analysis of docking score vs. conformational energy was made

[a] The 1R2S had the highest Glide XP score whereas the enantiomers had the highest Glide SP score

The second linker to be proposed was the ethenyl linker. Compounds incorporating this linker docked very well and the putative bioactive conformations had a low conformational energy penalty. Among the first compounds synthesized incorporating the ethenyl linker was compound **9** with an Aurora A IC_{50} of 180 nM. Two low energy conformations exist for this linker, shown in Fig. 3 for the model compounds **24a** and **24b**. The conformational energies are the same for the two conformations as calculated by DFT whereas the conformation **24a** is somewhat lower in energy calculated with MMFFs (Table 3). For most of the compounds the conformation represented by **24a** scored higher using both Glide SP and Glide XP. Consequently docked poses with this conformation were selected except for compounds **14** and **18** when docked into Aurora A. Only poses with the conformation represented by **24b** were obtained for these compounds.

Two more linkers were proposed represented by compounds **4**–**6**. Compounds incorporating the ethyl or ethynyl linkers docked well, but only the putative bioactive conformations of compounds with ethynyl linkers had a low

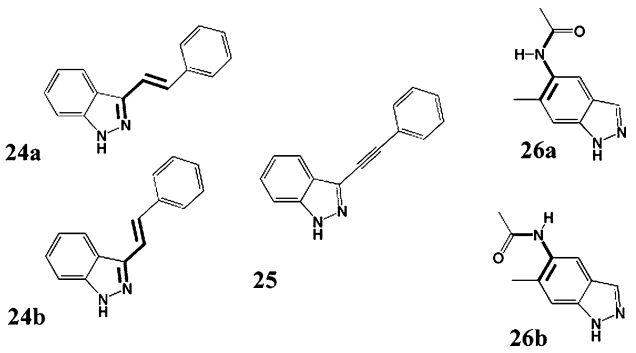


Fig. 3 The model systems used for DFT calculations

Table 3 Conformational energies in kJ/mol for the model compounds shown in Fig. 3

Compound	Torsion MM/DFT	MMFFs Conf E	DFT Conf E
24a	180/169.8	0.0	0.0
24b	32.2/1.6	10.4	0.0
26a	179.4/179.9	0.0	0.0
26b	28.9/74.8	29.9	6.5
26b	0/0	58.6	38.9

The torsion angle is the one highlighted in Fig. 3

conformational energy penalty. We predicted that compounds with ethynyl linkers would be equipotent with the ethenyl analogues and that compounds with a saturated linker should be poor Aurora A inhibitors. Compound **4** is an ethyl linked analogue of **7** and with a Aurora A IC_{50} of 300 nM it is 20 times less active than **7**. However contrary to our expectation compounds incorporating the ethynyl linker were 3–7 times less active than their ethenyl analogues. Figure 4 shows compounds **5** and **23** docked into Aurora A. The docked poses are very similar but the indazole of **5** is rotated slightly compared with **23**. The difference in activity between the ethenyl and ethynyl linkers may be explained by this slight difference in conformation. To investigate this we used the model compounds shown in Fig. 3. The difference in solvent accessible surface area and volume between compounds **24a** and **25** is less than 1% for both properties. This suggests that the extent of hydrophobic contact with the protein is similar for the two linker types. However the difference in calculated logP is 0.1 with the ethenyl linker compounds being the more lipophilic. The difference in solvation energy calculated by DFT with the PBF solvation model is 1.45 kJ/mol. Using Eq. 1 the difference in solvation energy translates into a difference in the binding constant by a factor of 2.

$$-\Delta G = RT \ln K \quad (1)$$

The difference in binding affinity between the two linkers may therefore be due to both the slightly different conformations and difference in solvation energy.

Lead optimization to increase Aurora A activity

The 1 and 2 position of the indazole incorporates the essential hydrogen bond donor and acceptor. The hydrogen in the 7 position points towards the gatekeeper residue Leu210 with which it makes a hydrophobic interaction (Fig. 4). Consequently there is no room for substitution at this position. With the styrene in the 3 position this leaves positions 4–6 for SAR exploration. Docking into Aurora A revealed that the protein could accommodate small substituents like methyl or chloro in the 6 position. Compounds with a methyl are equipotent or slightly less active than compounds without substitutents in this position. The methyl makes van der-Waals contacts with the side chains of the gatekeeper residue Leu210, the catalytic lysine Lys162 and Asp274. Since these favorable interactions do not increase the binding activity, the methyl probably changes the conformation of the substituent in the 5 position and this conformation is less favorable for binding to the protein. The highest scoring poses obtained from both Glide SP and Glide XP generally had the 5 position amide in the conformation shown for the model compound **26b** in Fig. 3. However a substituent in the 6 position penalizes this conformation and the lowest energy minima is that of compound **26a** (See Table 3). Consequently docked poses with the amide orientation like compound **26a** were chosen for compounds with substituents in the 6 position even though poses with different amide conformations were scored higher.

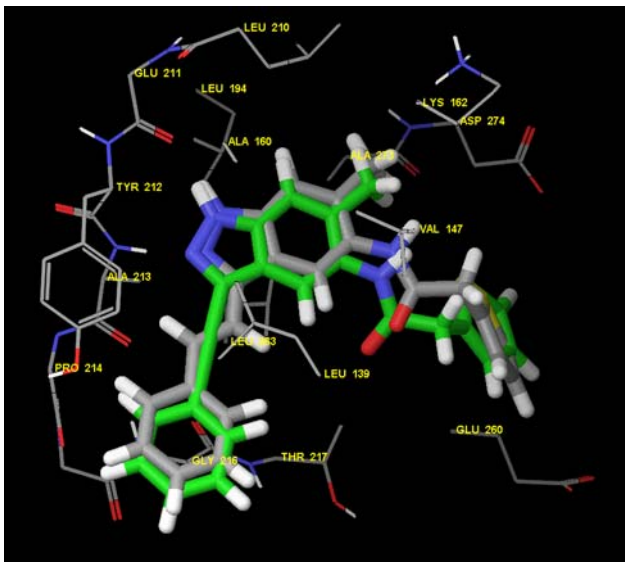


Fig. 4 Compounds **5** (tube with green carbon atoms) and **23** (tube grey carbon atoms) docked into Aurora A. The residues in the binding site are displayed in thin line with the alpha carbon labeled with residue type and number

The length of the substituent in the 5 position was explored. Compound **9** with a amino and compound **11** with a phenylamide are equipotent whereas **7** and **8** with a benzylamide are 10 times more potent. The phenylamide of **11** makes favorable van der-Waals contacts with hydrophobic atoms in the protein, but it also has close contacts to the carboxylic acid of Asp274, which may cancel the favorable interactions. Compounds **7** and **8** have an additional methylene unit between the amide and the phenyl making the side chain more flexible. In the docked poses of **7** and **8** the side chain does not make any close contacts with the protein. In compound **10** the substituent in the 5 position is one sulfur atom longer than that of the twice as potent compounds **7** and **8**. Although **10** is more lipophilic than **7** and **8** (calculated logP of 5.7 compared to 5.1 and 4.8) it has less hydrophobic contact with the protein as the phenyl is positioned in the phosphate binding region of the ATP binding site.

With an IC_{50} of 15 nM **7** has good Aurora A activity, however it is rather lipophilic with a calculated logP of 5.1. To decrease the logP and increase the solubility a solubility tag may be added to the *meta* or *para* position of the phenyl in the styrene. The *para* position points straight out of the binding site towards the solvent and a substituent there will

not interact with the protein (Fig. 4). However a substituent in the *meta* position may interact with the edge of the binding site and the protein surface. Compounds **13** and **15** with solubility tags are less lipophilic with calculated logPs of 4.6 and 4.8, respectively. For **13** the solubility tag makes a favorable interaction with the surface of the protein, resulting in single digit nanomolar affinity.

Lead optimization for Aurora subtype selectivity

A comparison of Aurora A and B X-ray structures revealed that the binding sites only differ by 2 residues. Tyr212 and Thr217 in Aurora A correspond to Phe172 and Glu177 respectively in Aurora B. Tyr212/Phe172 are located in the hinge region, and the side chain makes aromatic–aromatic interactions with the styrene in the inhibitor. The phenolic OH of Tyr212 is not involved in binding to the inhibitors. The distance between the oxygen and the closest carbon of the styrene (the *ortho* position) is 3.9Å, so considering protein flexibility it is not unreasonable to assume that a hydrogen bond may be formed between an inhibitor with a donor or acceptor in this position and Tyr212. However the backbone carbonyls of Ala213 and Pro214 are closer to the *ortho* position of the styrene than the phenol of Tyr212. This probably rules out use of Tyr212 to obtain Aurora A selectivity.

Thr217/Glu177 is a better target for obtaining Aurora subtype selectivity. In another chemical series [26] (data not shown) a basic nitrogen formed a hydrogen bond or salt bridge interaction with Thr217/Glu177 resulting in compounds that were approximately equipotent on Aurora A and B. When the basic nitrogen was changed for an ether oxygen the Aurora A activity remained in the mid-nanomolar range but the Aurora B activity was reduced to >10 μM [27]. While targeting Thr217/Glu177 to obtain Aurora A subtype selectivity proved useful, it is unlikely that this approach would yield Aurora B selective compounds as we observed that the Aurora B active compounds were generally much more promiscuous kinase inhibitors [26].

The 4 position of the indazole is pointing towards Thr217/Glu177 (Fig. 4). We synthesized compound **17** with a 4-methyl-piperazine in the 4 position, but the compound is less active and the subtype selectivity is only a factor of 3 towards Aurora B. We discovered that **16**, with a much smaller and neutral methyl substituent in the 4 position, is also Aurora B selective and equipotent with **17**. Further optimization yielded **22**, a potent Aurora B inhibitor with good physical properties (calculated logP is 4.2 and MW = 437) and 10-fold selectivity over Aurora A. Figure 5 shows compound **22** docked into Aurora B. The residues in the binding site around the glycine rich loop are displayed. Aurora A was superimposed on Aurora B and the same residues are shown for Aurora A. The superimposition

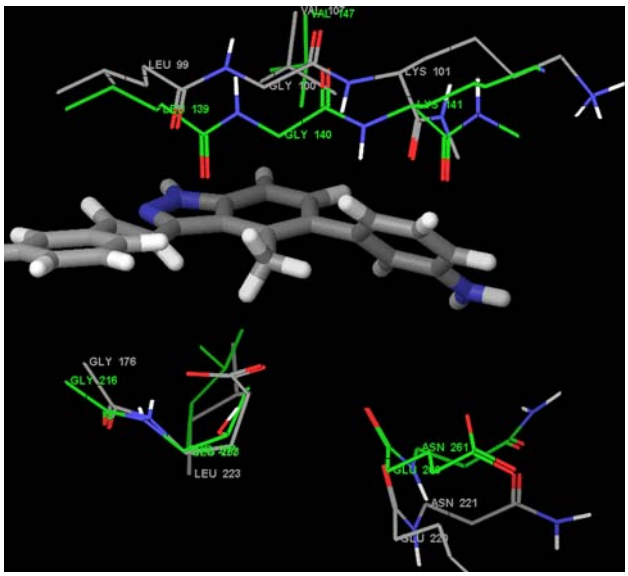


Fig. 5 A: Tube representation of compound **22** docked into Aurora B. Aurora A was subsequently superimposed on Aurora B. Only the residues in the binding site around the glycine rich loop are displayed (Aurora B with grey carbon and Aurora A with green carbon). Note that there is less space in the Aurora A binding site around the 4 position of the indazole

shows that there is less space in the Aurora A binding site around the 4 position of the indazole. This may explain the subtype selectivity of 4-substituted indazoles.

By superimposing all available Aurora A and B X-ray structures (PDB entries 1MQ4, 1MUO, 1OL6, 1OL6, 2BMC, 2C6E, 2J4Z, 2J50, 2NP8, 2BFX, 2BFY & 2VGP) and displaying the molecular surface of the binding site we discovered that the space around the 6 position of the indazole is larger in Aurora B than in Aurora A. There is some structural variation in the Aurora A X-ray structures with the DFG motive in different conformations. The Aurora A structure with the most room around the 6 position of the indazole is 1MQ4 which has the same DFG-in conformation as the Aurora B structures. The Aurora A Docking revealed that the largest substituent in the 6-position tolerated by Aurora B was 4-methyl-piperazine. Figure 6a and b display the surface of Aurora B and Aurora A respectively, with compound **14** docked into the binding site of Aurora B. Compounds **14** and **18** have good selectivity towards Aurora B (10 and >20 fold, respectively), however they are not particularly potent with Aurora B IC_{50}s of 180 nM and 500 nM, respectively.

Prediction of binding affinity

Both the Glide SP and Glide XP scores for the selected poses are displayed in Table 2 together with the conformational energy penalties. In Fig. 7a the Glide score for the docked poses are plotted against the logarithm of the

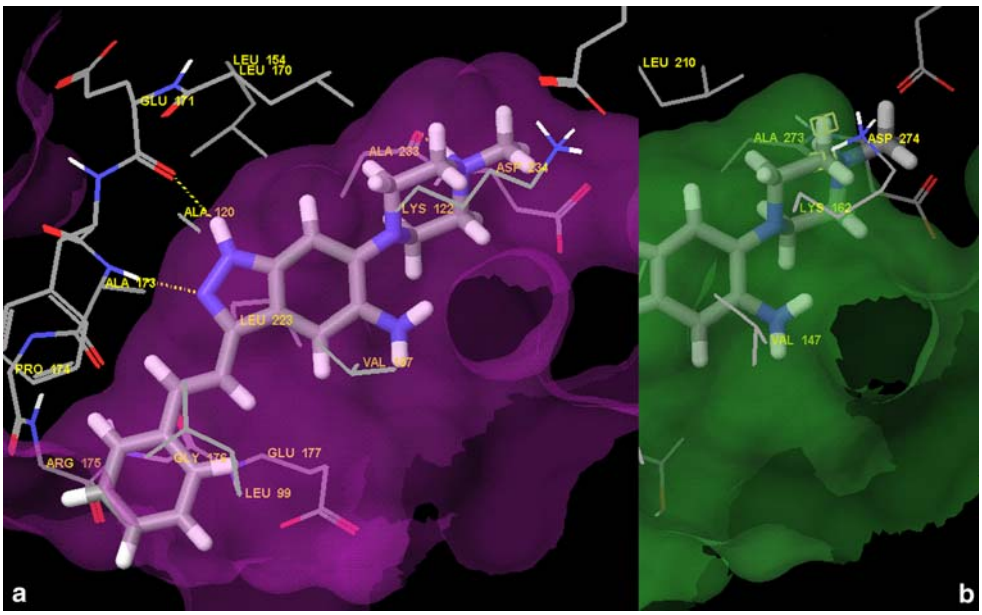


Fig. 6 (**a**) Tube representation of compound **14** docked into Aurora B. Hydrogen bonds are displayed with dotted yellow lines. The residues in the binding site are displayed in thin line with the alpha carbon labeled with residue type and number. The surface of the protein is shown in transparent magenta. (**b**) The Aurora A X-ray structure has been superimposed on Aurora B with the pose from Fig. 5a and the surface of Aurora A shown in transparent green. There is more room in the Aurora B binding site, which may be exploited to obtain Aurora B subtype selectivity

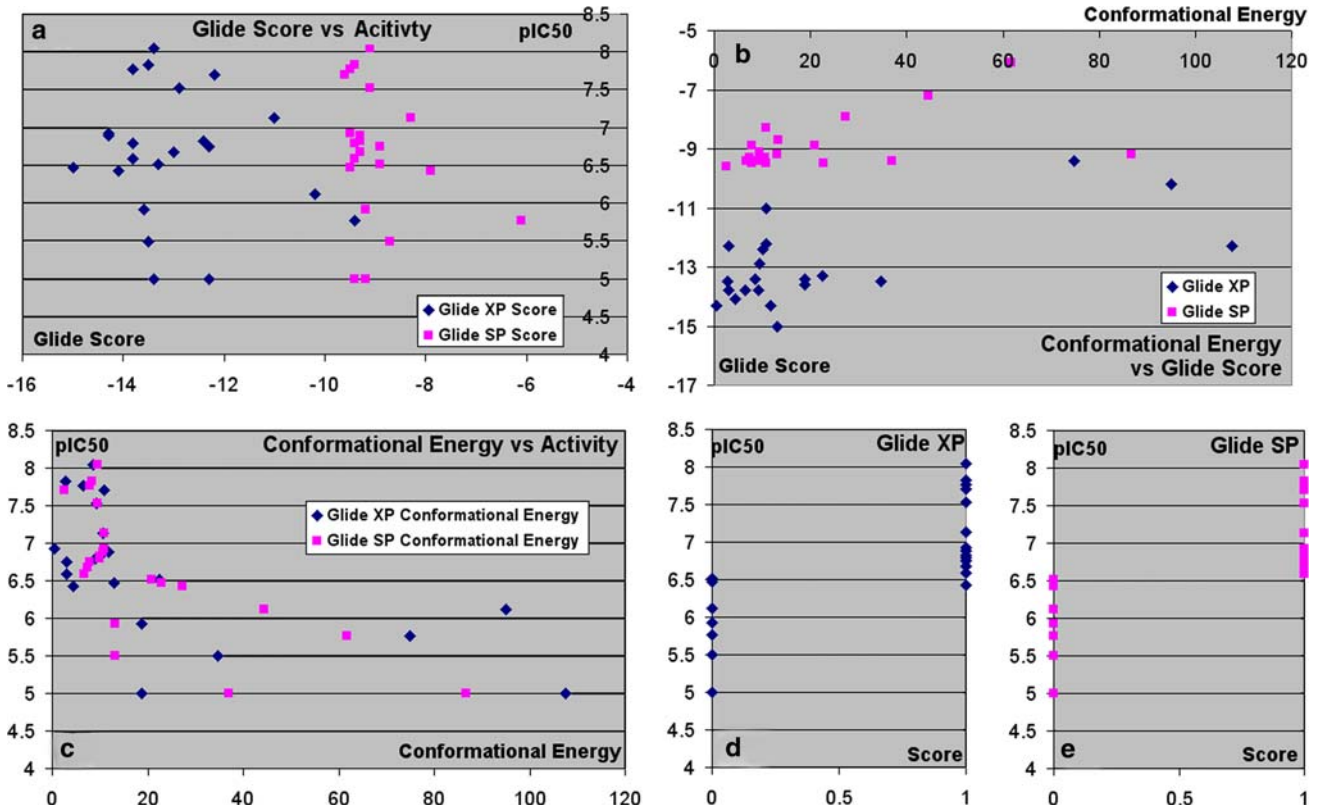


Fig. 7 (**a**) The Glide score (docking score) for the docked poses plotted against the logarithm of the Aurora A binding affinity. (**b**) The conformational energy penalty (kJ/mol) of the docked poses plotted against the Glide score. (**c**) The conformational energy penalty (kJ/mol) of the docked poses plotted against the logarithm of the Aurora A binding affinity. (**d**) and (**e**) The score of the scoring function (described in the "Results and discussion" section "Predicting binding affinity") plotted against the logarithm of the Aurora A binding affinity

Aurora A binding affinity. There is no correlation as the R square is 0.03 for Glide XP and 0.09 for Glide SP. This was expected and several papers have described similar results for various scoring functions [28–31]. In Fig. 7b the conformational energy penalty of the docked poses are plotted against the Glide score. The R square is 0.38 and 0.26 for Glide XP poses and Glide SP poses, respectively. Although the Glide scoring functions are proprietary and

have not been published it was expected to find this correlation (and that it should be better for Glide XP) as the Glide scoring functions is clamed to take conformational energies into account [32–34].

In Fig. 7c the conformational energy penalty of the docked poses are plotted against the logarithm of the Aurora A binding affinity. From visual inspection it is apparent that conformational energy provides some measure of separation between active and less active compounds. The correlation is reasonable with an R square of 0.36 for Glide XP poses and good for Glide SP with an R square of 0.49. A scoring function was constructed by assigning 1 to compounds for which the conformational energy penalty were less than 12 kJ/mol and the Glide XP score were more negative than −10. The remaining compounds were assigned 0. Similar for poses docked with Glide SP but a Glide score cut off value of −8 was used. This score is plotted against the logarithm of the Aurora A binding affinity in Fig. 7d and e. This score gives a very good separation of compounds into two groups with a cut off of 300 nM. The R square is 0.60 and 0.61 for Glide XP and Glide SP poses, respectively. Interestingly scoring of the poses from the computationally much less expensive Glide SP algorithm yielded a slightly better separation of binding affinity than the poses obtained from Glide XP. This is probably due to the way the poses were chosen. If the highest scoring pose had been selected for all compounds, the correlation between calculated conformational energy penalties and Aurora A binding affinity may have been poor.

Conclusions

Using structure-based design we have designed and synthesized a series of potent Aurora inhibitors. Starting from a HTS hit we modified the structure to obtain a novel lead structure. Subsequent lead optimization against Aurora A yielded compounds with desirable physical and chemical properties in the single digit nanomolar range. Comparison of Aurora A and B X-ray structures revealed subtle differences in the size and shape of the binding sites which were exploited in the design of Aurora B selective compounds.

We were able to rationalize the Aurora A binding activities by correlating these to the conformational energies of the putative bioactive conformations. However no correlation was found to the docking score. This demonstrates the importance of considering the conformational energies of docked poses during structure based design and not relying on the docking score alone.

References

1. Andrews PD, Knatko E, Moore WJ, Swedlow JR (2003) Curr Opin Cell Biol 15:672
2. Bischoff JR, Plowman GD (1999) Trends Cell Biol 9:454
3. Katayama H, Brinkley WR, Sen S (2003) Cancer Metastasis Rev 22:451
4. Fu J, Bian M, Jiang Q, Zhang C (2007) Mol Cancer Res 5:1
5. Nigg EA (2001) Nat Rev Mol Cell Biol 2:21
6. Goto H, Yasui Y, Nigg EA, Inagaki M (2002) Genes Cells 7:11
7. Hauf S, Cole RW, LaTerra S, Zimmer C, Schnapp G, Walter R, Heckel A, van Meel J, Rieder CL, Peters JM (2003) J Cell Biol 161:281
8. Li X, Sakashita G, Matsuzaki H, Sugimoto K, Kimura K, Hanaoka F, Taniguchi H, Furukawa K, Urano T (2004) J Biol Chem 279:47201
9. Yan X, Cao L, Li Q, Wu Y, Zhang H, Saiyin H, Liu X, Zhang X, Shi Q, Yu L (2005) Genes Cells 10:617
10. Giet R, Prigent C (1999) J Cell Sci 112(Pt 21):3591
11. Cheetham GM, Knegtel RM, Coll JT, Renwick SB, Swenson L, Weber P, Lippke JA, Austen DA (2002) J Biol Chem 277:42419
12. Nowakowski J, Cronin CN, McRee DE, Knuth MW, Nelson CG, Pavletich NP, Rogers J, Sang BC, Scheibe DN, Swanson RV, Thompson DA (2002) Structure 10:1659
13. Sen S, Zhou H, White RA (1997) Oncogene 14:2195
14. Ewart-Toland A, Briassouli P, de Koning JP, Mao JH, Yuan J, Chan F, MacCarthy-Morrogh L, Ponder BA, Nagase H, Burn J, Ball S, Almeida M, Linardopoulos S, Balmain A (2003) Nat Genet 34:403
15. Zhou H, Kuang J, Zhong L, Kuo WL, Gray JW, Sahin A, Brinkley BR, Sen S (1998) Nat Genet 20:189
16. Blanchard S, Deng W, Lee A, Poulsen A, Teo EL, Tu NP, William AD (2007) WO 2007/058626. S*BIO Pte Ltd., Singapore
17. Schrödinger (2007) Maestro, LigPrep, MarcoModel, Glide, QikProp and Jaguar are products of Schrödinger LLC, www.schrodinger.com
18. Chang G, Guida WC, Still WC (1989) J Am Chem Soc 111:4379
19. Halgren TA (1999) J Comput Chem 20:720
20. Hasel WH, Hendrickson TF, Still WC (1988) Tetrahedron Comput Method 1:103
21. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) J Am Chem Soc 112:6127
22. Sessa F, Mapelli M, Ciferri C, Tarricone C, Areces LB, Schneider TR, Stukenberg PT, Musacchio A (2005) Mol Cell 18:379
23. Bostrom J, Norrby PO, Liljefors T (1998) J Comput Aided Mol Des 12:383
24. Tannor DJ, Marten B, Murphy R, Friesner RA, Sitkoff D, Nicholls A, Honig B, Ringnalda M, Goddard WA (1994) J Am Chem Soc 116:11875
25. Renhowe PA, Shafer CM, McBride C, Silver J, Pecchi S, Machajewski T, McCrea B, Poon D, Thomas T (2003) WO2003/004488. Chiron Corporation, USA
26. Blanchard S, Ethirajulu K, Lee CHA, Nagaraj HKM, Poulsen A, Sun ET, Tan YLE, Teo EL, William AD (2006) WO2007/058628. S*BIO PTE LTD, Singapore
27. Blanchard S, Lee CHA, Nagaraj HKM, Poulsen A, Sun ET, Tan YLE, William AD (2006) WO2007/058627. S*BIO PTE LTD, Singapore
28. Bissantz C, Folkers G, Rognan D (2000) J Med Chem 43:4759
29. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Nat Rev Drug Discov 3:935
30. Tame JR (2005) J Comput Aided Mol Des 19:445
31. Wang R, Lu Y, Wang S (2003) J Med Chem 46:2287
32. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) J Med Chem 47:1739
33. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) J Med Chem 49:6177
34. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) J Med Chem 47:1750