

Identification of new potential *Mycobacterium tuberculosis* shikimate kinase inhibitors through molecular docking simulations

Carolina Pasa Vianna · Walter F. de Azevedo Jr.

Received: 7 January 2011 / Accepted: 28 April 2011 / Published online: 19 May 2011
© Springer-Verlag 2011

Abstract Tuberculosis (TB) is the major cause of human mortality from a curable infectious disease, attacking mainly in developing countries. Among targets identified in *Mycobacterium tuberculosis* genome, enzymes of the shikimate pathway deserve special attention, since they are essential to the survival of the microorganism and absent in mammals. The object of our study is shikimate kinase (SK), the fifth enzyme of this pathway. We applied virtual screening methods in order to identify new potential inhibitors for this enzyme. In this work we employed MOLDOCK program in all molecular docking simulations. Accuracy of enzyme-ligand docking was validated on a set of 12 SK-ligand complexes for which crystallographic structures were available, generating root-mean square deviations below 2.0 Å. Application of this protocol against a commercially available database allowed identification of new molecules with potential to become drugs against TB. Besides, we have identified the binding cavity residues that are essential to intermolecular interactions of this enzyme.

Keywords Molecular docking · Shikimate kinase · Shikimate pathway · Tuberculosis · Virtual screening

Introduction

Tuberculosis (TB) is the most important cause of human death from a curable infectious disease. It is estimated that, worldwide, one hundred million people are infected annually and about ten million develop the disease, with five million of those progressing to an infectious stage, culminating with approximately three million deaths. According to the World Health Organization [1], the overall incidence of TB increases approximately 0.3% per year. The resurgence of this health problem occurred mainly due to the proliferation of multi (MDR-TB), extensively (XDR-TB), and recently, totally-drug (TDR-TB) resistant Mt strains. Besides, the high susceptibility of HIV/AIDS infected patients to TB is also a health problem. Therefore, there is an urgent need for the discovery and development of new and better drugs for the TB treatment [2].

Enzymes of the shikimate pathway (SP) are promising targets for the development of antimicrobial agents [3] and herbicides [4], because they are essential to the survival of algae, higher plants, bacteria, fungi, apicomplexan parasites and absent in mammals [5]. It is a seven-step biosynthetic route that converts erythrose 4-phosphate to chorismate, a precursor of aromatic amino acids and many other essential compounds [6].

The object of our study is the fifth enzyme of the SP, shikimate kinase (SK) (EC 2.7.1.71), which catalyzes the specific phosphorylation of the 3-hydroxy group of shikimate using ATP as a co-substrate resulting in shikimate-3-phosphate and ADP [7, 8]. This enzyme is an established target against Mt, since Parish and Stoeker demonstrated that the SP is essential for the viability of Mt due to the disruption of the *aroK* gene, which codes for the SK enzyme [9].

C. P. Vianna · W. F. de Azevedo Jr. (✉)
Faculdade de Biociências, Instituto Nacional de Ciência e Tecnologia em Tuberculose (INCT-TB), Laboratório de Bioquímica Estrutural (LaBioQuest), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS),
Av. Ipiranga 6681,
Porto Alegre, RS 90619-900, Brazil
e-mail: walter@azevedolab.net

C. P. Vianna · W. F. de Azevedo Jr.
Programa de Pós Graduação em Biologia Celular e Molecular,
Pontifícia Universidade Católica do Rio Grande do Sul,
Porto Alegre, RS, Brazil

SK is a member of the nucleoside monophosphate kinases (NMP kinases) family, which suffer large conformational changes during catalysis (Fig. 1) [10]. The enzymes of this family are composed of three domains: the CORE, which contains a highly conserved phosphate-binding loop (P-loop), the LID domain, which undergoes substantial structural changes upon substrate binding, and the NMP-binding domain which is responsible for the recognition and binding of a specific substrate [11].

Drugs are usually discovered by trial and error by means of high-throughput screening approaches that use in vitro experiments to evaluate the activity of a large number of compounds against a known target. This procedure is very costly and time-consuming. If crystallographic information is available for the protein target, then molecular docking simulations can be a helpful computational approach in the drug-discovery process [12]. Molecular docking is a simulation method that predicts the conformation of a receptor-ligand complex, in which the receptor can be either a protein or a nucleic acid, and the ligand is a small molecule. This computer simulation can generate many possible positions for the ligand in the receptor-binding pocket. Therefore, a criterion is necessary that will allow comparisons of all possible positions of the ligand, and then a selection can be made for the best position.

Our goal here is to find potential inhibitors against shikimate kinase from *Mycobacterium tuberculosis* MtSK using virtual screening (VS). VS can decrease costs and improve hits rates for lead discovery. For this, we used the MtSK structure [13] as a target for the molecular docking simulations with MOLDOCK [14]. Our docking protocol was validated against an ensemble of 12 crystallographic structures available for complexes of MtSK. The VS was validated by inclusion of a known SK inhibitor in the small-molecule database with over 4500 structures. We describe the results obtained in terms of the MOLDOCK scores, modes of interaction and discuss the importance of the active site residues in the ligand binding process.

Materials and methods

Molecular docking simulations

One of the fundamental questions in structural biology is the study of protein-ligand interactions, particularly considering the pharmacological applications of such study in the design of drugs based on structure [15]. To simulate the interaction of MtSK with a library of ligands, we used the MOLDOCK program [14], an implementation of a variant of the evolutionary algorithm (EA). Recent evaluation of MOLDOCK strongly indicates that it is capable of finding the right position of a ligand. Furthermore, MOLDOCK exhibits better overall performance compared with SURFLEX, FLEXX, and GOLD [14]. In the present work, all simulations were performed in an iMac (Intel Processor Core 2 Duo, 2.66 GHz, 2 GB SDRAM DDR3 1066 MHz).

Re-docking and cross-docking

In molecular docking simulations, the best binary complex (protein-ligand) is the one closer to the crystallographic structure. For that reason we must establish a methodology that assesses the distance from the computer-generated solution (pose) to the crystallographic structure. This distance can be calculated using the root-mean-square deviation (RMSD), which is a measure of the differences between values predicted by a model and the values actually observed from the object being modeled or estimated (protein-ligand complex). The RMSD is calculated between two sets of atomic coordinates, in this case, one for the crystallographic structure ($x_{ctal}, y_{ctal}, z_{ctal}$; the object being modeled) and another for the atomic coordinates obtained from the docking simulations ($x_{pose}, y_{pose}, z_{pose}$; predicted model). A summation is then taken over all N atoms being compared, using the following equation:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{ctal,i} - x_{pose,i})^2 + (y_{ctal,i} - y_{pose,i})^2 + (z_{ctal,i} - z_{pose,i})^2}. \quad (1)$$

In docking simulations, it is expected that the best results generate RMSD values less than 2.0 Å compared with crystallographic structures [16]. This procedure of obtaining the crystallographic position of the ligand is often called “re-docking,” which is fundamentally a validation method that determines whether the molecular docking algorithm is able to recover the crystallographic position using computer

simulation. In this work, all RMSD calculations were calculated for non-hydrogen atoms.

In order to validate our docking protocol, we used the SK crystallographic coordinates available at the protein data bank (PDB), under the access code 2DFN [13]. We performed the docking simulation against the active site of MtSK and compared the docked poses with the crystallographic

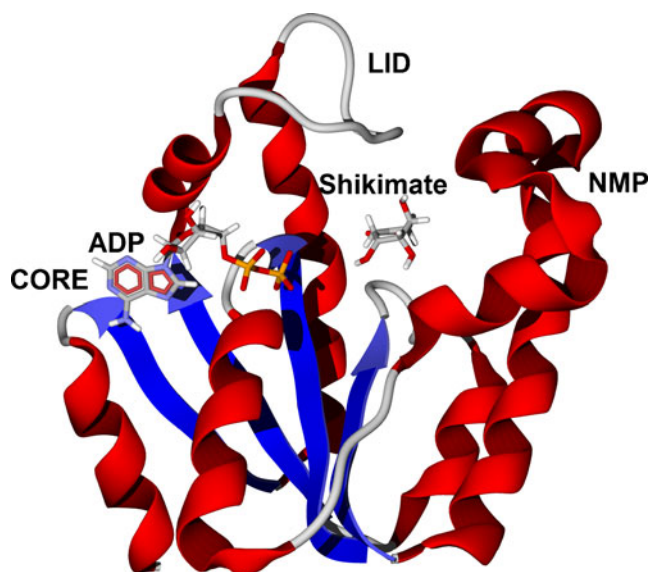


Fig. 1 Structure of shikimate kinase in complex with ADP and shikimate (PDB access code: 2DFN)

structure. We used the MOLDOCK default protocol with center at coordinates $x=(-15.23)$, $y=(-14.38)$, and $z=(14.88)$ Å, and a docking sphere radius of 9 Å. Figure 2 shows the docking sphere used in the simulations.

In the implementation of EA in MOLDOCK, computational approximations of an evolution course, called genetic operators, are applied to simulate the permanence of the most positive features. In a sample space, where there is a problem or a search routine and many different possible solutions (candidates), each option is ranked based on a set of parameters (scoring function or fitness function), and only the best ranked solutions are kept for the next iteration. This cycle is repeated until an optimal solution can be found. In the molecular docking simulations, the optimal solution is the one with the best scoring function, which should be the closest to the crystallographic structure. MOLDOCK presents two biological inspired algorithms to perform positional searches in docking simulations. One is

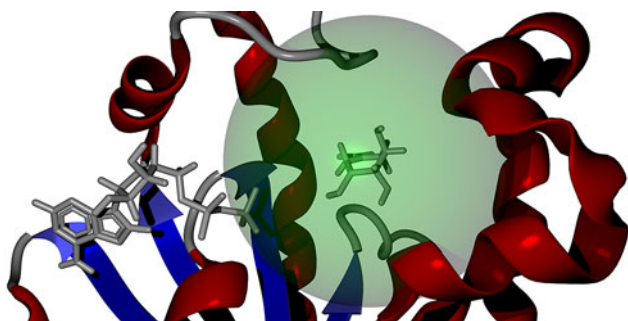


Fig. 2 Search space sphere (green) defined for molecular docking simulations

called the optimizer search algorithm (MOLDOCK Optimizer), which is based on an GDEA [14]. The second is a simplex evolution algorithm (SE) called MOLDOCK SE. GDEA is based on an EA adjustment called differential evolution (DE), which provides a distinct method for selecting and modifying candidate solutions (individuals). We used MOLDOCK Optimizer as search algorithm.

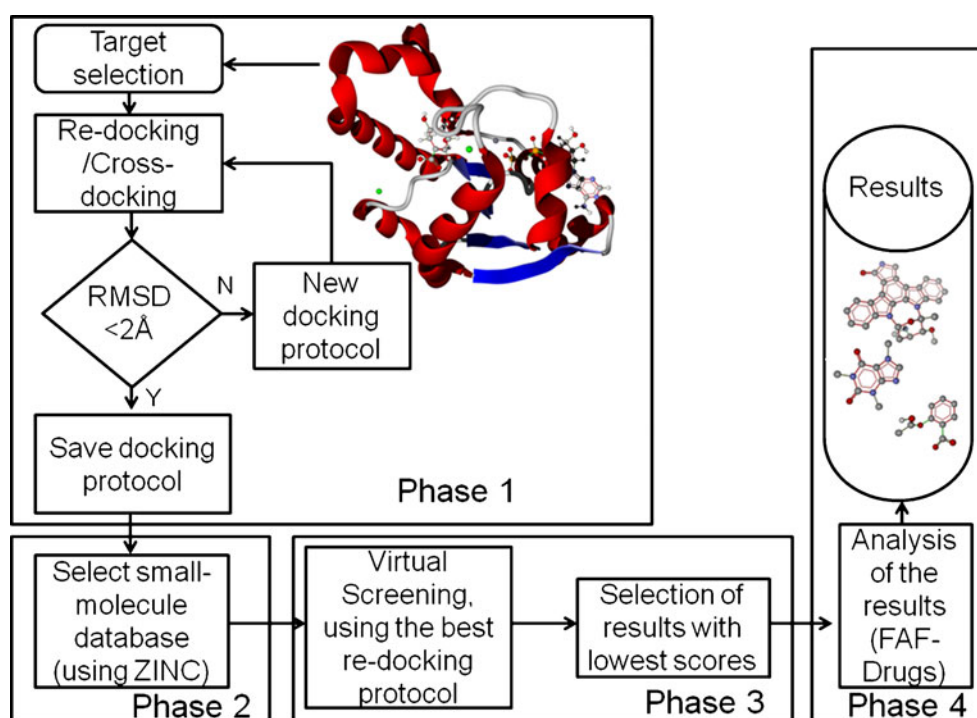
In addition to re-docking, a procedure called “cross-docking” can also be used to further validate a docking protocol. Considering that several crystallographic structures are available for the same protein, cross-docking can be applied. This procedure involves docking a number of ligands found in a variety of crystal structures of a protein identical to a single rigid protein crystallographic conformation [17]. When a protein target presents major conformational changes upon ligand binding, a significant difference is expected between the crystallographic and docked structures. We identified 12 MtSK structures in PDB with ligands in the shikimate-binding site (PDB access codes: 2DFN, 1U8A, 1WE2, 1ZYU, 2G1K, 2IYQ, 2IYR, 2IYS, 2IYX, 2IYY, 2IYZ, and 3BAF). This search was performed on February, 24th 2011. This validation procedures, re-docking and cross-docking, is the initial stage of a virtual screening protocol (**phase 1**) described in the next sections.

Virtual screening

Our virtual screening (VS) protocol is divided in four phases as shown in Fig. 3. **Phase 1** is focused on selection and validation of a docking protocol, as described earlier in the section re-docking and cross-docking. **Phase 1** ends when an adequate protocol is found (selection criterion $\text{RMSD} < 2.0$ Å). It should be pointed out that the RMSD criterion is dependent on the number of torsion angles, and a less demanding criterion may be adopted for re-docking of a ligand with a number of torsion angles higher than 10 [14]. Once a docking protocol is chosen we select a small-molecule database to be used in the screening (**phase 2**). Here we used a ligand library commercially obtainable at Acros Organics. The ligands (mol2 format) were downloaded from <http://zinc.dock.org> [18], with a total of 4579 small molecules. In addition to commercially oriented databases the ZINC database also provides an interface to build small-molecule databases based on molecular similarity, such as Tanimoto coefficient [19, 20].

In **phase 3**, we start docking simulations for each ligand present in the selected database. MOLDOCK program is the workhorse of the present protocol. It was used in all docking simulations described here. During a typical docking simulation several orientations can be obtained for each ligand. Here we selected the one with the lowest

Fig. 3 Flowchart of virtual screening process



scoring function. The scoring function used by MOLDOCK improves accuracy of scoring functions with a new hydrogen bonding term and new charge schemes. Four scoring functions are implemented in the MOLDOCK, including MOLDOCK score and PLANTS score [14, 21]. These two functions offer grid-based versions, in which hydrogen bond directionality is not considered. In the present protocol we employed grid-based MOLDOCK score since it offers approximately four-fold greater speed by performing a precalculation of potential-energy values on an equally spaced cubic grid.

The MOLDOCK score is based on the piecewise linear potential (PLP) scoring functions developed by Yang et al. [22, 23]. The docking scoring function $E_{MOLDOCK\ SCORE}$ is defined as the following:

$$E_{MOLDOCK\ SCORE} = E_{Intramol} + E_{intermol}, \quad (2)$$

where $E_{intermol}$ is the intermolecular interaction energy:

$$E_{intermol} = \sum_{i \in \text{ligand}} \sum_{j \in \text{protein}} \left[332 \frac{q_i q_j}{D r_{ij}} + E_{PLP}(r_{ij}) \right]. \quad (3)$$

All non-hydrogen atoms in the ligand and protein are taken in the summation. The first term accounts for electrostatic interactions, in which the factor 332 is used to obtain energy in kJ mol^{-1} . D represents the dielectric constant, which is the following: $D=4r_{ij}$. The second term (E_{PLP}) is a PLP, described elsewhere [22, 23]. To ensure that no energy term can be superior to the clash penalty, the electrostatic term is cut off at a level equivalent to the distance of 2.0 Å for distances less than 2.0 Å.

Intramolecular energy is given by the following equation:

$$E_{intramol} = E_{penalty} + \sum_{i \in \text{ligand}} \sum_{j \in \text{ligand}} E_{PLP}(r_{ij}) + \sum_{\text{singlebonds}} A[1 - \cos(n\phi - \phi_0)]. \quad (4)$$

The term $E_{penalty}$ is a penalty energy to be added to $E_{intramol}$ when two non-bonded atoms are closer than 2 Å (for non-hydrogen atoms). This term avoids unrealistic molecular topologies for the ligands. The second term is a PLP, already mentioned [22, 23]. The last term accounts for torsion energy, which is expressed as a periodic function. In this term, A , n , and ϕ_0 are empirically determined [22, 23]. MOLDOCK defines a limiting sphere where the search is focused. If a ligand non-hydrogen atom is positioned outside this limiting sphere (the search space sphere), then a constant penalty of 10000 is added to the total energy (implemented for the grid-based version of the MOLDOCK score).

After identification of potential inhibitors by molecular docking simulations, the best scored ligands were submitted to the web server FAF-Drugs [24], in order to assess physical-chemical properties (phase 4). These are key properties that need to be considered in early stages of the drug discovery process, and FAF-Drugs allows users to filter molecules via simple rules such as molecular weight, polar surface area, logP and number of rotatable bonds. The ligands were filtered following the Lipinski's rule of five (RO5). RO5 advocates that drugs which present oral

bioavailability, in general, follow: molecular weight less or equal to 500, $\text{Log}P$ less or equal to 5, number of hydrogen bond donor groups less or equal to 5 and number of hydrogen bond acceptor groups less or equal to 10 [25].

In addition to the 4579 small molecules present in the Acros database we added staurosporine (PubChem Compound Identification: CID 44259) to the database to be used in the VS. This molecule has been already tested directly on MtSK [26]. This addition allows testing whether this VS protocol is able to identify SK inhibitors present in a database with over 4,500 ligands.

Enrichment factor

In order to further validate the present VS protocol we calculated the enrichment factor (EF), which takes into account the improvement of the hit rate by a VS protocol compared to a random selection. EF is defined by the following equation,

$$EF = \frac{H_a/H_t}{A/N} \quad (5)$$

where H_a is the number of active compounds in the H_t top-ranked compounds of a total database of N compounds of which A are active [27, 28]. Successfully VS implies $EF \gg 1$. For this validation simulation, a sub-set of a previously described kinase decoy database (containing 627 molecules) was spiked with the four known MtSK inhibitors ($A=4$) [26]. This database for MtSK decoys and active MtSK inhibitors is available for downloading at: http://azevedolab.dominiotemporario.com/doc/mtsk_decoys_set.zip.

Results and discussion

Docking and cross-docking

Re-docking simulations (phase 1 of the VS protocol) using the structure 2DFN generated an RMSD of 1.6 Å. In addition, cross-docking simulations generated RMSD ranging from 1 to 2 Å further validating the present docking protocol. These two tests indicated that the docking simulation was successful, and that the protocol is good enough to be used for the virtual screening process.

Virtual screening

VS uses computational methodologies to identify biologically active molecules against a specific protein target. Two main methodologies are used in VS. Methods that search for similarity to validated ligands and molecular docking methods that require the use of crystallographic information

of the target. Here we made use of the second approach. VS studies performed by other research groups have been previously published on the identification of MtSK inhibitors as antitubercular drugs by similar molecular docking procedures [29, 30]. Nevertheless, very limited information (experimental data) on the direct effect of these compounds on MtSK is yet available. The novelty of the present work relies on the method used in the VS and the selection of compounds according to their pharmacological properties. The VS simulations were carried out using the MOLDOCK program, having as target the MtSK (PDB access code 2DFN). The ligand library comprises 4580 molecules (Acros database plus staurosporine). Addition of a known SK inhibitor allows testing the accuracy of the present protocol.

After docking simulations, we selected 20 top-scoring compounds from the initial set of 4580 compounds (selection based on MOLDOCK score). Staurosporine was present in the 20 top-scoring compounds obtained in the VS, with MOLDOCK score of -144.168 . Identification of a known SK inhibitor among the best VS results gives further validation for this VS protocol. These 20 potential inhibitors were submitted to filter tests, available at the web server FAF-Drugs [24], to exclude those compounds that have known undesirable physical-chemical features to oral bioavailability. We could have applied this filter analysis previous to docking simulations, since it would reduce simulation time. Nevertheless, we kept filtering analysis after docking simulation, since the MOLDOCK protocol was fast enough to be run in less than a week of CPU time of an iMac (Intel Processor Core 2 Duo, 2.66 GHz, 2 GB SDRAM DDR3 1066 MHz). In addition, application of filtering analysis previous to docking simulations could eliminate candidates that fail to filtering analysis but present promising MOLDOCK score, which could have toxicity reduced by small modification in the structure.

Especially interesting is the fact that staurosporine is a well-known cyclin-dependent kinase (CDK) inhibitor that has a plethora of structural and functional studies [31–35]. Staurosporine is non-selective and too toxic for use in therapy, but UCN-01, a hydroxylated form of staurosporine (7-hydroxystaurosporine), shows greater selectivity for CDK and is currently undergoing clinical trials in the United States and Japan [33]. This opens new possibilities to test new molecular moieties as potential SK inhibitors, the CDK inhibitors that have already shown low toxic effects make a promising dataset to be explored as potential SK inhibitors.

The FAF-Drugs parameters used were those of the Lipinski's role of five [25]. From the set with 20 selected molecules, nine fit the Lipinski's role of five, which includes staurosporine. Figures 4a–i show the molecular

Fig. 4 Molecular structures of the top-scoring compounds identified in the VS protocol. **a)** Staurosporine. **b)** ZINC15707201. **c)** ZINC20462780. **d)** ZINC15707234. **e)** ZINC15675581. **f)** ZINC15707188. **g)** ZINC22936889. **h)** ZINC20464408. **i)** ZINC22936937

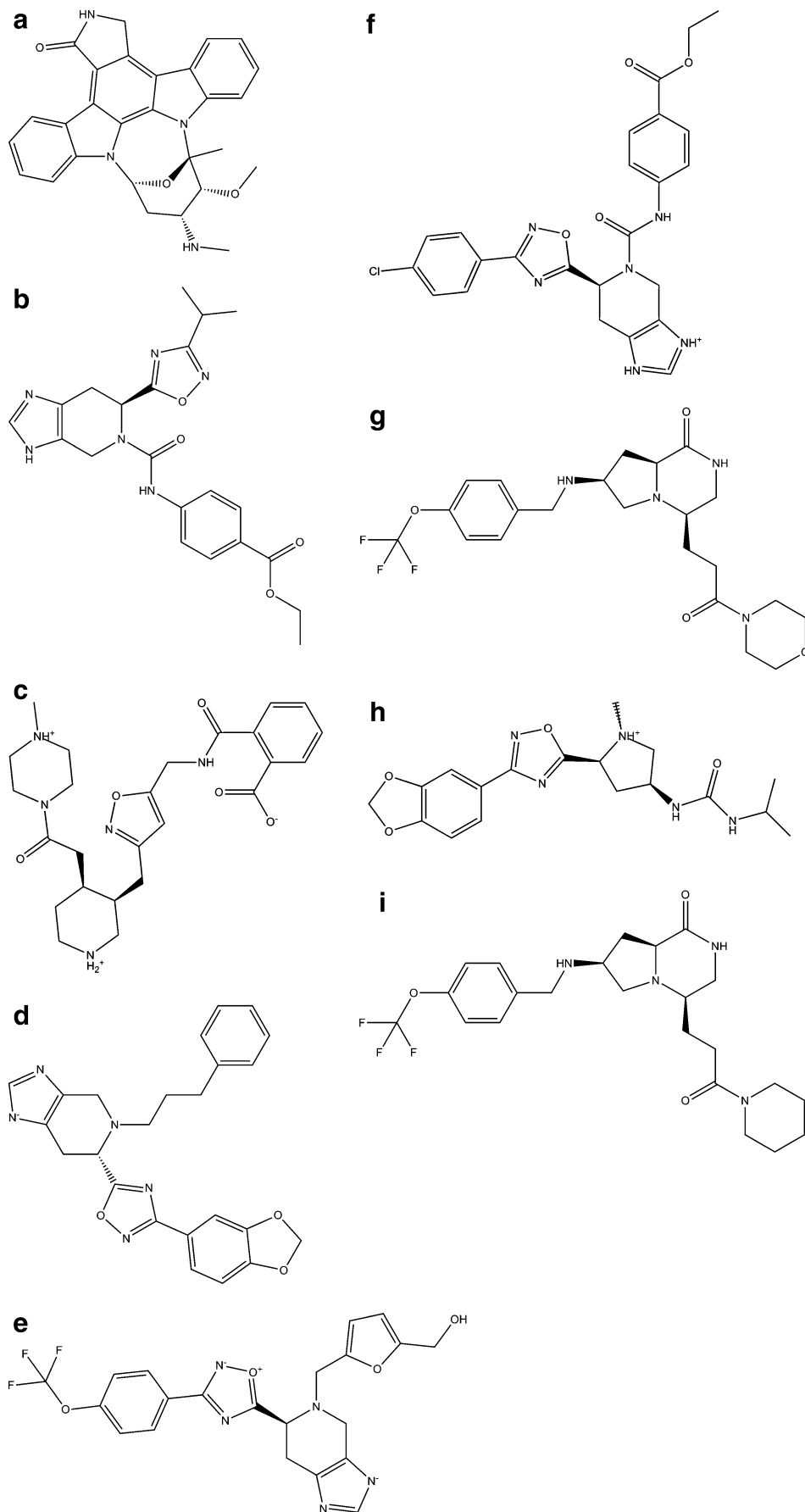


Table 1 Physical-chemical properties of ligands that fitted the Lipinski's role of five after analysis by FAF-Drugs

Ligand	ZINC code	Molecular weight (Da)	Number of HB acceptors	Number of HB donors	LogP
1	15707201	423.3	10	1	0.82
2	20462780	485.3	10	4	-0.3
3	15707234	428.3	8	0	2.62
4	15675581	460.3	9	2	1.57
5	15707188	490.8	10	1	2.61
6	22936889	471.3	8	3	1.27
7	20464408	373.2	9	2	1.5
8	22936937	469.3	7	3	2.54

structures for all nine ligands. Since staurosporine is already a known SK inhibitor we excluded it from the rest of the analysis. Staurosporine was included only to test the VS protocol. The selected ligands are shown in Table 1. The MOLDOCK scores for these eight molecules ranging from -144.208 to -151.943. All eight ligands show MOLDOCK scores better than staurosporine.

Enrichment factor

The complete database used to calculate the EF contains 627 decoys and four active molecules (N=631). After application of the present VS protocol against this database, six molecules ($H_t=6$) were retrieved as hits (0.95% of the database). Among these hits, two molecules were known MtSK inhibitors ($H_a=2$) [26]. Thus, the enrichment factor was found to be 52.6, indicating that it is 52.6 times more likely to pick an active compound from the database than an inactive one. Previously published benchmarking sets for molecular docking of kinases exhibited EF ranging from 1.2 to 54, indicating that the present VS protocol is adequate for our purposes.

Intermolecular interactions

In order to better understand the interactions of these eight molecules with MtSK, we used the program LIGPLOT [36] to access the atoms of both, the small molecules and the protein ones that are responsible to make hydrogen bonds and van der Waals contacts. A comparison among the MOLDOCK score values obtained for these ligands, is not enough yet to predict activity, since in vitro assays are necessary to conclude this. Therefore it is not possible to say that the selected compounds, the ones with the best MOLDOCK scores, would be the most potent ones. We could observe, only, that among the selected compounds the best scores mean a greater potential to interact with the shikimate-binding cavity.

The docking simulation results corroborate the importance of some shikimate-active site residues as responsible

to establish intermolecular interactions with the substrate as well as with the tested ligands. The binding of shikimate to its cavity, presents pivotal residues that make protein-ligand interactions possible, as shown in Fig. 5. These residues are essential to the ligand binding and, finally, to the reaction catalyzed by the enzyme. The SK residues that perform intermolecular hydrogen bonds (HB) with the shikimate are: Gly80, Arg136 and Arg58. SK makes van der Waals contacts with residues: Ile45, Asp34, Pro11, Pro118, Gly79, Phe57, Leu119 and Gly81.

Information about intermolecular interactions for all eight top-scoring compounds is summarized in Table 2. Analysis of shikimate-binding site indicated that all top-scoring compounds present interaction with residues Lys15, Ser16 and Arg117. Figure 6 shows the intermolecular interactions for the top-scoring compound (ligand 1, ZINC15707201). This figure is representative of the positioning of all top-scoring compounds in the shikimate-binding pocket. Ligands 1, 2, 4, 5, 6 and 8 highlight the presence of residue Val116, suggesting it is also relevant to

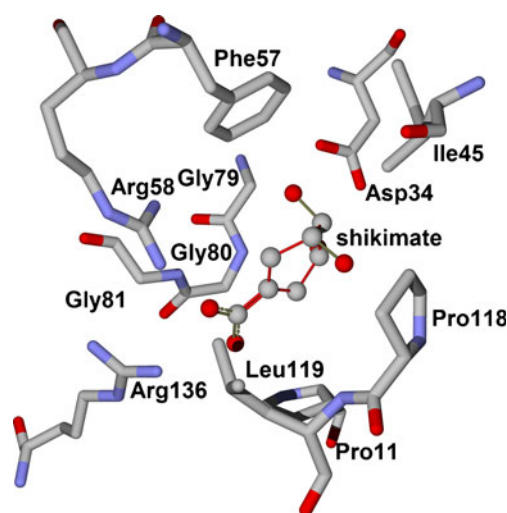


Fig. 5 Shikimate-binding pocket with main residues found in intermolecular interactions with shikimate

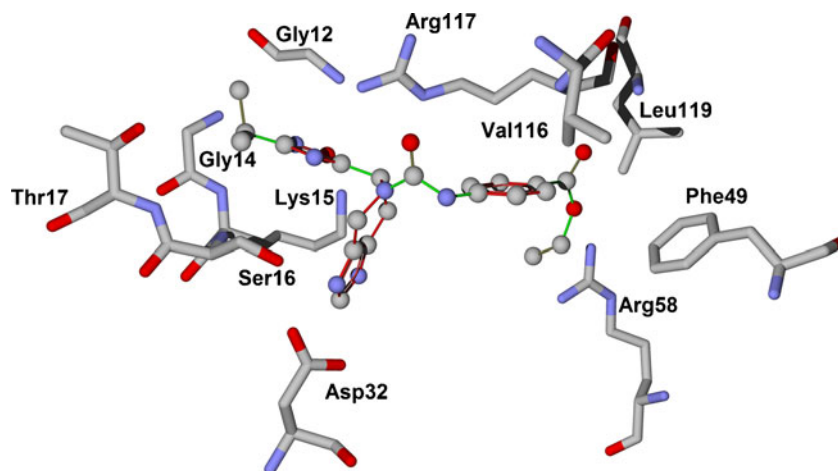
Table 2 Intermolecular interactions for the top-scoring ligands selected in the VS procedure. The presence of an X indicates that the interaction occurs. HB means hydrogen bonds and VDW means van der Waals contacts

Residues	Ligands							
	Zinc15675581	Zinc15707188	Zinc15707201	Zinc15707234	Zinc20462780	Zinc20464408	Zinc22936889	Zinc22936937
HB								
Gly12						X		
Ser13						X		
Gly14						X		
Lys15	X	X	X	X	X	X		
Ser16		X	X					X
Thr17	X					X	X	X
Asp32		X	X					
Asp34	X						X	X
Leu78		X						
Arg117	X	X		X	X	X		
Leu119		X	X	X		X	X	
VDW								
Leu10					X			
Gly12	X	X	X	X		X		
Gly14	X	X	X	X		X	X	X
Lys15							X	X
Ser16	X			X	X	X	X	
Thr17		X	X	X	X			
Val35	X							
Gly38	X			X				
Phe49	X	X	X	X		X		X
Arg58			X					X
Gly80					X			
Val116	X	X	X	X	X		X	X
Arg117			X				X	X

intermolecular interactions. Two previously published molecular docking studies focused on MtSK were able to identify intermolecular molecular interactions with the same residues [29, 30], further corroborating the pivotal

importance of these residues for ligand-binding affinity. Especially interesting is the fact that these previous molecular docking studies analyzed completely different molecular moieties, such as dipeptides (arginine-aspartate/

Fig. 6 Shikimate-binding pocket with main residues found in intermolecular interactions with the top-scoring compound (ZINC15707201)



lysine-aspartate) [30] and triazole/tetrazole heteroaromatic systems [29]. These molecular structures were not present in the database used in the present study.

Conclusions

Advanced molecular docking algorithms available nowadays make it possible to undertake larger virtual screening studies focused on small-molecules libraries up to millions of compounds. Here we described an efficient molecular docking protocol, which was able to recover crystallographic position of a ligand present in the active site of the SK. Re-docking and cross-docking simulations generated RMSD results below 2 Å. The virtual screening protocol was able to confirm a known SK inhibitor, staurosporine, as a top-scoring compound and presents an enrichment factor of 52.6. Furthermore, the present work indicates new molecules with the potential to become drugs against TB. Besides, we identified the MtSK binding-cavity residues that are essential to make possible the interactions of this enzyme with a variety of molecules. Analysis of the top-scoring compounds also indicates that MtSK has the ability to bind a variety of molecular moieties not previously identified.

Acknowledgments The authors would like to express their gratitude to the reviewers for their valuable comments and suggestions. This work was supported by National Institute of Science and Technology on Tuberculosis (Decit/SCTIE/MS-MCT-CNPq-FNDCTCAPES). W. F.A. Jr. is a research fellow of the National Council for Scientific and Technological Development of Brazil (CNPq). C.P.V. acknowledges a scholarship awarded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

References

- World Health Organization (2009) WHO Report. Geneva, Switzerland. WHO/CDS/TB/2009. <http://www.who.int/topics/tuberculosis/en/>. Accessed in October 2010
- Basso LA, Pereira Da Silva LH, Fett-Neto AG, De Azevedo Jr WF, Moreira IS, Palma MS, Calixto JB, Astolfi Filho S, dos Santos RR, Soares MBP, Santos DS (2005) Mem Inst Oswaldo Cruz 100:475–506
- Davies GM, Barrett-Bee KJ, Jude DA, Lehan M, Nichols WW, Pinder PE, Thain JL, Watkins WJ, Wilson RG (1994) Antimicrob Agents Chemother 38:403–406
- Coggins JR (1989) The shikimate pathway as a target for herbicides. In: Dodge AD (ed) Herbicides and Plant metabolism. Cambridge University Press, Cambridge, UK, pp 97–112
- Bentley R (1990) The shikimate pathway—a metabolic tree with many branches. Crit Rev Biochem Mol Biol 25:307–384
- Ratledge C (1982) Nutrition, growth and metabolism. In: Ratledge C, Stanford JL (eds) The Biology of the Mycobacteria, vol 1. Academic, London, pp 185–271
- Krell T, Coggins JR, Laphorn AJ (1998) The three-dimensional structure of shikimate kinase. J Mol Biol 278:983–997
- Pereira JH, Oliveira JS, Canduri F, Dias MVB, Palma MS, Basso LA, Santos DS, De Azevedo Jr WF (2004) Acta Crystallogr D Biol Crystallogr 60:2310–2319
- Parish T, Stoker NG (2002) The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. Microbiology 148:3069–3077
- Vonrhein C, Schlauderer GJ, Schulz GE (1995) Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. Structure 3:483–490
- Gu Y, Reshetnikova L, Li Y, Wu Y, Yan H, Singh S, Ji X (2002) Crystal structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the dynamic role of the LID domain in catalysis. J Mol Biol 319:779–789
- Peitsch MC (2004) Manuel Peitsch discusses knowledge management and informatics in drug discovery. Drug Discov Today Biocilico 02:94–96
- Dias MV, Vasconcelos IB, Prado AM, Fadel V, Basso LA, De Azevedo Jr WF, Santos DS (2007) Acta Crystallogr F Struct Biol Cryst Commun 63:1–6
- Thomsen R, Christensen MH (2006) MolDock: a new technique for high-accuracy molecular docking. J Med Chem 11:3315–3321
- De Azevedo Jr WF, Canduri F, Da Silveira NJF (2002) Biochem Biophys Res Commun 293:566–571
- Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shaw DE, Shelley M, Perry JK, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47:1739–1749
- Thilagavathi R, Mancera RL (2010) Ligand-protein cross-docking with water molecules. J Chem Inf Model 50:415–421
- Irwin JJ, Shoichet BK (2005) ZINC - a free database of commercially available compounds for virtual screening. J Chem Inf Model 45:177–182
- Timmers LFS, Pauli I, Caceres RA, De Azevedo Jr WF (2008) Drug-binding databases. Curr Drug Targets 9:1092–1099
- De Azevedo Jr WF (2010) Structure-based virtual screening. Curr Drug Targets 11:261–263
- De Azevedo Jr WF (2010) MolDock applied to structure-based virtual screening. Curr Drug Targets 11:327–334
- Yang JM (2004) Development and evaluation of a generic evolutionary method for protein-ligand docking. J Comput Chem 25:843–857
- Yang JM, Chen CC (2004) GEMDOCK: a generic evolutionary method for molecular docking. Proteins 55:288–304
- Miteva MA, Violas S, Montes M, Gomez D, Tuffery P, Villoutreix BO (2006) FAF-drugs: free adme/tox filtering of compound collections. Nucleic Acids Res 34:W738–W744
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 23:3–25
- Mulabagal V, Calderón AI (2010) Development of an ultrafiltration-liquid chromatography/mass spectrometry (UF-LC/MS) based ligand-binding assay and an LC/MS based functional assay for *Mycobacterium tuberculosis* shikimate kinase. Anal Chem 82:3616–3621
- Huang N, Shoichet BK, Irwin JJ (2006) Benchmarking sets for molecular docking. J Med Chem 49:6789–801
- Brooijmans N, Sharp KA, Kuntz ID (2003) Molecular recognition and docking algorithms. Annu Rev Biophys Biomol Struct 32:335–373
- Segura-Cabrera A, Rodriguez-Perez MA (2008) Structure-based prediction of *Mycobacterium tuberculosis* shikimate kinase inhibitors by high-throughput virtual screening. Bioorg Med Chem Lett 18:3152–3157

30. Kumar M, Verma S, Sharma S, Srinivasan A, Singh TP, Kaur P (2010) Structure-based in silico design of a high-affinity dipeptide inhibitor for novel protein drug target Shikimate kinase of *Mycobacterium tuberculosis*. *Chem Biol Drug Des* 76:277–284
31. Noble ME, Endicott JA, Johnson LN (2004) Protein kinase inhibitors: insights into drug design from structure. *Science* 303:1800–1805
32. Canduri F, De Azevedo Jr WF (2005) Structural basis for interaction of inhibitors with cyclin-dependent kinase 2. *Curr Comput-Aid Drug* 1:53–64
33. Lawrie AM, Noble ME, Tunnah P, Brown NR, Johnson LN, Endicott JA (1997) Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. *Nat Struct Biol* 4:796–801
34. De Azevedo Jr WF, Mueller-Dieckmann HJ, Schulze-Gahmen U, Worland PJ, Sausville E, Kim SH (1996) Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase. *Proc Natl Acad Sci USA* 93:2735–2740
35. De Azevedo Jr WF, Leclerc S, Meijer L, Havlicek L, Strnad M, Kim SH (1997) Inhibition of cyclin-dependent kinases by purine analogues: crystal structure of human cdk2 complexed with roscovitine. *Eur J Biochem* 243:518–526
36. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Prot Eng* 8:127–134