Bioorganic & Medicinal Chemistry Letters 21 (2011) 7064–7067

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Novel inhibitors of *Mycobacterium tuberculosis* dTDP-6-deoxy-L-*lyxo*-4-hexulose reductase (RmlD) identified by virtual screening

Yi Wang [a,b,†], Tamara Noelle Hess [c,†], Victoria Jones [c], Joe Zhongxiang Zhou [a,b], Michael R. McNeil [c,*], J. Andrew McCammon [a,b,*]

[a] Center for Theoretical Biological Physics, Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093, USA
[b] Center for Theoretical Biological Physics, Howard Hughes Medical Institute, Department of Pharmacology, University of California, San Diego, La Jolla, CA 92093, USA
[c] Department of Microbiology, Immunology and Pathology, Colorado State University, Ft. Collins, CO 80523, USA

ARTICLE INFO

Article history:
Received 24 August 2011
Revised 22 September 2011
Accepted 22 September 2011
Available online 29 September 2011

Keywords:
RmlD
M. tuberculosis
Virtual screening
L-Rhamnose synthesis
Anti-tuberculsois
Computer-aided drug design

ABSTRACT

The complex and highly impermeable cell wall of *Mycobacterium tuberculosis* (*Mtb*) is largely responsible for the ability of the mycobacterium to resist the action of chemical therapeutics. An L-rhamnosyl residue, which occupies an important anchoring position in the *Mtb* cell wall, is an attractive target for novel anti-tuberculosis drugs. In this work, we report a virtual screening (VS) study targeting *Mtb* dTDP-deoxy-L-*lyxo*-4-hexulose reductase (RmlD), the last enzyme in the L-rhamnosyl synthesis pathway. Through two rounds of VS, we have identified four RmlD inhibitors with half inhibitory concentrations of 0.9–25 μM, and whole-cell minimum inhibitory concentrations of 20–200 μg/ml. Compared with our previous high throughput screening targeting another enzyme involved in L-rhamnosyl synthesis, virtual screening produced higher hit rates, supporting the use of computational methods in future anti-tuberculosis drug discovery efforts.

Tuberculosis (TB) is an infectious disease caused by the pathogen *Mycobacterium tuberculosis* (*Mtb*), which claims nearly two million lives each year.[1] Since no new anti-TB drug has been introduced in the past 40 years, novel therapeutics are in urgent need to treat both drug-susceptible TB and the increasingly common drug-resistant strains.[2–4] The cell wall of *Mtb* is largely responsible for the ability of the mycobacterium to survive in a hostile environment.[5] This cellular envelope consists of three layers: an innermost peptidoglycan layer, an outermost mycolic acid layer, and the connecting arabinogalactan polysaccharide layer.[6,7] An L-rhamnosyl residue occupies an important anchoring position in this complex structure, connecting the arabinogalactan layer and the peptidoglycan layer.[8] Its synthesis has been shown to be an attractive target for novel anti-TB drugs.[7,9–12] In this work, we report a virtual screening (VS) study targeting dTDP-deoxy-L-*lyxo*-4-hexulose reductase (RmlD), the last enzyme involved in the conversion of glucose-1-phosphate to dTDP-L-rhamnose (Fig. 1).[13] To the best of our knowledge, these results represent the first virtual screening effort targeting RmlD.

As a member of the reductases/epimerases/dehydrogenases (RED) enzyme super family, RmlD is located in the short chain dehydrogenase/reductase (SDR) branch.[14] The N-terminal domain of the enzyme is dominated by a Rossmann-type fold (Fig. 2), which forms the cofactor binding site and contains a six-stranded β sheet sandwiched between six α helices. The C-terminal domain of RmlD forms the substrate binding site, containing three α helices and a double-stranded β sheet. Using either NADH or NADPH as a cofactor, RmlD catalyzes the sugar converting reaction at the interface of its two domains, where a hydride is transferred from the nicotinamide ring of NAD(P)H to the C4′-carbonyl of the substrate.

Since no crystal structure of *Mtb* RmlD is available, we first constructed a homology model using the program MODELLER,[15–17] with the RmlD structure from *Salmonella enterica* serovar Typhimurium (*S. typhimurium*)[14] as a template. While sequence alignment using ClustalX[18,19] and Bio3D[20] reveals a high similarity in the active sites of the two proteins (62% sequence identity, see Fig. S1), the *Mtb* RmlD homology model performed poorly in the redocking test of dTDP-L-rhamnose. As shown in Fig. S2, steric clash of dTDP-L-rhamnose with residue Arg224 from *Mtb* RmlD prevents the ligand from positioning its hexose ring inside the binding pocket. Additionally, the orientation of Thr104 in the conserved catalytic triad is altered in the homology model, precluding this key residue from forming a hydrogen bond with the ligand. Although the *Mtb* RmlD homology model might be improved through various modeling techniques, we decided to use the *S. typhimurium* RmlD structure in the remainder of the study. The similar active sites from the

* Corresponding authors.
E-mail addresses: M.Mcneil@colostate.edu (M.R. McNeil), jmccammon@mail.ucsd.edu (J.A. McCammon).
† These two authors contributed equally to the work.

0960-894X/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2011.09.094

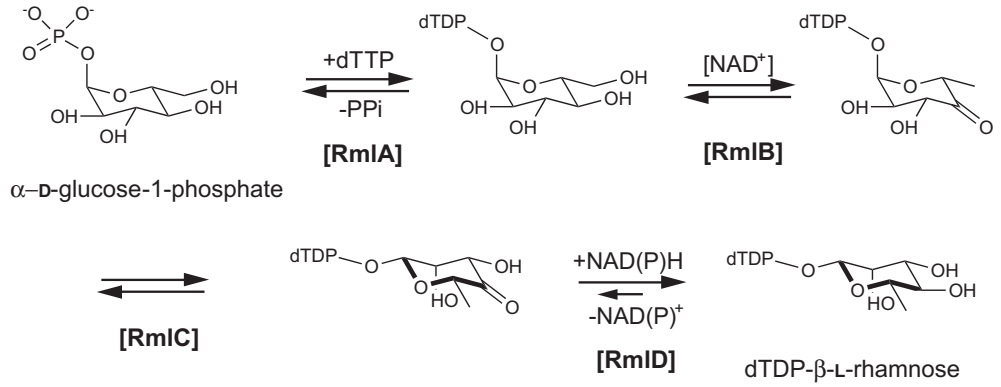

Figure 1. The biosynthesis pathway of L-rhamnose in *Mtb*. Four enzymes, RmlA to RmlD, catalyze the transition from glucose-1-phosphate to dTDP-L-rhamnose.

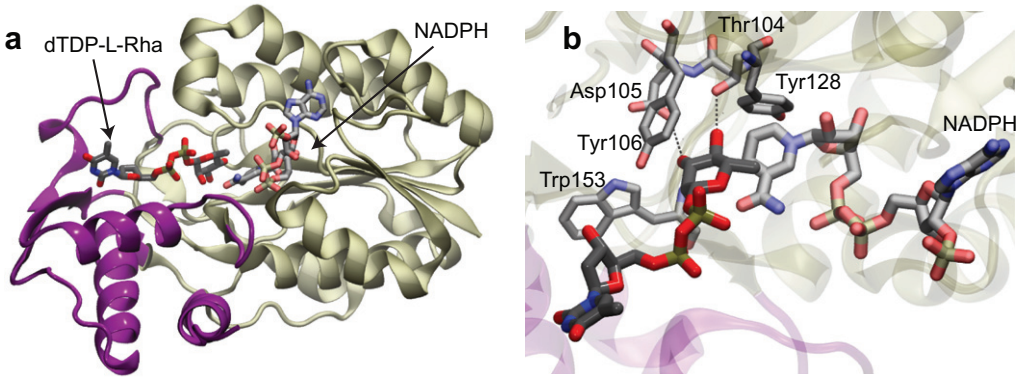

Figure 2. The crystal structure of RmlD from *S. typhimurium*. (a) RmlD in complex with the ligand dTDP-L-rhamnose and the cofactor NADPH. The N-terminal and C-terminal domains of the protein are colored in brown and purple, respectively. (b) The active site of RmlD. Key residues involved in the binding of dTDP-L-rhamnose are shown in stick representations.

two enzymes and their highly conserved reaction mechanism provide the basis of using the *S. typhimurium* structure in the virtual screening.

Altogether two rounds of VS were performed on RmlD, first using the relatively small NCI diversity set II and then using a subset of the larger NCI open database. The NCI diversity set II is a subset of ~140,000 compounds in the Developmental Therapeutics Program repository at the National Cancer Institute. The small size of this set (1364 compounds) allows fast initial screening for a target protein. Using the program GLIDE,[21–24] we performed altogether four VS runs: The apo-RmlD was used in the first three VS, where the grid box for docking was placed at the center of the cofactor binding site, the center of the ligand binding site, and the interface between the two binding sites, respectively; the fourth VS run was performed on RmlD in complex with NADPH, with the grid box placed at the ligand binding site. While in theory, the first three VS can be replaced by a single run with a large grid box covering the entire RmlD active site, in practice, a large grid box often increases the difficulty for docking programs to identify the correct binding poses. With four independent VS, we were able to focus the screening effort at the most relevant location in each run, and search for potential inhibitors with different modes of action, that is, binding with or without the cofactor NADPH.

Following the four VS runs, the top 80 poses from each VS were combined and examined visually. These poses correspond to 59 unique compounds, from which 31 were selected and tested using an enzymatic assay (see Supplementary data). Compounds interacting with key residues in the sugar converting reaction, for example, Asp105, Thr104 and Tyr128, were favored in the selection, since they may provide the highest specificity for RmlD and may be the most robust against development of resistance. Additionally, preference was given to compounds ranked high in two or more VS runs, based on the assumption that these compounds may have a higher probability to inhibit RmlD. Finally, the 'rule-of-five'[25] was used as a general guideline to exclude compounds with undesirable physico-chemical properties, although a less stringent criterion was used here (M_w >540 or log*P* >6), to avoid discarding hits that may be optimized in later stages of drug discovery. None of the final four experimentally confirmed hits (Table 1) has any violation of the 'rule-of-five'.

As described in Ma et al.,[26] the sugar reduction catalyzed by RmlD results in the oxidation of NADPH, and is accompanied by a decrease in absorption at OD_{340}. This decrease is used to monitor the progress of the reaction in our enzymatic assay.[27–29] All 31 compounds from the 1st-round VS (see Table S1) were tested using

Table 1
Activity of RmlD inhibitors in the enzymatic and whole-cell assays

Compound	IC_{50} (μM)	MIC (μg/ml)	M_w	Log*P*
1	2.1	133	336.29	1.16
2	0.9	200	320.30	1.79
3	15	62	258.23	0.63
4	25	20	252.31	3.16

Compounds are numbered as shown in Figure 3. The molecular weight (M_w) of each compound and its octane/water partition coefficient (log*P*) calculated using the program QikProp[38] are also listed. The IC_{50} and MIC values were obtained from at least two replicate experiments.

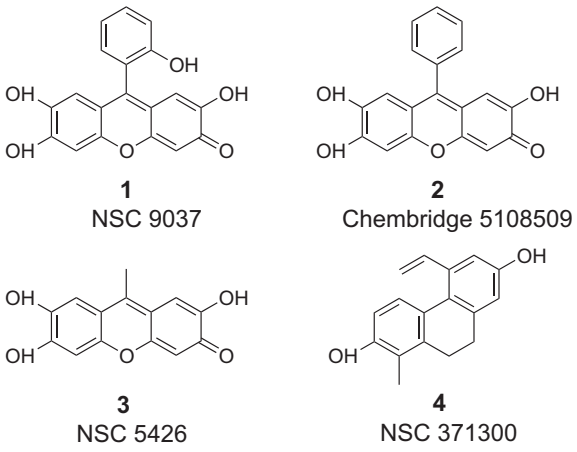

Figure 3. Structures of RmlD inhibitors identified through two rounds of virtual screening.

this assay at an initial concentration of 10 μg/ml. IC_{50} values were then determined for compounds that showed initial inhibitory activity. Following this protocol, one out of the 31 compounds, compound **1** (Fig. 3), was confirmed to be a RmlD inhibitor with an IC_{50} of 2.1 μM.

Given the binding mode of compound **1** (Fig. 4), three substructures of this molecule were used as seeds for a similarity search performed over 250,000 compounds in the NCI open database (Fig. S3). The 2nd-round virtual screening was then performed on the search results using the grid box centered at the dTDP-L-rhamnose binding site in the presence of NADPH. Following the similar selection criteria used in the 1st-round VS, eight high-ranking compounds were selected for experimental verification. Seven of these compounds were obtained from NCI, and compound **2** (Fig. 3) was obtained from Chembridge. Assay results indicate that three out of the eight compounds are active, with IC_{50} values ranging from 0.9 to 25 μM (Table 1). Two of the active compounds, compound **3** and **4**, were then used as seeds in another round of similarity search over the NCI open database. However, none of the 11 compounds selected from this additional VS calculation were active in the enzymatic assay. Taken together, the 2nd-round VS revealed three additional hit compounds (Fig. 3).

As shown in Figure 4, all four hit compounds bind to RmlD at the dTDP-L-rhamnose binding site in the presence of NADPH. A common structural feature shared by these compounds is a hydroxyl group that forms hydrogen bonds with Asp105 and/or Thr104. As part of a conserved catalytic triad, Thr104 has been found essential to the enzymatic activity of RmlD.[14] Therefore, the above hydrogen bonds are likely crucial to the inhibitory activity of the hit compounds. Another common feature shared by all the inhibitors is a rigid tricyclic ring that serves as the backbone of the structures. Part of the tricyclic ring replaces the hexose ring in dTDP-L-rhamnose, and is sandwiched between the nicotinamide group of the cofactor and the aromatic ring of Tyr106. Hydrophobic contacts are formed between the tricyclic ring and the nonpolar regions of Tyr106, Tyr128, Val67, Trp153, as well as the cofactor. These hydrophobic interactions occupy the perimeter of the active site and bury the hydrogen bond with Asp105 and/or Thr104 deep inside the binding pocket. Such 'hydrophobic enclosure' interactions have been found particularly favorable in receptor-ligand binding.[24]

It is worth noting that while the hit compounds identified in this study bind at the substrate binding site, potential RmlD inhibitors may also target the cofactor binding site of the enzyme: RmlD does not rely on a second substrate to reduce the oxidized cofactor, and to regain the enzymatic activity, NAD(P) has to be replaced by a new NAD(P)H molecule.[14] Compared with other cofactor-binding enzymes, binding of NAD(P)H is found to be relatively weak in RmlD.[14] Therefore, the enzyme may be a good target for Rossmann-fold inhibitors that bind at the cofactor binding site. Such inhibitors have been reported recently for the 17-β-hydroxysteroid dehydrogenase, another member of the SDR family.[30,31] Although they themselves may have selectivity issues, the Rossmann-fold inhibitors could provide the structural basis for designing potent inhibitors occupying both the substrate and the cofactor binding site.

The activity of the the four identified hits against whole *M. tuberculosis* growing in liquid culture was determined as the

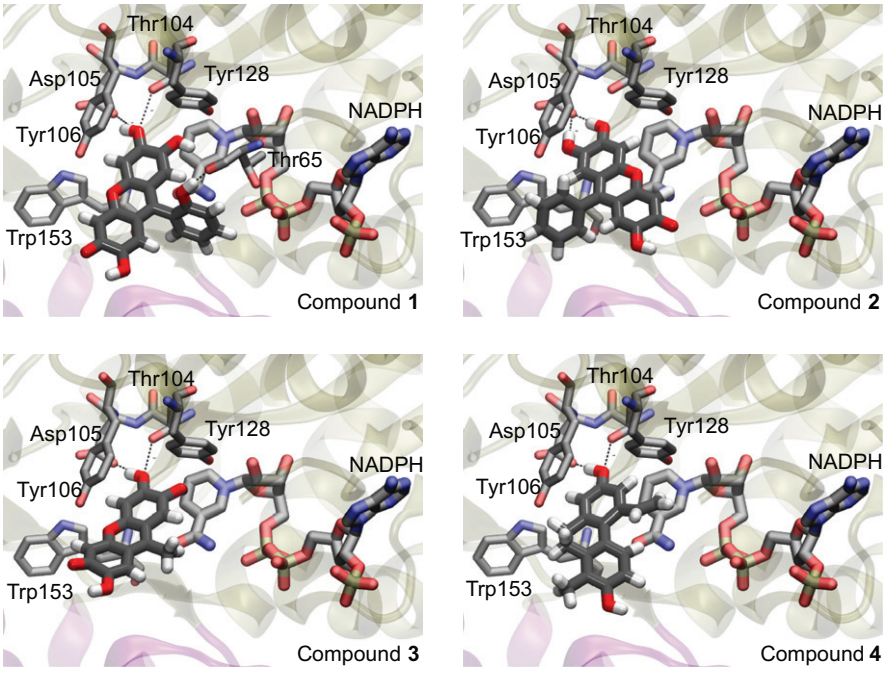

Figure 4. The binding poses of RmlD inhibitors identified through virtual screening. The cofactor NADPH and key residues involved in inhibitor binding are shown. Hydrogen bonds are highlighted by black dotted lines. The protein backbone is shown in transparent representations and colored as in Figure 2.

minimum inhibitory concentration (MIC) value using the microbroth dilution method described in Sun et al[32] and Brown et al.[33] The most potent RmlD inhibitors, compound **1** and compound **2** (IC_{50} = 2.1 and 0.9 μM), have modest activity in the whole-cell assay (MIC = 133 and 200 μg/ml, see Table 1). Compound **4**, which has an IC_{50} of 25 μM, shows the best whole-cell activity (MIC = 20 μg/ml). The former two compounds have a log*P* value of 1.16 and 1.79, respectively, whereas compound **4** has the highest log*P* (3.16) among the identified inhibitors. This result suggests that the low whole-cell activity of compounds **1** and **2** may be explained by their poor permeability through the *Mtb* cell wall. Compound **3**, which has a low log*P* (0.63) and a moderate IC_{50} (15 μM), is the second most potent compound in the whole-cell assay. This somewhat unexpected behavior might be related to the small size (M_w = 258.2) of compound **3**, which may provide it with a higher diffusion coefficient in the *Mtb* cell wall than compounds **1** and **2**. Analysis of more analogs of compounds **1**–**3** is required to fully elucidate the role of lipid permeability in the whole-cell activity of these RmlD inhibitors.

In summary, we performed two rounds of VS on RmlD and identified four novel inhibitors with a minimum IC_{50} of 0.9 μM and a minimum MIC of 20 μg/ml. Docking poses suggest that the identified inhibitors bind at the C-terminal domain of RmlD in the presence of the cofactor, and engage key residues required in enzyme catalysis, such as Tyr128 and Thr104, which have been found essential for the sugar converting reaction catalyzed by RmlD.[14] Common structural features of the inhibitors include a rigid tricyclic ring that serves as the backbone of the compounds, as well as a buried hydroxyl group forming H-bonds with key residues in the enzyme. Out of the four inhibitors, the smallest compounds (**3** and **4**) may serve as basic chemical scaffolds for further optimization.

Compared with antibiotics targeting other bacteria, lipophilicity may play a greater role in a compound's activity against *Mtb*. The outermost layer of *Mtb* cell wall contains a unique 70–90 carbon mycolic acid layer, which constitutes ~30% of the dry weight of the cell.[34] As a result of this layer, the mycobacterial cell wall is highly impermeable to small molecules, and can resist the action of a large number of chemical therapeutics.[6,35] For instance, the broad-spectrum antibiotic β-lactam has been found to be at least 100-fold less permeable in the cell wall of *Mtb* than the Gram-negative bacterium *E. coli*.[35,36] Therefore, future studies may be explore the optimization of the identified RmlD inhibitors through improving their permeability in the waxy cell wall of *Mtb*.

In our previous HTS work targeting the enzyme RmlC in the L-rhamnose synthesis pathway,[29] 201,368 compounds were screened and a 1.2% initial hit rate was obtained. Upon further test, 14 true hits were identified, corresponding to a 0.007% true hit rate. In this work, 31 compounds from the 1st-round VS were tested, and one compound was found to be active (3.2% initial hit rate). Through an additional round of VS based on similarity search, 19 more compounds were tested and three were found active, corresponding to a final 8.0% true hit rate. In comparison to the HTS work, VS produced a better hit rate by effectively enriching the database, and similarity search based on the identified inhibitor further improved its performance. These results support the use of computational methods in future anti-TB drug discovery efforts. Additionally, as demonstrated by a recent work examining the complementarity of HTS and VS,[37] the chance of identifying novel inhibitors may be further improved by combining these two approaches.

Acknowledgments

We would like to thank Dr. Jim Naismith for providing the *E. coli* strain over-expressing RmlB from *S. typhimurium*. We also thank Drs. Jacob Durrant, Sara Nichols and Barry Grant for their valuable suggestions, and Dr. Yongxuan Su for help with the compound purity test. Work at University of California is supported in part by the National Science Foundation, the National Institutes of Health, Howard Hughes Medical Institute, Center for Theoretical Biological Physics and the National Biomedical Computation Resource. Work at Colorado State University is supported by NIH grant P01 AI 057846.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2011.09.094.

References

1. World Health Organization, WHO report 2009, Geneva.
2. Ginsberg, A. M.; Spigelman, M. *Nat. Med.* **2007**, *13*, 290.
3. Ma, Z.; Lienhardt, C.; McIlleron, H.; Nunn, A. J.; Wang, X. *Lancet* **2010**, *375*, 60359.
4. Lienhardt, C.; Vernon, A.; Raviglione, M. C. *Curr. Opin. Pulm. Med.* **2010**, *16*, 186.
5. Nikaido, H. *Science* **1994**, *264*, 382.
6. Barry, C. E., III; Lee, R. E.; Mdluli, K.; Sampson, A. E.; Schroeder, B. G.; Slayden, R. A.; Yuan, Y. *Prog. Lipid Res.* **1998**, *37*, 143.
7. Barry, C., III; Crick, D.; McNeil, M. *Infect. Disord. Drug Targets* **2007**, *7*, 182.
8. McNeil, M.; Daffé, M.; Brennan, P. *J. Biol. Chem.* **1990**, *265*, 18200.
9. Deng, L.; Mikusova, K.; Robuck, K.; Scherman, M.; Brennan, P.; McNeil, M. *Antimicrob. Agents Chemother.* **1995**, *39*, 694.
10. Pan, F.; Jackson, M.; Ma, Y.; McNeil, M. *J. Bacteriol.* **2000**, *183*, 3991.
11. Ma, Y.; Pan, F.; McNeil, M. R. *J. Bacteriol.* **2002**, *184*, 3392.
12. Sassetti, C.; Boyd, D.; Rubin, E. *Mol. Microbiol.* **2003**, *48*, 77.
13. Graninger, M.; Nidetzky, B.; Heinrichs, D.; Whitfield, C.; Messner, P. *J. Biol. Chem.* **1999**, *274*, 25069.
14. Blankenfeldt, W.; Kerr, I.; Giraud, M.; McMiken, H.; Leonard, G.; Whitfield, C.; Messner, P.; Graninger, M.; Naismith, J. *Structure* **2002**, *10*, 773.
15. Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
16. Fiser, A.; Do, R.; Sali, A. *Protein Sci.* **2000**, *9*, 1753.
17. Martí-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sánchez, R.; Melo, F.; Šali, A. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291.
18. Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D. G.; Thompson, J. D. *Nucleic Acids Res.* **2003**, *31*, 3497.
19. Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McWilliam, P. A. M. H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23*, 2947.
20. Grant, B. J.; Rodrigues, A. P. C.; ElSawy, K. M.; McCammon, J. A.; Caves, L. S. D. *Bio3D Bioinformatics* **2006**, *22*, 2695.
21. Glide, version 5.6, Schrödinger, LLC New York, NY, 2010.
22. Friesner, R.; Banks, J.; Murphy, R.; Halgren, T.; Klicic, J.; Mainz, D.; Repasky, M.; Knoll, E.; Shelley, M.; Perry, J.; Shaw, D.; Francis, P.; Shenkin, P. *J. Med. Chem.* **2004**, *47*, 1739.
23. Halgren, T.; Murphy, R.; Friesner, R.; Beard, H.; Frye, L.; Pollard, W.; Banks, J. *J. Med. Chem.* **2004**, *47*, 1750.
24. Friesner, R.; Murphy, R.; Repasky, M.; Frye, L.; Greenwood, J.; Halgren, T.; Sanschagrin, P.; Mainz, D. *J. Med. Chem.* **2006**, *49*, 6177.
25. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **1997**, *46*, 3.
26. Ma, Y.; Stern, R.; Scherman, M.; Vissa, V.; Yan, W.; Jones, V.; Zhang, F.; Franzblau, S.; Lewis, W.; McNeil, M. *Antimicrob. Agents Chemother.* **2001**, *45*, 1407.
27. Stern, R.; Lee, T.; Lee, T.; Yan, W.; Scherman, M.; Vissa, V.; Kim, S.; Wanner, B.; McNeil, M. *Microbiology* **1999**, *145*(Pt 3), 663.
28. Hoang, T.; Ma, Y.; Stern, R.; McNeil, M.; Schweizer, H. *Gene* **1999**, *237*, 361.
29. Sivendran, S.; Jones, V.; Sun, D.; Wang, Y.; Grzegorzewicz, A. E.; Scherman, M. S.; Napper, A. D.; McCammon, J. A.; Lee, R. E.; Diamond, S. L.; McNeil, M. *Bioorg. Med. Chem.* **2010**, *18*, 896.
30. Vander Jagt, D. L.; Deck, L. M.; Royer, R. E. *Curr. Med. Chem.* **2000**, *7*, 479.
31. Brown, W. M.; Metzger, L. E.; Barlow, J. P.; Hunsaker, L. A.; Deck, L. M.; Royer, R. E.; Vander Jagt, D. L. *Chem. Biol. Interact.* **2003**, *143–144*, 481.
32. Sun, D.; Scherman, M.; Jones, V.; Hurdle, J.; Woolhiser, L.; Knudson, S.; Lenaerts, A.; Slayden, R.; McNeil, M.; Lee, R. *Bioorg. Med. Chem.* **2009**, *17*, 3588.
33. Brown, J. R.; North, E. J.; Hurdle, J. G.; Morisseau, C.; Scarborough, J. S.; Sun, D.; Korduláková, J.; Scherman, M. S.; Jones, V.; Grzegorzewicz, A.; Crew, R. M.; Jackson, M.; McNeil, M. R.; Lee, R. E. in press. doi:10.1016/j.bmc.2011.07.034.
34. Takayama, K.; Wang, C.; Besra, G. S. *Clin. Microbiol. Rev.* **2005**, *18*, 81.
35. Jarlier, V.; Nikaido, H. *J. Bacteriol.* **1990**, *172*, 1418.
36. Chambers, H.; Moreau, D.; Yajko, D.; Miick, C.; Wagner, C.; Hackbarth, C.; Kocagoz, S.; Rosenberg, E.; Hadley, W.; Nikaido, H. *Antimicrob. Agents Chemother.* **1995**, *39*, 2620.
37. Ferreira, R.; Simeonov, A.; Jadhav, A.; Eidam, O.; Mott, B.; Keiser, M.; McKerrow, J.; Maloney, D.; Irwin, J.; Shoichet, B. *J. Med. Chem.* **2010**, *53*, 4891.
38. QikProp, version 3.3, Schrödinger, LLC New York, NY, 2010.