Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 17 (2007) 1966–1970

Discovering new inhibitors of bacterial glucosamine-6P synthase (GlmS) by docking simulations

Nicolas Floquet,[a,b] Céline Richez,[a] Philippe Durand,[a] Bernard Maigret,[b] Bernard Badet[a] and Marie-Ange Badet-Denisot[a,*]

[a]*Institut de Chimie des Substances Naturelles (ICSN), UPR-2301, CNRS, Avenue de la Terrasse, 91198 Gif-sur-Yvette, France*
[b]*Equipe de Dynamique des Assemblages Membranaires (EDAM), CNRS/UHP N° 7565, Université Henri Poincaré-Nancy I, BP 239, 54506 Vandoeuvre-lès-Nancy Cedex, France*

Received 24 November 2006; revised 8 January 2007; accepted 9 January 2007
Available online 25 January 2007

Abstract—Results of an in silico screening of a freely accessible database encompassing 50,000 commercial compounds on bacterial glucosamine-6P synthase (Glms) are described. Each product was docked with the GOLD software in a region of 20 Å surrounding the sugar binding site and ranked according to its score. Among the 14 best-scored molecules, three molecules exhibited good experimental inhibition properties (IC_{50} = 70 μM) giving a high hit rate (H.R.: 0.23). Interestingly, these molecules are predicted to interact with a protein region that forms a pocket at the interface between the two enzyme monomers, opening the route to dimerization inhibitors.

Glucosamine-6P synthase (GlmS) is a dimeric enzyme that belongs to the family of glutamine-dependent amidotransferases and catalyses the first committed step of hexosamine metabolism.[1] It irreversibly converts D-fructose-6P into D-glucosamine-6P by using L-glutamine as a nitrogen source. The end product of this pathway, UDP-*N*-acetylglucosamine-6P, is an essential building block of bacterial and fungal cell walls. In mammalian cells, the equivalent protein (Gfat) is a tetrameric enzyme playing the role of a glucose sensor with crucial implications in the vascular complications of type II diabetes.[2] Therefore, GlmS and Gfat are considered as pharmacological targets and they have been intensively studied over the last decade. Based on biochemical and structural analyses, it has been demonstrated that the monomer of GlmS is composed of two independent domains,[3–5] namely the glutaminase and the synthase domains that are linked to each other by a hinge region of ~40 amino acid residues. The glutaminase domain catalyses the hydrolysis of L-glutamine into L-glutamate and ammonia (glutaminase site). The ammonia formed is then transferred through an 18 Å length channel[6] to the synthase domain which converts bound D-fructose-6P into D-glucosamine-6P (synthase site). Dimerization of the enzyme appears to be crucial for the closure of the sugar active site and the formation of the channel.[6] The amino acid sequences of Glms and Gfat share 40% identity, a value that increases to 80% in the functional regions (active sites, channel). The region that covers the active sites of Glms can therefore be considered as a good model for the study of Gfat. So far, most of the compounds reported to inhibit GlmS and Gfat are amino acids or sugar derivatives directed towards one of the enzyme active sites.[7,8] More recently, natural products such as amitrole[9] or aaptamine[10] or analogues[11,12] were also identified although the precise mode of action remains to be clarified. Two crystallographic structures describing the enzyme in a complex with D-fructose-6P and with both D-fructose-6P and 6-diazo-5-oxo-L-norleucine (DON), an affinity label bound to Cys_1 of the glutaminase site, were recently obtained[13]. These structures resolved at 2.05 and 2.35 Å, respectively, considerably improved the accuracy of the data previously reported at 3.1 Å,[6] allowing in particular the observation of water molecules in the enzyme active sites. We used the recent structures of the dimeric Glms to perform docking calculations.[14]

A prerequisite to any virtual screening experiment is the validation step. The docking protocol was therefore

* Corresponding author. Tel.: +33 (0) 1 69 82 31 10; fax: +33 (0) 1 69 07 72 47; e-mail: badetma@icsn.cnrs-gif.fr

0960-894X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2007.01.052

tested on a set of 33 molecules for which the experimental K_i was previously determined.[15] We first observed that docking into the empty water-free Glms structure led to unreliable results with no correlation between experimental K_i and calculated scores. This problem was eventually solved by including water molecules from crystallographic data. As reported in the initial X-ray structures of the isolated synthase domain (PDB: 1MOQ, 1MOR and 1MOS),[16] up to 11 water molecules connecting the ligand and the protein through hydrogen-bonding could be identified. This number varies according to the ability of the inhibitors to displace one or more of these water molecules. As an example, Figure 1a shows the water molecules that can be observed in the 2-deoxy-2-amino-D-glucitol-6P:enzyme complex.[16]

After screening the effect of removing different water molecules on the docking of this subset of inhibitors, five of them corresponding to those that are strictly conserved among all the available enzyme X-ray structures turned out to be important (Fig. 1b). An additional argument came from the results of a 5 ns molecular dynamics simulation performed on the Glms monomer where the fructose-6P site is fully accessible to the solvent and which confirmed their important residence time (data not shown). Accordingly, these water molecules were included in the docking calculations.

Using these 'hydrated' conditions all the ligands of known crystallographic structures were successfully replaced (Rmsd <1 Å) by the GOLD docking protocol (data not shown). Moreover, for each compound, the solution corresponding to the crystallographic model was found among the three best proposed by GOLD. The scores of the docking procedure for these 33 compounds are reported in Table 1, together with their experimental K_is.[15]

The correlation between the GOLD score values and the experimental K_is (Fig. 2) shows that the GOLD score is a good indicator of activity for this series of compounds with a remarkable Kendall rank correlation coefficient of $\tau = 3.4$ ($p < 5.10^{-4}$).[17]

Table 1. Experimental K_i(s) and scores of the 33 compounds of the docking validation set

Compounds	K_i (μM)	Score
Oxime of D-arabinose-5P	14	74.72
2-Amino-2-deoxy-D-glucitol-6P	25	73.16
2-Acetylamino-2-deoxy-D-glucitol-6P	28	74.95
2-Chloroacetylamino-2-deoxy-D-glucitol-6P	58	80.94
2-Amino-2-deoxy-6-phospho-D-gluconate	90	76.62
Oxime of D-glucosamine-6P	180	77.25
2-Hydroximo-6-deoxy-6-phosphono-D-fructose	200	77.36
N-iodoacetyl-D-glucosamine-6P*	220	61.62
Oxime of 5-methylphosphono-D-arabinose	360	67.28
5-Phosphonomethyl-D-arabinohydroxymolactone	370	61.11
2-Ethylamino-2-deoxy-D-glucitol-6P	500	77.23
Oxime of D-glucose-6P	580	74.79
2-Amino-2-deoxy-6-phosphonomethyl-D-glucitol-6P	650	65.59
1-Amino-1-deoxy-D-glucitol-6P	750	72.54
D-glucosamine-6P*	760	66.64
5-Phosphonomethyl-D-arabinolactone	1140	60.52
1,2-Anhydrohexitol-6P	1400	67.24
D-Glucitol-6 P	2400	71.01
6-D-Phosphogluconate	2500	76.91
6-Deoxyphosphono-D-fructose*	2500	61.97
D-Ribose-5P	3200	65.95
D-Arabinose-5P	8300	65.98
3-Amino-D-fructose-6P*	5000	63.27
2-Deoxy-D-glucitol-6P	15000	71.28

*For cyclic compounds both the α and the β anomers were docked; only the best was conserved for analysis.

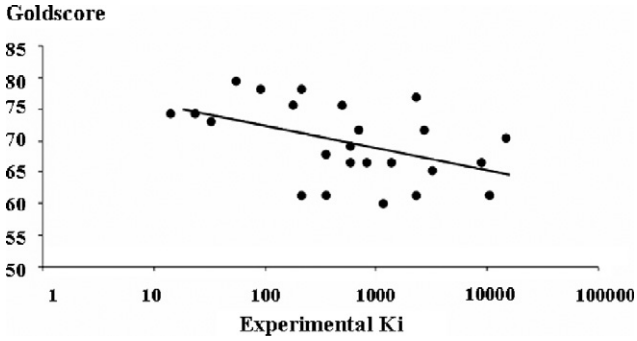


Figure 2. The relation between the scores and the experimental K_is of the 33 compounds of the docking validation set.

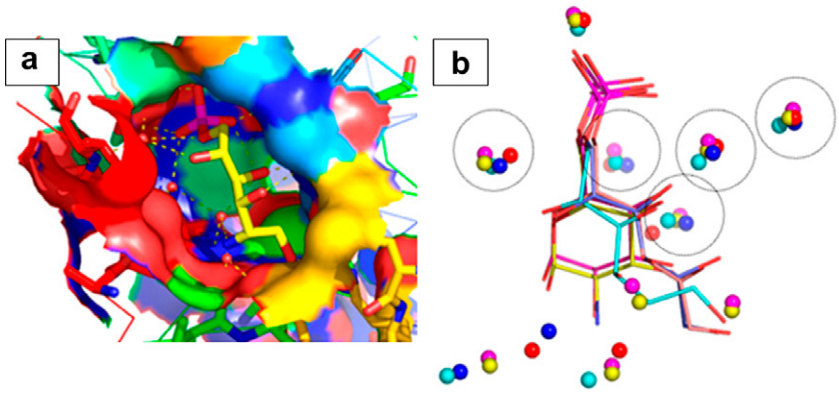
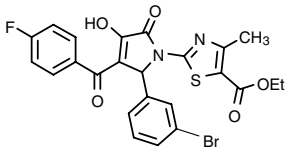
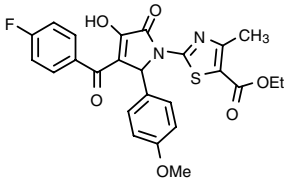
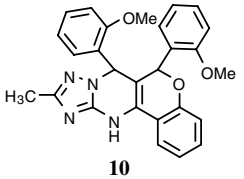
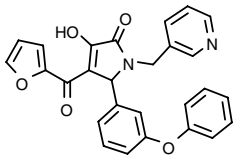
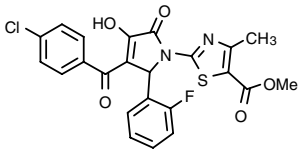
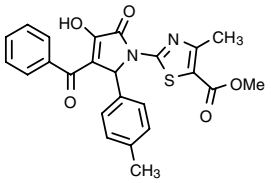
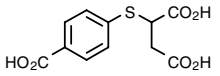
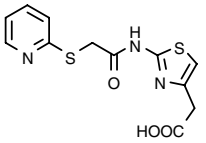
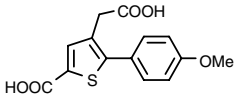
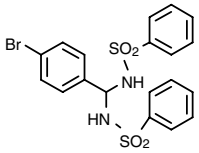
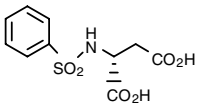
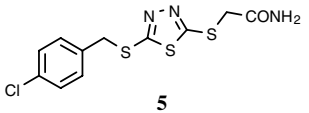
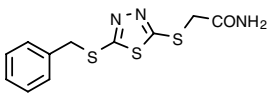
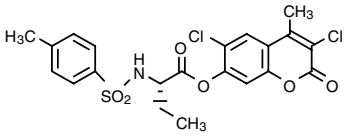
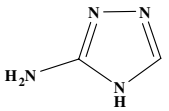
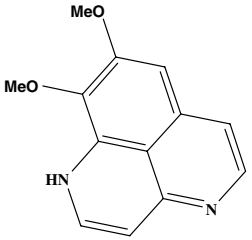


Figure 1. The role of water molecules in the binding of Glms ligands. (a) 2-deoxy-2-amino-D-glucitol-6P:enzyme complex[16] and (b) the five water molecules included for the docking calculations.

Table 2. Percentage of Glms activity inhibition by the compounds selected among the hits of the virtual screening

Structure	% inhibition (Score)	Structure	% inhibition (Score)	Structure	% inhibition (Score)
 **1**	**70** (71.0)	 **2**	**70** (70.6)	 **10**	**70** (67.1)
 **8**	**20** (67.5)	 **3**	**0** (69.9)	 **4**	**0** (69.8)
 **13**	**0** (69.3)	 **14**	**0** (69.3)	 **6**	**0** (68.0)
 **9**	**0** (67.3)	 **12**	**0** (67.0)	 **5**	**ND** (68.9)
 **7**	**ND** (67.7)	 **11**	**ND** (67.0)	 **Amitrole**	 **Aaptamine**

Enzyme activity was determined at 100 μM; ND: Not Determined (precipitation of the compound preventing analysis). Under the same conditions, the reference compound 2-amino-2-deoxy-D-glucitol-6P (score = 73.2) gave 85% inhibition (IC_{50} = 56 μM). The structures of amitrole (IC_{50} = 100 μM) and aaptamine (IC_{50} = 120 μM) are included for comparison purposes.

Following validation of the docking protocol, a bank of 50,000 commercially available compounds was screened in silico. The ligands from the databank were ranked according to their scores. After statistical analysis[18] of the whole scores distribution, 14 of the best scored molecules (corresponding to a score > mean (score) + 3 * SD. ($p > 99.9\%$)), which were predicted to be directed towards either the fructose-6P binding site or the interface between the two monomers of Glms, were tested (Table 2).[19]

Tested compounds were considered as 'active' when displaying an effect on activity at 100 μM, in the range of that exhibited by 2-amino-2-deoxy-D-glucitol-6P, one of the best known inhibitors of Glms (IC_{50} = 56 μM) and belonging to the validation set.

Three molecules (**1**, **2** and **10**) were found to significantly inhibit GlmS giving an IC_{50} of 70 μM, with no clear structural relationship with known inhibitors, including amitrole or aaptamine (see Table 2). Interestingly, these three molecules were predicted to bind at the interface between the two monomers of Glms. Compounds **1** and **2**, that share a common scaffold, were docked at exactly the same position forming H-bonds with residues Gly_{301}, Tyr_{304}, Leu_{484}, Lys_{487}, Glu_{495}, Ala_{496} and Tyr_{497} from one monomer and with residues Glu_{495}, Ala_{496}, Ala_{498} and Lys_{487} from the second monomer (Fig. 3a).

Compounds **3**, **4** and **8** which also share a common scaffold were predicted to interact within the same region, with an equivalent score, but were found inactive at the tested concentration of 100 μM, except for compound **8** for which an IC_{50} of 250 μM was measured. At the moment it is too early to rationalize these differences and obtain a clear structure activity relationship. Further studies taking into account the flexibility of the intersubunit site shall help refining our model.

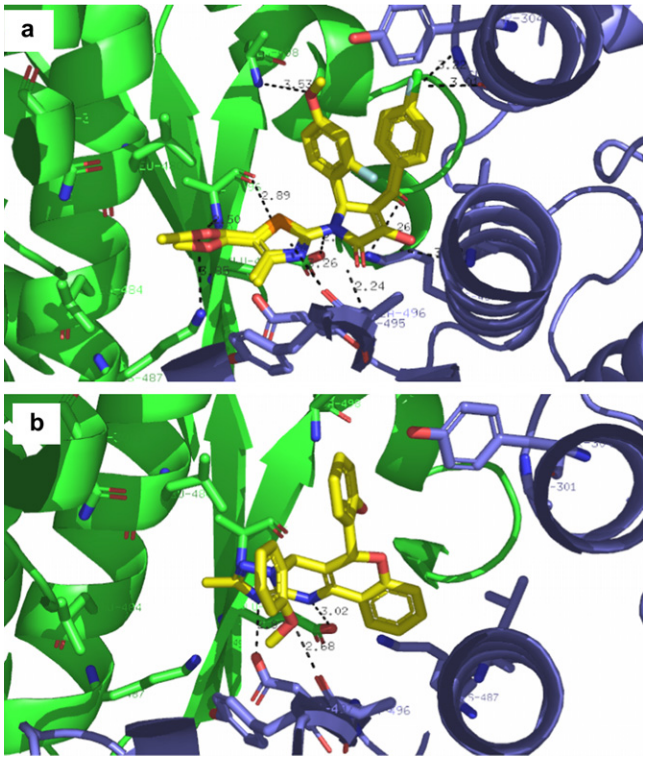


Figure 3. Models of compounds **1**, **2** (a) and **10** (b) docked at the interface of the two Glms monomers (ribbons in green and blue, respectively).

Compound **10** was predicted to interact within the same region with residues Glu_{495} and Ala_{496} of both monomers (Fig. 3b). Although we cannot exclude the possibility that this predicted interaction is favoured by the inclusion of water molecules in the active site, it could constitute an issue for the design of compounds that alter the dimerization process.

Unfortunately, the poor water solubility of the compounds prevented further experimentation to prove this new mechanism of inhibition. Efforts are actually in progress to design related compounds with an increased aqueous solubility to circumvent this problem.

In this work, we used the recent X-ray structures of bacterial glucosamine-6P synthase to dock a database of 50,000 commercially available compounds in the region around its active site. Among the best hits, 14 molecules were finally tested as putative inhibitors of Glms, and three displayed significant inhibitory properties giving a remarkable hit rate of 0.23. Interestingly, these three compounds were predicted to bind at the interface between the two Glms monomers, in a deep pocket located near the synthase site. These molecules display completely original structures as compared to the natural substrate and known inhibitors. We demonstrate here the possibility of rapidly discovering Glms ligands with new scaffolds by using a basic virtual screening strategy which takes into account the presence of crucial water molecules. This new feature may be relevant for the design of Glms inhibitors preventing or disrupting the association of its monomers, a strategy which presently remains difficult.

Acknowledgements

Financial support from ICSN to NF and CR is gratefully acknowledged. This work has been supported in part by the Indo-French program IFC/3003-A.

References and notes

1. McClain, D. A.; Crook, E. D. *Diabetes* **1996**, *45*, 1003.
2. Traxinger, R. R.; Marshall, S. *J. Biol. Chem.* **1991**, *266*, 10148.
3. Badet, B.; Vermoote, P.; Haumont, P. Y.; Lederer, F.; LeGoffic, F. *Biochemistry* **1987**, *26*, 1940.
4. Denisot, M. A.; Le Goffic, F.; Badet, B. *Arch. Biochem. Biophys.* **1991**, *288*, 225.
5. Leriche, C.; Badet-Denisot, M. A.; Badet, B. *J. Am. Chem. Soc.* **1996**, *118*, 1797.
6. Teplyakov, A.; Obmolova, G.; Badet, B.; Badet-Denisot, M. A. *J. Mol. Biol.* **2001**, *313*, 1093.

7. Janiak, A. M.; Hoffmann, M.; Milewska, M. J.; Milewski, S. *Bioorg. Med. Chem.* **2003**, *11*, 1653.
8. Yaouancq, L.; Rene, L.; Tran Huu Dau, M. E.; Badet, B. *J. Org. Chem.* **2002**, *67*, 5408.
9. Babykin, M. M.; Kovach, A.; Cerff, R.; Shestakov, S. V. *Genetika* **2000**, *36*, 1406.
10. Bobzin, S. C.; Yang, S.; Kasten, T. P. *J. Chromatogr. B. Biomed. Sci. Appl.* **2000**, *748*, 259.
11. Bolin, D.R.; Chen, S.; Mischke, S.G.; Qian, Y. PCT Int. Appl. WO 040101528, 2004.
12. Bolin, D.R.; Chen, S.; Mischke, S.G.; Qian, Y. PCT Int. Appl. WO 050401528, 2005.
13. Mouilleron, S.; Badet-Denisot, M. A.; Golinelli-Pimpaneau, B. *J. Biol. Chem.* **2006**, *281*, 4404.
14. An all-atom model of Glms (including hydrogens) was built from the crystallographic coordinates of the fructose-bound dimeric enzyme (PDB: 2BPJ) by using the Psfgen module distributed with NAMD[21] and the topology files from CHARMM (PARAM set 22).[22] The protein was converted to the SYBYL mol2 format before processing. For virtual screening, the ligands were downloaded from the ZINC databank,[23] a free database of 'ready to dock' commercial compounds for virtual screening. We obtained the subset corresponding to the compounds commercially available at Sigma Aldrich (49,961 molecules) in the SYBYL format. The docking simulations were performed on a Dell standard PC using the GOLD v2.1.1 software[24,25] and the GOLD score function. The method consists of flexibly adapting ligands to a rigid model of the binding site of a target by means of genetic algorithms. Each successive position of the ligand is evaluated by way of a scoring function composed of four energy components: protein–ligand hydrogen bonds, protein–ligand Van Der Waals, ligand internal VDW and a ligand torsional strain. The compounds were successively docked in the dimeric Glms in a sphere of 20 Å radius around Gln_{348}, a region that covers the fructose-6P binding site and its vicinity. To save computational time, only two dockings were performed for each ligand, using the default parameters of Gold (100,000 operations with a starting population of 100 and a selection pressure of 1.1). About 2 weeks were needed to obtain the results for all the compounds of the database. The docking protocol was first validated on a set of 33 molecules for which an experimental K_i is known.[15] Structures of these ligands were first constructed with Chem3D (CambridgeSoft) and converted to the SYBYL format using OpenBabel. Among these molecules, crystallographic data are also available for α-D-glucopyranose-6P,[16] 2-deoxy-2-amino-D-glucopyranose-6P,[26] 2-deoxy-2-amino-D-glucitol-6P[16] and D-fructose-6P.[6,13] For this part 50 dockings were performed for each compound. Images were obtained by using the Pymol Software available at http://pymol.sourceforge.net/.
15. Enzyme activity was measured following established protocols.[3] For glutamate determination, the production of glutamate was coupled to glutamate dehydrogenase (GDH) and followed by the appearance of APADH at 366 nm spectrometrically. Assay mixture contained 50 mM potassium phosphate, pH 7.5, 1 mM EDTA, 50 mM KCl, 1 mM APAD, 13 U of GDH, 0.5 mM F6P, Glms (0.25 μg) and a given inhibitor in appropriate concentration in a total volume of 1 mL. The reaction was started by adding 10 mM Gln and incubated at 37 °C. The production of GlcN-6P was measured by modified colorimetric method of Morgan-Elson.[3] The reaction mixture was incubated at 37 °C for 2 min before detection of GlcN-6P.
16. Teplyakov, A.; Obmolova, G.; Badet-Denisot, M. A.; Badet, B. *Protein Sci.* **1999**, *8*, 596.
17. Kendall, M. *Biometrica* **1938**, *30*, 81.
18. Copeland, R. A. In *Evaluation of Enzyme inhibitors in Drug Discovery*; John Wiley and Sons, Inc., 2005, pp 82–110.
19. The influence of the selected molecules on glucosamine-6P formation was analyzed using a modified microplate formated colorimetric Morgan-Elson assay.[20] ZINC database compounds *833975, 833967, 833912, 648353, 826316, 112038, 825816, 830844, 446637, 721204, 512680, 257796 and 355658* as well as all the other compounds used for the tests and purification were purchased from Sigma–Aldrich. The enzyme was purified as described.[27] Stocks of different compounds were prepared at 2.5 mM in dimethylsulfoxide (DMSO). The highest final DMSO concentration in the assay was 10% (vol/vol) and control experiments were run in each case. All reactions were initiated by addition of enzyme and run in triplicate. Incubation of the pure protein with a saturating concentration of L-glutamine (20 mM) and 0.5 mM of D-fructose-6P was performed (20 min, 37 °C) in the presence of different concentrations of the compound to be tested. The reaction was stopped by heating for 2 min at 90 °C and the glucosamine-6P was quantified by measurement of absorption at 585 nm, following derivatisation with *p*-dimethylaminobenzaldehyde.[28] This assay was adapted to a 96-well microplates format. The measurement was run on a spectra max M5 (Molecular Devices) using the IC_{50} determination protocol from the manufacturer.
20. Burghardt, C; Kochan, J.P., European Patent Application 03028086.1, **2004**.
21. Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comp. Phys.* **1999**, *151*, 283.
22. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
23. Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.* **2005**, *45*, 177.
24. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. *Proteins* **2003**, *52*, 609.
25. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. D. *J. Mol. Biol.* **1997**, *267*, 727.
26. Teplyakov, A.; Obmolova, G.; Badet-Denisot, M. A.; Badet, B.; Polikarpov, I. *Structure* **1998**, *6*, 1047.
27. Obmolova, G.; Badet-Denisot, M. A.; Badet, B.; Teplyakov, A. *J. Mol. Biol.* **1994**, *242*, 703.
28. Gosh, S.; Roseman, S. *Methods Enzymol.* **1962**, 414.