



รายงานวิจัยฉบับสมบูรณ์

โครงการเรื่อง การค้นหาและการออกแบบสารบัญยั้งที่ทำหน้าที่ได้สอง แบบเพื่อใช้เป็นสารตัวใหม่ในการต้านวัณโรค

โดย ผศ.ดร.อรดี พันธุ์กว้าง

เดือน ปี ที่เสร็จโครงการ

4 พฤษภาคม 2561

สัญญาเลขที่ MRG5980152

รายงานวิจัยฉบับสมบูรณ์

โครงการเรื่อง
การค้นหาและการออกแบบสารยับยั้งที่ทำหน้าที่ได้สองแบบ
เพื่อใช้เป็นสารตัวใหม่ในการต้านวัณโรค

ผู้วิจัย ผศ.ดร.อรดี พันธุ์กว้าง
คณะวิทยาศาสตร์ มหาวิทยาลัยนครพนม

นักวิจัยที่ปรึกษา รศ.ดร.พรพรรณ พึ่งโพธิ์
คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

สนับสนุนโดยสำนักงานคณะกรรมการการอุดมศึกษา
และสำนักงานกองทุนสนับสนุนการวิจัย

(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกอ. และ สกว. ไม่จำเป็นต้องเห็นด้วยเสมอไป)

Abstract

Project Code : MRG5980152

Project Title : Identification and rational design of the bi-functional inhibitor as novel potent anti-tuberculosis agent through molecular modelling approaches

Investigator : Auradee Punkvang

E-mail Address : auradee.punkvang@npu.ac.th

Project Period : 2 years

Tuberculosis caused by *M. tuberculosis* remains as a major world health problem due to drug resistance of *M. tuberculosis*. The enzymes of shikimate pathway are potential drug targets for the development of anti-TB agents because they are essential for growth of *M. tuberculosis* and absent in humans. The third and fourth enzymes of the shikimate pathway (DHQ dehydratase and shikimate dehydrogenase, respectively) have similar substrates. Accordingly, it is possible to identify the bi-functional inhibitors that could block both of DHQ dehydratase and shikimate dehydrogenase. The bi-functional inhibitors should be the promising anti-TB agents with the new inhibition mechanism different from the existing TB drugs. Therefore, this work aims to rational design and identify bi-functional inhibitors as novel potent anti-TB agents against both of DHQ dehydratase and shikimate dehydrogenase. The combination of MD simulations and CoMSIA studies were employed in order to investigate the structural insight for rational design of bi-functional inhibitors. Furthermore, virtual screening based on the hybrid ligand and structural base approach was used to identify the promising compound from the commercial database as the bi-functional inhibitors. The integrated results obtained from this work could guide us to propose potent bi-functional inhibitors of DHQ dehydratase and shikimate dehydrogenase.

Keywords : Tuberculosis, DHQ dehydratase, shikimate dehydrogenase, MD simulations, QSAR

บทคัดย่อ

รหัสโครงการ: MRG5980152

ชื่อโครงการ: การค้นหาและการออกแบบสารยับยั้งที่ทำหน้าที่ได้สองแบบเพื่อใช้เป็นสารตัวใหม่ในการต้านวัณโรค

ชื่อนักวิจัย: อรดี พันธุ์กว้าง สถาบัน: มหาวิทยาลัยนครพนม

E-mail Address: auradee.punkvang@npu.ac.th

ระยะเวลาโครงการ : สองปี

วัณโรคมีสาเหตุมาจากเชื้อไมโคแบคทีเรียทูปเบอร์คิวโลซิสและยังคงเป็นปัญหาหลักทางด้านสุขภาพในทั่วโลก เอนไซม์ที่เกี่ยวข้องกับกระบวนการ shikimate เป็นเอนไซม์เป้าหมายที่มีศักยภาพสำหรับการพัฒนาสารต้านวัณโรค เนื่องจากเอนไซม์เหล่านี้มีความสำคัญต่อการเจริญเติบโตของเชื้อไมโคแบคทีเรียทูปเบอร์คิวโลซิสและเป็นเอนไซม์ที่ไม่พบในคน เอนไซม์ในขั้นตอนที่สามและสี่ของกระบวนการ shikimate (เอนไซม์ DHQ dehydratase และเอนไซม์ shikimate dehydrogenase ตามลำดับ) มี substrate ที่คล้ายคลึงกัน ดังนั้นจึงมีความเป็นไปได้ที่จะค้นหาสารยับยั้งที่สามารถยับยั้งได้ทั้งสองเอนไซม์ (bi-functional inhibitor) ซึ่งสารยับยั้งนี้น่าจะเป็นสารต้านวัณโรคที่มีกลไกยับยั้งแตกต่างจากยาต้านวัณโรคที่มีอยู่ ดังนั้นในงานวิจัยนี้จึงมีวัตถุประสงค์เพื่อออกแบบและระบุ bi-functional inhibitor เพื่อใช้เป็นสารตัวใหม่ในการต้านวัณโรค โดยได้นำเอาการศึกษาการจำลองแบบพลวัต (MD simulation) และการศึกษาความสัมพันธ์ระหว่างโครงสร้างและกัมมันตภาพในการยับยั้งในเชิงปริมาณ (CoMSIA) มาใช้ในการศึกษาเพื่อให้เข้าใจถึงลักษณะสำคัญทางโครงสร้างที่จะสามารถนำไปออกแบบ bi-functional inhibitor ได้ และนอกจากนี้ยังได้ทำการคัดสรรเสมือนจริง (virtual screening) เพื่อคัดสรรสารจากฐานข้อมูลเพื่อให้ได้สารที่มีคุณสมบัติ bi-functional inhibitor จากข้อมูลทั้งหมดที่ได้จากการศึกษาทำให้สามารถระบุโครงสร้างสารที่มีความเป็นไปได้ในการยับยั้งได้ทั้งเอนไซม์ DHQ dehydratase และ เอนไซม์ shikimate dehydrogenase

คำสำคัญ : วัณโรค, เอนไซม์ DHQ dehydratase, เอนไซม์ shikimate dehydrogenase, การจำลองแบบพลวัต

Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

1. Auradee Punkvang, Pharit Kamsri, Patchareenart Saparpakorn, Supa Hannongbua and Pornpan Pungpo. Assessing structural concepts to balance the activity of 2,3-anhydroquinone derivatives against both type II dehydroquinase and *Mycobacterium tuberculosis*. Manuscript
2. Auradee Punkvang, Pharit Kamsri, Adrian Mulholland, Jim Spencer, Supa Hannongbua, and Pornpan Pungpo. Simulations of Shikimate Dehydrogenase from *Mycobacterium tuberculosis* in Complex with 3-dehydroshikimate and NADPH Suggest the Structural Concept of a Hybrid MtbSDH Inhibitor. Manuscript

CONTENTS

	PAGE
EXECUTIVE SUMMARY	1
OBJECTIVES	4
LITERATURE REVIEW	5
METHODOLOGY	12
Molecular dynamics simulations	12
CoMSIA method	12
Molecular docking calculations	16
Virtual screening	17
RESULT AND DISCUSSION	18
<i>Mtb</i>DHQ inhibitors	18
Quantitative investigation for <i>Mtb</i> DHQ-inhibitor interaction	18
The influence of R substituent on the binding affinity	20
The structural requirement of R substituent	21
The positional change of <i>Mtb</i> DHQ for inhibitor binding	24
The structural concept to design <i>Mtb</i> DHQ inhibitors	26
The removal of negatively charged carboxylate group	27
Addition of a non-sterically encumbered amine	27
<i>Mtb</i>SDH inhibitors	28
Stability of <i>Mtb</i> SDH on MD simulations	28
Binary DHS/ <i>Mtb</i> SDH complex structure	29
K69A and D105N mutations	31
Ternary DHS/NADPH/ <i>Mtb</i> SDH complex structure	33
NADPH binding in <i>Mtb</i> SDH	35
RMSF of <i>Mtb</i> SDH	36
Positional change of <i>Mtb</i> SDH	37
Rational design guideline for the hybrid <i>Mtb</i> SDH inhibitor	38
Core structural design of bi-funtional inhibitor	40
Virtual screening of bi-funtional inhibitor	40
CONCLUSION	44
APPENDIX	46

EXECUTIVE SUMMARY

Tuberculosis (TB) is an infection disease caused by bacteria named *Mycobacterium tuberculosis* (*M. tuberculosis*), which most commonly affects the lungs. TB is treatable with the standard regimens using the combination of existing TB drugs such as isoniazid and rifampicin, two of the most effective TB drugs. However, *M. tuberculosis* strains could develop the resistance to TB drugs. It means that these drugs did not effective for TB treatment. Accordingly, there is a need for the development of new drugs as well as new targets to treat drug resistant TB. The potential strategy for the development of new drugs against *M. tuberculosis* is to target essential biosynthetic pathways of this bacteria that are absent in humans. This is very important for specific drug design with low toxicity. The shikimate pathway is the biosynthetic pathway essential for growth of *M. tuberculosis*. It produces the chorismate, an important precursor for biosynthesis of aromatic compounds such as aromatic amino acids (phenylalanine, tyrosine, and tryptophan), vitamins E and K in bacteria. Remarkably, the shikimate pathway is absent in humans. Therefore, the enzymes of this pathway are potential drug targets for the development of anti-TB agents. The shikimate pathway of *M. tuberculosis* comprises seven different enzymes, each of which catalyzes a separate step of the pathway that converts phosphoenolpyruvate and D-erythrose to chorismate. The third and fourth enzymes of the shikimate pathway (DHQ dehydratase and shikimate dehydrogenase, respectively) catalyze the chemical reactions of 3-dehydroquinate (DHQ) and 3-dehydroshikimate (DHS) substrates, respectively. These substrates have similar chemical structures. Accordingly, it could be possible to identify or design the bi-functional inhibitors that could block both of DHQ dehydratase (*MtbDHQ*) and shikimate dehydrogenase (*MtbSDH*). The bi-functional inhibitors should be the promising anti-TB agents with the new inhibition mechanism different from the existing TB drugs. Therefore, this work aims to gain the structural insight for rational design and identification of bi-functional inhibitors as new potent anti-TB agents against both of *MtbDHQ* and *MtbSDH*. The structural concepts for designing 2,3-anhydroquinat

derivatives that can accumulate inside *M. tuberculosis* and target mtDHQ were reported here. These will assist in the rational design of potent compounds against *M. tuberculosis* targeting *MtbDHQ*. Integration of various computational approaches (per-residue free energy decomposition, Comparative Molecular Similarity Index Analysis (CoMSIA) and molecular dynamics (MD) simulations) into the current guiding principles for compound accumulation in gram-negative bacteria is our strategy to develop structural concepts. The carboxylate moiety is crucial for binding affinity of *MtbDHQ* inhibitors, but it impedes accumulation inside the bacterial cell. The replacement of the negatively charged carboxylate moiety with a neutral hydrogen bond acceptor should balance the activity of *MtbDHQ* inhibitors against both *M. tuberculosis* and *MtbDHQ*. Due to the capability of changing the position of *MtbDHQ* residues, the R substituent attached to the core structure requires a small linker and bulky end cap. Importantly, the presence of NH_3^+ at the end of R substituent might enhance the activity of 2,3-anhydroquinone derivatives against both *M. tuberculosis* and *MtbDHQ*. In the case of *MtbSDH*, the missing of binary DHS/*MtbSDH* complex and ternary DHS/NADPH/*MtbSDH* complex conceal insight into the binding of DHS substrate and NADPH cofactor as well as the catalytic reactions of *MtbSDH*. Here, molecular dynamics (MD) simulations were performed to create these *MtbSDH* complex structures. Several hydrogen bonds created by Ser18, Ser20, Thr65, Lys69, Asn90, Gln243 and Gln247 have prominent contributions to DHS binding in *MtbSDH*. Mutations of highly conserved Lys69 and Asp105 in *MtbSDH* did not greatly influence the binding of DHS substrate. Asp105 holds the ϵ -ammonium group of Lys69 in a suitable position for *MtbSDH* catalysis via a salt bridge. The coordinates of DHS and NADPH in the ternary *MtbSDH* complex obtained from our work are consistent with a C4-*proS* hydride transfer from NADPH to DHS and proton transfer from Lys69 to DHS. The substrate binding pocket size of *MtbSDH* is highly specific for its substrate and the catalytic residue is held in a suitable position for reaction. The structural details observed from the *MtbSDH* complex structures in this work allow us to provide a rational design guideline for a hybrid *MtbSDH* inhibitor. The integrated results obtained from this work could guide us

to propose core structures of bi-functional inhibitors of *Mtb*DHQ and *Mtb*SDH. Then, these core structures were utilized for virtual screening to identify the promising compound from Specs database as the bi-functional inhibitors for both *Mtb*DHQ and *Mtb*SDH.

OBJECTIVES

1. To investigate the structural concept for rational design of bi-functional inhibitors against both of *Mtb*DHQ and *Mtb*SDH through CoMSIA and MD simulation approaches
2. To design the core structure of bi-functional inhibitors of *Mtb*DHQ and *Mtb*SDH on the basis of the obtained structural concept
3. To identify bi-functional inhibitors from Specs database using virtual screening

LITERATURE REVIEW

Tuberculosis (TB) caused by *M. tuberculosis* remained one of the top 10 causes of death worldwide in 2015. An estimated 10.4 million new (incidences) TB cases worldwide and an estimated 480 000 new cases of multidrug-resistant TB (MDR-TB) have been reported by the World Health Organization (WHO) [1]. New patients with drug susceptible TB can be cured in six months using the standard regimens with a combination of first line TB drugs including isoniazid, rifampicin, pyrazinamide, ethambutol and streptomycin. However, *M. tuberculosis* strains have developed resistance to the available TB drugs. Drug-resistant TB is classified as multi drug-resistant tuberculosis (MDR-TB), extensively drug-resistant tuberculosis (XDR-TB) and totally drug-resistant tuberculosis (TDR-TB) [2-4]. Due to drug resistant TB, the standard regimens using existing TB drugs are often ineffective. To combat the drug resistant TB, intensive research efforts are being directed towards identification and development of new anti-TB agents. The potential strategy for the development of new drugs against *M. tuberculosis* is to target essential biosynthetic pathways of this bacteria that are absent in humans. This is very important for design of specific drugs with low toxicity. The shikimic acid pathway is present in bacteria, fungi, plants, and in certain apicomplexan parasites, but absent in humans. Therefore, it is an attractive target for the development of new antimicrobials [5-8]. Seven enzymes involved in this pathway sequentially catalyze a series of chemical reactions for biosynthesis of chorismate [9]. This is a precursor for the biosynthesis of aromatic amino acids, folate, ubiquinone, and vitamins E and K [7, 10]. Type II dehydroquinase of *M. tuberculosis* (*MtbDHQ*), the third enzyme in the shikimic acid pathway, has generated the most interest for rational design of anti-TB agents. It is a dodecameric enzyme assembled from four trimeric units, the minimum active unit of type II dehydroquinase [11]. *MtbDHQ* catalyzes the reversible dehydration of 3-dehydroquinate (DHQ) to form 3-dehydroshikimate (DHS). The atomic details of the catalytic mechanism of this enzyme have been elucidated using computational and biochemical studies. The enzymatic process is initiated by the

direct deprotonation of Tyr24 by Asp88^a from a neighboring subunit to generate the catalytic tyrosinate. The abstraction of the more acidic axial hydrogen of the substrate by tyrosinate is followed by generation an enolate intermediate. For product formation, the spontaneous abstraction of the His101 N δ proton was followed by formation and elimination of a water molecule [12]. A large number of competitive reversible inhibitors of *Mtb*DHQ have been developed that are comprised of analogues of the natural substrate [13] and mimic of the intermediate of the enzyme-catalyzed reaction [14-24]. The several crystal structures of *Mtb*DHQ–inhibitor binary complexes have also been reported [14, 15, 25, 26]. These crystal structures are very important since they reveal the role of key residues responsible for catalysis and binding of *Mtb*DHQ inhibitors. The previous development of 2,3-anhydroquinone derivatives targeting *Mtb*DHQ was confounded by their poor activity against *M. tuberculosis*. The *in vitro* antibacterial activity of developed 2,3-anhydroquinone derivatives was not observed at the minimum inhibitory concentration (MIC), i.e., below 200 μ g/ml [14].

SDH encoded by the *aroE* gene participates in the fourth reaction step of the shikimic acid pathway. It catalyzes the reversible NADPH dependent reduction of DHS to SKM. The overall structure of SDH is comprised two α/β domains linked centrally by two α -helices. A deep groove created between these two domains is the active site for the binding of substrate and cofactor [27-32]. The crystal structures of apo *Mtb*SDH (PDB code 4P4N) and the binary SKM/*Mtb*SDH complex (PDB code 4P4G) have shown that it shares a three-dimensional structure with all members of the SDH family. The kinetic isotope effect and proton inventory studies showed that the kinetic mechanism of *Mtb*SDH is in agreement with a steady-state ordered bi-bi mechanism. First, DHS is bound to the *Mtb*SDH active site and this is followed by NADPH binding. The C4-*proS* hydride (B side) of NADPH is transferred to DHS in the oxy-reduction reaction of *Mtb*SDH [33]. However, in the case of SDH from *E. coli* and *T. thermophilus* HB8, the C4-*proR* hydride (A side) of NADPH is transferred in this reaction [34-35]. Both the hydride transfer of NADPH and the protonation of DHS in the catalytic reaction of *Mtb*SDH proceed in the same step (a concerted mechanism) [33]. An amino acid

residue with an apparent pKa value of 8.9 participates in the catalytic activity of *MtbSDH* [33]. Site-directed mutagenesis shows that the conserved Lys69 plays a catalytic role in *MtbSDH*. The catalytic constant value for the wild-type *MtbSDH* is 68-fold larger than that of the mutant K69A *MtbSDH* [36].

References

1. World Health Organization (WHO). 2016. WHO Global TB Report 2016. Available at http://www.who.int/tb/publications/global_report/en/.
2. World Health Organization (WHO). 2010. Multidrug and Extensively Drug-Resistant TB (M/XDR-TB):2010 Global Report on Surveillance and Response. Available at http://www.who.int/tb/features_archive/m_xdrtb_facts/en/.
3. Velayati A.A., P. Farnia, M.R. Masjedi, T.A. Ibrahim, P. Tabarsi, R.Z. Haroun, H.O. Kuan, J. Ghanavi, P. Farnia, and M. Varahram. 2009. Totally drug-resistant tuberculosis strains: evidence of adaptation at the cellular level. *Eur. Respir. J.* 34:1202-1203.
4. Udawadia Z.F., R.A. Amale, K.K. Ajbani, and C. Rodrigues. 2012. Totally drug-resistant tuberculosis in India. *Clin. Infect. Dis.* 54:579-581.
5. Roberts F., C.W. Roberts, J.J. Johnson, D.E. Kyle, T. Krell, J.R. Coggins, G.H. Coombs, W.K. Milhous, S. Tzipori, D.J. Ferguson, D. Chakrabarti, and R. McLeod. 1998. Evidence for the shikimate pathway in apicomplexan parasites. *Nature.* 393:801-805.
6. Keeling P.J., J.D. Palmer, R.G. Donald, D.S. Roos, R.F. Waller, and G.I. McFadden. 1999. Shikimate pathway in apicomplexan parasites. *Nature.* 397:219-220.
7. Bentley R. 1990. The shikimate pathway--a metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol.* 25:307-384.
8. Ducati R.G., L.A. Basso, and D.S. Santos. 2007. Mycobacterial shikimate pathway enzymes as targets for drug design. *Curr. Drug Targets.* 8:423-435.
9. Sassetti C.M., D.H. Boyd, and E.J. Rubin. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* 48:77-84.

10. Dosselaere F., and J. Vanderleyden. 2001. A metabolic node in action: chorismate-utilizing enzymes in microorganisms. *Crit. Rev. Microbiol.* 27:75-131.
11. Price N.C., D.J. Boam, S.M. Kelly, D. Duncan, T. Krell, D.G. Gourley, J.R. Coggins, R. Virden, and A.R. Hawkins. 1999. The folding and assembly of the dodecameric type II dehydroquinases. *Biochem. J.* 338:195-202.
12. Coderch C., E. Lence, A. Peón, H. Lamb, A.R. Hawkins, F. Gago, and C. González-Bello. 2014. Mechanistic insight into the reaction catalysed by bacterial type II dehydroquinases. *Biochem. J.* 458:547-557.
13. Howard N.I., M.V. Dias, F. Peyrot, L. Chen, M.F. Schmidt, T.L. Blundell, and C. Abell. 2015. Design and structural analysis of aromatic inhibitors of type II dehydroquinase from *Mycobacterium tuberculosis*. *ChemMedChem.* 10:116-133.
14. Tizón L., J.M. Otero, V.F. Prazeres, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, J.A. Ainsa, L. Castedo, and C. González-Bello. 2011. A prodrug approach for improving antituberculosis activity of potent *Mycobacterium tuberculosis* type II dehydroquinase inhibitors. *J. Med. Chem.* 54:6063-6084.
15. Dias M.V., W.C. Snee, K.M. Bromfield, R.J. Payne, S.K. Palaninathan, A. Ciulli, N.I. Howard, C. Abell, J.C. Sacchettini, and T.L. Blundell. 2011. Structural investigation of inhibitor designs targeting 3-dehydroquinase dehydratase from the shikimate pathway of *Mycobacterium tuberculosis*. *Biochem. J.* 436:729-739.
16. Sánchez-Sixto C., V.F. Prazeres, L. Castedo, H. Lamb, A.R. Hawkins, and C. González-Bello. 2005. Structure-based design, synthesis, and biological evaluation of inhibitors of *Mycobacterium tuberculosis* type II dehydroquinase. *J. Med. Chem.* 48:4871-4881.
17. Prazeres V.F., C. Sánchez-Sixto, L. Castedo, H. Lamb, A.R. Hawkins, A. Riboldi-Tunnicliffe, J.R. Coggins, A.J. Lapthorn, and C. González-Bello. 2007. Nanomolar competitive inhibitors of *Mycobacterium tuberculosis* and *Streptomyces coelicolor* type II dehydroquinase. *ChemMedChem.* 2:194-207.

18. Toscano M.D., R.J. Payne, A. Chiba, O. Kerbarh, and C. Abell. 2007. Nanomolar inhibition of type II dehydroquinase based on the enolate reaction mechanism. *ChemMedChem*. 2:101-112.
19. Payne R.J., A. Riboldi-Tunncliffe, O. Kerbarh, A.D. Abell, A.J. Lapthorn, and C. Abell. 2007. Design, synthesis, and structural studies on potent biaryl inhibitors of type II dehydroquinases. *ChemMedChem*. 2:1010-1013.
20. Payne R.J., F. Peyrot, O. Kerbarh, A.D. Abell, and C. Abell. 2007. Rational design, synthesis, and evaluation of nanomolar type II dehydroquinase inhibitors. *ChemMedChem*. 2:1015-1029.
21. Tran A.T., K.M. Cergol, N.P. West, E.J. Randall, W.J. Britton, S.A. Bokhari, M. Ibrahim, A.J. Lapthorn, and R.J. Payne. 2011. Synthesis and evaluation of potent ene-yne inhibitors of type II dehydroquinases as tuberculosis drug leads. *ChemMedChem*. 6:262-265.
22. Blanco B., A. Sedes, A. Peón, H. Lamb, A.R. Hawkins, L. Castedo, and C. González-Bello. 2012. Synthesis of 3-alkyl enol mimics inhibitors of type II dehydroquinase: factors influencing their inhibition potency. *Org. Biomol. Chem*. 10:3662-3676.
23. Frederickson M., A.W. Roszak, J.R. Coggins, A.J. Lapthorn, and C. Abell. 2004. (1R,4S,5R)-3-Fluoro-1,4,5-trihydroxy-2-cyclohexene-1-carboxylic acid: the fluoro analogue of the enolate intermediate in the reaction catalyzed by type II dehydroquinases. *Org. Biomol. Chem*. 2:1592-1596.
24. Paz S., L. Tizón, J.M. Otero, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, A.J. Lapthorn, L. Castedo, and C. González-Bello. 2011. Tetrahydrobenzothiophene derivatives: conformationally restricted inhibitors of type II dehydroquinase. *ChemMedChem*. 6:266-272.
25. Peón A., J.M. Otero, L. Tizón, V.F. Prazeres, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, F. Gago, L. Castedo, and C. González-Bello. 2010. Understanding the key factors that control the inhibition of type II

- dehydroquinase by (2R)-2-benzyl-3-dehydroquinic acids. *ChemMedChem*. 5:1726-1733.
26. Lence E., L. Tizón, J.M. Otero, A. Peón, V.F. Prazeres, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, and C. González-Bello. 2013. Mechanistic basis of the inhibition of type II dehydroquinase by (2S)- and (2R)-2-benzyl-3-dehydroquinic acids. *ACS Chem. Biol.* 8:568-577.
 27. Padyana, A.K.; Burley, S.K. Crystal Structure of Shikimate 5-dehydrogenase (SDH) Bound to NADP: Insights into Function and Evolution. *Structure* **2003**, *11*, 1005–10013
 28. Ye, S.; Von Delft, F.; Brooun, A.; Knuth, M.W.; Swanson, R.V.; McRee, D.E. The Crystal Structure of Shikimate Dehydrogenase (AroE) Reveals a Unique NADPH Binding Mode. *J. Bacteriol.* **2003**, *185*, 4144–4151.
 29. Michel, G.; Roszak, A.W.; Sauve, V.; Maclean, J.; Matte, A.; Coggins, J.R.; Cygler, M.; Laphorn, A.J. Structures of Shikimate Dehydrogenase AroE and Its Paralog YdiB. A Common Structural Framework for Different Activities. *J. Biol. Chem.* **2003**, *278*, 19463–19472.
 30. Singh, S.; Korolev, S.; Koroleva, O.; Zarembinski, T.; Collart, F.; Joachimiak, A.; Christendat, D. Crystal Structure of a Novel Shikimate Dehydrogenase from *Haemophilus influenzae*. *J. Biol. Chem.* **2005**, *280*, 17101–17108.
 31. Peek, J.; Lee, J.; Hu, S.; Senisterra, G.; Christendat, D. Structural and Mechanistic Analysis of a Novel Class of Shikimate Dehydrogenases: Evidence for a Conserved Catalytic Mechanism in the Shikimate Dehydrogenase Family. *Biochemistry* **2011**, *50*, 8616–8627.
 32. Peek, J.; Garcia, C.; Lee, J.; Christendat, D. Insights into the Function of RifI2: Structural and Biochemical Investigation of a New Shikimate Dehydrogenase Family Protein. *Biochim. Biophys. Acta* **2013**, *1834*(2), 516–523.
 33. Fonseca, I.O.; Silva, R.G.; Fernandes, C.L.; de Souza, O.N.; Basso, L.A.; Santos, D.S.; Arch. Biochem. Biophys. Kinetic and Chemical Mechanisms of Shikimate Dehydrogenase from *Mycobacterium tuberculosis*. **2007**, *457*(2), 123-133.

34. Dansette, P.; Azerad, R. The Shikimate Pathway: II. Stereospecificity of Hydrogen Transfer Catalyzed by NADPH-dehydroshikimate Reductase of *E. coli*. *Biochimie* **1974**, 56(5), 751-755.
35. Bagautdinov, B.; Kunishima, N. Crystal Structures of Shikimate Dehydrogenase AroE from *Thermus thermophilus* HB8 and Its Cofactor and Substrate Complexes: Insights into the Enzymatic Mechanism. *J. Mol. Biol.* **2007**, 373(2), 424-438.
36. Rodrigues, V.S. Jr.; Breda, A.; Santos, D.S.; Basso, L.A. The Conserved Lysine69 Residue Plays a Catalytic Role in *Mycobacterium tuberculosis* Shikimate Dehydrogenase. *BMC Res Notes*. **2009**, 2, 227.
37. Tizón L., J.M. Otero, V.F. Prazeres, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, J.A. Ainsa, L. Castedo, and C. González-Bello. 2011. A prodrug approach for improving antituberculosis activity of potent *Mycobacterium tuberculosis* type II dehydroquinase inhibitors. *J. Med. Chem.* 54:6063-6084.
38. Sánchez-Sixto C., V.F. Prazeres, L. Castedo, H. Lamb, A.R. Hawkins, and C. González-Bello. 2005. Structure-based design, synthesis, and biological evaluation of inhibitors of *Mycobacterium tuberculosis* type II dehydroquinase. *J. Med. Chem.* 48:4871-4881.
39. Prazeres V.F., C. Sánchez-Sixto, L. Castedo, H. Lamb, A.R. Hawkins, A. Riboldi-Tunnicliffe, J.R. Coggins, A.J. Lapthorn, and C. González-Bello. 2007. Nanomolar competitive inhibitors of *Mycobacterium tuberculosis* and *Streptomyces coelicolor* type II dehydroquinase. *ChemMedChem*. 2:194-207.
40. Morris G.M., D.S. Goodshell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, and A.J. Olson. 1998. Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* 19:1639-1662.
41. Richter M.F., B.S. Drown, A.P. Riley, A. Garcia, T. Shirai, R.L. Svec, and P.J. Hergenrother. 2017. Predictive compound accumulation rules yield a broad-spectrum antibiotic. *Nature*. 545:299-304.
42. Niederweis M. 2003. Mycobacterial porins – new channel proteins in unique outer membranes. *Mol. Microbiol.* 49:1167-1177.

METHODOLOGY

MD simulations

All MD simulations were performed using AMBER12 program. The X-ray crystal structures of *Mtb*DHQ and *Mtb*SDH were used as initial coordinates for MD simulations. The complex structures of inhibitors with *Mtb*DHQ and *Mtb*SDH were initially modelled by molecular docking calculations. Then, the complex structure was solvated by TIP3 waters in an octahedral box extending up to 10 Å from each solute species. Na⁺ ions were added to neutralize the charge of system. The added water molecules and ions in the solvated systems were relaxed using the Sander program to relieve bad steric interactions. Thereafter, the whole system was minimized without restraint condition. The systems were then gradually warmed up from 0 to 300 K. The position-restrained dynamics simulation was performed to relax the positions of the solute molecules. Finally, MD simulations without the position restraints were performed for 10 ns. MD trajectories were evaluated in terms of the structural flexibility and the interaction energy between inhibitor and residual residues of *Mtb*DHQ and *Mtb*SDH binding pockets.

CoMSIA method

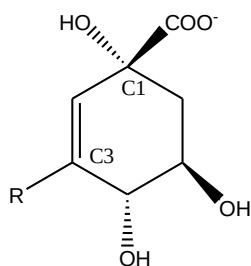
Data set and molecular alignment

Chemical structures and biological activities expressed in term of the K_i values of 2,3-anhydroquinone derivatives (Table 1) were collected from the published literature [37-39]. These K_i values were transformed into $\log (1/ K_i)$. 2,3-anhydroquinone derivatives were divided into sets of 24 training compounds for development of a CoMSIA model and 4 compounds that were used as a test set for model validation. The test set compounds were selected based on their structural diversity and wide range of activities. The binding modes of the data set obtained from molecular docking calculations were used for molecular alignment to set up the CoMSIA model.

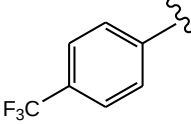
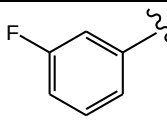
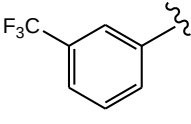
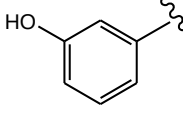
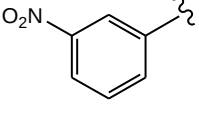
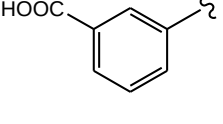
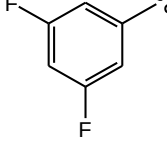
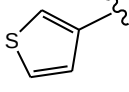
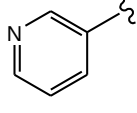
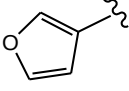
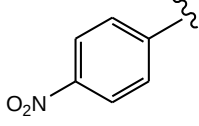
CoMSIA model development

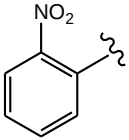
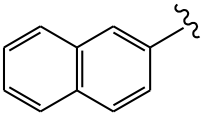
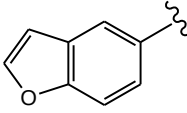
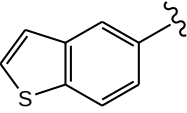
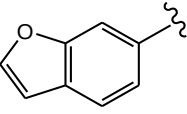
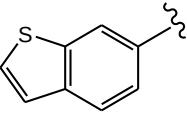
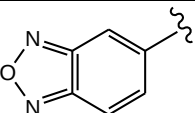
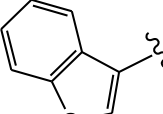
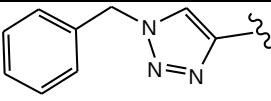
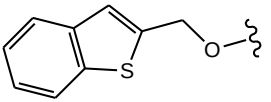
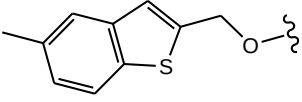
SYBYL-X 2.0 molecular modeling software was used to develop the CoMSIA model. Five CoMSIA descriptor fields including steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields were calculated for each compound. Partial least square (PLS) analysis was employed to derive a linear relationship between CoMSIA descriptor fields and activities. PLS analysis using the leave-one-out (LOO) cross-validation method was performed to determine the optimal number of components. Sequentially, a final analysis with the optimal number of components was performed to develop a CoMSIA model. The non-cross-validated correlation coefficient (r^2) and the leave-one-out cross-validated correlation coefficient (q^2) were used to evaluate the predictive capability of the CoMSIA model. To evaluate the external predictive ability of the selected CoMSIA model, it were employed to predict $\log(1/K_i)$ values of test set compounds that were not used to develop the CoMSIA model.

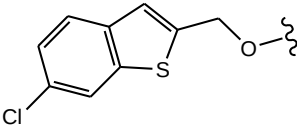
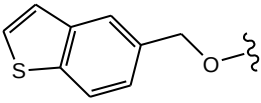
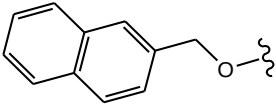
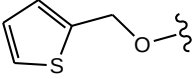
Table 1 Chemical structures and $\log(1/K_i)$ values of *Mtb*DHQ inhibitors (2,3-anhydroquinate derivative)



Code	Compound ^a	R	$\log(1/K_i)$
1	4a		2.82
2	4b		2.74

3	4c		1.77
4	4d		2.82
5	4e		2.67
6	4f		1.02
7	4g		4.27
8 ^b	4h		0.82
9	4i		2.61
10	4j		3.23
11	4k		1.35
12 ^b	4l		3.08
13	1-5a		2.19

14	2-5b		0.96
15	6		2.92
16	10a		2.70
17	10b		2.55
18	11a		3.01
19	11b		3.07
20 ^b	12		2.71
21	13		1.88
22	14		2.43
23	5a		4.55
24	5b		4.37

25	5c		4.46
26 ^b	5d		4.51
27	5e		4.46
28	5f		3.63

^a The compound name indicated in the published literature

^b indicates test set compound

Molecular docking calculations

Molecular docking calculations were performed using the Autodock 4.2 program [40]. Residues of *Mtb*DHQ and *Mtb*SDH were kept rigid, whereas the structure of the inhibitor was flexible during the molecular docking calculations. The number of grid points in the x, y and z dimensions with a value of 40 and a grid point spacing of 0.375 Å were used to define the 3D grid box size. The grid box was centered on the coordinates found in the crystal structure of the inhibitor contained in chain A. The conformations of inhibitor in *Mtb*DHQ and *Mtb*SDH pockets were generated using the Lamarckian Genetic Algorithm (LGA) with the GA runs of 100. The RMSD value between the docked and observed X-ray conformations of inhibitor lower than 1 Å were used to verify the docking calculations. The conformation showing the lowest binding energy was selected as the representative binding mode of each inhibitor obtained from molecular docking calculations.

Designing of core bi-functional inhibitors

The obtained results from this work would be utilized to design the core structures of novel bi-functional inhibitors of DHQ dehydratase and shikimate dehydrogenase. Then, all designed core structures were used for virtual screening.

Virtual screening

Virtual screening approaches using ligand and structure bases were utilized to screen the potential compound from Specs database as bi-functional inhibitors of DHQ dehydratase and shikimate dehydrogenase. Initially, designed core structures of bi-functional inhibitors were used to screen compounds from Specs database. Subsequently, drug-likeness screening was performed to filter compounds that have properties suitable to be drug. Compound which did not follow the criteria of the Lipinski's rule (hydrogen bond donor less than 5, hydrogen bond acceptors less than 10, a molecular mass less than 500 daltons and an octanol-water partition coefficient log P less than 5) was removed. Finally, remaining compounds will be docked to the binding pockets of DHQ dehydratase and shikimate dehydrogenase using molecular docking calculations. Compounds that give the calculated binding energies are comparable in both binding pockets would be selected as bi-functional inhibitors of DHQ dehydratase and shikimate dehydrogenase.

RESULT AND DISCUSSION

1. *Mtb*DHQ inhibitors

1.1 Quantitative investigation for *Mtb*DHQ-inhibitor interaction

The high resolution crystal structures of *Mtb*DHQ with its inhibitors [14-15, 25-26] were solved to provide insight into inhibitor binding. However, the quantitative contribution of each *Mtb*DHQ residue on inhibitor binding could not be determined. Accordingly, per-residue free energy decomposition was performed to investigate this quantitative contribution. Most of the solved crystal structures of *Mtb*DHQ in complex with their inhibitor were found to be monomeric, and the flexible-loop residues 19–24 were mostly missing [14, 25-26]. Therefore, the crystal structure of trimeric *Mtb*DHQ (chains A, K and L) in complex with compound **4**, a promising *Mtb*DHQ inhibitor with a K_i value of 2.3 μ M [15], was selected. This structure shows all residues in the active site, including the flexible-loop residues 19–24. The quantitative contribution of individual residues in chain A and adjacent chain L responsible for the binding of compound **4** was visualized through its decomposition energy (Figure 1). Most of the residues in an adjacent chain L had a weak contribution to the binding of this compound. Only Asp88^a (^a indicates the residue from an adjacent chain L) had an attractive interaction energy of -9.45 kcal/mol, consistent with a hydrogen bond between its side chain and the OH moiety of compound **4** (Figure 2). The lowest interaction energy of -15.00 kcal/mol was observed between His101 and compound **4**, indicating that His101 had the greatest contribution to the binding of this compound. It was responsible for the electrostatic interaction of the carboxylate (COO⁻) moiety of compound **4** with the positive charge of His101 and one hydrogen bond of the OH with an NH group of His101 (Figure 2). These results indicate that Asp88^a and His101 are not only important for the enzymatic process of *Mtb*DHQ [12], but they are also crucial for inhibitor binding. Other prominent attractive interaction energies were observed for Asn75, Ile102, Ser103 and Arg112 (Figure 1) consistent with hydrogen bonds formed between these residues with compound **4** (Figure 2). It is notable that the carboxylate moiety played the most

important role in creating these remarkable attractive interaction energies. Obviously, it interacts with four key residues, Asn75, His101, Ile102 and Ser103 (Figure 1). Accordingly, the carboxylate moiety was crucial for binding of compound **4** in the *Mtb*DHQ pocket. The R substituent had a small contribution to binding of this compound. It generated two prominent attractive interaction energies with the residues of Asn12 and Tyr24 (Figure 1).

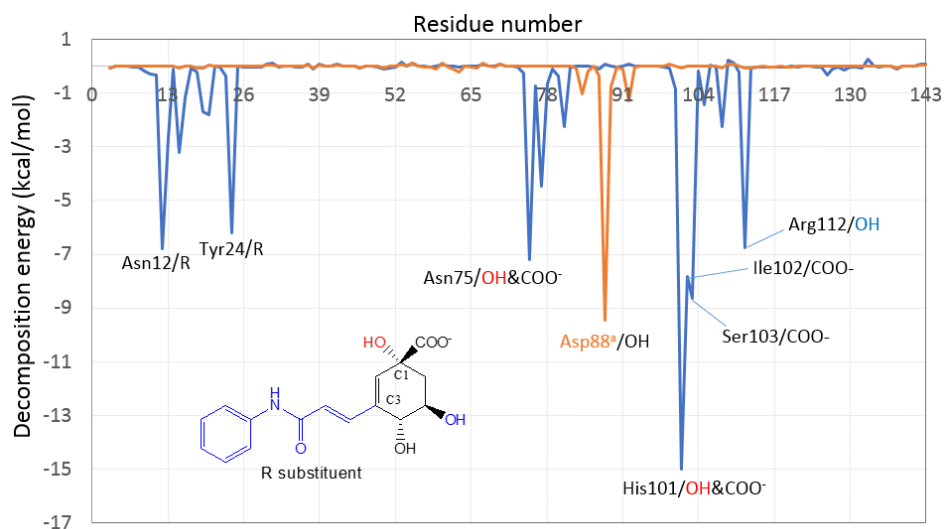


Figure 1 The per-residue free energy decomposition plot for all residues except those of residues 1-3 of the *Mtb*DHQ chain A (blue line) and adjacent chain L (orange line) in complex with compound **4**

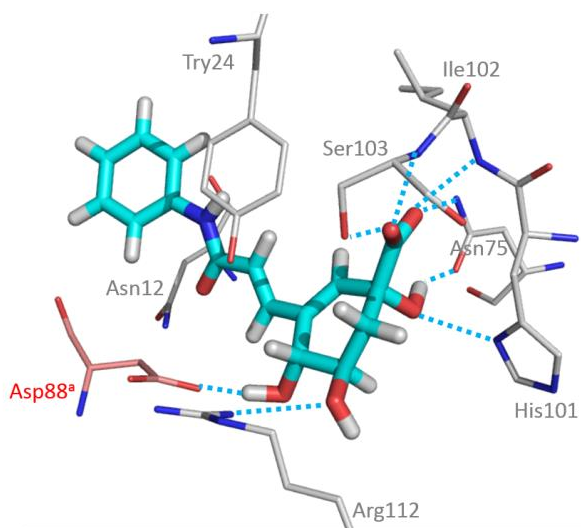


Figure 2 The binding mode of compound **4** (cyan) in the pocket of *Mtb*DHQ (PDB code 3N86)

1.2 The influence of R substituent on the binding affinity

Most *Mtb*DHQ inhibitors were designed by the modification of the R substituent attached to the C3 atom of the 2,3-anhydroquinone core (Table 1) [14, 16-17]. It has been substantiated that the R substituent helps to form a stabilizing interaction between inhibitors and the flexible active-site loop residues 19-24. They are a significant factor for increased potency of the inhibitor [15]. The potency of compound **4** ($K_i=2.3 \mu\text{M}$) increased approximately 87 fold compared with compound **2** ($K_i=200 \mu\text{M}$) [15]. Therefore, this work aimed to quantitatively investigate the influence of the R substituent on the binding interaction of *Mtb*DHQ inhibitors. In this case, compounds **2** and **4** (Figure 3) were selected for per-residue free energy decomposition to determine the interaction energy per residue of *Mtb*DHQ bound with these compounds. The contribution of each *Mtb*DHQ residue for inhibitor binding is quantitatively visualized in Figure 4. As expected, the R substituent of compound **4** had more attractive interaction energies with the surrounding residues (Asn12, Leu13, Arg15, Leu16, Arg19, Glu20 and Tyr24) than those of compound **2**. Remarkably, Asp88^a had a large contribution for binding of compound **4**, whereas it showed a small contribution for binding of compound **2**. This indicates that the R substituent not only produced a greater degree of complementarity to the *Mtb*DHQ pocket, but it is also enhanced the binding affinity of the 2,3-anhydroquinone core with Asp88^a.

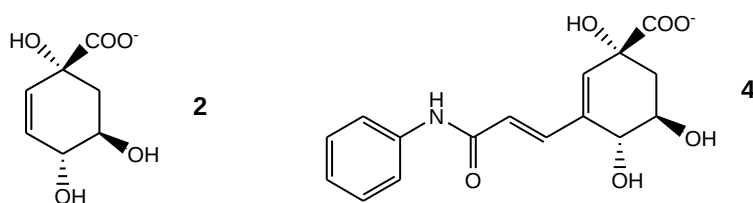


Figure 3 The chemical structures of compounds **2** and **4** [15]

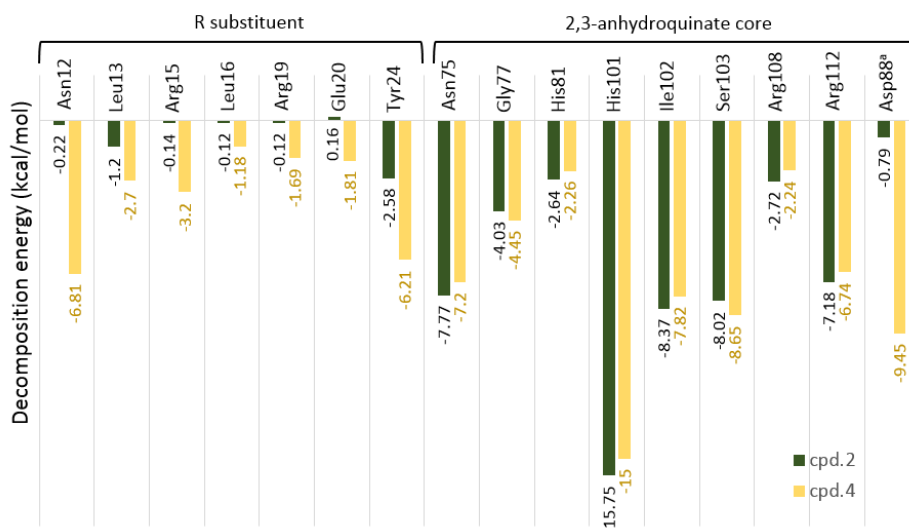


Figure 4 The per-residue free energy decomposition plot for the binding residues of *MtbDHQ* in complex with compounds **2** and **4**

1.3. The structural requirement of R substituent for the better binding affinity

The R substituent attached to the C3 atom of the 2,3-anhydroquinone core was the only substituent utilized to design potent *MtbDHQ* inhibitors. Accordingly, the CoMSIA method combined with MD simulations was employed in this work with the aim of elucidating the structural requirements of the R substituent for rational design of new *MtbDHQ* inhibitors with better binding affinity.

1.3.1 CoMSIA model and contour map

The CoMSIA model is valid with the q^2 and r^2 values of 0.68 and 0.97 respectively, with six optimal components. They indicate the high power of this model to estimate the $\log(1/K_i)$ values of training set data. Additionally, the CoMSIA model showed good capability in predicting $\log(1/K_i)$ values of the test set data with an r^2 value of 0.86. The CoMSIA model was developed from a combination of three molecular descriptors including steric, hydrophobic and hydrogen donor fields. The contributions of these fields to the $\log(1/K_i)$ value of *MtbDHQ* inhibitor were 18.7%, 56.8% and 24.5%, respectively. This indicated a greater contribution of the hydrophobic field on the activity for *MtbDHQ* inhibition. The steric, hydrophobic and hydrogen donor fields important for *MtbDHQ* inhibition could be visualized in CoMSIA contour maps (Figure 5). Green and yellow

contours indicate areas of favorable and unfavorable steric bulk, respectively (Figure 5a). Magenta and white contours represent areas where the hydrophobic and hydrophilic groups were predicted to favor biological activities (Figure 5b). The cyan and orange contours indicate regions that favor the hydrogen donor group and do not favor the hydrogen donor group, respectively (Figure 5c). The following interpretation of CoMSIA contour maps reveals the structural requirements of an R substituent for better binding affinity of *Mtb*DHQ inhibitors.

1.3.2 The structural requirement interpretation based on CoMSIA contours

Two yellow contours located near the C3 atom of 2,3-anhydroquinone core, in which the R substituent is directly attached, indicate unfavorable steric bulk. Obviously, most of compounds with the bulky R substituent directly attached to the C3 atom showed the least activity for *Mtb*DHQ inhibition, such as compound codes **1-22** (Table 1). Another steric requirement is indicated by the green contour at the end of the R substituent of compound code **5a**, far from the C3 atom. This contour indicates a favorable steric bulk in the green region. Based on the presence of steric contours, the R substituents should contain a small linker near the C3 atom and the bulky end cap, such as the R substituent of compound codes **23-28**. This feature of the R substituent might be specified as the capability of changing the position of *Mtb*DHQ residues.

The large purple contour located at the edge of the R substituent of compound code **23** indicates a favorable hydrophobic moiety at this site (Figure 5b). This favorable hydrophobic region is surrounded by the CH₂ side chains of two negatively charged residues of an adjacent subunit (Glu92^a and Asp88^a). The presence of a hydrophobic moiety in this region might be preferable for interactions with the alkyl moieties of Glu92^a and Asp88^a. All hydrogen donor contours are present near the 2,3-anhydroquinone core, the general structure of all compounds in the data set. There is no significant hydrogen donor contours present for the R substituent. Therefore, the hydrogen donor atom may be ignored in the design of an R substituent.

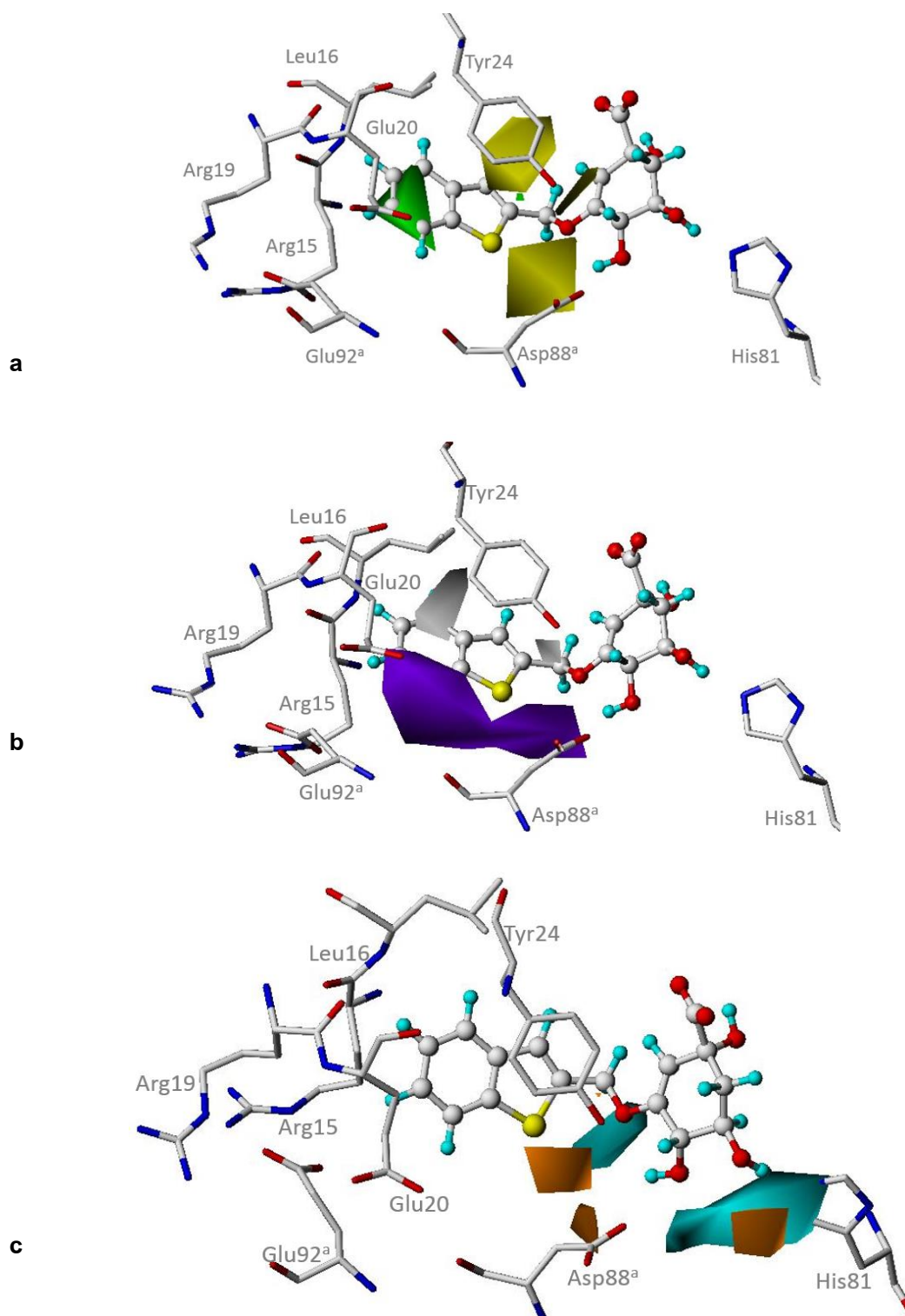


Figure 5 Steric (a), hydrophobic (b) and hydrogen donor (d) contour maps of CoMSIA model in combination with the most active compound code **23** (ball and stick) and *Mtb*DHQ residues (stick)

1.4 The positional change of *Mtb*DHQ for inhibitor binding

To reveal the positional change of *Mtb*DHQ necessary to accommodate a different inhibitor, two complex structures of compound codes **23** and **21** (Table 1) with *Mtb*DHQ were modeled by MD simulations. Then, root-mean-square deviation (RMSD) values between the coordinates of each residue of *Mtb*DHQ complexed with compound codes **23** and **21** were calculated to measure their positional change. The R substituent of compound code **23**, the most active compound, coincided with the steric contours, whereas that of compound code **21** violated the steric requirement of an R substituent. The binding mode of the 2,3-anhydroquinone cores of compound codes **23** and **21** did not show significant alteration, but that of the R substituent was different (Figure 6). Five residues of *Mtb*DHQ including Pro11, Asn12, Leu13, Tyr24 and Asp88^a were located near the C3 atom (Figure 6). Except Tyr24, the positions of these residues were not significantly changed in either of the complexed structures of compound codes **23** and **21**. Their RMSD values were in the range of 0.32-1.02 Å (Figure 7). It is clear that the positions of Pro11, Asn12, Leu13 and Asp88^a were not rearranged to accommodate the different sized R substituent near the C3 atom. The bulky moiety at this site might cause steric hindrance with Pro11, Asn12, Leu13 and Asp88^a. To avoid the steric hindrance of these residues, the bulky R substituent of compound code **21** significantly altered the position of the key catalytic residue (Tyr24) from that found for the binding of compound code **23** (Figure 6) with a RMSD value of 6.54 Å (Figure 7). The position of Tyr24 found for the binding of compound code **23** was similar to that observed for the binding of the 3-dehydroshikimate substrate in the catalytic reaction of DHQase (Figure 8). This implies that binding of the inhibitor causing the large change of the Tyr24 position observed in the catalytic reaction results the loss of binding affinity in the DHQase pocket. Accordingly, the small linker near the C3 atom was required for the R substituent related to the unfavorable yellow contour (Figure 5a). The end of the R substituent, which is buried in the favorable green contour (Figure 5a), was surrounded by Arg15, Leu16, Gly17, Arg18, Arg19, Glu20, Pro21, His63^a, Ala89^a and Glu92^a. This fragment of compound code **21** was absent in this pocket, whereas the R substituent

fragment of compound code **23** was present (Figure 5). A large difference in the positions of these residues was found for binding of compound codes **21** and **23** with RMSD values in range of 1.15-6.71 Å, except for His63^a and Ala89^a (Figure 7). Therefore, the pocket formed by Arg15, Leu16, Gly17, Arg18, Arg19, Glu20, Pro21 and Glu92^a could be rearranged to accommodate a different sized R substituent.

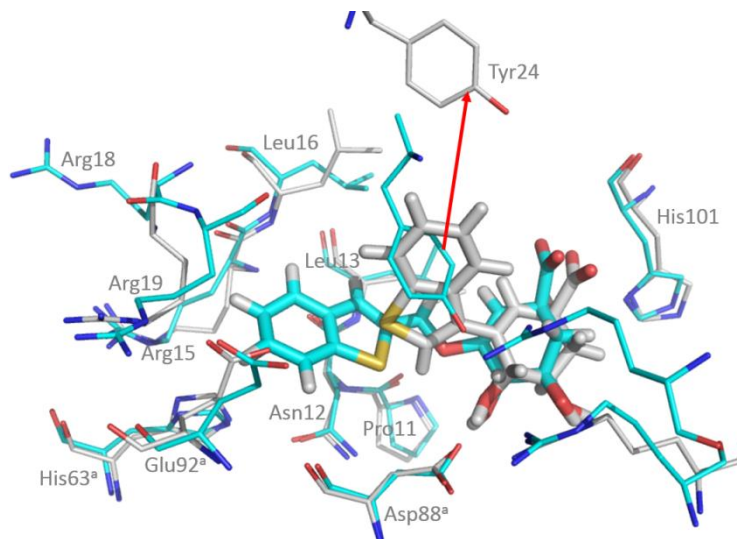


Figure 6. Superimposition of compound codes **21** (carbon atom colored in grey) and **23** (carbon atom leveled in cyan) in the *Mtb*DHQ binding pocket. The carbon atom of the residue complexed with compound codes **21** and **23** are represented in grey and cyan, respectively

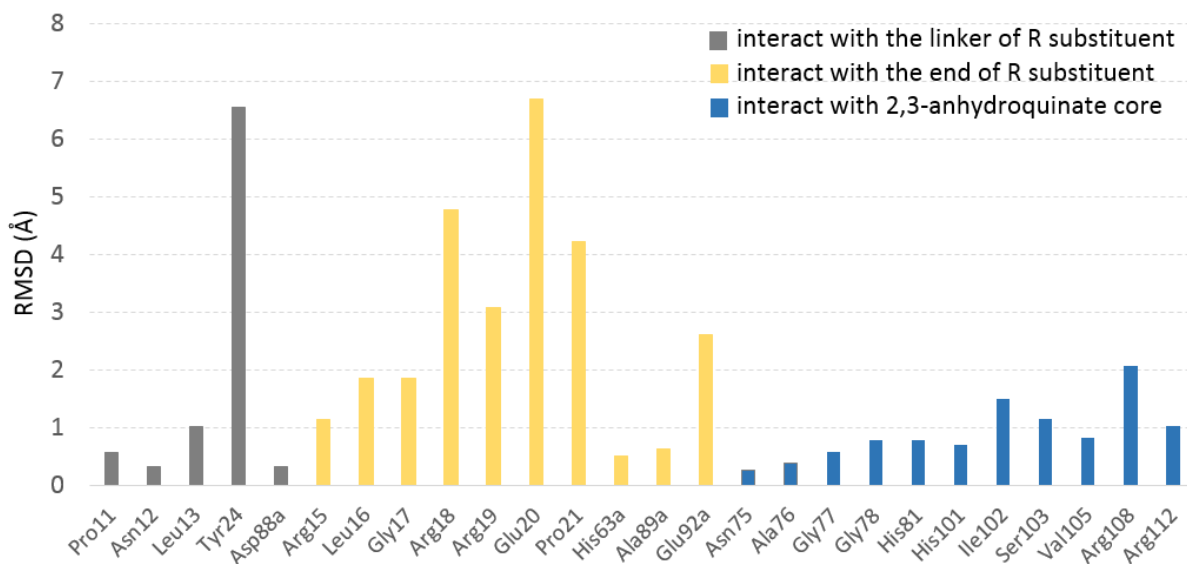


Figure 7. RMSD values of the coordinates of the *Mtb*DHQ binding residues complexed with compound codes **23** and **21**

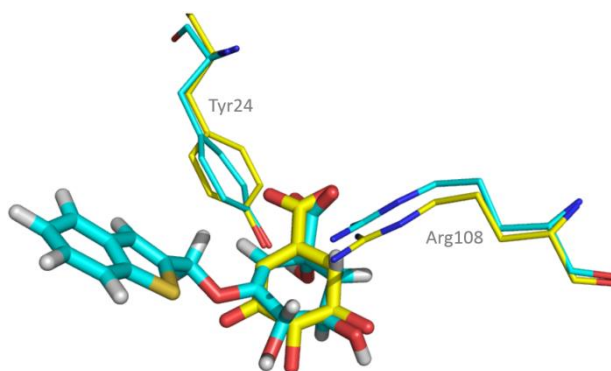


Figure 8. The positions of Tyr24 found for binding of compound code **23** (carbon atoms in cyan) and 3-dehydroshikimate substrate, 3N59 PDB code (carbon atoms in yellow)

1.5 The structural concept to design *Mtb*DHQ inhibitors with potent activity against *M. tuberculosis*

In vitro activity against *M. tuberculosis* was not observed for the compounds listed in Table 1 at the minimum inhibitory concentration (MIC), below 200 µg/ml. It was assumed that the high hydrophilic nature of these compounds caused this lack activity because hydrophilic compounds traverse the mycobacterial cell wall slowly [14]. Recently, systematic analysis of the accumulation of small molecules in gram-negative

bacteria (*E. coli*) provided the guiding principle for compound accumulation in *E. coli*. Compounds that are most likely to accumulate contain a non-sterically encumbered amine, are amphiphilic and rigid, and have low globularity. Additionally, carboxylic acid compounds with a strong negative charge do not accumulate in *E. coli* [41]. Although, *M. tuberculosis* is classified as gram-positive bacteria, it has a closer relationship to gram-negative bacteria [42]. Therefore, the recent guiding principle for compound accumulation in *E. coli* might be used for compounds that antagonize *M. tuberculosis*. Indeed, all four current first line TB drugs, isoniazid, rifampin, ethambutol and pyrazinamide, contain an amine group. Moreover, transformation of the carboxylate groups of compounds listed in Table 1 into the stable esters dramatically increases their efficacy against *M. tuberculosis*. For example, the propyl ester in compound code **23** is more active than the corresponding carboxylated compound code **23** with MIC values of 5 and 200 $\mu\text{g/ml}$, respectively [14]. This implies that compounds with efficacy against *M. tuberculosis* seem to have these features that promote for accumulation in *E. coli*. This finding could provide a strategy to design *Mtb*DHQ inhibitors with better activity against *M. tuberculosis*. On the basis of our strategy, the following guidelines for designing *Mtb*DHQ inhibitors with the potential activity against *M. tuberculosis* were developed.

1.6 The removal of negatively charged carboxylate group of *Mtb*DHQ inhibitors

Based on the guiding principles for compound accumulation in *E. coli*, the negatively charged carboxylate group of some compounds impedes their accumulation inside bacterial cells. However, our finding reveals that the carboxylate moiety is crucial for the binding of *Mtb*DHQ inhibitors. It can form four hydrogen bonds with Asn75, Ile102 and Ser103 and is electrostatically attractive to His101. Therefore, the replacement of the negatively charged carboxylate moiety with a neutral hydrogen bond acceptor should balance the activity of compounds against both *M. tuberculosis* and *Mtb*DHQ.

1.7 Addition of a non-sterically encumbered amine into *Mtb*DHQ inhibitors

A non-sterically encumbered primary amine aids in the accumulation of small molecules in *E. coli* [41]. Therefore, we aimed to introduce $-\text{NH}_3^+$ into the *Mtb*DHQ inhibitor to increase efficacy against *M. tuberculosis*. To avoid the electrostatic repulsion

that occurred from the introduction of -NH_3^+ , positively charged residues located within $8\text{\AA}$ of the inhibitor were considered (Figure 9). Three positive charges of His101, Arg108 and Arg112 were present near the anhydroquinone core of compound code **23**. Therefore, a NH_3^+ substituent could not be introduced into this moiety. There are two negative charges of Glu20 and Glu92^a located near the end of the R substituent. Accordingly, the NH_3^+ substituent can be added at the end of the R substituent. It would induce the electrostatic attraction Glu20 and Glu92^a leading to increased binding affinity of the *Mtb*DHQ inhibitor. Therefore, the presence of NH_3^+ might enhance efficacy against both *M. tuberculosis* and *Mtb*DHQ.

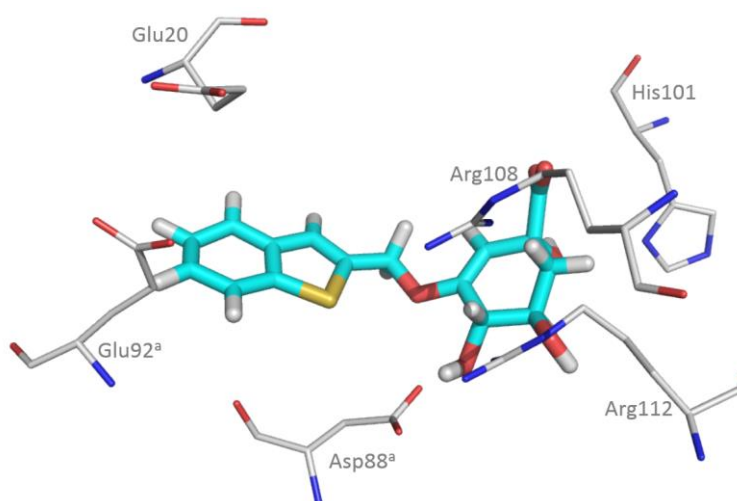


Figure 9 Positively charged residues located within $8\text{\AA}$ of compound code **23**

2. *Mtb*SDH inhibitors

2.1 Stability of *Mtb*SDH on MD simulations

The RMSD of the distance between atoms of each solute species in five MD simulation systems with respect to the initial structure of each system was calculated to reveal the system stability. *Mtb*SDH in three wild-type systems (apo *Mtb*SDH, binary DHS/*Mtb*SDH and ternary DHS/NADPH/*Mtb*SDH) was equilibrated for ~ 20 ns, ~ 10 ns and ~ 5 ns, respectively, and kept stable over the rest of the simulation time for three repeated MD simulations. These indicated that wild-type systems of *Mtb*SDH reached their equilibrium state easily when they bound the DHS substrate and NADPH cofactor.

Two mutant systems of *Mtb*SDH (DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH) reached their equilibrium state after ~ 10 ns and ~ 40 ns respectively, in three repeated MD simulations. The DHS/K69A *Mtb*SDH system reached an equilibrium state faster than the DHS/D105N *Mtb*SDH system. This indicated that the D105N mutation had a greater effect on the structural rearrangement of *Mtb*SDH than the K69A mutation. The coordinate structures of each system obtained from three repeated MD simulations were slightly different. Three backbone structures of *Mtb*SDH obtained from three repeated MD simulations in apo *Mtb*SDH, DHS/*Mtb*SDH, DHS/NADPH/*Mtb*SDH, DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH showed small average RMSD values of 1.36 Å, 1.46 Å, 1.00 Å, 1.04 Å and 1.19 Å, respectively.

2.2 Binary DHS/*Mtb*SDH complex structure

The binding conformation of DHS in *Mtb*SDH obtained from MD simulation was very similar to that of the SKM found in the crystal structure of *Mtb*SDH (Figure 10). The carboxylate group of DHS was anchored by four hydrogen bonds formed by highly conserved Thr65, Ser18, Ser20 and Tyr215. Two hydroxyl groups (C3 and C4) formed hydrogen bonds with the side chains of Gln243, Gln247, Asn90 and Asp105. The carbonyl oxygen at the C5 atom interacted with the side chains of Lys69 to create a hydrogen bond. Hydrogen bonds created between DHS and Thr65, Ser18, Ser20, Gln243, Gln247, Asn90 and Lys69 had a prominent contribution to the binding of DHS substrate in the DHS/*Mtb*SDH complex as evaluated by the decomposition energy (Figure 11). Lys69 showed the greatest contribution to the DHS binding with the lowest energy of -7.76 kcal/mol. The position of Lys69 observed in our investigation relates to the proposed mechanism of *Mtb*SDH. It donated a proton to the carbonyl oxygen of DHS and removed a proton from SKM during the catalytic reaction.

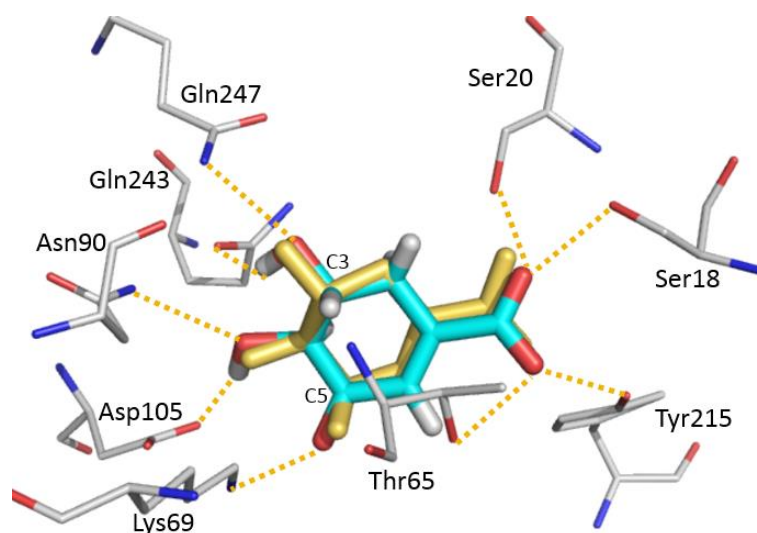
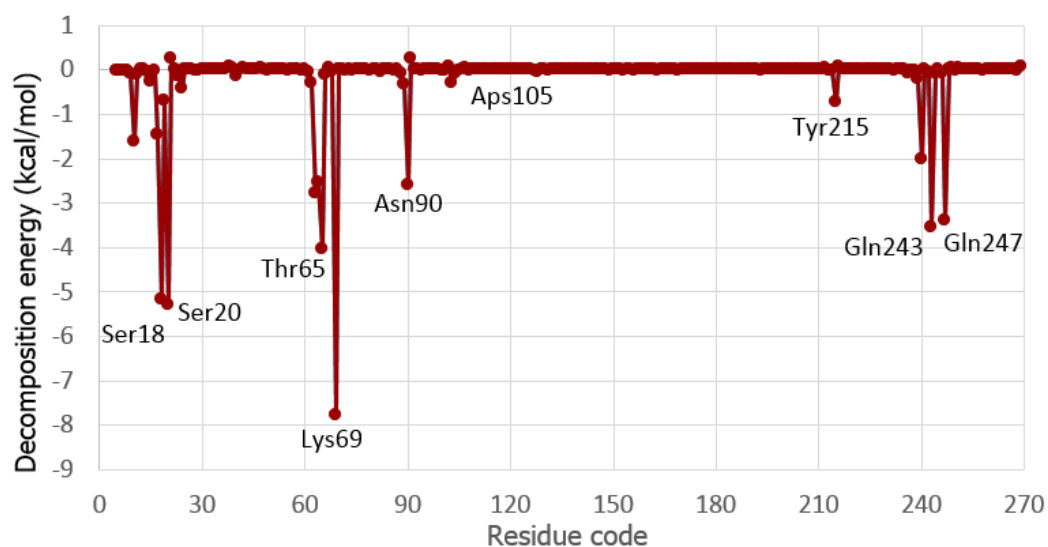


Figure 10 Superposition of a binary DHS/*Mtb*SDH complex structure modeled by MD simulation with the SKM found in the crystal structure of *Mtb*SDH (PDB code 4P4G). Carbon atoms of DHS are cyan colored and the SKM are orange. Dotted lines indicate



hydrogen bond interactions.

Figure 11 The decomposition energy visualizing the quantitative contribution of each *Mtb*SDH residue on the binding of DHS substrate in a binary DHS/*Mtb*SDH complex structure.

2.3 K69A and D105N mutations

Lys69 and Asp105 numbered by *M. tuberculosis* code were conserved in *Tt*SDH, *E. coli* SDH, *H. influenzae* SDH, *A. thaliana* SDH-like and *M. jannaschii* SDH. The crystal structure of *Tt*SDH (PDB code 2EV9) showed that these invariant residues were coupled by a salt bridge. In the catalysis of *Tt*SDH and *A. thaliana* SDH-like, the invariant lysine and aspartate residues were proposed to be a catalytic dyad. Mutation of these invariant residues in *A. thaliana* SDH-like resulted in the loss of enzyme activity. The catalysis role of Lys69 in *Mtb*SDH was investigated using the K69A mutation that dramatically decreases the catalytic constant value for *Mtb*SDH. However, this mutation modestly increased the Michaelis-Menten constant for DHS (K69A = 76 μ M, wild-type = 29 μ M) [36]. In the present work, the complex structures of DHS in K69A and D105N *Mtb*SDH were modeled by MD simulation to visualize the influence of these mutations on DHS binding. Our results showed that the MM-GBSA free energies of DHS in the wild-types, K69A and D105N *Mtb*SDH, were comparable with values of -26.35 ± 3.61 , -24.46 ± 4.67 and -21.34 ± 4.79 kcal/mol, respectively. This indicated that K69A and D105N mutations of *Mtb*SDH did not greatly influence the binding of the DHS substrate. Residues that formed important hydrogen bonds found in the wild-type DHS/*Mtb*SDH complex, except of Lys69, still showed a prominent contribution to DHS binding in K69A and D105N *Mtb*SDH. The decomposition energies found in DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH complexes were comparable to those found in the wild-type DHS/*Mtb*SDH complex (Figure 12). This implies that the remaining hydrogen bonds helped to retain the binding of DHS in K69A and D105N *Mtb*SDH. Although, K69A and D105N mutations did not greatly influence DHS binding, these mutations significantly changed the position of Lys69 for the catalytic reaction of *Mtb*SDH. In the wild-type *Mtb*SDH, Lys69 was located in a suitable position for proton transfer during the catalytic reaction. The average distance between the carbonyl oxygen at the C5 atom of DHS and the nitrogen atom in the ϵ -ammonium group ($-\text{NH}_3^+$) of Lys69 was 2.82 ± 0.13 Å (Figure 13). This distance is close to that found in the crystal structure of the SKM/*Mtb*SDH complex (pdb code 4P4G) between the $-\text{NH}_3^+$ of Lys69 and the hydroxyl

oxygen of the C5 atom of SKM (2.94 Å). The K69A mutation in *Mtb*SDH resulted a missing ϵ -ammonium group that obviously decreased the catalytic constant value for *Mtb*SDH [20]. The D105N mutation in *Mtb*SDH certainly broke the salt bridge created between Lys69 and Asp105. Since the salt bridge was broken, the ϵ -ammonium group of Lys69 was far from the carbonyl oxygen at the C5 atom of DHS (Figure 13). The average distance between the carbonyl oxygen of the C5 atom of DHS and the ϵ -ammonium group of Lys69 was 4.45 ± 1.09 Å. This distance is longer than that found in wild-type *Mtb*SDH, so proton transfer might not have been supported during the catalytic reaction of *Mtb*SDH. This indicates that Asp105 plays a crucial rule to fix the ϵ -ammonium group of Lys69 in a suitable position for the catalytic reaction of *Mtb*SDH via a salt bridge. Accordingly, mutation of Asp105 in *Mtb*SDH might result the loss of enzyme activity.

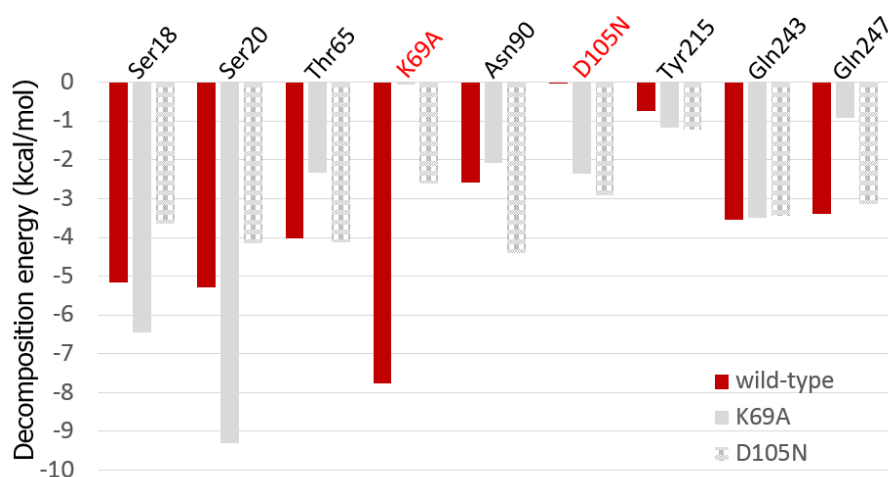


Figure 12 Comparison of the decomposition energies found in the complex structures of DHS in the wild-type, K69A and D105N *Mtb*SDH. Residues that formed important hydrogen bonds found in the wild-type DHS/*Mtb*SDH complex were selected for comparison.

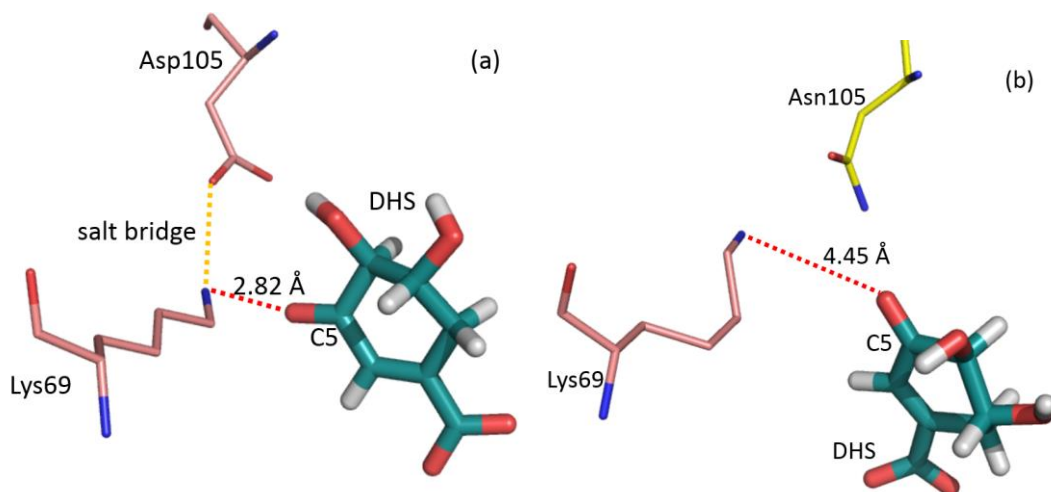


Figure 13 The average distances between the carbonyl oxygen at the C5 atom of DHS and the nitrogen atom of -NH_3^+ of Lys69 found in the wild-type (a) and D105N (b) *MtbSDH*. These distances were averaged over the equilibrium state of each simulation system. D105N mutation broke the salt bridge created between Lys69 and Asp105 (a) and resulted in the long distance between Lys69 and DHS (b).

2.4 Ternary DHS/NADPH/*MtbSDH* complex structure

DHS substrate and NADPH cofactor were bound in a deep groove between the N-terminal and C-terminal domains. The binding of NADPH did not significantly change the binding position of the DHS substrate in *MtbSDH* (Figure 14). The DHS substrate was located close to the nicotinamide ring of NADPH with a *proS* conformation. The average distance between the *proS* hydrogen of the C4 atom and the C5 atom of DHS was 3.95 ± 0.37 Å (Figure 15). This coordinate is consistent with a C4 *proS* hydride transfer from NADPH to the C5 atom of DHS in the oxy-reduction reaction catalyzed by *MtbSDH*. The NH_3^+ of Lys69 and the C5 carbonyl oxygen of DHS were separated by a distance of 2.89 ± 0.13 Å (Figure 15), which is consistent proton transfer between them. The negatively charged groups of Asp105 and NADPH sandwiched the NH_3^+ of Lys69 (Figure 16), which may have facilitated proton transfer by these negative charges. The coordinates of NADPH, DHS and Lys69 found in our work (Figure 14) are in the

agreement with the concerted mechanism of *MtbSD*, both hydride and proton transfer taking place in the same step.

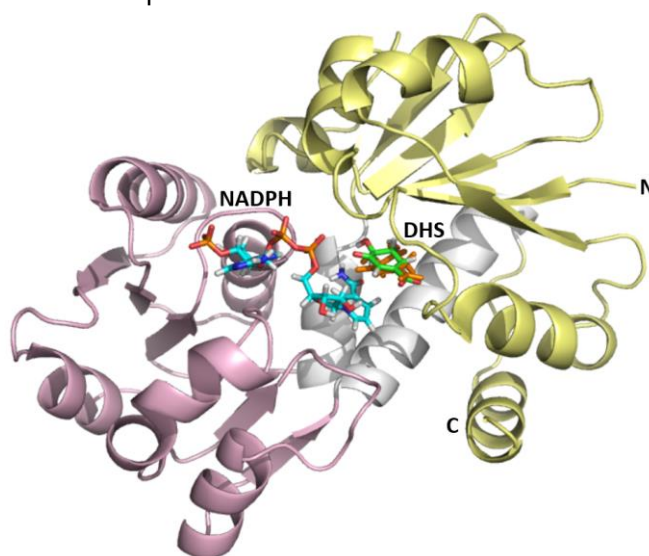


Figure 14 Superposition of the ternary DHS/NADPH/*MtbSDH* complex structure obtained from MD simulations onto DHS (orange) found in the binary DHS/*MtbSDH* complex. N-terminal (residues 5-103 and 254-269) and C-terminal (residues 114-235) domains are shown in yellow and purple, respectively. Two α -helix linkers (residues 104-113 and 236-253) are shown in gray.

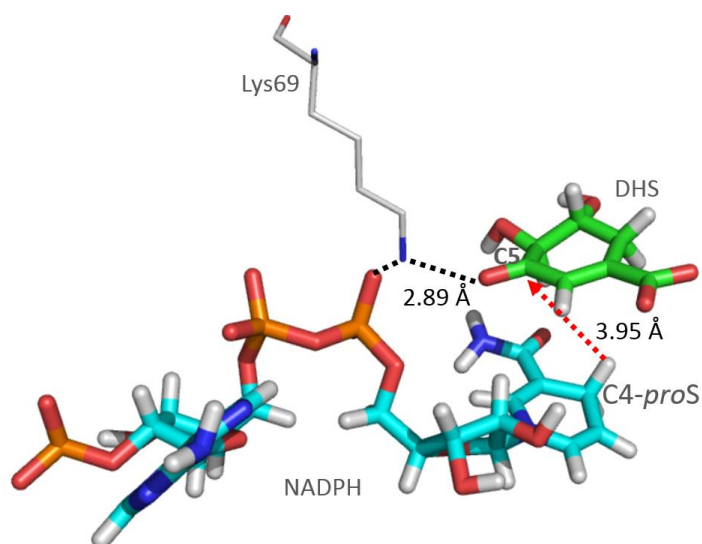


Figure 15 The reaction coordinates of NADPH, DHS and Lys69 found in the ternary DHS/NADPH/*MtbSDH* complex structure.

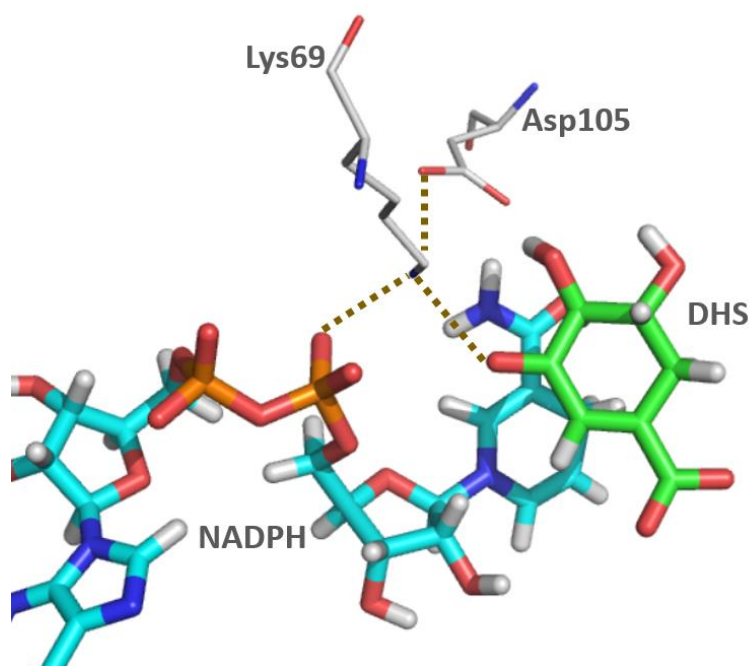


Figure 16 Interactions of the NH_3^+ group of Lys69 with NADPH, DHS and Asp105 in the ternary DHS/NADPH/*Mtb*SDH complex structure.

2.5 NADPH binding in *Mtb*SDH

Two charged residues of Arg145 and Lys153 create an “electrostatic clamp” to bind the adenine phosphate of NADPH in DHS/NADPH/*Mtb*SDH (Figure 17). The electrostatic clamp formed by the charged residues plays a crucial role in the adenine phosphate binding of NADP^+ in the SDH of *E. coli* and *T. thermophilus* HB8. The pyrophosphate of NADPH interacts the diphosphate-binding loop (Gly124-Ala129), which is conserved in the SDH family [35]. It forms hydrogen bonds with the NH backbone and the hydroxyl sidechain of Thr128 and the NH backbone of Gly127. The NH backbones of Ser125 and Gly126 create two hydrogen bonds to the 3'-hydroxyl group of the adenosine ribose of NADPH (Figure 17). The interactions formed by the diphosphate-binding loop are very similar to those found in the NADP^+ binding in SDH of *E. coli* and *T. thermophilus* HB8 (pdb code 1NYT and 2EV9, respectively).

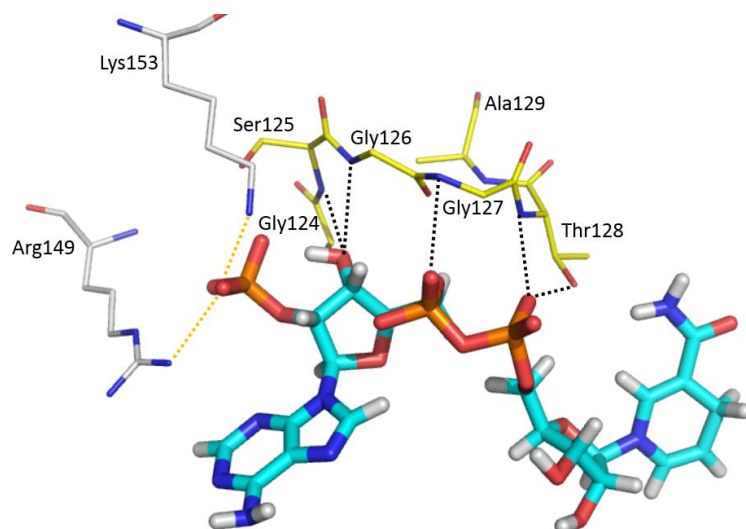


Figure 17 Interactions of NADPH with electrostatic clamp residues (carbon atoms are presented in grey) and the diphosphate-binding loop (carbon atoms are yellow).

2.6 RMSF of *Mtb*SDH

RMSF values visualizing the flexibility of apo *Mtb*SDH residues were in the range of 1.20–6.00 Å except for that of Asp105. This residue had the lowest RMSF of 0.91 Å. Apo *Mtb*SDH structures obtained from the crystal structure (PDB code 4P4N) and our MD simulation showed that the negative side chain of Asp105 was coupled by the positive side chain of Lys69 to form a salt bridge. Therefore, this salt bridge was responsible for the low flexibility of Asp105. RMSF values of residues for DHS binding were in range of 0.91–3.30 Å. The binding pocket of NADPH showed more flexible than that of DHS with RMSF values in range of 1.20–5.15 Å (Figure 18). Particularly, the electrostatic clamp residues, Arg149 and Lys153, had high flexibility with RMSF values of 5.15 and 4.12 Å, respectively. The high flexibility of these residues might serve to capture the adenine phosphate of NADPH.

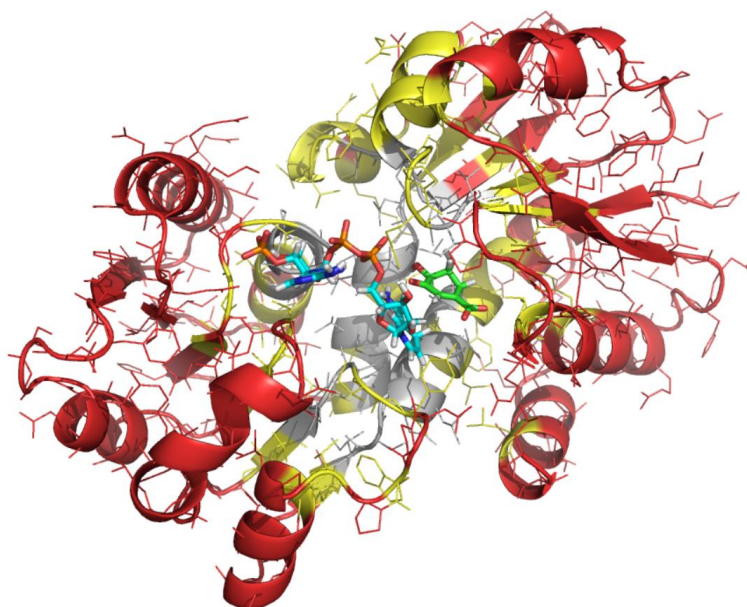


Figure 18 RMSF visualization of *apo MtbSDH* residues on DHS/NADPH/*MtbSDH* obtained from MD simulation. RMSF values were calculated for all atoms of each residue in *apo MtbSDH* over the last 20 ns. RMSF values were classified in the range of 1.2-1.99 Å (grey), 2.01-3.99 Å (yellow) and 4.01-6.00 Å (red).

2.7 Positional change of *MtbSDH*

The conformations of SDH from *T. thermophilus* HB8 observed upon cofactor binding in the binary and ternary complex structures were not significantly changed [35]. Similarly, small conformational changes induced by NADP binding were observed in *MtbSDH* [43]. In agreement with our results, the superposition of an alpha carbon backbone of *apo MtbSDH* onto the ternary DHS/NADPH/*MtbSDH* complex showed a small RMSD value of 1.30 Å. The RMSD value of each *MtbSDH* residue was in the range of 0.22–3.69 Å, except for that of His17. It had the highest RMSD value of 5.81 Å (Figure 19). The position of each residue in the substrate binding pocket was not significantly changed upon DHS binding as evaluated by RMSD values in range of 0.48-1.93 Å (Figure 19). Two residues acting as catalytic residues (Lys69 and Asp105) had small RMSD values of 1.22 and 1.27 Å, respectively. These results suggest that the size of the substrate binding pocket in *MtbSDH* is highly specific for its substrate and the catalytic residue is fixed in a suitable position for the catalytic reaction. In the cofactor

binding pocket, the RMSD value of each residue was in the range of 0.63-2.75 Å. Met66, Gly127, Arg149, Val196 and Met239 responsible for binding of pyrophosphate and adenine phosphate of NADPH had significant conformational changes with RMSD values in the range of 2.25-2.75 Å (Figure 19). Alternatively, the remaining residues had no significant conformational changes as indicated by RMSD values in the range of 0.63-1.91 Å.

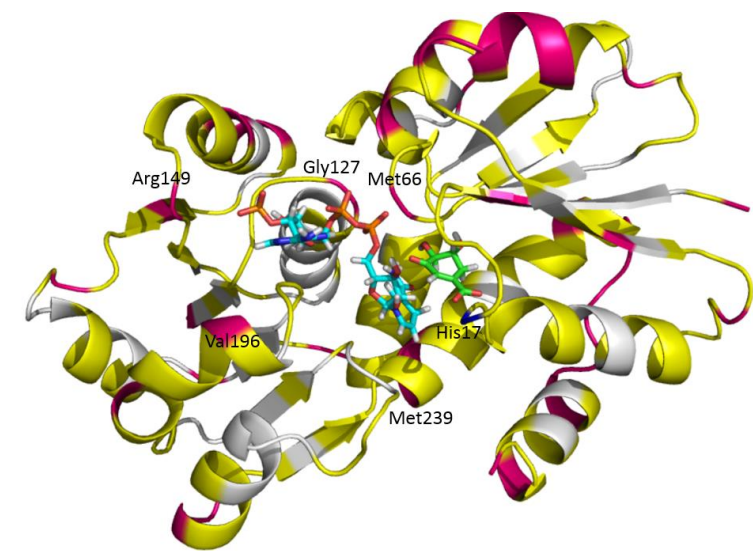


Figure 19 RMSD visualization of each *MtbSDH* residue on DHS/NADPH/*MtbSDH* obtained from MD simulation. RMSD values were calculated between the heavy atom position of each residue in apo *MtbSDH* superpositioned onto DHS/NADPH/*MtbSDH*. RMSD values of residue were classified in the range of 0.22-1.00 Å (grey), 1.01-1.99 Å (yellow) and 2.01-3.69 Å (pink).

2.8 Rational design guideline for the hybrid *MtbSDH* inhibitor

The pocket for DHS binding, as well as the adjacent pocket for the nicotinamide ring and ribose, seemed to be specific as evidenced by their small RMSF and RMSD values (Figure 18 and 19). An inhibitor that could bind in both DHS and NADPH pockets and did not induce the conformational change of *MtbSDH* should be a promising *MtbSDH* inhibitor. Here, it is called a hybrid inhibitor that mimics the binding of DHS and the nicotinamide ring and ribose of NADPH. The head and tail of the hybrid inhibitor were planned so that they bind in the DHS pocket and the ribose and nicotinamide pocket

(NADPH pocket), respectively (Figure 20). The carboxylate and oxygen carbonyl groups should be presented at the head of the hybrid inhibitor to mimic DHS binding. The carboxylate group is important for DHS binding in *MtbSDH*. It formed four hydrogen bonds with Thr65, Ser18, Ser20, and Tyr215. However, the negatively charged carboxylate group of some compounds impeded their accumulation inside bacterial cells. Therefore, replacement of the negatively charged carboxylate group with a neutral hydrogen bond acceptor should balance the activity against both *M. tuberculosis* and *MtbSDH*. The tail of the hybrid inhibitor should be attached to head via a linker at a position located between the carboxylate and oxygen carbonyl groups. This position facilitates the binding of the hybrid inhibitor in the NADPH binding pocket. It is possible to present a hydrophobic moiety at the tail of the hybrid inhibitor due to the surrounding of hydrophobic sidechains of Met66, Ala213, Ile214, Tyr215 and Met239 (Figure 20).

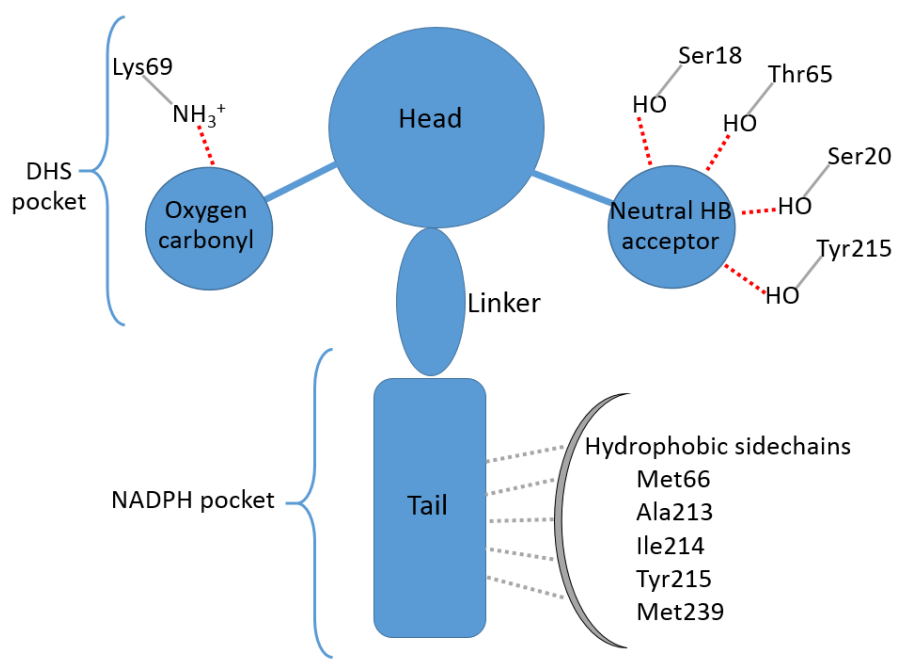


Figure 20. The skeleton of a hybrid inhibitor designed based on the binding of DHS substrate, nicotinamide and ribose of NADPH found in ternary DHS/NADPH/*MtbSDH* complex structure. Red and grey dotted lines respectively indicate the possible hydrogen bonds (HB) and hydrophobic interactions.

3. Core structural design of bi-funtional inhibitor against *Mtb*SDH and *Mtb*DHQ

The carboxylate group is important for inhibitor binding in both of *Mtb*DHQ and *Mtb*SDH. However, the negatively charged carboxylate group impeded the accumulation inside bacterial cells. Therefore, it was replaced with a neutral hydrogen bond acceptor to balance the activity against *M. tuberculosis* and the target enzymes including *Mtb*DHQ and *Mtb*SDH. The core structures of bi-funtional inhibitor designed based on the results integrated from this work are presented in Figure 21. The carboxylate group was replaced by the carbonyl, sulfonyl and nitro groups.

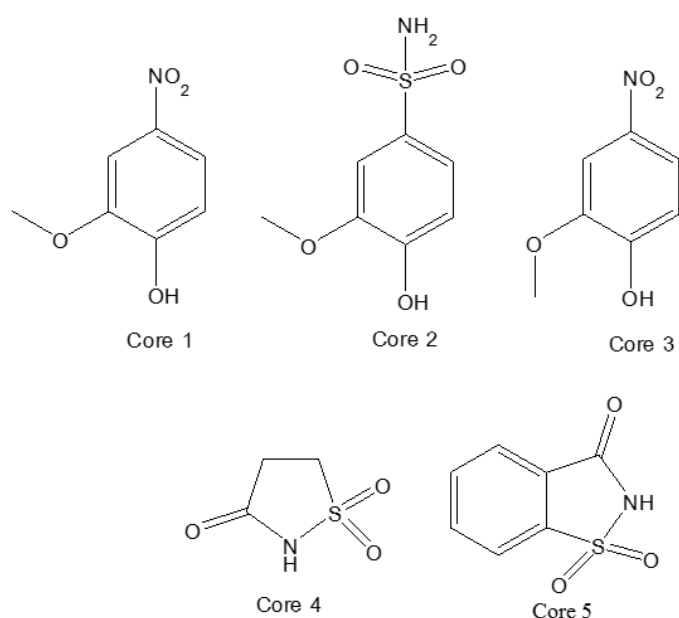


Figure 21. Core structures of bi-funtional inhibitor designed based on the structural requirement of *Mtb*DHQ and *Mtb*SDH inhibitors

4. Virtual screening of bi-funtional inhibitor

Virtual screening flowchart for screening the potential compound from Specs database as bi-funtional inhibitors of *Mtb*DHQ and *Mtb*SDH inhibitors is shown in Figure 22. Initially, five designed core structures of bi-funtional inhibitors were used for fragment search to collect compounds from Specs database. Subsequently, drug-likeness screening and Lipinski's rule were performed to filter compounds that have properties suitable to be drug. 252 compounds were yielded from this step. Then, 42

compounds that have the feature are similar to that required for *Mtb*DHQ or *Mtb*SDH inhibitors were selected. These compounds were docked to the binding pockets of *Mtb*DHQ and *Mtb*SDH. Finally, 22 compounds show good calculated binding energies in both binding pockets with the residual lower than 2 kcal/mol (Table 2). However, only five compounds (Figure 23) display the calculated binding energies comparable to that of known *Mtb*DHQ inhibitor (compound 23). Therefore, these compounds were selected as hit bi-functional inhibitors.

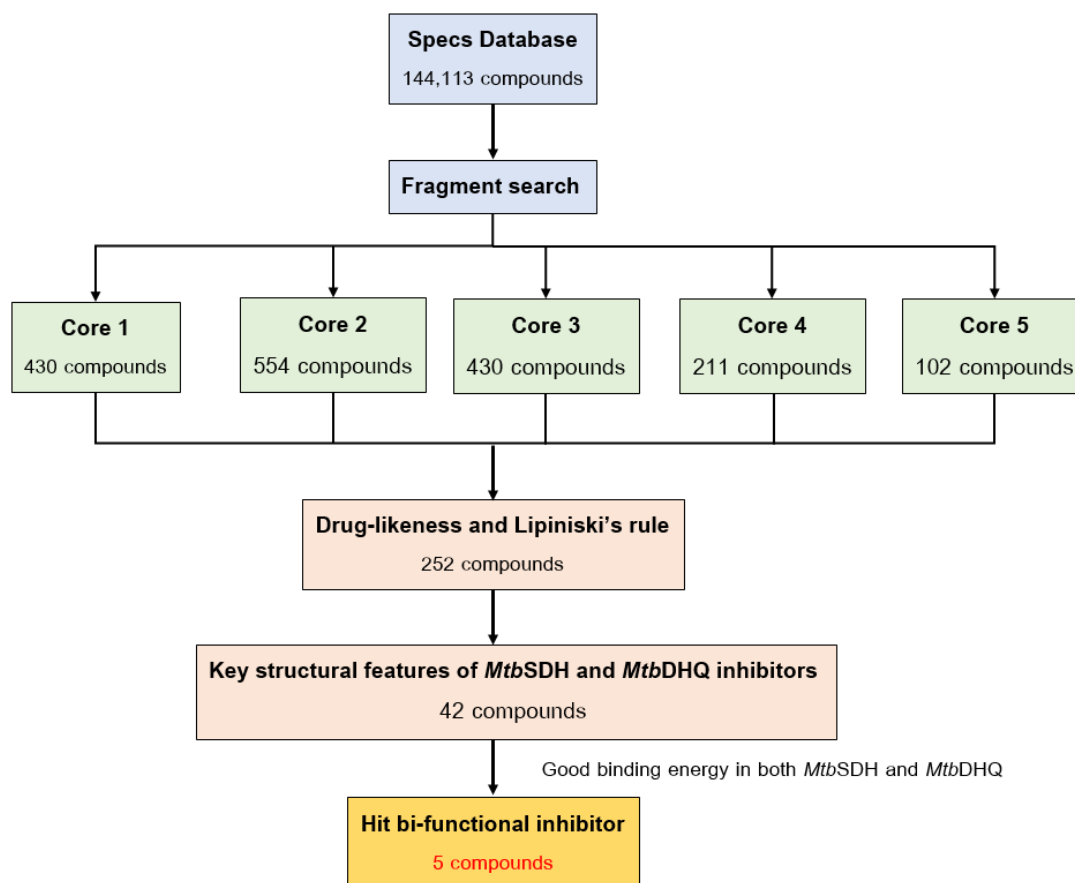


Figure 22. Virtual screening workflow of bi-functional inhibitors

Table 2. The calculated binding energy of 42 compounds in the binding pockets of *Mtb*DHQ and *Mtb*SDH. Five hit compounds are highlighted by bold letters.

Specs ID	Calculated binding energy (kcal/mol)		Residue
	<i>Mtb</i> DHQ	<i>Mtb</i> SDH	
AS-871/43476603	-9.84	-9.85	0.01
AG-205/07697006	-8.18	-8.07	0.11
AK-968/37044065	-7.70	-7.47	0.23
AS-871/43476540	-8.36	-7.91	0.45
AS-871/43476476	-8.24	-7.67	0.57
AG-205/34705002	-8.99	-8.34	0.65
AN-989/40821717	-6.06	-5.33	0.73
AH-487/40936930	-5.99	-6.76	0.77
AN-655/15260058	-7.57	-8.35	0.78
AE-641/14714036	-7.74	-6.94	0.8
AH-487/15274228	-8.73	-7.91	0.82
AG-690/11449018	-8.67	-7.70	0.97
AG-690/12764118	-9.39	-8.42	0.97
AN-329/11481754	-9.02	-7.91	1.11
AO-081/42097078	-10.75	-9.57	1.18
AK-968/40709741	-5.43	-6.72	1.29
AQ-405/42300507	-11.14	-9.63	1.51
AK-778/11118018	-8.40	-6.73	1.67
AS-871/43475827	-11.07	-9.37	1.70
AK-968/11565041	-6.09	-7.95	1.86
AS-871/43476492	-10.75	-8.86	1.89
AS-871/43476475	-10.86	-8.91	1.95
AK-968/40340372	-9.94	-7.93	2.01
AK-968/41026130	-6.44	-8.51	2.07
AN-989/15295003	-10.84	-8.77	2.07
AP-853/43367956	-10.21	-8.07	2.14
AP-853/42400248	-10.67	-8.51	2.16
AP-853/42400279	-12.09	-9.92	2.17

AS-871/43477986	-11.81	-9.6	2.21
AS-871/43476340	-10.41	-8.15	2.26
AS-871/43476333	-11.36	-9.01	2.35
AS-871/43478006	-11.58	-9.23	2.35
AE-508/36402010	-4.78	-7.25	2.47
AA-768/31453027	-10.93	-8.43	2.5
AP-263/40819873	-10.07	-7.44	2.63
AP-263/43502899	-9.02	-6.35	2.67
AS-871/43476588	-12.12	-9.11	3.01
AS-871/43476589	-11.87	-8.68	3.19
AG-205/10698018	-9.29	-5.82	3.47
AS-871/43477996	-12.56	-8.88	3.68
AF-399/41003900	-10.89	-7.11	3.78
AG-690/34682049	-7.59	7.49	15.08
Compound 23	-9.95		

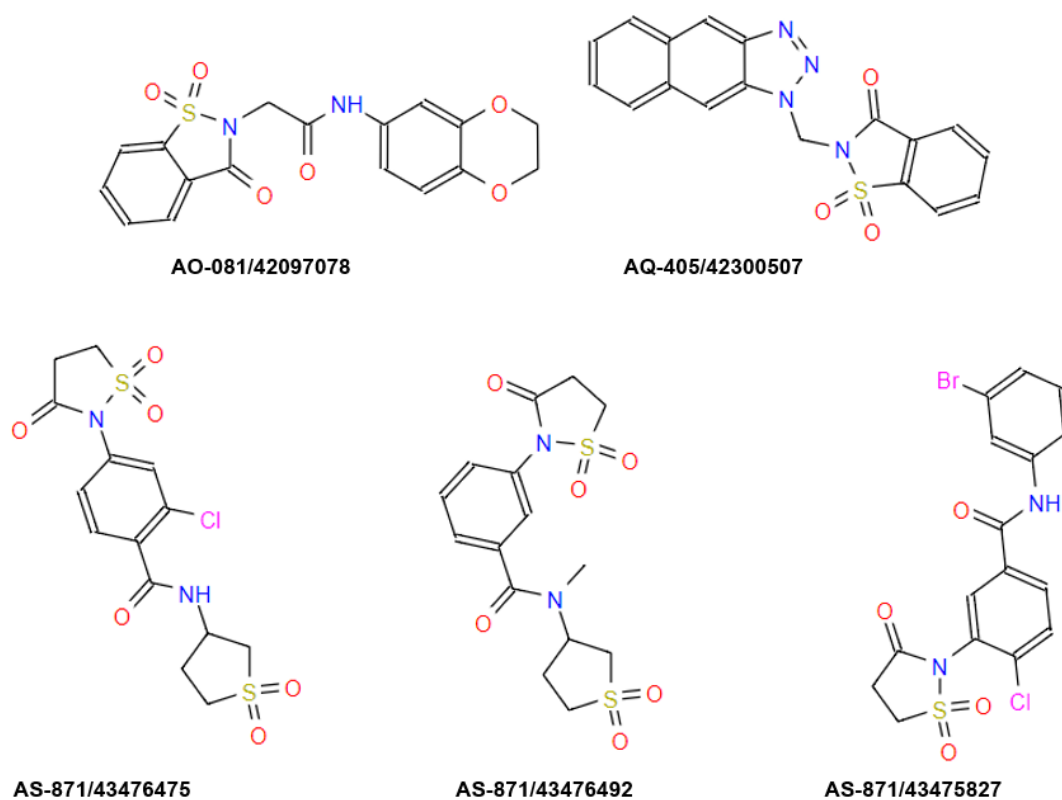


Figure 23. Chemical structures of five hit bi-functional inhibitors

CONCLUSION

The combination of MD simulations and CoMSIA studies were employed to investigate the structural insight for rational design of bi-functional inhibitors. Initially, a combination of CoMSIA and MD simulations was used to investigate the structural features of 2,3-anhydroquinone derivatives required for targeting *Mtb*DHQ. The quantitative contribution of each residue of *Mtb*DHQ responsible for inhibitor binding was visualized using per-residue free energy decomposition. This revealed that His101 had the greatest contribution for binding of 2,3-anhydroquinone derivatives in the *Mtb*DHQ pocket. The carboxylate moiety is crucial for binding of 2,3-anhydroquinone derivatives, whereas the R substituent has a small contribution. The CoMSIA method was used to visualize the structural requirement of an R substituent for better binding affinity of 2,3-anhydroquinone derivatives. The required feature of an R substituent was related to the positional change in *Mtb*DHQ residues as determined through MD simulation. A small linker attached to the C3 atom of the 2,3-anhydroquinone core is preferable for the R substituent because the positions of Pro11, Asn12, Leu13 and Asp88^a located near the C3 atom are not rearranged to accommodate a different sized R substituent. Alternatively, a bulky group could be introduced to the end of an R substituent, since it located in the pocket formed by Arg15, Leu16, Gly17, Arg18, Arg19, Glu20, Pro21 and Glu92^a. This can be rearranged to accommodate a different sized R substituent. Then, the structural features for targeting *mt*DHQ obtained for 2,3-anhydroquinone derivatives were mapped into the known guiding principles for compound accumulation in gram-negative bacteria. This guides the replacement of a negatively charged carboxylate moiety with a neutral hydrogen bond acceptor. The presence of NH_3^+ at the end of an R substituent may allow *Mtb*DHQ inhibitors to accumulate inside *M. tuberculosis* cells. In the case of *Mtb*SDH, binary DHS/*Mtb*SDH and ternary DHS/NADPH/*Mtb*SDH structures were determined using MD simulations. Furthermore, two mutant *Mtb*SDH complex structures, including ones for DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH, were investigated to visualize the influence of these mutations on DHS binding. Structural

analysis of *Mtb*SDH complex structures in terms of binding interaction, flexibility and positional change yielded insights into substrate and cofactor binding in *Mtb*SDH and a structural framework for rational drug design. The skeleton of a hybrid inhibitor planned to bind both in DHS and NADPH pockets has been elucidated here. The results integrated from this work were used to design five core structures of bi-functional inhibitor. Then, these core structures were utilized for virtual screening of bi-functional inhibitors. Finally, five hit bi-functional inhibitors were obtained from virtual screening approach. These hit compounds would be purchased and evaluated as anti-TB agents.

APPENDIX

Assessing structural concepts to balance the activity of 2,3-anhydroquinone derivatives against both type II dehydroquinase and *Mycobacterium tuberculosis*

Auradee Punkvang^{1,*}, Pharit Kamsri¹, Patchareenart Saparpakorn², Supa Hannongbua² and Pornpan Pungpo³

¹Faculty of Science, Nakhon Phanom University, Nakhon Phanom, 48000, Thailand

²Department of Chemistry, Faculty of Science, Kasetsart University, Chatuchak, Bangkok, 10900, Thailand

³Department of Chemistry, Faculty of Science, Ubon Ratchathani University, 85 Sthollmark Rd., Warinchamrap, Ubon Ratchathani, 34190, Thailand

*e-mail: auradee.punkvang@npu.ac.th

Abstract

Mycobacterium tuberculosis type II dehydroquinase (mtDHQ), one of the enzymes in the shikimate pathway, is an attractive target because it is essential for *Mycobacterium tuberculosis* (*M. tuberculosis*), but absent from humans. All of currently developed 2,3-anhydroquinone derivatives targeting mtDHQ show poor activity against *M. tuberculosis* due to the permeability barrier of the mycobacterial cell wall. In the present work, we report structural concepts for designing 2,3-anhydroquinone derivatives that can accumulate inside *M. tuberculosis* and target mtDHQ. These will assist in the rational design of potent compounds against *M. tuberculosis* targeting mtDHQ. Integration of various computational approaches (per-residue free energy decomposition, Comparative Molecular Similarity Index Analysis (CoMSIA) and molecular dynamics (MD) simulations) into the current guiding principles for compound accumulation in gram-negative bacteria is our strategy to develop structural concepts. The carboxylate moiety is crucial for binding affinity of 2,3-anhydroquinone derivatives, but it impedes accumulation inside the bacterial cell. The replacement of the negatively charged carboxylate moiety with a neutral hydrogen bond acceptor should balance the activity of 2,3-anhydroquinone derivatives against both *M. tuberculosis* and mtDHQ. Due to the capability of changing the position of mtDHQ residues, the R substituent attached to the C3 atom of the 2,3-anhydroquinone core requires a small linker and bulky end cap. Importantly, the presence of NH_3^+ at the end of R substituent might enhance the activity of 2,3-anhydroquinone derivatives against both *M. tuberculosis* and mtDHQ.

Keywords

Mycobacterium tuberculosis, type II dehydroquinase, shikimate pathway, MD simulation, QSAR

1. Introduction

Tuberculosis (TB) caused by *M. tuberculosis* remained one of the top 10 causes of death worldwide in 2015. An estimated 10.4 million new (incidences) TB cases worldwide and an estimated 480 000 new cases of multidrug-resistant TB (MDR-TB) have been reported by the World Health Organization (WHO) [1]. New patients with drug susceptible TB can be cured in six months using the standard regimens with a combination of first line TB drugs including isoniazid, rifampicin, pyrazinamide, ethambutol and streptomycin. However, *M. tuberculosis* strains have developed resistance to the available TB drugs. Drug-resistant TB is classified as

multi drug-resistant tuberculosis (MDR-TB), extensively drug-resistant tuberculosis (XDR-TB) and totally drug-resistant tuberculosis (TDR-TB) [2-4]. Due to drug resistant TB, the standard regimens using existing TB drugs are often ineffective. To combat the drug resistant TB, intensive research efforts are being directed towards identification and development of new anti-TB agents. The potential strategy for the development of new drugs against *M. tuberculosis* is to target essential biosynthetic pathways of this bacteria that are absent in humans. This is very important for design of specific drugs with low toxicity. The shikimic acid pathway is present in bacteria, fungi, plants, and in certain apicomplexan parasites, but absent in humans. Therefore, it is an attractive target for the development of new antimicrobials [5-8]. Seven enzymes involved in this pathway sequentially catalyze a series of chemical reactions for biosynthesis of chorismate [9]. This is a precursor for the biosynthesis of aromatic amino acids, folate, ubiquinone, and vitamins E and K [7, 10]. Type II dehydroquinase of *M. tuberculosis* (mtDHQ), the third enzyme in the shikimic acid pathway, has generated the most interest for rational design of anti-TB agents. It is a dodecameric enzyme assembled from four trimeric units, the minimum active unit of type II dehydroquinase [11]. mtDHQ catalyzes the reversible dehydration of 3-dehydroquinate (DHQ) to form 3-dehydroshikimate (DHS). The atomic details of the catalytic mechanism of this enzyme have been elucidated using computational and biochemical studies. The enzymatic process is initiated by the direct deprotonation of Tyr24 by Asp88^a from a neighboring subunit to generate the catalytic tyrosinate. The abstraction of the more acidic axial hydrogen of the substrate by tyrosinate is followed by generation an enolate intermediate. For product formation, the spontaneous abstraction of the His101 N δ proton was followed by formation and elimination of a water molecule [12]. A large number of competitive reversible inhibitors of mtDHQ have been developed that are comprised of analogues of the natural substrate [13] and mimic of the intermediate of the enzyme-catalyzed reaction [14-24]. The several crystal structures of mtDHQ-inhibitor binary complexes have also been reported [14, 15, 25, 26]. These crystal structures are very important since they reveal the role of key residues responsible for catalysis and binding of mtDHQ inhibitors. However, the crystal structures do not reveal the quantitative contribution of each mtDHQ residue on the binding of its inhibitors. Accordingly, in this work, per-residue free energy decomposition were calculated to visualize the quantitative contribution of each residue. This will provide clear knowledge of the key residues and fragments responsible for binding of mtDHQ inhibitor. Comparative Molecular Similarity Index Analysis (CoMSIA) [27] combined with molecular dynamics (MD) simulations were employed in this work to determine the structural requirements of the R substituent attached to the C3 atom of the core structure of 2,3-anhydroquinone derivatives. This is the substituent mostly utilized to design potent mtDHQ inhibitors. The integrated results obtained from our work provide structural concepts for rational design of potent mtDHQ inhibitors. The previous development of 2,3-anhydroquinone derivatives targeting mtDHQ was confounded by their poor activity against *M. tuberculosis*. The *in vitro* antibacterial activity of developed 2,3-anhydroquinone derivatives was not observed at the minimum inhibitory concentration (MIC), i.e., below 200 $\mu\text{g/ml}$ [14]. To overcome this problem, the structural concepts for designing mtDHQ inhibitors with a potent MIC value against *M. tuberculosis* are developed in the present work.

2. Materials and methods

2.1 3D structural generation of 2,3-anhydroquinone derivatives

The chemical structures of 2,3-anhydroquinone derivatives were developed using the standard tools available in the GaussView 3.07 program. Then, these structures were fully optimized using the HF/6-31G* method implemented in the Gaussian 03 program.

2.2 Molecular docking calculations

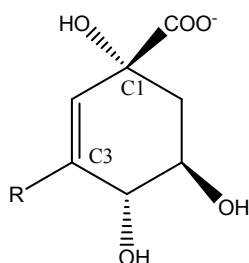
Molecular docking calculations were performed using the Autodock 4.2 program [28]. The coordinates found in the crystal structure of trimeric mtDHQ units designated as chains A, K and L (PDB code 3N87) [15] were used for molecular docking calculations. Residues of mtDHQ were kept rigid, whereas the structure of the inhibitor was flexible during the molecular docking calculations. The number of grid points in the x, y and z dimensions with a value of 40 and a grid point spacing of 0.375 Å were used to define the 3D grid box size. The grid box was centered on the coordinates found in the crystal structure of the inhibitor contained in chain A. The conformations of inhibitor in the mtDHQ pocket were generated using the Lamarckian Genetic Algorithm (LGA) with the GA runs of 100. The RMSD value between the docked and observed X-ray conformations of inhibitor lower than 1 Å were used to verify the docking calculations. Then, molecular docking calculations with validated parameters were performed for all of 2,3-anhydroquinone derivatives. The conformation showing the lowest binding energy was selected as the representative binding mode of each inhibitor obtained from molecular docking calculations.

2.3 CoMSIA method

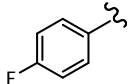
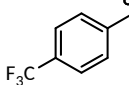
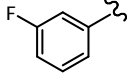
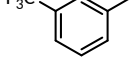
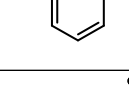
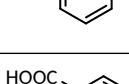
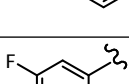
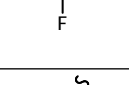
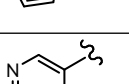
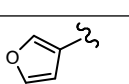
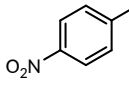
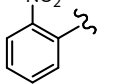
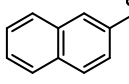
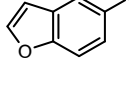
2.3.1 Data set and molecular alignment

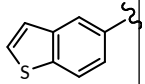
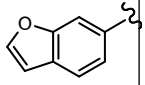
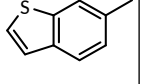
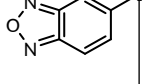
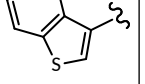
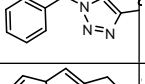
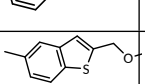
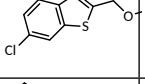
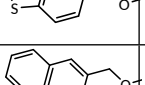
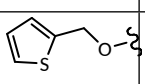
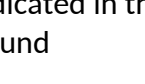
Chemical structures and biological activities expressed in term of the K_i values of 2,3-anhydroquinone derivatives (Table 1) were collected from the published literature [14, 16, 17]. These K_i values were transformed into $\log(1/K_i)$. 2,3-anhydroquinone derivatives were divided into sets of 24 training compounds for development of a CoMSIA model and 4 compounds that were used as a test set for model validation. The test set compounds were selected based on their structural diversity and wide range of activities. The binding modes of the data set obtained from molecular docking calculations were used for molecular alignment to set up the CoMSIA model.

Table 1 Chemical structures and $\log(1/K_i)$ values of 2,3-anhydroquinone derivatives



Code	Compound ^a	R	$\log(1/K_i)$
1	4a		2.82

2	4b		2.74
3	4c		1.77
4	4d		2.82
5	4e		2.67
6	4f		1.02
7	4g		4.27
8 ^b	4h		0.82
9	4i		2.61
10	4j		3.23
11	4k		1.35
12 ^b	4l		3.08
13	1-5a		2.19
14	2-5b		0.96
15	6		2.92
16	10a		2.70

17	10b		2.55
18	11a		3.01
19	11b		3.07
20 ^b	12		2.71
21	13		1.88
22	14		2.43
23	5a		4.55
24	5b		4.37
25	5c		4.46
26 ^b	5d		4.51
27	5e		4.46
28	5f		3.63

^a The compound name indicated in the published literature

^b indicates test set compound

2.3.2 CoMSIA model development

SYBYL-X 2.0 molecular modeling software was used to develop the CoMSIA model. Five CoMSIA descriptor fields including steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields were calculated for each compound. Partial least square (PLS) analysis was employed to derive a linear relationship between CoMSIA descriptor fields and activities. PLS analysis using the leave-one-out (LOO) cross-validation method was performed to determine the optimal number of components. Sequentially, a final analysis with the optimal number of components was performed to develop a CoMSIA model. The non-cross-validated correlation coefficient (r^2) and the leave-one-out cross-validated correlation coefficient (q^2) were used to evaluate the predictive capability of the CoMSIA model. To evaluate the external predictive ability of the selected CoMSIA model, it were employed to predict $\log(1/K_i)$ values of test set compounds that were not used to develop the CoMSIA model.

2.4 Molecular dynamics simulations

Molecular dynamics simulations were carried out using AMBER12 software [29-31]. The chain A bound inhibitor obtained from molecular docking calculations was superimposed with chains K and L to create the coordinates of the inhibitor in the active sites of chains K and L. Then, trimeric mtDHQ (chains A, K and L) in complex with inhibitors was used as the initial structure for MD simulations. All missing hydrogen atoms of trimeric mtDHQ were added using the LEaP program [31]. The delta and epsilon nitrogens of His101 were protonated on the basis of the reaction mechanism [12]. The Amber ff12SB force field was used for the physical description of mtDHQ. The general Amber force field (GAFF) [32] and restrained electrostatic potential (RESP) partial charges [33-35] of the mtDHQ inhibitors were generated using the antechamber module implemented in the AMBER12 package. The initial complex structure was solvated with TIP3P water [36] in a truncated octahedral box extending up to 10 Å from the solute species. Na⁺ ions were added to neutralize the system charge. Energy minimization was performed in two steps, each using a steepest decent algorithm followed by the conjugate gradient algorithm. In the first step, water molecules and counter ions were minimized by restraining all atoms of the solute with a restraint weight of 500 kcal/molÅ². In the second step, the whole system was minimized without any restraint. Then, the system was gradually warmed from 0 K to 300 K in 30 ps by restraining all atoms of the complex with a restraint weight of 2 kcal/molÅ². This was followed by 70 ps of position-restrained dynamics simulation with a restraining weight of 2 kcal/molÅ² at 300 K under an isobaric condition. Finally, 10 ns MD simulations without any restraints were performed using the same conditions. Long-range electrostatic interactions were applied using Particle Mesh Ewald (PME) [37] during the simulations. The cut-off distance for the long-range van der Waals interaction was set to 8 Å. The SHAKE method [38] was applied to constrain the bond lengths of hydrogen atoms attached to heteroatoms. Coordinates and energy outputs during MD simulation were every 20 ps.

2.5 Per-residue free energy decomposition

The coordinates found in the crystal structure of trimeric mtDHQ (chains A, K and L) in complex with compounds **2** and **4** (PDB codes 3N7A and 3N86, respectively) were selected for per-residue free energy decomposition. Before this calculation, energy minimization for each system was performed using the SANDER module of AMBER 12. The energy minimization and the system preparation were similarly prepared as described for MD simulations, except for the remaining water molecules found in each crystal structure. The coordinates of trimeric mtDHQ in complex with compounds **2** and **4** obtained from the energy minimization were taken as the per-residue free energy decomposition. The MM-GBSA method in AMBER 12 was used to perform the per-residue free energy decomposition. The per-residue free energy included van der Waals, electrostatic, polar solvation and nonpolar solvation contributions.

3. Results and Discussion

3.1 Quantitative investigation for mtDHQ-inhibitor interaction

The high resolution crystal structures of mtDHQ with 2,3-anhydroquinone derivatives [14, 15] were solved to provide insight into inhibitor binding. However, the quantitative contribution of each mtDHQ residue on inhibitor binding could not be determined. Accordingly, per-residue free energy decomposition was performed to investigate this

quantitative contribution. Most of the solved crystal structures of mtDHQ in complex with their inhibitor were found to be monomeric, and the flexible-loop residues 19–24 were mostly missing [14, 25, 26]. Therefore, the crystal structure of trimeric mtDHQ (chains A, K and L) in complex with compound **4**, a promising mtDHQ inhibitor with a K_i value of 2.3 μM [15], was selected. This structure shows all residues in the active site, including the flexible-loop residues 19–24. The quantitative contribution of individual residues in chain A and adjacent chain L responsible for the binding of compound **4** was visualized through its decomposition energy (Fig. 1). Most of the residues in an adjacent chain L had a weak contribution to the binding of this compound. Only Asp88^a (^a indicates the residue from an adjacent chain L) had an attractive interaction energy of -9.45 kcal/mol, consistent with a hydrogen bond between its side chain and the OH moiety of compound **4** (Fig. 2). The lowest interaction energy of -15.00 kcal/mol was observed between His101 and compound **4** (Fig. 1), indicating that His101 had the greatest contribution to the binding of this compound. It was responsible for the electrostatic interaction of the carboxylate (COO^-) moiety of compound **4** with the positive charge of His101 and one hydrogen bond of the OH with an NH group of His101 (Fig. 2). These results indicate that Asp88^a and His101 are not only important for the enzymatic process of mtDHQ, but they are also crucial for inhibitor binding. Other prominent attractive interaction energies were observed for Asn75, Ile102, Ser103 and Arg112 (Fig. 1) consistent with hydrogen bonds formed between these residues with compound **4** (Fig. 2). It is notable that the carboxylate moiety played the most important role in creating these remarkable attractive interaction energies. Obviously, it interacts with four key residues, Asn75, His101, Ile102 and Ser103 (Fig. 1). Accordingly, the carboxylate moiety was crucial for binding of compound **4** in the mtDHQ pocket. The R substituent had a small contribution to binding of this compound. It generated two prominent attractive interaction energies with the residues of Asn12 and Tyr24 (Fig. 1).

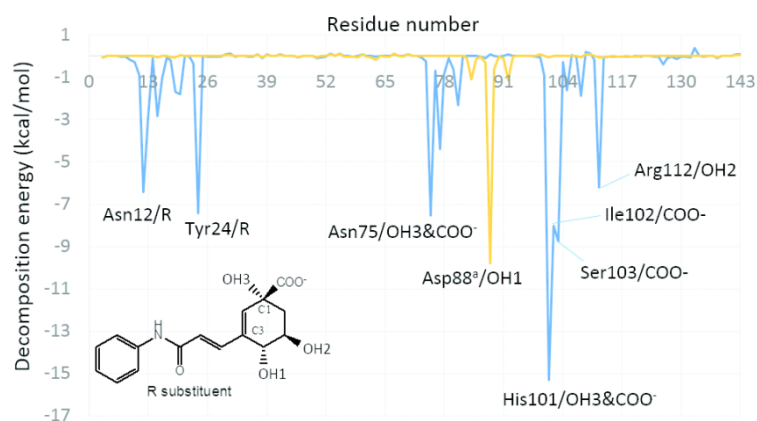


Fig. 1. The per-residue free energy decomposition plot for all residues except those of residues 1-3 of the mtDHQ chain A (blue line) and adjacent chain L (orange line) in complex with compound **4**

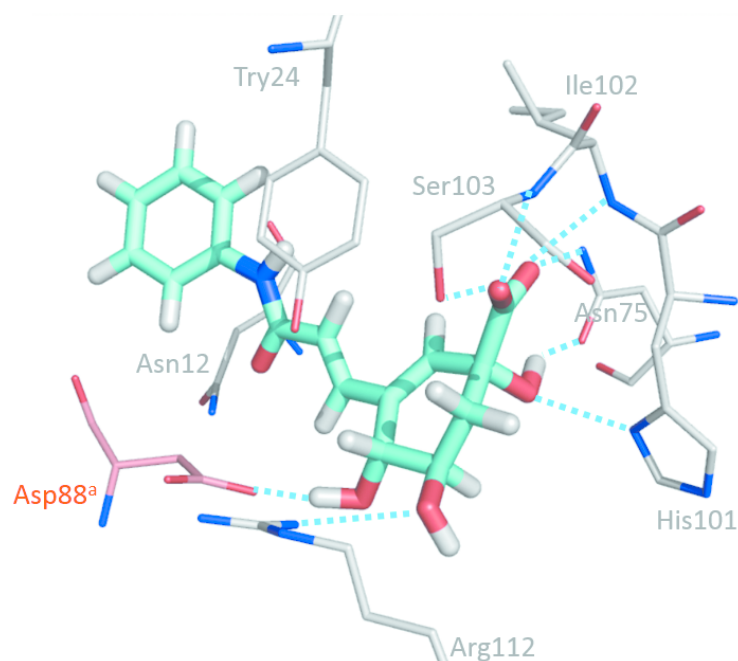


Fig. 2. The binding mode of compound **4** (cyan) in the pocket of mtDHQ (PDB code 3N86)

3.2 The influence of R substituent on the binding affinity

2,3-anhydroquinone derivatives were designed by the modification of the R substituent attached to the C3 atom of the core structure (Table 1). It has been substantiated that the R substituent helps to form a stabilizing interaction between inhibitors and the flexible active-site loop residues 19-24. They are a significant factor for increased potency of the inhibitor [15]. The potency of compound **4** ($K_i=2.3 \mu\text{M}$) increased approximately 87 fold compared with compound **2** ($K_i=200 \mu\text{M}$) [15]. Therefore, this work aimed to quantitatively investigate the influence of the R substituent on the binding interaction of 2,3-anhydroquinone derivatives. In this case, compounds **2** and **4** (Fig. 3) were selected for per-residue free energy decomposition to determine the interaction energy per residue of mtDHQ bound with these compounds. The contribution of each mtDHQ residue for inhibitor binding is quantitatively visualized in Fig. 4. As expected, the R substituent of compound **4** had more attractive interaction energies with the surrounding residues (Asn12, Leu13, Arg15, Leu16, Arg19, Glu20 and Tyr24) than those of compound **2**. Remarkably, Asp88^a had a large contribution for binding of compound **4**, whereas it showed a small contribution for binding of compound **2**. This indicates that the R substituent not only produced a greater degree of complementarity to the mtDHQ pocket, but it is also enhanced the binding affinity of the 2,3-anhydroquinone core with Asp88^a.

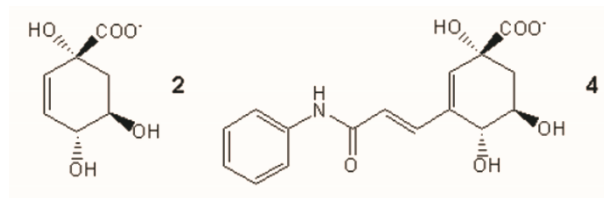


Fig. 3. The chemical structures of compounds **2** and **4** [15]

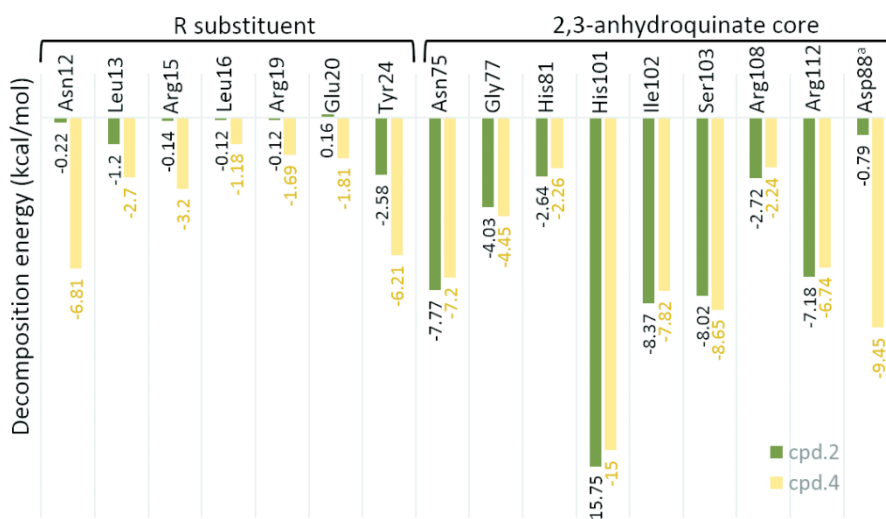


Fig. 4. The per-residue free energy decomposition plot for the binding residues of mtDHQ in complex with compounds **2** and **4**

3.3 The structural requirement of R substituent for the better binding affinity

The R substituent attached to the C3 atom of the core structure was the only substituent utilized to design potent 2,3-anhydroquinone derivatives. Accordingly, the CoMSIA method combined with MD simulations was employed in this work with the aim of elucidating the structural requirements of the R substituent for rational design of new mtDHQ inhibitors with better binding affinity.

3.3.1 CoMSIA model and contour map

The CoMSIA model is valid with the q^2 and r^2 values of 0.68 and 0.97 respectively, with six optimal components. They indicate the high power of this model to estimate the $\log(1/K_i)$ values of training set data. Additionally, the CoMSIA model showed good capability in predicting $\log(1/K_i)$ values of the test set data with an r^2 value of 0.86. The CoMSIA model was developed from a combination of three molecular descriptors including steric, hydrophobic and hydrogen donor fields. The contributions of these fields to the $\log(1/K_i)$ value of 2,3-anhydroquinone derivatives were 18.7%, 56.8% and 24.5%, respectively. This indicated a greater contribution of the hydrophobic field on the activity for mtDHQ inhibition. The steric, hydrophobic and hydrogen donor fields important for mtDHQ inhibition could be visualized in CoMSIA contour maps (Fig. 5). Green and yellow contours indicate areas of favorable and unfavorable steric bulk, respectively (Fig.5a). Magenta and white contours represent areas where the hydrophobic and hydrophilic groups were predicted to favor biological activities (Fig.5b). The cyan and orange contours indicate regions that favor the hydrogen donor group and do not favor the hydrogen donor group, respectively (Fig.5c). The following interpretation of CoMSIA contour maps reveals the structural requirements of an R substituent for better binding affinity of 2,3-anhydroquinone derivatives.

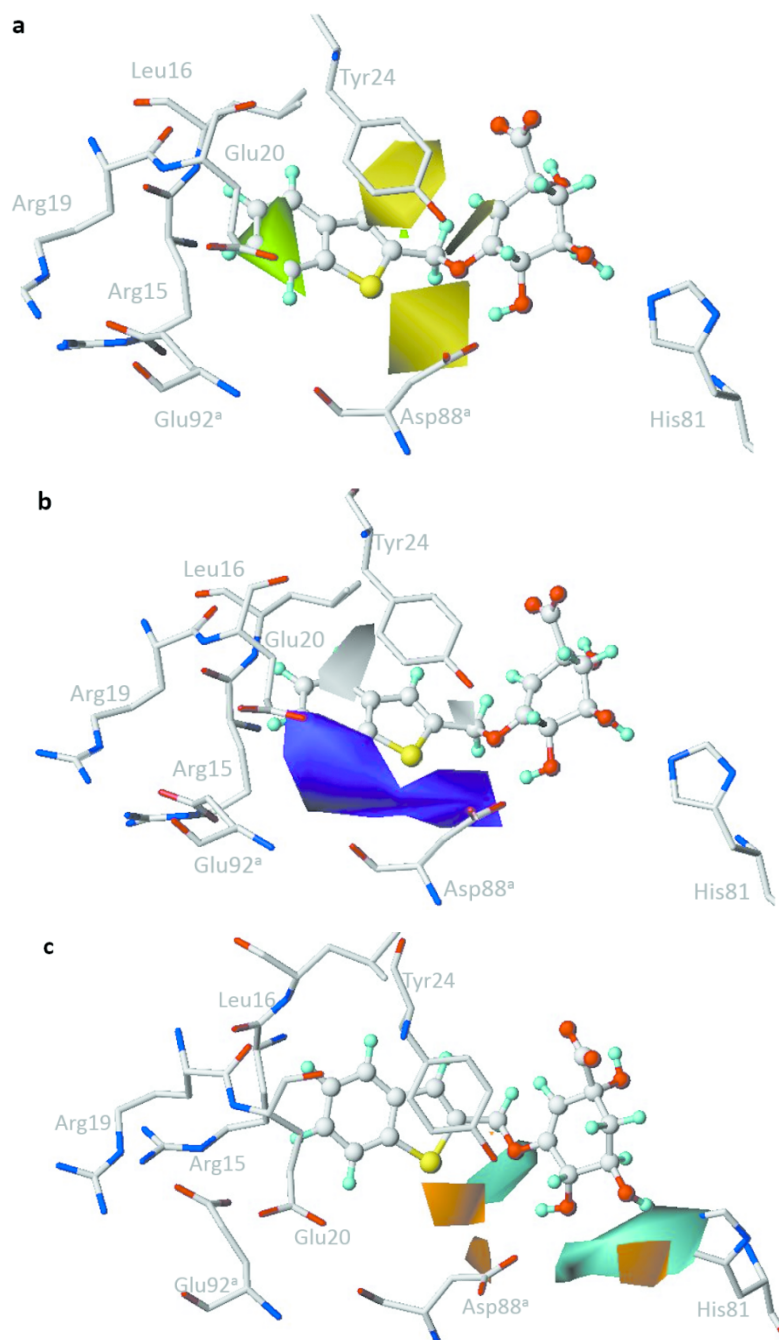


Fig. 5. Steric (a), hydrophobic (b) and hydrogen donor (c) contour maps of CoMSIA model in combination with the most active compound code **23** (ball and stick) and mtDHQ residues (stick)

3.3.2 The structural requirement interpretation based on CoMSIA contours

Two yellow contours located near the C3 atom of 2,3-anhydroquinone core, in which the R substituent is directly attached, indicate unfavorable steric bulk. Obviously, most of compounds with the bulky R substituent directly attached to the C3 atom showed the least activity for mtDHQ inhibition, such as compound codes **1-22** (Table 1). Another steric requirement is indicated by the green contour at the end of the R substituent of compound code **23**, far from the C3 atom. This contour indicates a favorable steric bulk in the green

region. Based on the present of steric contours, the R substituents should contain a small linker near the C3 atom and the bulky end cap, such as the R substituent of compound codes **23-28**. This feature of the R substituent might be specified as the capability of changing the position of mtDHQ residues (described in Section 3.4).

The large purple contour located at the edge of the R substituent of compound code **23** indicates a favorable hydrophobic moiety at this site (Fig. 5b). This favorable hydrophobic region is surrounded by the CH₂ side chains of two negatively charged residues of an adjacent subunit (Glu92^a and Asp88^a). The presence of a hydrophobic moiety in this region might be preferable for interactions with the alkyl moieties of Glu92^a and Asp88^a. All hydrogen donor contours are present near the 2,3-anhydroquinone core, the general structure of all compounds in the data set. There is no significant hydrogen donor contours present for the R substituent. Therefore, the hydrogen donor atom may be ignored in the design of an R substituent.

3.4 The positional change of mtDHQ for inhibitor binding

To reveal the positional change of mtDHQ necessary to accommodate a different inhibitor, two complex structures of compound codes **23** and **21** (Table 1) with mtDHQ were modeled by MD simulations. Then, root-mean-square deviation (RMSD) values between the coordinates of each residue of mtDHQ complexed with compound codes **23** and **21** were calculated to measure their positional change. The R substituent of compound code **23**, the most active compound, coincided with the steric contours, whereas that of compound code **21** violated the steric requirement of an R substituent. The binding mode of the 2,3-anhydroquinone cores of compound codes **23** and **21** did not show significant alteration, but that of the R substituent was different (Fig. 6). Five residues of mtDHQ including Pro11, Asn12, Leu13, Tyr24 and Asp88^a were located near the C3 atom. Except Tyr24, the positions of these residues were not significantly changed in either of the complexed structures of compound codes **23** and **21** (Fig. 6). Their RMSD values were in the range of 0.32-1.02 Å (Fig. 7). It is clear that the positions of Pro11, Asn12, Leu13 and Asp88^a were not rearranged to accommodate the different sized R substituent near the C3 atom. The bulky moiety at this site might cause steric hindrance with Pro11, Asn12, Leu13 and Asp88^a. To avoid the steric hindrance of these residues, the bulky R substituent of compound code **21** significantly altered the position of the key catalytic residue (Tyr24) from that found for the binding of compound code **23** (Fig. 6) with a RMSD value of 6.54 Å (Fig. 7). The position of Tyr24 found for the binding of compound code **23** was similar to that observed for the binding of the 3-dehydroshikimate substrate in the catalytic reaction of DHQase (Fig. 8). This implies that binding of the inhibitor causing the large change of the Tyr24 position observed in the catalytic reaction results the loss of binding affinity in the DHQase pocket. Accordingly, the small linker near the C3 atom was required for the R substituent related to the unfavorable yellow contour (Fig. 5a). The end of the R substituent, which is buried in the favorable green contour (Fig. 5a), was surrounded by Arg15, Leu16, Gly17, Arg18, Arg19, Glu20, Pro21, His63^a, Ala89^a and Glu92^a. This fragment of compound code **21** was absent in this pocket, whereas the R substituent fragment of compound code **23** was present (Fig. 6). A large difference in the positions of these residues was found for binding of compound codes **21** and **23** with RMSD values in range of 1.15-6.71 Å, except for His63^a and Ala89^a (Fig. 7). Therefore, the

pocket formed by Arg15, Leu16, Gly17, Arg18, Arg19, Glu20, Pro21 and Glu92^a could be rearranged to accommodate a different sized R substituent.

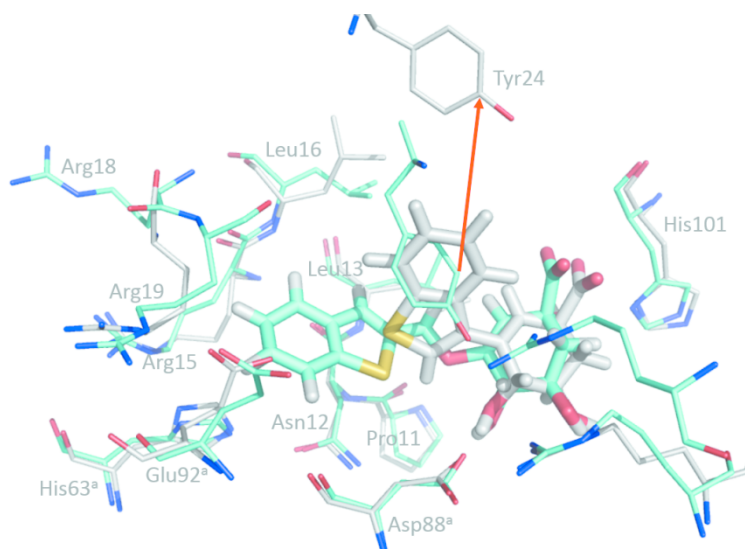


Fig. 6. Superimposition of compound codes **21** (carbon atom colored in grey) and **23** (carbon atom leveled in cyan) in the mtDHQ binding pocket. The carbon atom of the residue complexed with compound codes **21** and **23** are represented in grey and cyan, respectively

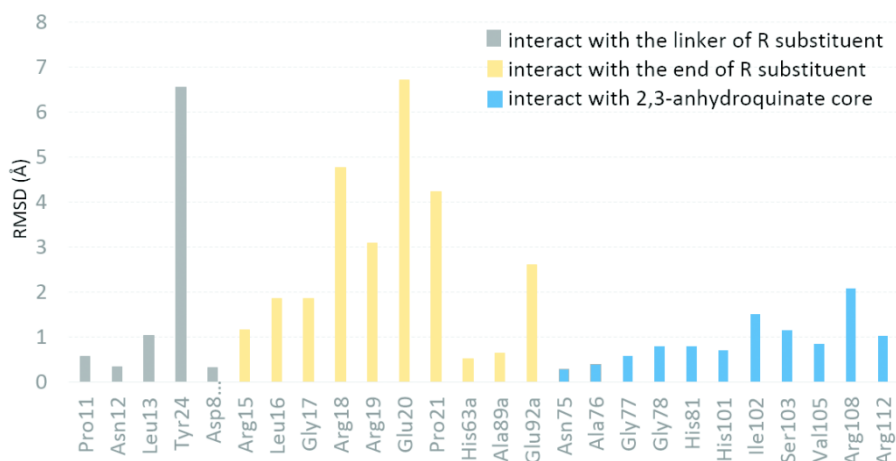


Fig. 7. RMSD values of the coordinates of the mtDHQ binding residues complexed with compound codes **23** and **21**

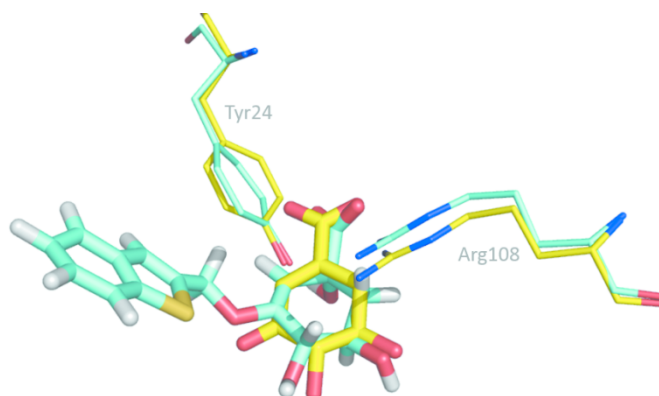


Fig. 8. The positions of Tyr24 found for binding of compound code **23** (carbon atoms in cyan) and 3-dehydroshikimate substrate, 3N59 PDB code (carbon atoms in yellow)

3.5 The structural concept to design mtDHQ inhibitors with potent activity against *M. tuberculosis*

In vitro activity against *M. tuberculosis* was not observed for the compounds listed in Table 1 at the minimum inhibitory concentration (MIC), below 200 µg/ml. It was assumed that the high hydrophilic nature of these compounds caused this lack activity because hydrophilic compounds traverse the mycobacterial cell wall slowly [14]. Recently, systematic analysis of the accumulation of small molecules in gram-negative bacteria (*E. coli*) provided the guiding principle for compound accumulation in *E. coli*. Compounds that are most likely to accumulate contain a non-sterically encumbered amine, are amphiphilic and rigid, and have low globularity. Additionally, carboxylic acid compounds with a strong negative charge do not accumulate in *E. coli* [39]. Although, *M. tuberculosis* is classified as gram-positive bacteria, it has a closer relationship to gram-negative bacteria [40]. Therefore, the recent guiding principle for compound accumulation in *E. coli* might be used for compounds that antagonize *M. tuberculosis*. Indeed, all four current first line TB drugs, isoniazid, rifampin, ethambutol and pyrazinamide, contain an amine group. Moreover, transformation of the carboxylate groups of compounds listed in Table 1 into the stable esters dramatically increases their efficacy against *M. tuberculosis*. For example, the propyl ester in compound code **23** is more active than the corresponding carboxylated compound code **23** with MIC values of 5 and 200 µg/ml, respectively [14]. This implies that compounds with efficacy against *M. tuberculosis* seem to these features that promote for accumulation in *E. coli*. This finding could provide a strategy to design mtDHQ inhibitors with better activity against *M. tuberculosis*. On the basis of our strategy, the following guidelines for designing mtDHQ inhibitors with the potential activity against *M. tuberculosis* were developed.

3.5.1 The removal of negatively charged carboxylate group

Based on the guiding principles for compound accumulation in *E. coli*., the negatively charged carboxylate group of some compounds impedes their accumulation inside bacterial cells. However, our finding reveals that the carboxylate moiety is crucial for the binding of 2,3-anhydroquinone derivatives in mtDHQ. It can form four hydrogen bonds with Asn75, Ile102 and Ser103 and is electrostatically attractive to His101. Therefore, the replacement of the negatively charged carboxylate moiety with a neutral hydrogen bond acceptor should balance the activity of 2,3-anhydroquinone derivatives against both *M. tuberculosis* and mtDHQ.

3.5.2 Addition of a non-sterically encumbered amine

A non-sterically encumbered primary amine aids in the accumulation of small molecules in *E. coli* [39]. Therefore, we aimed to introduce -NH₃⁺ into 2,3-anhydroquinone derivatives to increase efficacy against *M. tuberculosis*. To avoid the electrostatic repulsion that occurred from the introduction of -NH₃⁺, positively charged residues located within 8Å of compound code **23** were considered (Fig. 9). Three positive charges of His101, Arg108 and Arg112 were present near the anhydroquinone core of compound code **23**. Therefore, a NH₃⁺ substituent could not be introduced into this moiety. There are two negative charges of Glu20 and Glu92^a located near the end of the R substituent. Accordingly, the NH₃⁺ substituent can be added at the end of the R substituent. It would induce the electrostatic attraction Glu20

and Glu92^a leading to increased affinity for inhibitor binding in mtDHQ. Therefore, the presence of NH₃⁺ might enhance efficacy against both *M. tuberculosis* and mtDHQ.

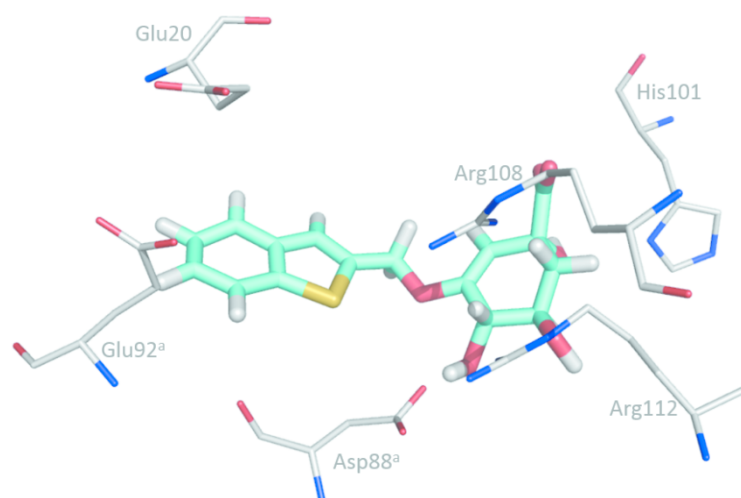


Fig. 9. Charged residues located within 8Å of compound code **23**

Conclusions

Three computational approaches, including per-residue free energy decomposition, CoMSIA and MD simulations combined with the known guiding principle for compound accumulation in gram-negative bacteria could provide structural concepts for designing compounds that can accumulate inside *M. tuberculosis* cells to target mtDHQ. Initially, a combination of three computational approaches was used to investigate the structural features of 2,3-anhydroquinone derivatives required for targeting mtDHQ. The quantitative contribution of each residue of mtDHQ responsible for inhibitor binding was visualized using per-residue free energy decomposition. This revealed that His101 had the greatest contribution for binding of 2,3-anhydroquinone derivatives in the mtDHQ pocket. The carboxylate moiety is crucial for binding of 2,3-anhydroquinone derivatives, whereas the R substituent has a small contribution. The CoMSIA method was used to visualize the structural requirement of an R substituent for better binding affinity of 2,3-anhydroquinone derivatives. The required feature of an R substituent was related to the positional change in mtDHQ residues as determined through MD simulation. A small linker attached to the C3 atom of the 2,3-anhydroquinone core is preferable for the R substituent because the positions of Pro11, Asn12, Leu13 and Asp88^a located near the C3 atom are not rearranged to accommodate a different sized R substituent. Alternatively, a bulky group could be introduced to the end of an R substituent, since it located in the pocket formed by Arg15, Leu16, Gly17, Arg18, Arg19, Glu20, Pro21 and Glu92^a. This can be rearranged to accommodate a different sized R substituent. Then, the structural features for targeting mtDHQ obtained for 2,3-anhydroquinone derivatives were mapped into the known guiding principles for compound accumulation in gram-negative bacteria. This guides the replacement of a negatively charged carboxylate moiety with a neutral hydrogen bond acceptor. The presence of NH₃⁺ at the end of an R substituent may allow mtDHQ inhibitors to accumulate inside *M. tuberculosis* cells. The structural concepts provided here will assist in the rational design of potent compounds against *M. tuberculosis* by targeting mtDHQ.

Acknowledgements

This research was supported by the Thailand Research Fund (MRG5980152). Nakhon Phanom University, Faculty of Science, Ubon Ratchathani University and NECTEC are gratefully acknowledged for supporting this research.

References

1. World Health Organization (WHO). 2016. WHO Global TB Report 2016. Available at http://www.who.int/tb/publications/global_report/en/.
2. World Health Organization (WHO). 2010. Multidrug and Extensively Drug-Resistant TB (M/XDR-TB):2010 Global Report on Surveillance and Response. Available at http://www.who.int/tb/features_archive/m_xdrtb_facts/en/.
3. Velayati A.A., P. Farnia, M.R. Masjedi, T.A. Ibrahim, P. Tabarsi, R.Z. Haroun, H.O. Kuan, J. Ghanavi, P. Farnia, and M. Varahram. 2009. Totally drug-resistant tuberculosis strains: evidence of adaptation at the cellular level. *Eur. Respir. J.* 34:1202-1203.
4. Udawadia Z.F., R.A. Amale, K.K. Ajbani, and C. Rodrigues. 2012. Totally drug-resistant tuberculosis in India. *Clin. Infect. Dis.* 54:579-581.
5. Roberts F., C.W. Roberts, J.J. Johnson, D.E. Kyle, T. Krell, J.R. Coggins, G.H. Coombs, W.K. Milhous, S. Tzipori, D.J. Ferguson, D. Chakrabarti, and R. McLeod. 1998. Evidence for the shikimate pathway in apicomplexan parasites. *Nature*. 393:801-805.
6. Keeling P.J., J.D. Palmer, R.G. Donald, D.S. Roos, R.F. Waller, and G.I. McFadden. 1999. Shikimate pathway in apicomplexan parasites. *Nature*. 397:219-220.
7. Bentley R. 1990. The shikimate pathway--a metabolic tree with many branches. *Crit. Rev. Biochem. Mol. Biol.* 25:307-384.
8. Ducati R.G., L.A. Basso, and D.S. Santos. 2007. Mycobacterial shikimate pathway enzymes as targets for drug design. *Curr. Drug Targets.* 8:423-435.
9. Sassetti C.M., D.H. Boyd, and E.J. Rubin. 2003. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* 48:77-84.
10. Dosselaere F., and J. Vanderleyden. 2001. A metabolic node in action: chorismate-utilizing enzymes in microorganisms. *Crit. Rev. Microbiol.* 27:75-131.
11. Price N.C., D.J. Boam, S.M. Kelly, D. Duncan, T. Krell, D.G. Gourley, J.R. Coggins, R. Virden, and A.R. Hawkins. 1999. The folding and assembly of the dodecameric type II dehydroquinases. *Biochem. J.* 338:195-202.
12. Coderch C., E. Lence, A. Peón, H. Lamb, A.R. Hawkins, F. Gago, and C. González-Bello. 2014. Mechanistic insight into the reaction catalysed by bacterial type II dehydroquinases. *Biochem. J.* 458:547-557.
13. Howard N.I., M.V. Dias, F. Peyrot, L. Chen, M.F. Schmidt, T.L. Blundell, and C. Abell. 2015. Design and structural analysis of aromatic inhibitors of type II dehydroquinase from *Mycobacterium tuberculosis*. *ChemMedChem*. 10:116-133.
14. Tizón L., J.M. Otero, V.F. Prazeres, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, J.A. Ainsa, L. Castedo, and C. González-Bello. 2011. A prodrug approach for improving antituberculosis activity of potent *Mycobacterium tuberculosis* type II dehydroquinase inhibitors. *J. Med. Chem.* 54:6063-6084.
15. Dias M.V., W.C. Snee, K.M. Bromfield, R.J. Payne, S.K. Palaninathan, A. Ciulli, N.I. Howard, C. Abell, J.C. Sacchettini, and T.L. Blundell. 2011. Structural investigation of

inhibitor designs targeting 3-dehydroquinate dehydratase from the shikimate pathway of *Mycobacterium tuberculosis*. *Biochem. J.* 436:729-739.

16. Sánchez-Sixto C., V.F. Prazeres, L. Castedo, H. Lamb, A.R. Hawkins, and C. González-Bello. 2005. Structure-based design, synthesis, and biological evaluation of inhibitors of *Mycobacterium tuberculosis* type II dehydroquinase. *J. Med. Chem.* 48:4871-4881.
17. Prazeres V.F., C. Sánchez-Sixto, L. Castedo, H. Lamb, A.R. Hawkins, A. Riboldi-Tunnicliffe, J.R. Coggins, A.J. Lapthorn, and C. González-Bello. 2007. Nanomolar competitive inhibitors of *Mycobacterium tuberculosis* and *Streptomyces coelicolor* type II dehydroquinase. *ChemMedChem.* 2:194-207.
18. Toscano M.D., R.J. Payne, A. Chiba, O. Kerbarh, and C. Abell. 2007. Nanomolar inhibition of type II dehydroquinase based on the enolate reaction mechanism. *ChemMedChem.* 2:101-112.
19. Payne R.J., A. Riboldi-Tunnicliffe, O. Kerbarh, A.D. Abell, A.J. Lapthorn, and C. Abell. 2007. Design, synthesis, and structural studies on potent biaryl inhibitors of type II dehydroquinases. *ChemMedChem.* 2:1010-1013.
20. Payne R.J., F. Peyrot, O. Kerbarh, A.D. Abell, and C. Abell. 2007. Rational design, synthesis, and evaluation of nanomolar type II dehydroquinase inhibitors. *ChemMedChem.* 2:1015-1029.
21. Tran A.T., K.M. Cergol, N.P. West, E.J. Randall, W.J. Britton, S.A. Bokhari, M. Ibrahim, A.J. Lapthorn, and R.J. Payne. 2011. Synthesis and evaluation of potent ene-yne inhibitors of type II dehydroquinases as tuberculosis drug leads. *ChemMedChem.* 6:262-265.
22. Blanco B., A. Sedes, A. Peón, H. Lamb, A.R. Hawkins, L. Castedo, and C. González-Bello. 2012. Synthesis of 3-alkyl enol mimics inhibitors of type II dehydroquinase: factors influencing their inhibition potency. *Org. Biomol. Chem.* 10:3662-3676.
23. Frederickson M., A.W. Roszak, J.R. Coggins, A.J. Lapthorn, and C. Abell. 2004. (1R,4S,5R)-3-Fluoro-1,4,5-trihydroxy-2-cyclohexene-1-carboxylic acid: the fluoro analogue of the enolate intermediate in the reaction catalyzed by type II dehydroquinases. *Org. Biomol. Chem.* 2:1592-1596.
24. Paz S., L. Tizón, J.M. Otero, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, A.J. Lapthorn, L. Castedo, and C. González-Bello. 2011. Tetrahydrobenzothiophene derivatives: conformationally restricted inhibitors of type II dehydroquinase. *ChemMedChem.* 6:266-272.
25. Peón A., J.M. Otero, L. Tizón, V.F. Prazeres, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, F. Gago, L. Castedo, and C. González-Bello. 2010. Understanding the key factors that control the inhibition of type II dehydroquinase by (2R)-2-benzyl-3-dehydroquinic acids. *ChemMedChem.* 5:1726-1733.
26. Lence E., L. Tizón, J.M. Otero, A. Peón, V.F. Prazeres, A.L. Llamas-Saiz, G.C. Fox, M.J. van Raaij, H. Lamb, A.R. Hawkins, and C. González-Bello. 2013. Mechanistic basis of the inhibition of type II dehydroquinase by (2S)- and (2R)-2-benzyl-3-dehydroquinic acids. *ACS Chem. Biol.* 8:568-577.
27. Klebe G., U. Abraham, and T. Mietzner. 1994. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* 37:4130-4146.

28. Morris G.M., D.S. Goodshell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, and A.J. Olson. 1998. Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* 19:1639-1662.
29. Case D.A., T.A. Darden, III T.E. Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman. 2012. AMBER 12 University of California. San Francisco.
30. Pearlman D.A., D.A. Case, J.W. Caldwell, W.S. Ross, III T.E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, and P. Kollman. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.* 91:1-41.
31. Case D.A., T. Cheatham, T. Darden, H. Gohlke, R. Luo, Jr.K.M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. Woods. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26:1668-1688.
32. Wang J., R.M. Wolf, J.W. Caldwell, P.A. Kollman, and D.A. Case. 2004. Development and testing of a general amber force field. *J. Comput. Chem.* 25:1157-1174.
33. Cieplak P., W.D. Cornell, C. Bayly, and P.A. Kollman. 1995. Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem.* 16:1357-1377.
34. Fox T., and P.A. Kollman. 1998. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B.* 102:8070-8079.
35. Bayly C.I., P. Cieplak, W.D. Cornell, and P.A. Kollman. 1993. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* 97:10269-10280.
36. Jorgensen W.L., J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein. 1983. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 79:926-935.
37. Darden T., D. York, and L.J. Pedersen. 1993. Particle Mesh Ewald-an NLog(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089-10092.
38. Ryckaert J.P., G. Ciccotti, H.J.C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327-341.
39. Richter M.F., B.S. Drown, A.P. Riley, A. Garcia, T. Shirai, R.L. Svec, and P.J. Hergenrother. 2017. Predictive compound accumulation rules yield a broad-spectrum antibiotic. *Nature.* 545:299-304.
40. Niederweis M. 2003. Mycobacterial porins – new channel proteins in unique outer membranes. *Mol. Microbiol.* 49:1167-1177.

1
2
3
4
5
6
7 Simulations of Shikimate Dehydrogenase from
8
9
10 *Mycobacterium tuberculosis* in Complex with 3-
11
12
13
14 dehydroshikimate and NADPH Suggest the Structural
15
16
17
18 Concept of a Hybrid *Mtb*SDH Inhibitor
19
20
21
22

23 *Auradee Punkvang^{a,*}, Pharit Kamsri^a, Adrian Mulholland^b, Jim Spencer^c, Supa Hannongbua^d,*
24
25 *and Pornpan Pungpo^e*
26
27

28 ^aFaculty of Science, Nakhon Phanom University, 48000 Nakhon Phanom, Thailand
29
30

31
32 ^bSchool of Chemistry, University of Bristol, Clifton, BS8 1TS Bristol, United Kingdom
33
34

35 ^cSchool of Cellular and Molecular Medicine, University of Bristol, Clifton, BS8 1TD Bristol,
36
37 United Kingdom
38
39

40
41 ^dDepartment of Chemistry, Faculty of Science, Kasetsart University, Chatuchak, 10900
42
43 Bangkok, Thailand
44
45

46 ^eDepartment of Chemistry, Faculty of Science, Ubon Ratchathani University, Warinchamrap,
47
48 34190 Ubon Ratchathani, Thailand
49
50

51
52 **KEYWORDS:** *Mycobacterium tuberculosis*, Shikimate dehydrogenase, shikimate pathway, MD
53
54 simulation
55
56
57
58
59
60

ABSTRACT: Shikimate dehydrogenase (SDH) from *M. tuberculosis* (*MtbSDH*), encoded by the *aroE* gene, is essential for *M. tuberculosis*, but absent from humans. Therefore, it is a promising target for anti-tuberculosis agent development. *MtbSDH* catalyzes the reversible NADPH-dependent reduction of 3-dehydroshikimate (DHS) to shikimate (SKM) in the fourth reaction step of the shikimic acid pathway. The missing of binary DHS/*MtbSDH* complex and ternary DHS/NADPH/*MtbSDH* complex conceal insight into the binding of DHS substrate and NADPH cofactor as well as the catalytic reactions of *MtbSDH*. Here, molecular dynamics (MD) simulations were performed to create these *MtbSDH* complex structures. Several hydrogen bonds created by Ser18, Ser20, Thr65, Lys69, Asn90, Gln243 and Gln247 have prominent contributions to DHS binding in *MtbSDH*. Mutations of highly conserved Lys69 and Asp105 in *MtbSDH* did not greatly influence the binding of DHS substrate. Asp105 holds the ϵ -ammonium group of Lys69 in a suitable position for *MtbSDH* catalysis via a salt bridge. The coordinates of DHS and NADPH in the ternary *MtbSDH* complex obtained from our work are consistent with a C4-*proS* hydride transfer from NADPH to DHS and proton transfer from Lys69 to DHS. The substrate binding pocket size of *MtbSDH* is highly specific for its substrate and the catalytic residue is held in a suitable position for reaction. The structural details observed from the *MtbSDH* complex structures in this work allow us to provide a rational design guideline for a hybrid *MtbSDH* inhibitor.

INTRODUCTION

Tuberculosis (TB), caused by *M. tuberculosis*, remains one of the top 10 causes of death worldwide. In 2016, the WHO reported 6.3 million new cases of TB (up from 6.1 million in 2015).¹ New patients with drug susceptible TB can be cured in six months using the standard

regimens with a combination of first-line TB drugs, including isoniazid, rifampicin, pyrazinamide, ethambutol and streptomycin. However, *M. tuberculosis* strains have developed resistance to the available TB drugs. Drug-resistant TB is classified as multi drug-resistant tuberculosis (MDR-TB), extensively drug-resistant tuberculosis (XDR-TB) or totally drug-resistant tuberculosis (TDR-TB).²⁻⁴ Due to drug resistant TB, the standard regimens using existing TB drugs are now often ineffective. To combat drug resistant TB, intensive research efforts are being directed towards identification and development of new anti-TB agents. The potential strategy for the development of new drugs against *M. tuberculosis* is to target essential biosynthetic pathways of this bacteria that are absent in humans. This is very important for design of specific drugs with low toxicity. The shikimic acid pathway is present in bacteria, fungi, plants, and in certain apicomplexan parasites, but it is absent in humans. Therefore, it is an attractive target for the development of new antimicrobials.⁵⁻⁸ Seven enzymes involved in this pathway sequentially catalyze a series of chemical reactions for biosynthesis of chorismate.⁹ This is a precursor for the biosynthesis of aromatic amino acids, and folate, ubiquinone, as well as vitamins E and K.^{7,10} SDH encoded by the *aroE* gene participates in the fourth reaction step of the shikimic acid pathway. It catalyzes the reversible NADPH dependent reduction of DHS to SKM (Figure 1). The overall structure of SDH is comprised two α/β domains linked centrally by two α -helices. A deep groove created between these two domains is the active site for the binding of substrate and cofactor.¹¹⁻¹⁶ The crystal structures of apo *Mtb*SDH (PDB code 4P4N) and the binary SKM/*Mtb*SDH complex (PDB code 4P4G) have shown that it shares a three-dimensional structure with all members of the SDH family. The kinetic isotope effect and proton inventory studies showed that the kinetic mechanism of *Mtb*SDH is in agreement with a steady-state ordered bi-bi mechanism. First, DHS is bound to the *Mtb*SDH active site and this is

followed by NADPH binding. The C4-*proS* hydride (B side) of NADPH is transferred to DHS in the oxy-reduction reaction of *MtbSDH*.¹⁷ However, in the case of SDH from *E. coli* and *T. thermophilus* HB8, the C4-*proR* hydride (A side) of NADPH is transferred in this reaction.¹⁸⁻¹⁹ Both the hydride transfer of NADPH and the protonation of DHS in the catalytic reaction of *MtbSDH* proceed in the same step (a concerted mechanism).¹⁷ An amino acid residue with an apparent pKa value of 8.9 participates in the catalytic activity of *MtbSDH*.¹⁷ Site-directed mutagenesis shows that the conserved Lys69 plays a catalytic role in *MtbSDH*. The catalytic constant value for the wild-type *MtbSDH* is 68-fold larger than that of the mutant K69A *MtbSDH*.²⁰ However, molecular level information for binding of the DHS substrate and NADPH cofactor in *MtbSDH* is lacking due to the absence of their crystal structures. Here, to gain more insight into the substrate and cofactor binding in *MtbSDH*, we determined the binary *MtbSDH* complex with DHS substrate (DHS/*MtbSDH*) and the ternary *MtbSDH* complex with both NADPH cofactor and DHS substrate (DHS/NADPH/*MtbSDH*) using molecular dynamics (MD) simulations. Furthermore, two mutant *MtbSDH* complexes, including Lys69Ala (K69A) *MtbSDH* and Asp105Asn (D105N) *MtbSDH* in complex with DHS (DHS/K69A *MtbSDH* and DHS/D105N *MtbSDH*, respectively) were investigated to visualize the influence of these mutations on DHS binding. The structural details observed from several complex structures of *MtbSDH* in this work could be used to provide a rational design guideline for a hybrid *MtbSDH* inhibitor.

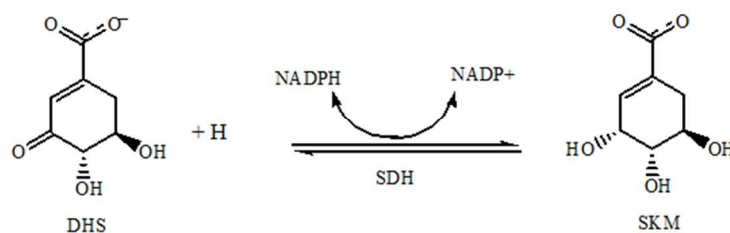


Figure 1. The reaction catalyzed by shikimate dehydrogenase in the fourth reaction step of the shikimic acid pathway.

MATERIALS AND METHODS

Preparation of *Mtb*SDH complex structure. An initial binary DHS/*Mtb*SDH complex structure was generated by molecular docking calculations, as detailed below. Then, it was used to set up a ternary DHS/NADPH/*Mtb*SDH complex structure. The superposition of a binary DHS/*Mtb*SDH complex structure on the x-ray structure of SDH from *Helicobacter pylori* in complex with SKM and NADPH (PDB code 3PHI) was performed to generate the coordinates of NADPH in a ternary DHS/NADPH/*Mtb*SDH complex structure. The nicotinamide ring of NADPH was manually rotated to the *pro*-S conformation. The resulting binary and ternary *Mtb*SDH complex structures obtained were used as the initial coordinates for MD simulations.

Molecular docking calculations. Molecular docking calculations were performed using the Autodock 4.2 program.²¹ The coordinates of *Mtb*SDH from the crystal structure of the SKM/*Mtb*SDH complex structure (PDB code 4P4G) were used for these calculations. Forty grid points in the x, y and z dimensions with a grid point spacing of 0.375 Å were used to define the 3D grid box. The grid box was centered on the coordinates of the SKM found in the crystal structure. Residues of *Mtb*SDH were kept rigid, whereas the structures of small molecules (SKM and DHS) were flexible during the molecular docking calculations. The conformations of small molecules were generated using the Lamarckian Genetic Algorithm (LGA) with 100 GA runs. The RMSD value between the docked and observed x-ray conformations of SKM smaller than 1 Å were used to verify the docking calculations. Then, molecular docking calculations with

validated parameters were performed for DHS. The conformation showing the lowest binding energy was selected as the representative binding mode of DHS in *Mtb*SDH.

Molecular dynamics simulations. Five *Mtb*SDH systems were simulated and are reported here. Three of these systems are wild-type *Mtb*SDH, namely apo *Mtb*SDH, binary DHS/*Mtb*SDH and ternary DHS/NADPH/*Mtb*SDH. The remaining simulations are of mutant *Mtb*SDH systems: two binary structures of DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH. All MD simulations were done using AMBER16 software.²² All missing hydrogen atoms of *Mtb*SDH were added using the LEaP program.²³⁻²⁴ The Amber ff12SB force field was used for the physical description of *Mtb*SDH. The general Amber force field (GAFF)²⁵ and restrained electrostatic potential (RESP) partial charges²⁶⁻²⁸ of DHS and NADPH were generated using the antechamber module implemented in the AMBER16 package. The initial complex structure was solvated with TIP3P water²⁹ in a truncated octahedral box extending 10 Å from the solute species. Na⁺ ions were added to neutralize the system charge. Energy minimization was performed in two steps, each using a steepest descent algorithm followed by a conjugate gradient algorithm. In the first step, water molecules and counter ions were minimized by restraining all atoms of the solute with a restraint weight of 500 kcal/molÅ². In the second step, the whole system was minimized with no restraints. Then, the system was gradually warmed from 0 K to 300 K over 30 ps by restraining all atoms of the complex with a restraint weight of 2 kcal/molÅ². This was followed by 70 ps of position-restrained dynamics simulation with a restraining weight of 2 kcal/molÅ² at 300 K under an isobaric condition. Finally, 60 ns and 100 ns MD simulations for wild-type and mutant *Mtb*SDH structures, respectively, were performed with no restraints using the same condition. Three MD simulations of each *Mtb*SDH system were performed to assess the reproducibility of the results. Long-range electrostatic interactions were applied using Particle Mesh Ewald

(PME)³⁰ during the simulations. The cut-off distance for the long-range van der Waals interaction was set to 8 Å. The SHAKE method³¹ was applied to constrain the bond lengths of hydrogen atoms attached to heteroatoms. Coordinates and energy outputs were recorded every 20 ps during MD simulation. The cpptraj program³² in AMBER16 was employed to calculate the root mean square deviation (RMSD), root mean square fluctuations (RMSF) and distance between atom pairs of interest. The MM GBSA free energy and the decomposition energy using snapshots collected every 40 ps from the equilibrium state of each system were calculated using the python script, MMPBSA.py³³, in the AMBER16 program. The ptraj module in the AMBER12 program³⁴ was used for cluster analysis of the snapshots collected from the equilibrium state of each system. The representative snapshot in the most populated cluster was selected as the representative structure of each system.

RESULTS AND DISCUSSION

Stability of *Mtb*SDH. The RMSD of the distance between atoms of each solute species in five MD simulation systems with respect to the initial structure of each system was calculated to reveal the system stability. *Mtb*SDH in three wild-type systems (apo *Mtb*SDH, binary DHS/*Mtb*SDH and ternary DHS/NADPH/*Mtb*SDH) was equilibrated for ~20 ns, ~10 ns and ~5 ns, respectively, and kept stable over the rest of the simulation time for three MD simulations (see Supporting Information Figures S1-S3). These indicated that wild-type systems of *Mtb*SDH reached their equilibrium state easily when they bound the DHS substrate and NADPH cofactor. Two mutant systems of *Mtb*SDH (DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH) reached their equilibrium state after ~10 ns and ~40 ns respectively, in three MD simulations (see Supporting Information Figures S4-S5). The DHS/K69A *Mtb*SDH system reached an

equilibrium state faster than the DHS/D105N *MtbSDH* system. This indicated that the D105N mutation had a greater effect on the structural rearrangement of *MtbSDH* than the K69A mutation. The structures of each system obtained from three MD simulations were slightly different. Three backbone structures of *MtbSDH* obtained from three MD simulations of apo *MtbSDH*, DHS/*MtbSDH*, DHS/NADPH/*MtbSDH*, DHS/K69A *MtbSDH* and DHS/D105N *MtbSDH* showed small average RMSD values of 1.36 Å, 1.46 Å, 1.00 Å, 1.04 Å and 1.19 Å, respectively.

Binary DHS/*MtbSDH* complex structure. The binding conformation of DHS in *MtbSDH* obtained from MD simulation was very similar to that of shikimate found in the crystal structure of *MtbSDH* (Figure 2). The carboxylate group of DHS was anchored by four hydrogen bonds formed by highly conserved Ser18, Ser20, Thr65 and Tyr215. Two hydroxyl groups (C3 and C4) formed hydrogen bonds with the side chains of Asn90, Asp105, Gln243 and Gln247. The carbonyl oxygen at the C5 atom interacted with the side chain of Lys69 to create a hydrogen bond. Hydrogen bonds created between DHS and Ser18, Ser20, Thr65, Lys69, Asn90, Gln243 and Gln247 had a prominent contribution to the binding of DHS substrate in the DHS/*MtbSDH* complex as evaluated by the decomposition energy (Figure 3). Lys69 showed the greatest contribution to the DHS binding with the lowest energy of -7.76 kcal/mol. The position of Lys69 observed in our investigation relates to the proposed mechanism of *MtbSDH*. It donated a proton to the carbonyl oxygen of DHS and removed a proton from SKM during the catalytic reaction.¹⁷

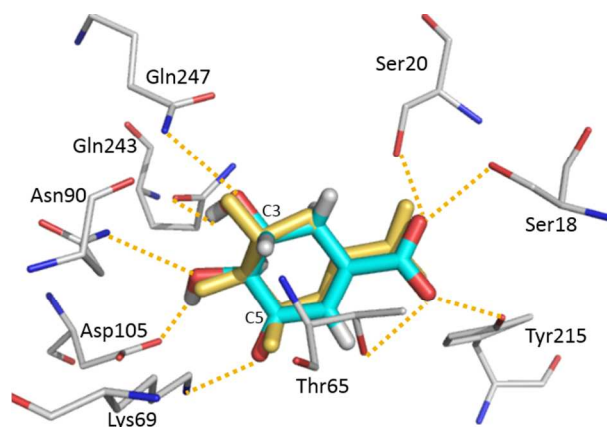


Figure 2. Superposition of a binary DHS/*Mtb*SDH complex structure modeled here by MD simulation with shikimate found in the crystal structure of *Mtb*SDH (PDB code 4P4G). Carbon atoms of DHS are cyan colored and shikimate are orange. Dotted lines indicate hydrogen bond interactions.

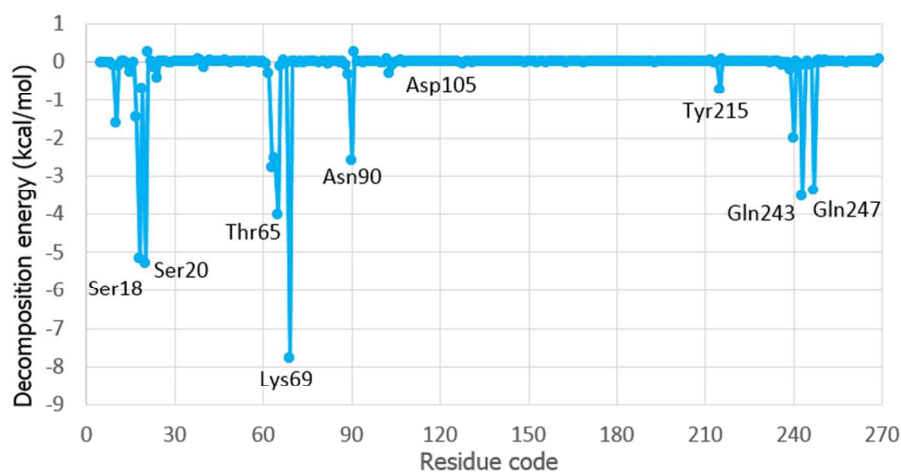


Figure 3. The MM-GBSA decomposition energy showing the quantitative contribution of each *Mtb*SDH residue on the binding of DHS substrate in a binary DHS/*Mtb*SDH complex structure.

K69A and D105N mutations. Lys69 and Asp105 numbered by *M. tuberculosis* code are conserved in *Tt*SDH, *E. coli* SDH, *H. influenzae* SDH, *A. thaliana* SDH-like and *M. jannaschii*

SDH.¹⁹ The crystal structure of *Tt*SDH (PDB code 2EV9) showed that these invariant residues are coupled by a salt bridge. In the catalysis of *Tt*SDH and *A. thaliana* SDH-like, the invariant lysine and aspartate residues were proposed to be a catalytic dyad.^{19,35} Mutation of these invariant residues in *A. thaliana* SDH-like resulted in the loss of enzyme activity.³⁵ The catalysis role of Lys69 in *Mtb*SDH was investigated using the K69A mutation that dramatically decreases the catalytic constant value for *Mtb*SDH. However, this mutation modestly increased the Michaelis-Menten constant for DHS (K69A = 76 μ M, wild-type = 29 μ M).²⁰ In the present work, the complex structures of DHS in K69A and D105N *Mtb*SDH were modeled by MD simulations to visualize the influence of these mutations on DHS binding. Our results showed that the MM-GBSA free energies of DHS in the wild-type, K69A and D105N *Mtb*SDH, were comparable with values of -26.4 ± 3.6 , -24.5 ± 4.7 and -21.3 ± 4.8 kcal/mol, respectively. This indicated that K69A and D105N mutations of *Mtb*SDH did not greatly influence the binding of the DHS substrate. Residues that formed important hydrogen bonds found in the wild-type DHS/*Mtb*SDH complex, except of Lys69, still showed a prominent contribution to DHS binding in K69A and D105N *Mtb*SDH. The decomposition energies found in DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH complexes were comparable to those found in the wild type DHS/*Mtb*SDH complex (Figure 4). This implies that the remaining hydrogen bonds helped to retain the binding of DHS in K69A and D105N *Mtb*SDH. Although, K69A and D105N mutations did not greatly influence DHS binding, these mutations significantly changed the position of Lys69 for the catalytic reaction of *Mtb*SDH. In the wild-type *Mtb*SDH, Lys69 was located in a suitable position for proton transfer during the catalytic reaction. The average distance between the carbonyl oxygen at the C5 atom of DHS and the nitrogen atom in the ϵ -ammonium group ($-\text{NH}_3^+$) of Lys69 was 2.82 ± 0.13 Å (Figure 5). This distance is close to that found in the crystal structure of the

SKM/*Mtb*SDH complex (PDB code 4P4G) between the -NH_3^+ of Lys69 and the hydroxyl oxygen of the C5 atom of SKM (2.94 Å). The K69A mutation in *Mtb*SDH resulted a missing ϵ -ammonium group that obviously decreased the catalytic constant value for *Mtb*SDH.²⁰ The D105N mutation in *Mtb*SDH certainly broke the salt bridge created between Lys69 and Asp105. Since the salt bridge was broken, the ϵ ammonium group of Lys69 was far from the carbonyl oxygen at the C5 atom of DHS (Figure 5). The average distance between the carbonyl oxygen of the C5 atom of DHS and the ϵ ammonium group of Lys69 was 4.45 ± 1.09 Å. This distance is longer than that found in wild type *Mtb*SDH, so proton transfer might not have been supported during the catalytic reaction of *Mtb*SDH. This indicates that Asp105 plays a crucial rule to fix the ϵ -ammonium group of Lys69 in a suitable position for the catalytic reaction of *Mtb*SDH via a salt bridge. Accordingly, mutation of Asp105 in *Mtb*SDH might result the loss of enzyme activity.

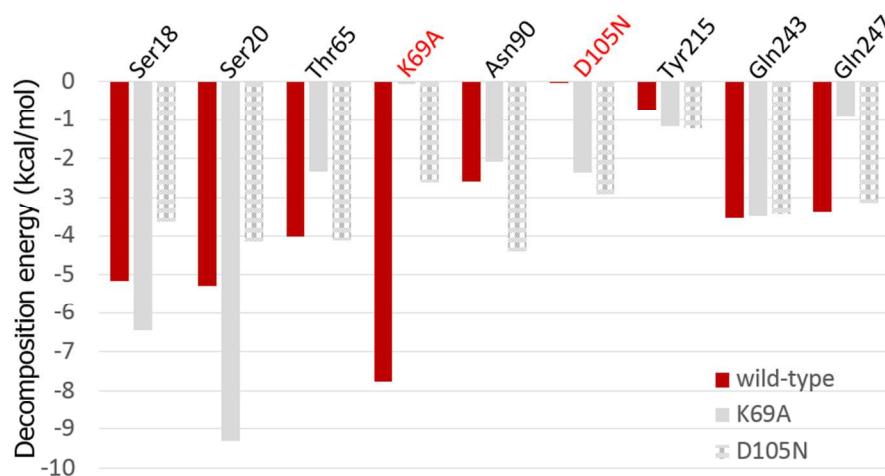


Figure 4. Comparison of the MM-GBSA decomposition energies found in the complex structures of DHS in the wild-type, K69A and D105N *Mtb*SDH. Residues that formed important hydrogen bonds found in the wild-type DHS/*Mtb*SDH complex were selected for comparison.

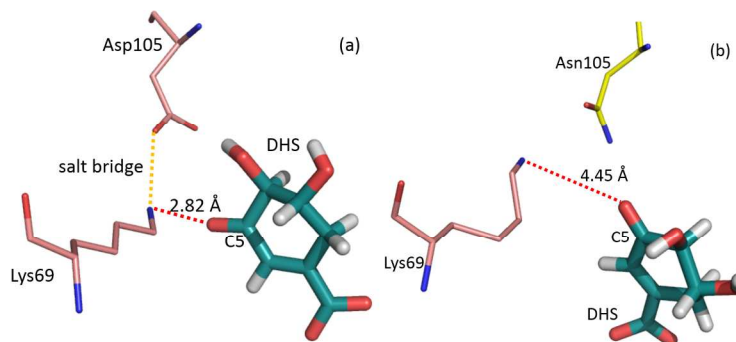


Figure 5. The average distances between the carbonyl oxygen at the C5 atom of DHS and the nitrogen atom of -NH_3^+ of Lys69 found in the wild-type (a) and D105N (b) *MtbSDH*. These distances were averaged over the equilibrium state of each simulation system. D105N mutation broke the salt bridge created between Lys69 and Asp105 (a) and resulted in the long distance between Lys69 and DHS (b).

Ternary DHS/NADPH/*MtbSDH* complex structure. DHS substrate and NADPH cofactor were bound in a deep groove between the *N*-terminal and *C*-terminal domains. The binding of NADPH did not significantly change the binding position of the DHS substrate in *MtbSDH* (Figure 6). The DHS substrate was located close to the nicotinamide ring of NADPH with a *proS* conformation. The average distance between the *proS* hydrogen of NADPH and the C5 atom of DHS was 3.95 ± 0.37 Å (Figure 7). This coordinate is consistent with a *proS* hydride transfer from NADPH to the C5 atom of DHS in the oxy-reduction reaction catalyzed by *MtbSDH*.¹⁷ The NH_3^+ of Lys69 and the C5 carbonyl oxygen of DHS were separated by a distance of 2.89 ± 0.13 Å (Figure 7), which is consistent proton transfer between them. The negatively charged groups of Asp105 and NADPH sandwiched the NH_3^+ of Lys69 (Figure 8), which may have facilitated proton transfer by these negative charges. The coordinates of NADPH, DHS and Lys69 found in our work (Figure 7) are in the agreement with the concerted mechanism of *MtbSD*, both hydride and proton transfer taking place in the same step.¹⁷

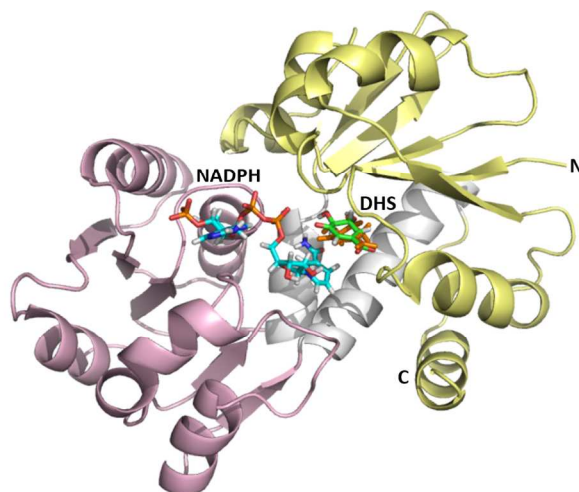


Figure 6. Superposition of the ternary DHS/NADPH/*Mtb*SDH complex structure obtained from MD simulations onto DHS (orange) found in the binary DHS/*Mtb*SDH complex. *N*-terminal (residues 5-103 and 254-269) and C-terminal (residues 114-235) domains are shown in yellow and purple, respectively. Two α -helix linkers (residues 104-113 and 236-253) are shown in gray.

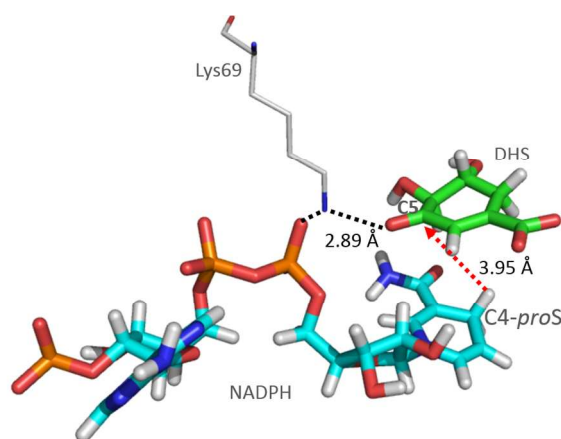


Figure 7. The position of NADPH, DHS and Lys69 found in the ternary DHS/NADPH/*Mtb*SDH complex structure.

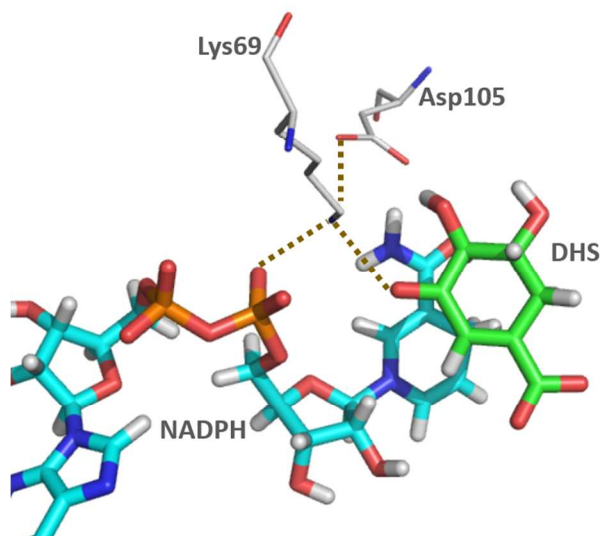


Figure 8. Interactions of the NH_3^+ group of Lys69 with NADPH, DHS and Asp105 in the ternary DHS/NADPH/*Mtb*SDH complex structure.

NADPH binding in *Mtb*SDH. Two charged residues of Arg149 and Lys153 create an “electrostatic clamp” to bind the adenine phosphate of NADPH in DHS/NADPH/*Mtb*SDH (Figure 9). The electrostatic clamp formed by the charged residues plays a crucial rule in the adenine phosphate binding of NADP^+ in the SDH of *E. coli* and *T. thermophilus* HB8.^{13,19} The pyrophosphate of NADPH interacts the diphosphate-binding loop (Gly124-Ala129), which is conserved in the SDH family.¹⁹ It forms hydrogen bonds with the NH backbone and the hydroxyl sidechain of Thr128 and the NH backbone of Gly127. The NH backbones of Ser125 and Gly126 create two hydrogen bonds to the 3'-hydroxyl group of the adenosine ribose of NADPH (Figure 9). The interactions formed by the diphosphate-binding loop are very similar to those found in the NADP^+ binding in SDH of *E. coli* and *T. thermophilus* HB8 (PDB code 1NYT and 2EV9, respectively).

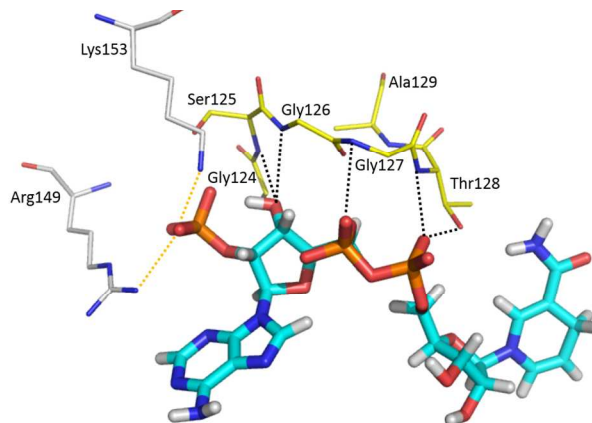


Figure 9. Interactions of NADPH with electrostatic clamp residues (carbon atoms are presented in grey) and the diphosphate-binding loop (carbon atoms are yellow).

RMSF of *Mtb*SDH. RMSF values visualizing the flexibility of apo *Mtb*SDH residues were in the range of 1.20–6.00 Å except for that of Asp105. This residue had the lowest RMSF of 0.91 Å. Apo *Mtb*SDH structures obtained from the crystal structure (PDB code 4P4N) and our MD simulation showed that the negative side chain of Asp105 was coupled by the positive side chain of Lys69 to form a salt bridge. Therefore, this salt bridge was responsible for the low flexibility of Asp105. RMSF values of residues for DHS binding were in range of 0.91–3.30 Å. The binding pocket of NADPH showed more flexible than that of DHS with RMSF values in range of 1.20–5.15 Å (Figure 10). Particularly, the electrostatic clamp residues, Arg149 and Lys153, had high flexibility with RMSF values of 5.15 and 4.12 Å, respectively. The high flexibility of these residues might serve to capture the adenine phosphate of NADPH.



Figure 10. RMSF visualization of apo *MtbSDH* residues on DHS/NADPH/*MtbSDH* obtained from MD simulation. RMSF values were calculated for all atoms of each residue in apo *MtbSDH* over the last 20 ns. RMSF values are illustrated in the range of 1.2-1.99 Å (grey), 2.01-3.99 Å (yellow) and 4.01-6.00 Å (red).

Structural comparison of *MtbSDH*. The conformation of SDH from *T. thermophilus* HB8 in the binary and ternary complex structures are very similar, indicated that there are no large structural changes associated with cofactor binding.¹⁹ Similarly, small conformational changes induced by NADP binding were observed in *MtbSDH*.³⁶ In agreement with our results, the superposition of the alpha carbon backbone of apo *MtbSDH* onto the ternary DHS/NADPH/*MtbSDH* complex showed a small RMSD value of 1.30 Å. The RMSD value of each *MtbSDH* residue was in the range of 0.22–3.69 Å, except for that of His17. It had the highest RMSD value of 5.81 Å (Figure 11). The position of each residue in the substrate binding pocket was not significantly changed upon DHS binding as evaluated by RMSD values in range of 0.48-1.93 Å (Figure 11). Two catalytic residues (Lys69 and Asp105) had small RMSD values of 1.22 and 1.27 Å, respectively. These results suggest that the size of the substrate binding pocket in *MtbSDH* is highly specific for its substrate and the catalytic residue is fixed in a

suitable position for the catalytic reaction. In the cofactor binding pocket, the RMSD value of each residue was in the range of 0.63-2.75 Å. Met66, Gly127, Arg149, Val196 and Met239 responsible for binding of pyrophosphate and adenine phosphate of NADPH had significant conformational changes with RMSD values in the range of 2.25-2.75 Å (Figure 11). The remaining residues had no significant conformational changes as indicated by RMSD values in the range of 0.63-1.91 Å.

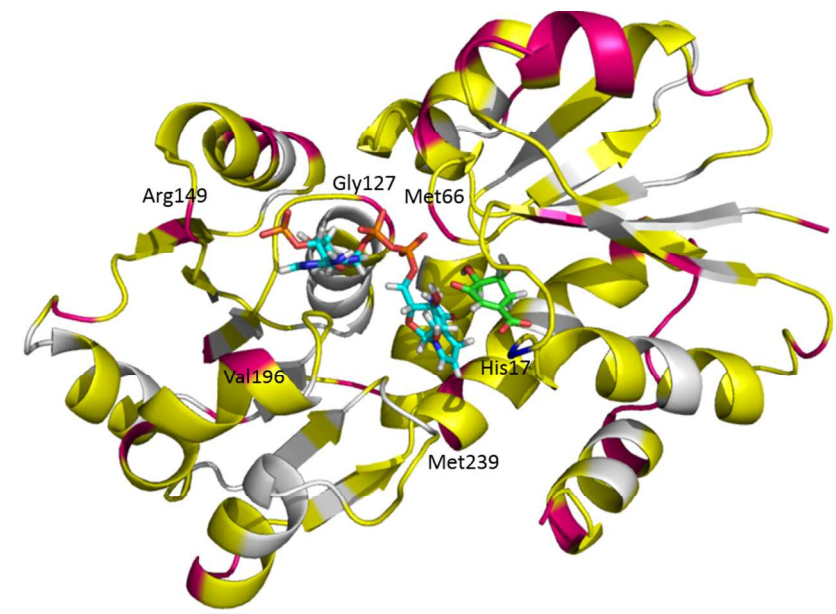


Figure 11. RMSD visualization of apo *MtbSDH* residues on DHS/NADPH/*MtbSDH* obtained from MD simulation. RMSD values were calculated between the heavy atom position of each residue in apo *MtbSDH* superpositioned onto DHS/NADPH/*MtbSDH*. RMSD values of residue were classified in the range of 0.22-1.00 Å (grey), 1.01-1.99 Å (yellow) and 2.01-3.69 Å (pink).

Rational design guidelines for hybrid *MtbSDH* inhibitors. The pocket for DHS binding, as well as the adjacent pocket for the nicotinamide ring and ribose, seemed to be specific as evidenced by their small RMSF and RMSD values as detailed above. An inhibitor that could

bind in both DHS and NADPH pockets and did not induce the conformational change of *Mtb*SDH should be a promising *Mtb*SDH inhibitor. Here, it is called a hybrid inhibitor that mimics the binding of DHS and the nicotinamide ring and ribose of NADPH. The head and tail of the hybrid inhibitor were planned so that they bind in the DHS pocket and the ribose and nicotinamide pocket (NADPH pocket), respectively (Figure 12). The carboxylate and oxygen carbonyl groups should be presented at the head of the hybrid inhibitor to mimic DHS binding. The carboxylate group is important for DHS binding in *Mtb*SDH. It formed four hydrogen bonds with Ser18, Ser20, Thr65 and Tyr215 (Figure 2). However, the negatively charged carboxylate group of some compounds impeded their accumulation inside bacterial cells.³⁷⁻³⁸ Therefore, replacement of the negatively charged carboxylate group with a neutral hydrogen bond acceptor should balance the activity against both *M. tuberculosis* and *Mtb*SDH. The tail of the hybrid inhibitor should be attached to head via a linker at a position located between the carboxylate and oxygen carbonyl groups. This position facilitates the binding of the hybrid inhibitor in the NADPH binding pocket. It is possible to present a hydrophobic moiety at the tail of the hybrid inhibitor due to the surrounding of hydrophobic sidechains of Met66, Ala213, Ile214, Tyr215 and Met239 (Figure 12).

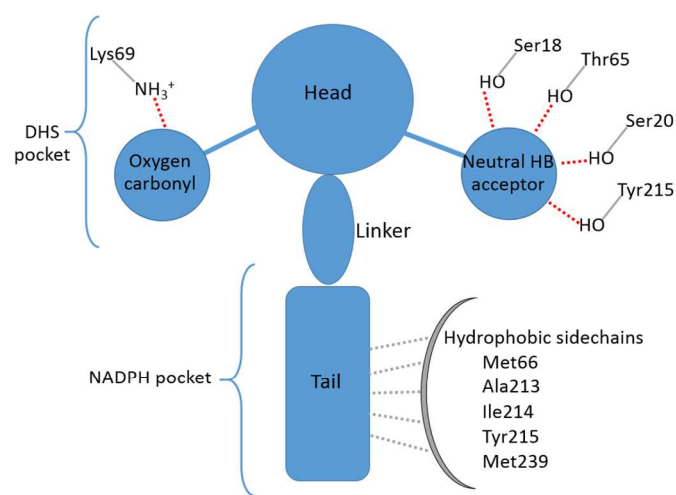


Figure 12. The skeleton of a proposed hybrid inhibitor designed based on the binding of DHS substrate, nicotinamide and ribose of NADPH found in ternary DHS/NADPH/*Mtb*SDH complex structure. Red and grey dotted lines respectively indicate the possible hydrogen bonds (HB) and hydrophobic interactions.

CONCLUSION

Binary DHS/*Mtb*SDH and ternary DHS/NADPH/*Mtb*SDH structures were determined using MD simulations. Furthermore, two mutant *Mtb*SDH complex structures, including ones for DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH, were investigated to visualize the influence of these mutations on DHS binding. Structural analysis of *Mtb*SDH complex structures in terms of binding interaction, flexibility and positional change yielded insights into substrate and cofactor binding in *Mtb*SDH and a structural framework for rational drug design. The skeleton of a hybrid inhibitor planned to bind both in DHS and NADPH pockets has been elucidated here.

ASSOCIATED CONTENT

Supporting Information. RMSD plots of five *Mtb*SDH systems: apo *Mtb*SDH, binary DHS/*Mtb*SDH, ternary DHS/NADPH/*Mtb*SDH, DHS/K69A *Mtb*SDH and DHS/D105N *Mtb*SDH.

AUTHOR INFORMATION

Corresponding Author

*E-mail: auradee.punkvang@npu.ac.th (Auradee Punkvang)

ACKNOWLEDGMENT

This research was supported by the Thailand Research Fund (MRG5980152). Faculty of Science, Ubon Ratchathani University and the University of Bristol are gratefully acknowledged for computational resource support of this research. We thank EPSRC for funding via BristolBridge (grant number EP/M027546/1).

REFERENCES

- (1) World Health Organization (WHO). 2016. Global tuberculosis report 2017. Available at http://www.who.int/tb/publications/global_report/en/.
- (2) World Health Organization (WHO). 2010. Multidrug and Extensively Drug-Resistant TB (M/XDR-TB):2010 Global Report on Surveillance and Response. Available at http://www.who.int/tb/features_archive/m_xdrtb_facts/en/.
- (3) Velayati, A.A.; Farnia, P.; Masjedi, M.R.; Ibrahim, T.A.; Tabarsi, P.; Haroun, R.Z.; Kuan, H.O.; Ghanavi, J.; Farnia, P.; Varahram, M. Totally Drug-resistant Tuberculosis Strains: Evidence of Adaptation at the Cellular Level. *Eur. Respir. J.* **2009**, *34*, 1202-1203.
- (4) Udawadia, Z.F.; Amale, R.A.; Ajbani, K.K.; Rodrigues, C. Totally Drug-resistant Tuberculosis in India. *Clin. Infect. Dis.* **2012**, *54*, 579-581.
- (5) Roberts, F.; Roberts, C.W.; Johnson, J.J.; Kyle, D.E.; Krell, T.; Coggins, J.R.; Coombs, G.H.; Milhous, W.K.; Tzipori, S.; Ferguson, D.J.; Chakrabarti, D.; McLeod, R. Evidence for the Shikimate Pathway in Apicomplexan Parasites. *Nature* **1998**, *393*, 801-805.

- (6) Keeling, P.J.; Palmer, J.D.; Donald, R.G.; Roos, D.S.; Waller, R.F.; McFadden, G.I. Shikimate Pathway in Apicomplexan Parasites. *Nature* **1999**, *397*, 219-220.
- (7) Bentley, R. The Shikimate Pathway--A Metabolic Tree with Many Branches. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 307-384.
- (8) Ducati, R.G.; Basso, L.A.; Santos, D.S. Mycobacterial Shikimate Pathway Enzymes as Targets for Drug Design. *Curr. Drug Targets* **2007**, *8*, 423-435.
- (9) Sassetti, C.M.; Boyd, D.H.; Rubin, E.J. Genes Required for Mycobacterial Growth Defined by High Density Mutagenesis. *Mol. Microbiol.* **2003**, *48*, 77-84.
- (10) Dosselaere, F.; Vanderleyden, J. A Metabolic Node in Action: Chorismate-utilizing Enzymes in Microorganisms. *Crit. Rev. Microbiol.* **2001**, *27*, 75-131.
- (11) Padyana, A.K.; Burley, S.K. Crystal Structure of Shikimate 5-dehydrogenase (SDH) Bound to NADP: Insights into Function and Evolution. *Structure* **2003**, *11*, 1005-10013
- (12) Ye, S.; Von Delft, F.; Brooun, A.; Knuth, M.W.; Swanson, R.V.; McRee, D.E. The Crystal Structure of Shikimate Dehydrogenase (AroE) Reveals a Unique NADPH Binding Mode. *J. Bacteriol.* **2003**, *185*, 4144-4151.
- (13) Michel, G.; Roszak, A.W.; Sauve, V.; Maclean, J.; Matte, A.; Coggins, J.R.; Cygler, M.; Lapthorn, A.J. Structures of Shikimate Dehydrogenase AroE and Its Paralog YdiB. A Common Structural Framework for Different Activities. *J. Biol. Chem.* **2003**, *278*, 19463-19472.

(14) Singh, S.; Korolev, S.; Koroleva, O.; Zarembinski, T.; Collart, F.; Joachimiak, A.; Christendat, D. Crystal Structure of a Novel Shikimate Dehydrogenase from *Haemophilus influenzae*. *J. Biol. Chem.* **2005**, *280*, 17101–17108.

(15) Peek, J.; Lee, J.; Hu, S.; Senisterra, G.; Christendat, D. Structural and Mechanistic Analysis of a Novel Class of Shikimate Dehydrogenases: Evidence for a Conserved Catalytic Mechanism in the Shikimate Dehydrogenase Family. *Biochemistry* **2011**, *50*, 8616–8627.

(16) Peek, J.; Garcia, C.; Lee, J.; Christendat, D. Insights into the Function of RifI2: Structural and Biochemical Investigation of a New Shikimate Dehydrogenase Family Protein. *Biochim. Biophys. Acta* **2013**, *1834*(2), 516–523.

(17) Fonseca, I.O.; Silva, R.G.; Fernandes, C.L.; de Souza, O.N.; Basso, L.A.; Santos, D.S.; Arch. Biochem. Biophys. Kinetic and Chemical Mechanisms of Shikimate Dehydrogenase from *Mycobacterium tuberculosis*. **2007**, *457*(2), 123–133.

(18) Dansette, P.; Azerad, R. The Shikimate Pathway: II. Stereospecificity of Hydrogen Transfer Catalyzed by NADPH-dehydroshikimate Reductase of *E. coli*. *Biochimie* **1974**, *56*(5), 751–755.

(19) Bagautdinov, B.; Kunishima, N. Crystal Structures of Shikimate Dehydrogenase AroE from *Thermus thermophilus* HB8 and Its Cofactor and Substrate Complexes: Insights into the Enzymatic Mechanism. *J. Mol. Biol.* **2007**, *373*(2), 424–438.

(20) Rodrigues, V.S. Jr.; Breda, A.; Santos, D.S.; Basso, L.A. The Conserved Lysine69 Residue Plays a Catalytic Role in *Mycobacterium tuberculosis* Shikimate Dehydrogenase. *BMC Res Notes*. **2009**, *2*, 227.

(21) Morris, G.M.; Goodshell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. Automated Docking using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639-1662.

(22) Case, D.A.; Betz, R.M.; Cerutti, D.S.; Cheatham, III, T.E.; Darden, T.A.; Duke, R.E.; Giese, T.J.; Gohlke, H.; Goetz, A.W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T.S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K.M.; Monard, G.; Nguyen, H.; Nguyen, H.T.; Omelyan, I.; Onufriev, A.; Roe, D.R.; Roitberg, A.; Sagui, C.; Simmerling, C.L.; Botello-Smith, W.M.; Swails, J.; Walker, R.C.; Wang, J.; Wolf, R.M.; Wu, X.; Xiao, L.; Kollman, P.A. AMBER 2016, University of California, San Francisco, CA, 2016.

(23) Pearlman, D.A.; Case, D.A.; Caldwell, J.W.; Ross, W.S.; Cheatham, III, T.E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, A Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comp. Phys. Commun.* **1995**, *91*, 1-41.

(24) Case, D.A.; Cheatham, T.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K.M. Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668-1688.

(25) Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and Testing of A General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157-1174.

(26) Cieplak, P.; Cornell, W.D.; Bayly, C.; Kollman, P.A. Application of the Multimolecule and Multiconformational RESP Methodology to Biopolymers: Charge Derivation for DNA, RNA, and Proteins. *J. Comput. Chem.* **1995**, *16*, 1357-1377.

(27) Fox, T.; Kollman, P.A. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B.* **1998**, *102*, 8070-8079.

(28) Bayly, C.I.; Cieplak, P.; Cornell, W.D.; Kollman, P.A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269-10280.

(29) Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.

(30) Darden, T.; York, D.; Pedersen, L.J. Particle Mesh Ewald-an NLog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.

(31) Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C. Numerical Integration of the Cartesian Equations of Motion of A System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327-341.

(32) Roe, D. R.; Cheatham, T. III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.

(33) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.*, **2012**, 8, 3314–3321.

(34) Case, D.A.; Darden, T.A.; Cheatham, T.E., III; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Walker, R.C.; Zhang, W.; Merz, K.M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A.W.; Kolossváry, I.; Wong, K.F.; Paesani, F.; Vanicek, J.; Wolf, R.M.; Liu, J.; Wu, X.; Brozell, S.R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D.R.; Mathews, D.H.; Seetin, M.G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P.A. AMBER 12, University of California, San Francisco, CA, 2012.

(35) Singh, S.A.; Christendat, D. Structure of Arabidopsis dehydroquinase Dehydratase-Shikimate Dehydrogenase and Implications for Metabolic Channeling in the Shikimate Pathway. *Biochemistry* **2006**, 45(25), 7787-7796.

(36) Arcuri, H.A.; Borges, J.C.; Fonseca, I.O.; Pereira, J.H.; Neto, J.R.; Basso, L.A.; Santos, D.S.; de Azevedo, W.F. Jr. Structural Studies of Shikimate 5-Dehydrogenase from *Mycobacterium tuberculosis*. *Proteins* **2008**, 72(2), 720-730.

(37) Richter, M.F.; Drown, B.S.; Riley, A.P.; Garcia, A.; Shirai, T.; Svec, R.L.; Hergenrother, P.J. Predictive Compound Accumulation Rules Yield a Broad-Spectrum Antibiotic. *Nature* **2017**, 545, 299-304.

(38) Tizón, L.; Otero, J.M.; Prazeres, V.F.; Llamas-Saiz, A.L.; Fox, G.C.; van Raaij, M.J.; Lamb, H.; Hawkins, A.R.; Ainsa, J.A.; Castedo, L.; González-Bello, C. A Prodrug Approach for

1
2
3 Improving Antituberculosis Activity of Potent *Mycobacterium tuberculosis* Type II
4
5 Dehydroquinase Inhibitors. *J. Med. Chem.* **2011**, *54*, 6063-6084.
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60