Structural investigation of inhibitor designs targeting 3-dehydroquinate dehydratase from the shikimate pathway of Mycobacterium tuberculosis.

Marcio V.B. Dias, William Snee, Karen M Bromfield, Richard Payne, Satheesh K Palaninathan, Alessio Ciulli, Nigel I Howard, Chris Abell, James C Sacchettini, Tom L Blundell

Structural investigation of inhibitor designs targeting 3-dehydroquinate dehydratase from the shikimate pathway of *Mycobacterium tuberculosis*.

Marcio V. B. Dias[*1], William C. Snee[†1], Karen M. Bromfield[‡], Richard J. Payne[‡§], Satheesh K. Palaninathan[†], Alessio Ciulli[‡], Nigel I. Howard[‡], Chris Abell[‡], James C. Sacchettini[†2], and Tom L. Blundell[*2]

From Department of Biochemistry[*], University Chemical Laboratory[‡], University of Cambridge, UK and Department of Biochemistry and Biophysics[†], Texas A&M University, USA, [§]Current address: School of Chemistry, The University of Sidney, Australia

Running head: Inhibitors of type II 3-dehydroquinate dehydratase

[1]These authors contributed equally

[2]Address correspondence to: Tom L. Blundell, PhD, Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge, CB2 1GA, UK. Tel.: +44 1223 333628, E-mail: tom@cryst.bioc.cam.ac.uk and James C. Sacchettini, PhD, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA. Tel.: 979-862-7636, E-mail: sacchett@tamu.edu

The abbreviations used are: DHQase, 3-dehydroquinate dehydratase; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; Hepes, 4-(hydroxyethyl)-1-piperazineethanesulfonic acid; HpDHQase, 3-dehydroquinate dehydratase from *Helicobacter pylori*; ITC, isothermal titration calorimetry; IPTG, isopropyl β-D-1-thiogalactopyranoside; KBr, potassium bromide; NMR, nuclear magnetic resonance; MtDHQase, 3-dehydroquinate dehydratase from *Mycobacterium tuberculosis*; PDB, protein data bank; PEG, polyethylene glycol; ScDHQase, 3-dehydroquinate dehydratase from *Streptocymes coelicolor*; STD, saturation transfer difference; Tris, (hydroxymethyl) aminomethane;

Key words: dehydroquinate dehydratase, shikimate pathway, *Mycobacterium tuberculosis*, drug discovery, transition state analogues, crystal structure.

The shikimate pathway is essential in *Mycobacterium tuberculosis* and its absence in humans makes the enzymes of this pathway potential drug targets. In this report, we provide structural insights into ligand and inhibitor binding to 3-dehydroquinate dehydratase (dehydroquinase) from *Mycobacterium tuberculosis* (MtDHQase), the third enzyme of the shikimate pathway. The enzyme has been crystallized in complex with its reaction product, 3-dehydroshikimate, and with six different competitive inhibitors. The inhibitor 2,3-anhydroquinate mimics the flattened enol/enolate reaction intermediate and serves as an anchor molecule for four of the inhibitors investigated. MtDHQase also forms a complex with citrazinic acid, a planar analog of the reaction product. The structure of MtDHQase in complex with a 2,3-anhydroquinate moiety attached to a biaryl group shows that this group extends to an active site subpocket inducing significant structural re-arrangement. The flexible extensions of inhibitors designed to form π-stacking interactions with the catalytic Tyr24 residue have been investigated. The high resolution crystal structures of the MtDHQase complexes provide structural evidence for the role of the loop residues 19-24 in MtDHQase ligand binding and catalytic mechanism and provide rationale for the design and efficacy of inhibitors.

One strategy for the development of new therapies against pathogenic bacteria such as *Mycobacterium tuberculosis* is to target essential biosynthetic pathways of the microorganism that are absent in humans, such as the shikimate pathway. The shikimate pathway produces chorismate, an important precursor of aromatic compounds in bacteria, fungi, plants, and apicomplexan parasites [1-4]. This pathway comprises seven different enzymes, each of which catalyzes a separate step of the pathway that converts erythrose-4-phosphate and phosphoenol pyruvate to chorismate [1,5]. Chorismate is the substrate for five branching pathways involved in the production of menaquinones, siderophores, aromatic amino acids (phenylalanine, tyrosine, and tryptophan), vitamins E and K, *p*-aminobenzoic acid, and other aromatic compounds [3]. Studies of gene disruption have shown that the shikimate pathway is essential for growth of *Mycobacterium tuberculosis* [6].

The third enzyme of the shikimate pathway is 3-dehydroquinate dehydratase (dehydroquinase, DHQase), which is the product of the *aroD* gene. There are two forms of DHQase (type I and type II) with different structures and mechanisms, both of which catalyze the reversible conversion of 3-dehydroquinate to 3-dehydroshikimate [7]. DHQase functions in two metabolic pathways: the catabolic quinate pathway, enabling certain organisms to convert quinate to protocatechuate for use as a carbon source, and the biosynthetic shikimate pathway [8-10]. Type I DHQase exists as a heat-labile homodimer and is involved only in the biosynthetic pathway. Its mechanism involves a covalent iminimium intermediate to catalyze the dehydration of 3-dehydroquinate through a *syn* elimination [11]. Type II DHQase exists as a heat-stable homododecamer that can function in both biosynthetic and catabolic pathways, probably involving dehydration through an enol/enolate intermediate (Figure 1) [5]. The catalysis of type II DHQase involves an *anti* elimination of water processed through a stepwise E_1CB mechanism. A base-catalyzed abstraction of the axial proton at C2 is required for type II DHQase to form the enolate intermediate and in a second step the hydroxyl group of the C1 position is removed by acid catalysis to form the product. A conserved tyrosine of active site is responsible to remove the *pro*-S proton in the first step of the reaction and also a conversed histidine finalizes the reaction acting as proton donor in the acid-catalyzed reaction (figure 1) [7].

Mycobacterium tuberculosis uses the type II DHQase in a biosynthetic role. The crystallographic structures of type II DHQase from *M. tuberculosis* (MtDHQase) [7] and other bacteria [12, 13] reveal a dodecamer formed from a tetramer of trimers with 23 symmetry and the trimer being the minimal active oligomeric form [7]. The parallel β-sheet of each subunit has strands in the following order: 2, 1, 3, 4, and 5, with two α-helices on each side of the sheet as in the flavodoxin type α/β fold [7,14]. The active site in type II DHQases is located in a cleft formed near the carboxy-terminal regions of strands β1 and β3 of the parallel β-sheet, which is common among proteins with the flavodoxin-like fold [14]. Two of the key residues, Arg19 and Tyr24 in MtDHQase, are conserved in type II DHQases from other organisms (Supplementary Figure 1, Supplementary Figure 2a-b, Arg23 and Tyr28 in *Streptomyces coelicolor* ScDHQase, and Arg17 and Tyr22 in *Helicobacter pylori* HpDHQase) and have been identified by chemical modification and site-directed mutagenesis as being essential for DHQase enzyme activity [15]. Both of these residues are located on a flexible loop (residues 19-24 in MtDHQase) that is completely disordered in the *apo*-DHQase structure (PDB accession code 2DHQ) [7]. Structural studies of ScDHQase [12,16-18] and HpDHQase [13,19-20] were able to capture the full loop region in the presence of ligands and demonstrate that ligand binding causes the flexible loop residues to form a lid that closes over the active site.

Structures of MtDHQase have been solved with and without ligands, including structures with the reaction intermediate analogues 2,3-anhydroquinate (**2**; PDB accession code: 1H0R, unpublished) and 3-hydroxyimino-quinic acid (PDB accession code: 1H0S, unpublished), but only recently has the full closure of the flexible catalytic loop residues 19-24 been reported [20].

This flexible loop region is essential for substrate binding and catalysis, and has been an important element for DHQase inhibitor design. More importantly, this region plays a major role in the formation of a subpocket located adjacent to the enzyme active site of ScDHQase. The formation of this subpocket greatly increased the druggable space for ScDHQase and was the basis for the design of several nM inhibitors [17-19, 21-27]. The inhibitors designed to target this subpocket of ScDHQase also showed potent inhibition of MtDHQase and many of them show ~10-fold or greater differences in inhibition between the two enzymes [19, 21, 22, 24]. These studies strongly suggest a general

conservation of features between the active sites of these two enzymes, but also that there are structural differences between them that need to be characterized.

In this paper we present structures of MtDHQase that show all the residues in the active site, including the flexible loop residues 19-24, in complex with the enzymatic reaction product, 3-dehydroshikimate, and six competitive inhibitors. The inhibitors include 2,3-anhydroquinate, an analogue based on the enol/enolate reaction intermediate, four compounds that extend from the 2,3-anhydroquinate template, and citrazinic acid, whose planarity represents a novel class of MtDHQase inhibitor. The binary complex structures presented herein provide structural details for the lid closure mechanism in MtDHQase, and also provide structural evidence and characterization of a subpocket located adjacent to the MtDHQase active site and its influence on MtDHQase inhibitor binding.

EXPERIMENTAL PROCEDURES

Cloning, overexpression and purification of MtDHQase

The *aroD* gene from *Mycobacterium tuberculosis* (Rv2537c; UniProtKB/Swiss-Prot P0A4Z6 (AROQ_MYCTU)) was cloned into either the pET28a with a thrombin cleavage site or the pET28b vector with a TEV cleavage site (modified from the original) after the N-terminal hexahistidine tag. Both plasmids containing the *aroD* gene were transformed in to BL21 (DE3) competent cells (Novagen) by heat shock. The cells were grown in LB or 2YT medium at 37 ºC until OD = 0.6 and then induced with 1 mM IPTG for 4-6 hours at 37 ºC. The cells were then harvested by centrifugation at 4000 rpm for 30 minutes. The pellet was suspended in approximately 30 mL of buffer composed of 50 mM Tris, 250 mM NaCl, pH 7.8 and was supplemented with one Roche Inhibitor Cocktail tablet (EDTA free). The cells were then lysed by either sonication or French press and clarified by centrifugation at 21,000 *g* for 40 minutes. The 0.2 micron filtered supernatant was loaded onto a nickel column and the pure His-tagged protein was eluted with a gradient concentration from 20 to 500mM imidazole. MtDHQase structures with inhibitors **1** and **6-7** were obtained from purified His-tagged protein expressed using the pET28b vector. MtDHQase expressed using the pET28a vector was used for structures with inhibitors **2-5**. The crystals for the complexes MtDHQase with inhibitors **2-5** were obtained after cleavage of His-tag. To remove the His-tag, MtDHQase was incubated overnight at 4°C with thrombin (restriction grade; Invitrogen). The tag-free MtDHQase was further purified by size-exclusion chromatography using a Superdex 75 26/60 column (GE Healthcare) with a buffer containing 50 mM Tris, 250 mM NaCl, pH 7.8.

MtDHQase ligands

The ligands (Figure 2) are **1** (3-dehydroshikimate); **2** (2,3 anhydroquinate); **3** ((1*S*,4*R*,5*R*)-3-(3-benzoylphenyl)-1,4,5-trihydroxycyclohex-2-enecarboxylic acid); **4** (1*R*,4*R,5R*)-1,4,5-trihydroxyl-3-(2-phenylcarbamoyl-vinyl)-cyclo-hex-2-enecarboxylic acid); **5** ((1*R*,4*R*,5*R*)-1,4,5-trihydroxy-3-(3-phenoxy-prop-(*E*)-enyl)-cyclohex-2-enecarboxylic acid); **6** ((1*R*,4*R*,5*R*)-3-(tert-butylcarbamoyl)-1,4,5-trihydroxycyclohex-2-enecarboxylic acid); **7** (citrazinic acid represented in both tautomeric forms). The syntheses of inhibitors **2-6** were performed according to previously published protocols [19,21,22,26]. Inhibitor **7** was purchased from Sigma Aldrich (St. Louis, MO, USA).

Crystallization with ligands, data collection, and processing

MtDHQase expressed using the pET28a vector was concentrated to 10 or 40 mg/mL in a buffer composed of 50 mM Tris, 250 mM NaCl, pH 7.8 and MtDHQase expressed using the pET28b vector was concentrated to 10 mg/mL in a buffer composed of 50 mM Tris, 100 mM NaCl, 1mM EDTA, 1mM DTT, pH 7.5 using Amicon Ultra MWCO 5 KDa protein concentrators. MtDHQase in complex with different compounds was crystallized using different conditions and techniques due to the difficulty of reproducing the crystals. The crystallization protocols and data collection/processing details for each crystal structure are described in Supplementary Table 1.

Structure determination and refinement

The MtDHQase:**1** structure was solved by molecular replacement using MOLREP [28] implemented in CCP4 [29] with the protein atomic coordinates for MtDHQase from the Protein Data Bank (accession code 2DHQ). Non-protein atoms were removed from MtDHQase:**1,** which was then used as a molecular replacement model for structures MtDHQase:**2-4, 6, 7**. The MtDHQase:**5** structure was solved by molecular replacement using the MtDHQase PDB coordinates 1H0R. Refinement was performed using REFMAC 5.2 from the CCP4 suite [29,30]. TLS refinement was applied to structures of MtDHQase:**1** and **7** (files were generated by TLSMD server) [31]. Visual inspection and manual rebuilding were performed using the programs XtalView/Xfit [32] and Coot 0.3.1 [33]. The water molecules were added manually and automatically using the program XtalView/Xfit [32] or with Coot 0.3.1 [33] and checked on the basis of B-factor values. The stereochemistry was checked using PROCHECK [34]. The figures were prepared using UCSF-Chimera [35], Raster3D [36], XtalView/Xfit [32], and Pymol [37].

MtDHQase enzyme assay

The enzyme activity was assayed by monitoring the formation of the product, 3-dehydroshikimate. The initial rate of the reaction was measured by the increase in absorbance at 234 nm, from the formation of the enone-carboxylate chromophore of 3-dehydroshikimate ($\varepsilon = 1.2\times 10^4$ $M^{-1}cm^{-1}$). The assays were performed in duplicate at 25 °C in 50 mM Tris, pH 7.0. A final enzyme concentration of 102 nM was used. The assay was initiated by the addition of the substrate (3-dehydroquinate), after incubating the buffer, inhibitor, and enzyme at 25 °C for two minutes.

Kinetic parameters (K_M and k_{cat}) were obtained by measuring the initial rates of reaction over a range of substrate concentrations (typically 0.25 K_M – 10 K_M). The data were fitted to the Michaelis-Menton equation using least-squares fitting in GraFit [38]. The values of K_M and V_{max} were determined using this software, and k_{cat} was calculated from the latter value and the total enzyme concentration in the assay.

The kinetic data for inhibition studies were obtained by measuring the initial rates of reaction over a range of 4-5 inhibitor concentrations at 4-5 different substrate concentrations (0.5 K_M - 5 K_M). The inhibition constants (K_I) and the standard deviations associated with these values were determined using a least-squares fitting using GraFit [38]. GraFit was also used to carry out an F-test statistical analysis on the data, to confirm that the data satisfied a competitive inhibition model.

Isothermal titration calorimetry

ITC experiments were performed using a MicroCal VP-ITC instrument and all data were analyzed with the software implemented in Origin (version 7). In all titrations, *M. tuberculosis* His_6-DHQase was at a concentration of 80-160 μM and was buffered in 100 mM Tris, pH 8.0, 50-100 mM NaCl, in the presence of 5% v/v DMSO. Titrations of inhibitor **7** (1.6 mM) and (3*R*,5*R*,6*R*)-3,5,6-trihydroxycyclohex-1-ene-1,3-dicarboxylic acid (780 μM) were performed at 30 °C and 15 °C, respectively.

In a typical experiment, 19 injections of 15 μL were made at 4 minute intervals from a 300 μL syringe rotating at 300 rpm and loaded with inhibitor solution. In all titrations, an initial injection of 2 μL inhibitor was made and the corresponding data discarded during data analysis. Control titrations of inhibitor to buffer were performed and subtracted from the inhibitor to protein titrations. The thermodynamic parameters were obtained by fitting the data to a single-site binding model.

Saturation transfer difference (STD) NMR spectroscopy

^{1}H NMR spectroscopic experiments were performed at 278 K on a 700 MHz Bruker NMR spectrometer equipped with a 5 mm triple TXI cryoprobe with *z* gradients. STD experiments [39] employed a 40 ms selective Gaussian 180° shaped pulse at a frequency alternating between 'on resonance' (1.0 ppm) and 'off resonance' (~80 ppm) after every scan. Water suppression was achieved by using a W5 Watergate gradient spin-echo pulse sequence [40]. The resulting spectra were analysed with TopSpin. A sample of 0.35 mM citrazinic acid (**7**) was prepared in 25 mM Tris pH 7.8, 20 mM NaCl, 10% (v/v) D_2O, in the absence of protein, and presence of 10 μM His_6-DHQase to a total volume of 200 μL. The inhibitor (3*R*,5*R*,6*R*)-3,5,6-trihydroxycyclohex-1-ene-1,3-dicarboxylic acid was added to a final concentration of 400 μM to displace **7** and show specific binding in the active site. (Trimethylsilyl)-propionic acid-d_4 (TSP; 20 μM) was present for calibration purposes.

RESULTS

The high resolution crystal structures of MtDHQase in complex with the reaction product, 3-dehydroshikimate (**1**) and six competitive inhibitors (**2**-**7**), described herein (Figure 3) define key interactions associated with the active site residues including the flexible loop residues 19-24 (Figures 3-7). The inhibitor 2,3-anhydroquinate (**2**) was designed to mimic the flattened enol/enolate intermediate in the reaction mechanism and inhibitors **3-6** were evolved from **2,** extending to occupy a subpocket adjacent to the active site (Figure 3b-e), based on the ScDHQase:**2**:glycerol structure (PDB accession code 1GU1) [13]. The inhibitor, citrazinic acid (**7**) has similar substituents to the product but lacks the ring puckering. The MtDHQase:**1, 4, 6-7** and MtDHQase:**2-3** structures were solved in space groups $P2_1$ and P1 respectively and both have two dodecamers in the asymmetric unit (Table 2). No significant differences were observed in the three-dimensional structure and oligomeric state for MtDHQase in complex with these ligands compared to the previously reported *apo*-MtDHQase structure (PDB accession code 2DHQ) [8]. Table 1 provides the resolution and statistics of refinement of the structures (A complete statistics of the structures can be obtained in the Supplementary Table 2).

Ligand induced stabilization of MtDHQase catalytic loop residues

The type II DHQase catalyzed conversion of 3-dehydroquinate to 3-dehydroshikimate requires the abstraction of the *pro*-S hydrogen from the C2 carbon atom by the negatively charged phenolate group of Tyr24 in *M. tuberculosis* and the overall *anti*-elimination of water by a stepwise E_1CB mechanism resulting in the C1-C2 conjugation observed in **1** (Figure 1) [13]. The binding of **1** and substrate or analogues of substrates are assumed to stabilize the flexible active site loop residues 19-24 into a catalytic conformation (Figures 3 and 4), even though this has not been observed with previous MtDHQase binary complex structures (PDB accession codes 1H0R and 1H0S, unpublished). A recently reported MtDHQase structure (PDB accession code 2XB8) has captured the full loop region extended into a conformation similar to the crystal structures reported in this manuscript (*vide infra*) [20].

The structure of MtDHQase in complex with **1** was solved to 2.5 Å resolution, and electron density (3σ) that is presumed to belong to the catalytic reaction product, 3-dehydroshikimate (**1**) was observed in subunits C and G (Figure 4b) (the substrate was converted to product during crystallization). Electron density for the flexible active site loop residues 19-24 in subunits C and G was also visible. However, in the other subunits no discernible additional electron density for the ligand or the loop was observed.

The binding of ligand **1** is stabilized by hydrogen bonding interactions with the backbone amides of Ile102 and Ser103, with the side-chain atoms Arg19, Asn75, His81, Arg112, and Asp88* (* indicates residues from an adjacent subunit) and by hydrophobic interactions of its cyclohexene moiety with the side-chain atoms of residues Tyr24, His101, Leu13, Val105, Arg112, and the Cα's of Gly77 and Gly78 (Figure 4c and Figure 5a and Supplementary table 3). The binding of **1**, as expected, seems to induce closure of the loop residues 19-24 through its interaction with the side-chain atoms of the loop residues, Tyr24 and Arg19. The side-chain phenol of Tyr24 forms a hydrophobic interaction with the cyclohexene ring of **1** while the side-chain of Arg19 is positioned to form hydrogen bond interactions with the C3 carbonyl group of **1** (Supplementary table 3). In addition, the side-chain atoms of Arg19 appear to interact with Tyr24 through electrostatic interaction.

Loop residues 19-24 are involved in a network of water-mediated hydrogen bonds that coordinate the position of the catalytic Tyr24, orienting its phenolic oxygen atom 3.5 Å from the C2 carbon atom of **1** (Figure 4c). The Tyr24 side-chain is further positioned through hydrophobic interactions with the side-chain carbon atoms of Leu16 (3.8 Å), Val23 (4.3 Å), and Ile102 (4.3 Å) (figure not shown). Residues Arg19 and Arg108 form electrostatic and hydrogen bonding interactions with Tyr24 respectively. These interactions are presumed to stabilize the electron rich π-system of the Tyr24 side-chain. The Tyr24 phenol oxygen atom makes a hydrogen bond interaction with Arg108 Nη1 (3.1 Å) and with a water molecule (WAT1) (2.8 Å) (Figure 4c). These groups are likely to be involved in proton abstraction from Tyr24 to generate the active phenolate.

The close proximity of the Tyr24 side-chain to the C2 atom of **1** and its requirement for substrate conversion to **1** indicate that it may also play an important role in stabilizing reaction

intermediates during catalysis. Inhibitors that are mimics of reaction intermediates can be strong binders and this was the basis for the design of 2,3-anhydroquinate (**2),** which has a structure with features assumed to be in the transition state, e.g. C2-C3 conjugation. It acts as a competitive inhibitor with a K_i of 200 μM [26].

A new pocket for MtDHQase drug targeting

The structure of the MtDHQase:**2**:glycerol complex was solved to 2.0 Å resolution (Figures 3b and 6a). All subunits contained inhibitor **2** bound in the active site and a glycerol molecule bound in 13 of the 24 subunits. Inhibitor **2** is positioned in a similar mode as observed in previous structures in complex with MtDHQase and ScDHQase (PDB accession codes 1H0R and 1GU1 respectively) however this structure successfully captured the full ordering of the flexible catalytic loop residues 19-24. Hydrogen bonding interactions are observed with side-chain atoms of Arg112, His81, His101, Ser103, and Asp88* and with the main-chain atoms of Ile102 and Ser103 (Supplementary Figure 3a). The binding of inhibitor **2** is also stabilized by hydrophobic interactions similar to those observed for MtDHQase:**1** (Figure 5b).

The structure also contains a glycerol molecule, presumed to be from the cryoprotectant solution, which is bound in a subpocket adjacent to the enzyme active site as previously observed for the structure of ScDHQase**:2**:glycerol (PDB accession code 1GU1). The glycerol molecule replaces the position of the side-chain atoms of Arg19 observed in the MtDHQase:**1** structure (Supplementary Figure 3b).

The side-chain atoms of Arg19 undergo significant structural re-arrangement, rotating approximately 96° compared to the position of the Arg19 side-chain in MtDHQase:**1**, resulting in the formation of several new intra-protein interactions and causing a small shift (1.4 Å - 1.1 Å) in the backbone atoms of residues 20-22 (Supplementary Figure 3b). The new position of Arg19 forms a direct hydrogen bond between its side-chain Nη2 and Asp67* Oδ1 (3.0Å) and forms several water mediated hydrogen bonds with main-chain and side-chain atoms of Glu92* and the main-chain atoms of Arg15 and Arg18. The new position of Arg19's side-chain also makes a hydrophobic interaction with the carbon side-chain atoms of Glu92*.

Insights into inhibitor design targeting the MtDHQase active site subpocket

To investigate the importance of the MtDHQase subpocket on inhibitor binding, we have solved four structures of MtDHQase in complex with competitive type II DHQase inhibitors that were designed based on the binding of **2** in the ScDHQase-glycerol structure [13, 19, 21, 22].

We solved the 2.4 Å resolution crystal structure of MtDHQase in complex with inhibitor **3**, a biaryl derivative of inhibitor **2** containing a phenyl ring directly attached to the C3 atom of the 2,3-anhydroquinate core that is sequentially connected to a terminal phenyl ring through a rigid carbonyl linker (Figure 2 and 3c). Ligand **3** inhibits both MtDHQase and ScDHQase with a K_i value of 11 μM and 4.7 μM respectively [21]. Electron density was observed for inhibitor **3** in all subunits and for the flexible active site loop residues 19-24 in most of the subunits of the asymmetric unit with the exception of the side-chain of Arg19 in several subunits. The structure of MtDHQase:**3** reveals that its biaryl extension binds in different conformations for several MtDHQase subunits and that the overall binding mode occupies only the edge position of the glycerol molecule in MtDHQase:**2** (Figure 6b). Despite the multiple conformations observed for the terminal phenyl ring of **3**, the binding mode of the anhydroquinate core and its primary phenyl ring are conserved in all subunits and the binding of inhibitor **3** does not cause significant alterations to loop residues 19-24 compared to MtDHQase:**2**:glycerol (Figure 6b) .

The primary phenyl ring connected to the C3 atom of the 2,3-anhydroquinate core of **3** forms weak π-π stacking with Tyr24, and the terminal phenyl ring does not form additional stacking interactions (Figure 5c and 6b). The carbonyl linker of inhibitor **3** is positioned within hydrogen bonding distance of the side-chain oxygen atom of Asp88* (figure 5c). The biaryl extension of inhibitor **3** makes hydrophobic interactions with the side-chain and main-chain carbon atoms Tyr24, Asn12, Leu13, Arg19, Glu20, Glu92*, Asp88*, and Ala91* (figure 5c). The terminal phenyl ring of **3** in several subunits is positioned ~4.5 Å from the glycerol molecule of MtDHQase:**2** and occupies a hydrophobic cleft near the solvent exposed region of the active site composed by the carbon atoms of Arg19, Glu20, Glu92*, Asp88*, and Ala91* (Figure not shown).

These interactions observed in the crystal structure of MtDHQase:**3** might contribute to its increased inhibitor potency (~20 fold) compared to **2**, however the flexibility of the extension from the anhydroquinate anchor of **3** and its corresponding complementarity to the MtDHQase binding pocket, is limited due to rigidity induced by the carbonyl linkage of the two aromatic rings.. This is further supported by kinetic studies with other biaryl inhibitors from this series that show significant improvements in binding affinity using linkers with greater flexibility than **3** (Supplementary table 4) [21]. The rigidity of the biaryl extension prevents inhibitor **3** from reaching the position taken by the glycerol molecule in MtDHQase:**2** and forming stacking interactions between the terminal phenyl ring of **3** and the side-chain of Tyr24. To gain insight into the influence of the flexibility of different linkers in MtDHQase inhibitors we have solved the structures for MtDHQase in complex with compounds **4** and **5** (Figure 3d, e).

Compounds **4** and **5** closely resemble one another, the main difference being in the flexibility of the linker. Inhibitor **4** was designed to occupy the subpocket by extending the structure of **2** with a rigid olefinic amide linker attached to a terminal phenyl ring (figure 2) and has a K_i of 2.3 μM and 2.1 μM against MtDHQase and ScDHQase respectively [19]. Alternatively, inhibitor **5** was designed to have greater flexibility of its side-chain phenyl moiety. Inhibitor **5** has a phenyl ring connected via a terminal ether linkage attached to the C3 atom of inhibitor **2** yielding one of the most potent DHQase inhibitors reported to date (Figure 2, Supplementary table 4, [19]). Inhibitor **5** inhibits both MtDHQase and ScDHQase with a K_i of 0.14 μM and 0.01 μM respectively [19].

The 1.9 Å resolution crystal structures of MtDHQase:**4** and MtDHQase:**5** reveal that both phenyl rings fully occupy the position of the MtDHQase:**2** glycerol molecule (Figure 3c-d). Electron density was observed for inhibitor **4** and for all flexible active site loop residues 19-24 in all subunits of the asymmetric unit. No significant differences in binding conformation were observed for inhibitor **4** in different subunits of the two dodecamers of the asymmetric unit. The complex of MtDHQase:**5** has been crystallized in space group F23 and has only a monomer in the asymmetric unit. In this structure, electron density is observed for the active site loop.

In MtDHQase:**4**, the amide linker of inhibitor **4** rigidifies its C3 side-chain hindering the π-stacking interaction between the Tyr24 side-chain phenol and the terminal phenyl ring of **4**. The rotationally limited terminal phenyl ring of inhibitor **4** also makes hydrophobic interactions with the side-chain carbon atoms of Asn12, Arg15, Arg19, Glu20, Ala89*, and Glu92* (Figure 5d) and causes the expulsion of a water molecule previously observed to link the glycerol molecule of MtDHQase:**2** with Arg15 and Gly25 (figure not shown). The carbon linker atoms of **4** make hydrophobic interactions with the side-chain atoms of Asn12, Leu13, Tyr24, and Asp88* and the Gly77(Figure 5d). In addition, the amide linker of **4** forms hydrogen bond interactions with its carbonyl oxygen atom and the side-chain nitrogen atom of Asn12 and a water mediated hydrogen bond with Gly78 and Pro11 probably potentiating its binding affinity (Figure 5d).

In MtDHQase:**5**, the more flexible terminal phenyl ring of **5** is shifted 1.2 Å from the terminal phenyl ring of **4** positioning it further into a hydrophobic portion of the subpocket (Figure 6c). The positioning of the terminal phenyl ring of **5** is 0.7 Å closer to the side-chain phenol of Tyr24 compared to the phenyl ring of inhibitor **4** and makes stronger edge-on π-stacking interactions with the side-chain phenol group. In addition, the terminal phenyl ring of **5** makes new hydrophobic interactions with the side-chain atoms of Leu13 and Leu16. The carbon linker atoms of **5** make similar interactions to those observed in MtDHQase:**4** but does not make any hydrogen bond interactions between its ether oxygen atom and the active site residues (figure 5e). The optimized positioning of the terminal phenyl ring of **5** into the hydrophobic patch comprised of Leu13, Leu16, Tyr24 is enabled through the increased flexibility of its ether linkage compared to the more rigid linker of **4** and results in a 16-fold increase in potency between these two inhibitors and a 1,400 fold increase from the template inhibitor **2** (Figure 5d, e and Supplementary table 4) [19].

The three crystal structures (MtDHQase:**3-5**) indicate that the dramatic increases in potency for each of these compounds [19, 21] correspond with the degree of complementarity for their aromatic extensions from the 2,3-anhydroquinate anchor with the side-chain of Tyr24. The formation of stabilizing interactions between inhibitors and the flexible loop region appears to be a significant factor in the increase of inhibitor potency which can be further substantiated by the crystal structure of MtDHQase:**6**.

We solved the 2.5 Å resolution crystal structure of MtDHQase in complex with **6** (figure 3e), which contains a *tert*-butyl amide extension from the C3 atom of **2** (figure 2). Inhibitor **6** competitively inhibits MtDHQase and ScDHQase with a K_i of 27 μM and 29 μM respectively (Supplementary table 4) [22].

Inhibitor **6** is bound in the active sites of 14 of 24 MtDHQase subunits per asymmetric unit. The active site loop residues 19-24 for the MtDHQase:**6** complex were predominantly disordered with only some subunits containing electron density for backbone atoms of this region. Inhibitor **6** binds in similar orientations in 11 of the subunits and takes up an alternate conformation for its *tert*-butyl extension in the remaining 3 occupied subunits.

Analysis of the MtDHQase:**6** structure reveals that the *tert*-butyl amide group does not cause significant changes in the position of the anhydroquinate core of inhibitor **6** relative to its position in MtDHQase:**5** or MtDHQase:**2** (Figure 6d). The structure of MtDHQase:**6** is unique in that the inhibitor only partly reaches the site of the glycerol molecule in MtDHQase:**2**. The *tert*-butyl amide group of inhibitor **6** extends towards the MDHQase glycerol binding site and makes hydrophobic interactions with the side-chain carbon atoms of Asn12 and Leu13, as well as Asp88* (figure not shown). The oxygen atom from the C3 amide linker group is within hydrogen bonding distance to the side-chain of Asp88*. The side-chain atoms of the active site flexible loop residues 19-24 are fully disordered indicating a lack of stabilizing interaction between these atoms and the *tert*-butyl moiety of inhibitor **6**. The superposition of MtDHQase:**6** and MtDHQase**:5** shows that the *tert*-butyl group is 1.6 Å from the position occupied by the side-chain of Tyr24 in MtDHQase:**1-2** and indicates that the flexible loop residues remain disordered due to steric interference (Figure 6d). The disordering of the flexible loop in MtDHQase:**6** appears to significantly reduce binding affinity by approximately 200 fold compared to inhibitor **5,** which binds with the greatest degree of complementarity to the active site subpocket.

Discovery and characterization of a planar inhibitor of MtDHQase

The inhibitor series derived from inhibitor **2** are excellent probes for investigating the active site subpocket of DHQase, but their relatively complex structures and non-trivial syntheses led us to search for more synthetically tractable chemical templates. We therefore screened a small collection of planar product analogs to identify new candidate anchor molecules. Kinetic studies identified compound **7** as a competitive inhibitor of MtDHQase with a K_i of 300 μM (Supplementary figure 4). Inhibitor **7** (citrazinic acid), is an analogue of **1** that lacks the ring puckering due to its planar pyridine core (Figure 2) and contains two hydroxyl groups, which can undergo keto-enol tautomerization (Figure 2) possibly enabling it to mimic the molecular conversion of the C3 keto group of substrate (3-dehydroquinate) to the C3 enol/enolate group of the reaction intermediate.

STD-NMR spectroscopy was used to confirm the competitive binding of inhibitor **7** to MtDHQase followed by the determination of its ligand efficiency [42] using ITC (Supplementary figures 5-6). The ligand efficiency is defined as a measure of the binding energy per non-hydrogen atom of a ligand [42]. Numerous studies have led to the establishment that successful lead molecules typically have ligand efficiency values $\geq$ 0.3 kcal mol^{-1} per non-hydrogen atom [43]. Inhibitor **7** has a very high ligand efficiency of 0.51 kcal mol^{-1} per non-hydrogen atom making it an attractive scaffold for development of future series of MtDHQase inhibitors.

We solved the crystal structure of MtDHQase:**7** to 2.25 Å resolution providing structural details into its binding at the MtDHQase active site. Electron density for inhibitor **7** and for the active site loop residues 19-24 is observed in 22 of the 24 subunits of the asymmetric unit (Figure 3f and Supplementary Figure 7). It is unclear which citrazinic acid tautomer is bound in the crystal structure of MTDHQase:**7**, because the resolution is insufficient to discern between 1.4 Å (C-OH) and 1.2 Å (C=O) bond lengths. However, based on the interactions observed between MtDHQase and the tautomeric groups of inhibitor **7** (*vide infra*, Supplementary table 3), we speculate that the keto-tautomer form is bound in the MtDHQase:**7** crystal structure. The remainder of the discussion will refer to the keto-tautomer of **7** for purposes of describing and evaluating binding interactions.

MtDHQase:**7** reveals that inhibitor **7** is predominantly bound using interactions similar to those observed in MtDHQase:**1-2** except that the core of **7** is shifted towards the flexible loop region by 0.8 Å possibly due to new hydrogen bond interactions formed between the pyridine nitrogen atom of **7** and the subunit residue Asp88* (Supplementary table 3, Figure 7a-b). The flexible loop region

shows similar hydrogen bond interactions with inhibitor **7** to those observed in MtDHQase:**1** (Supplementary table 3), however the hydrogen bond interactions of the C3 carbonyl oxygen atom of **7** with Arg19 Nη2 (2.4Å) and with a conserved water molecule (2.5 Å) (Figure 7a) appear stronger than those observed in the MtDHQase:**1** structure.

Structural comparison of MtDHQase:**1** and MtDHQase:**7** reveals that binding of the more planar **2** does not cause significant movements of the flexible loop region or the active site residues upon binding of **7**, with the exception of rearrangements for the side-chain atoms His101 and Val105 (Figure 7b). In MtDHQase:**7**, His101 Cδ2 makes hydrophobic interactions with Val105 Cγ1, which has adopted an alternate side-chain conformation (figure not shown). His101 Nδ1 moves to 3.1 Å from Asn75 Oδ, and makes a new intra-protein hydrogen bond not observed in the MtDHQase:**1** structure (Figure 7b). Furthermore Glu99 Oε2 also moves 0.8 Å closer to His101 Nδ2 making a stronger hydrogen bonding interaction. In addition to these movements, the hydrogen bond formed by the pyridine nitrogen atom of **7** and the side-chain carboxylate of Asp88* of a neighboring subunit causes subtle alteration of the salt bridge interaction between Asp88* and Arg112 in comparison to MtDHQase:**1** (Figure 7a-b).

The structural movement of the catalytic His101 side-chain results in the formation of new intra-protein interactions that are not observed in MtDHQase:**1-2.** This new orientation of His101 in MtDHQase:**7** is presumed to be non-catalytic due to its increased intra-protein interactions (Figure 7a-c) and appears to be a key feature in the ability of **7** to inhibit MtDHQase.

DISCUSSION

The MtDHQase crystallographic structures presented herein reveal structural details for the flexible catalytic loop residues 19-24 in MtDHQase, and also provide structural evidence and characterization of a pocket adjacent to the MtDHQase active site and its influence on MtDHQase inhibitor binding. MtDHQase:**1** reveals a series of intra-protein interactions involved in the coordination of the Tyr24 side-chain into a presumed catalytically relevant position, which is further coordinated by the formation of electrostatic interactions with the side-chain of Arg19. The requirement of these two residues for catalytic activity makes them attractive focal points for inhibitor development. Our investigation of MtDHQase in complex with inhibitors that occupy both the MtDHQase active site and the adjacent subpocket (**3-6**) and inhibitors that occupy only the MtDHQase active site (**2** and **7**) reveals the most potent MtDHQase inhibitors replace the electrostatic interaction between the side-chain atoms of Arg19 and Tyr24 through the expulsion of the Arg19 side-chain to the solvent as was observed by Peón *et. al.* [20]. This structural perturbation of the MtDHQase loop residues 19-24 effectively expands the available area for creating protein-inhibitor binding interactions.

The lack of structural evidence and information for this active site adjacent subpocket in MtDHQase previously limited our understanding of differences in inhibitor complementarity between MtDHQase and ScDHQase. These structural investigations against MtDHQase and their comparison with crystals structures available for ScDHQase reveal key differences in the active site subpockets of the enzymes from these two organisms. The flexibility of the terminal phenyl rings for **3** and **4** are both constrained by their rigid linkers and both display negligible differences in inhibition potency (Figure 1, Supplementary table 4). However inhibitor **5**, which has a much more flexible linker, displays a significant (~14 fold) difference in potency against MtDHQase and ScDHQase (Figure 1, Supplementary table 4).

Two structures of ScDHQase in complex with inhibitors containing terminal phenyl rings with linkers of similar length and flexibility to **5** have been deposited in the PDB (accession codes: 2CJF and 2BT4) [18, 19]. One of these structures is in complex with Biaryl-B, a biphenyl inhibitor that has a flexible thioether linkage and is a 5-fold more potent inhibitor of MtDHQase than ScDHQase (PDB access code 2CJF) (Supplementary table 4) [19]. The other is a structure of ScDHQase in complex with CA2, a single ring extension inhibitor (PDB accession code 2BT4). The superposition of MtDHQase:**5** and ScDHQase:Biaryl-B reveals several differences in the positioning of their respective terminal phenyl rings providing insights into unique features of their subpockets. Although the terminal phenyl ring of Biaryl-B makes similar interactions to those of **5**, the π-stacking interaction between it and the phenol group of ScDHQase Tyr28 appears to be stronger than is observed for

MtDHQase:**5**. The terminal phenyl ring of Biaryl-B is rotated ~50° and is shifted 1.4 Å further into the pocket relative to **5** (Supplementary Figure 8c). It is possible that the difference in the positioning of Biaryl-B is caused by its own biaryl system, however this is likely not the case as the crystal structure of ScDHQase in complex with the single ring extension inhibitor CA2, also positions its terminal phenyl ring exactly as observed in ScDHQase:Biaryl-B (Supplementary figure 8a-b).

The superposition of MtDHQase:**5** and ScDHQase:Biaryl-B shows that backbone atoms of the loop region are shifted by ~2.2 Å between these two structures and the main-chain carbonyl group of MtDHQsae Arg19 is positioned similarly to that of ScDHQase Gln22. The two peptide bonds are flipped 180° relative to each other, positioning the MtDHQase Arg19 backbone carbonyl group into the subpocket in a similar position to that observed in MtDHQase:**2**:glycerol (Supplementary Figure 8c). Furthermore two residues forming hydrophobic interactions with the terminal phenyl ring in ScDHQase:Biaryl-B are replaced with polar residues in MtDHQase. ScDHQase Leu19 is equivalently positioned to MtDHQase Arg15 and ScDHQase Thr96 is equivalently positioned to MtDHQase Glu92. The differences in polarity for these residues in addition to the flipped peptide bond of Arg19 presumably contribute to the differences in inhibitor potencies.

The seven MtDHQase:ligand structures presented in this report show the complete active site structure of MtDHQase, including its full catalytic loop region, MtDHQase**1**,**2**,**7** structures, and in an extended conformation, MtDHQase:**3**,**4**,**5** structures, similar to the recent report by Peón *et. al.* [20]. This study clearly shows the formation of a stacking interaction between Tyr24 and an aromatic system such as the primary phenyl ring of inhibitor **3** or the terminal phenyl ring of inhibitor **4** and **5**, is important in the gain of the affinity and that its engagement might be further explored using different linkers from the anhydroquinate core to increase complementarity. We also report the first structure of MtDHQase in complex with a planar molecule bound in the active site, representing a novel scaffold for MtDHQase inhibitor development. Very few studies have tried to optimize interactions of the anchor molecule to MtDHQase and no significant progress has been reported to the best of our knowledge [8,22]. Citrazinic acid (**7**) represents a stepping stone in this direction and synthetic efforts are currently in progress to develop novel MtDHQase inhibitors based on the planarity of citrazinic acid that target the MtDHQase subpocket as guided by the x-ray crystal structures reported herein.

REFERENCES

1. Abell, C. (1999) *Enzymology and molecular biology of the shikimate pathway*. In Comprehensive natural products chemistry, U. Sankawa, ed. (Amsterdam, Netherlands: Elsevier, 573-607)
2. Coggins, J. R., Abell, C., Evans, L. B., Frederickson, M., Robinson, D. A ., Roszak, A. W., and Lapthorn, A. P. (2003) Experiences with the shikimate-pathway enzymes as targets for rational drug design. Biochem. Soc. Trans. **31**, 548-552
3. Dosselaere, F., and Vanderleyden, J. (2001) A metabolic node in action: chorismate-utilizing enzymes in microorganisms. Crit. Rev. Microbio. **27**, 75-131
4. 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., and McLeod, R. (1998) Evidence for the shikimate pathway in apicomplexan parasites. Nature *393*, 801-805
5. Harris, J., Kleanthous, C., Coggins, J. R., Hawkins, A. R., and Abell, C. (1993) Different mechanistic and stereochemical courses for the reactions catalysed by type II dehydroquinases. J. Chem. Soc. Chem. Commum. **13**, 1080-1081
6. Parish, T., and Stoker, N.G. (2002) The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis.* Microbiology **48**, 3069-3077
7. Gourley, D. G., Shrive, A. K., Polikarpov, I., Krell, T., Coggins, J. R., Hawkins, A. R., Isaacs, N. W., and Sawyer, L. (1999) The two types of 3-dehydroquinase have distinct structures but catalyze the seme overall reaction. Nat. Struct. Biol. **6**, 521-525
8. Gonzalez-Bello, C., and Castedo, L. (2007) Progress in type II dehydroquinase inhibitors: from concept to practice. Med. Res. Rev. **27**, 177-208
9. Giles, N. H., Case, M. E., Baum, J., Geever, R., Huiet, L., Patel, V., and Tyler, B. (1985) Gene organization and regulation in the qa (quinic acid) gene cluster of *Neurospora crassa.* Microbiol. Rev. **49**, 338-358
10. Hawkins, A. R., Lamb, H. K., Moore, J. D., Charles, I. G., and Roberts, C. F. (1993) The pre-chorismate (shikimate) and quinate pathways in filamentous fungi: theoretical and practical aspects. J. Gen. Microbiol. **139**, 2891-2899
11. Chaudhuri, S., Ducan, K., Graham, L. D., and Coggins, J. R. (1991) Identification of the active-site lysine residues of two biosynthetic 3-dehydroquinases. Biochem. J. **275**, 1-6
12. Roszak, A. W., Robinson, D. A., Krell, T., Hunter, I. S., Fredrickson, M., Abell, C., Coggins, J. R., and Lapthorn, A. J. (2002) The structure and mechanism of the type II dehydroquinase from *Streptomyces coelicolor.* Structure **10**, 493-503
13. Robinson, D. A., Stewart, K. A., Price, N. C., Chalk, P. A., Coggins, J. R., and Lapthorn, A. J. (2006) Crystal structure of *Helicobacter pylori* type II dehydroquinase inhibitor complexes: new directions for inhibitor design. J. Med. Chem. **49**, 1282-1290
14. Burnett, R. M., Darling, G. D., Kendall, D. S., LeQuense, M. E., Mayhew, S. G., Smith, W. W., and Ludwig, M. L. (1974*)* The structure of the oxidized form of clostridial flavodoxin at 1.9 Å resolution. J. Bio. Chem. **249**, 4383-492
15. Krell, T., Horsburgh, M. J., Cooper, A., Kelly, S. M., and Coggins, J. R. (1996) Localization of the active site of type II dehydroquinases. Identification of a common arginine-containing motif in the two classes of dehydroquinases. J. Biol. Chem. **271**, 24492-24497
16. Frederickson, M., Roszak, A. W., Coggins, J. R., Lapthorn, A. J., and Abell, C. (2004) (1*R*,4*S*,5*R*)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
17. Toscano, M. D., Stewart, K. A., Coggins, J. R., Lapthorn, A. J., and Abell, C. (2005) Rational design of new bifunctional inhibitors of type II dehydroquinase. Org. Biomol. Chem. **3**, 3102-3104
18. Payne, R. J., Peyrot, F., Kerbarh, O., Abell, A. D. , and Abell, C. (2007*)* Rational design, synthesis, and evaluation of nanomolar type II dehydroquinase inhibitors. ChemMedChem **2**, 1015-1029
19. Lee, B. I., Kwak, J. E., and Suh, S. W. (2003) Crystal structure of the type II 3-dehydroquinase from *Helicobacter pylori.* Proteins **51**, 616-617

20. Peón, A., Otero, J. M., Tizón, L., Prazeres, V. F., Llamas-Saiz, A. L. Fox, G. C., van Raaij, M. J., Lamb, H., Hawkins, A. R., Gago, F., Castedo, L., Gonzáles-Bello, C. (2010) Understanding the key factors that control the inhibition of type II dehydroquinase by (2R)-2-benzyl-3-dehydroquinic acid. ChemMedChem **5**, 1726-1733.
21. Payne, R. J., Riboldi-Tunnicliffe, A., Kerbarh, O., Abell, A. D., Lapthorn, A. J., and Abell, C. (2007) Design, synthesis, and structural studies on potent biaryl inhibitors of type II dehydroquinases. ChemMedChem **2**, 1010-1013
22. Toscano, M. D., Payne, R. J., Chiba, A., Kerbarh, O., and Abell, C. (2007) Nanomolar inhibition of type II dehydroquinase based on the enolate reaction mechanism. ChemMedChem **2**, 101-112
23. González-Bello, C., Lence, E., Toscano, M.D., Castedo, L., Coggins, J.R., and Abell, C. (2003) Parallel solid-phase synthesis and evaluation of inhibitors of *Streptomyces coelicolor* type II dehydroquinase. J. Med. Chem. **46**, 5735-5744
24. Prazeres, V. F., Sánchez-Sixto, C., Castedo, L., Lamb, H., Hawkins, H.R., Riboldi-Tunnicliffe, A., Coggins, J.R., Lapthorn, A.J., and Gonzáles-Bello, C. (2007) Nanomolar competitive inhibitors of *Mycobacterium tuberculosis* and *Streptomyces coelicolor* type II dehydroquinase. *ChemMedChem* **2**, 194-207
25. Sánchez-Sixto, C., Prazeres, V.F., Castedo, L., Lamb, H., Hawkins, A.R., and González-Bello, C. (2005) Structure-based design, synthesis, and biological evaluation of inhibitors of *Mycobacterium tuberculosis* type II dehydroquinase. J. Med. Chem. **48**, 4871-4881
26. 38 Frederickson, M., Parker, E. J., Hawkins, A. R., Coggins, J. R., and Abell, C. (1999) Selective inhibition of type II dehydroquinase. J. Org. Chem. **64**, 2612-2613
27. Prazeres, V. F., Castedo, L., Lamb, H., Hawkins, A. R., and Gonzalez-Bello, C. (2009) 2-substituted-3-dehydroquinic acids as potent competitive inhibitors of type II dehydroquinase. ChemMedChem **4**, 1980-1984
28. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. J. Appl. Cryst. **30**, 1022-1025
29. Collaborative computational project, number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Cryst. Sect. D **50**, 760-763
30. Murshudov, G. N., Vagin, A. A., and Dodson (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. Sect. D **53**, 240-255
31. Painter, J., and Merritt, E. A. (2006) TLSMD web server for the generation of multi-group TLS models. *J Appl Cryst* **39**, 109-111
32. McRee, D. E. (1999) XtalView/Xfit – A versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. **125**, 156-165
33. Ernsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D **60**, 2126-2132
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornthon, J. M. (1993) PROCKECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. **26**, 283-291
35. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera – a visualization system for exploratory research and analyss. J. Comput. Chem. **25**, 1605-1612
36. Merritt, E. A., and Murphy, M. E. (1994) Raster3D version 2.0. A program for photorealistc molecular graphics. Acta Crystallogr. Sect. D **50**, 869-873
37. DeLano, W. L. (2002) *The* PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA, USA
38. GraFit®: Erithacus Software, London PO Box 274, Horley, Surrey RH6 9YJ, UK
39. Mayer, M., and Meyer, B. (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angew. Chem. Int. Ed. **38**, 1784-1788
40. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailores excitation for single-quantum NMR spectroscopy of aqueous solutions. J. Biomol. NMR **2**, 661-665
41. Shneier, A., Kleanthous, C., Deka, R., Coggins, J. R., and Abell, C. (1991) Observation of an imine intermediate on dehydroquinase by electrospray mass spectrometry. J. Am. Chem. Soc. **113**, 9416–9418

42. Hopkins, A. L., Groom, C. R., and Alex, A. (2004) Ligand efficiency: a useful metric for lead selection. Drug Discov. Today **9**, 430-431
43. Hajduk, P. J. (2006) Fragment-based drug design: how big is too big? J. Med. Chem. **49**, 6972-6976
44. Sánchez-Sixto, C., Prazeres, V. F., Castedo, L., Suh, S. W., Lamb, H., Hawkins, A. R., Canada, F. J., Jimenez-Barbero, J., and Gonzalez-Bello, C. (2008) Competitive inhibitors of *Helicobacter pylori* type II dehydroquinase: synthesis, biological evaluation, and NMR studies. ChemMedChem **3**, 756-770
45. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected on oscillation mode. Methods Enzymol**. 276**, 307-326
46. Leslie, A. G. W. (2006) The integration of macromolecular diffraction data. Acta Cryst. Sect. D **62**, 48-57
47. Evans, P. (2006) Scaling and assessment of data quality. Acta Cryst. Sect. D 62, 72-82

FOOTNOTES

This work was funded by a grant from the Bill and Melinda Gates Foundation (sub-contract on *Integrated Methods for Tuberculosis Drug Discovery* grant to the Seattle Biomedical Research Institute). M. V. B. D. thanks the CNPq (Brazil) for a post doctoral fellowship. We appreciate the support of staff scientists at beamline 19-ID and 23-ID of the Advanced Photon Source, Argonne National Laboratory for help in data collection. We also thank Justin Roberts, Joel Freundlich, and Hilary Baird for excellent technical assistance and Siaska Castro and Tracey Musa for comments on the manuscript.

The abbreviations used are: DHQase, 3-dehydroquinate dehydratase; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; Hepes, 4-(hydroxyethyl)-1-piperazineethanesulfonic acid; HpDHQase, 3-dehydroquinate dehydratase from *Helicobacter pylori*; ITC, isothermal titration calorimetry; IPTG, isopropyl β-D-1-thiogalactopyranoside; KBr, potassium bromide; NMR, nuclear magnetic resonance; MtDHQase, 3-dehydroquinate dehydratase from *Mycobacterium tuberculosis*; PDB, protein data bank; PEG, polyethylene glycol; ScDHQase, 3-dehydroquinate dehydratase from *Streptocymes coelicolor*; STD, saturation transfer difference; Tris, (hydroxymethyl) aminomethane;

TABLE

Table 1 Crystallographic statistics of MtDHQase:**1-7** structures

	MtDQHase:1	**MtDQHase:2: Glycerol**	**MtDHQase:3**	**MtDHQase:4**	**MtDHQase:5**	**MtDHQase:6**	**MtDQHase:7**
PDB ID	3N59	3N7A	3N87	3N86	3N76	3N8N	3N8K
Space group	$P2_1$	P1	P1	$P2_1$	F23	$P2_1$	$P2_1$
Maximum resolution	2.5	2.0	2.4	1.9	1.9	2.5	2.3
R_{cryst} [a]	19.8	17.7	19.1	15.5	17.5	20.5	20.5
R_{free} [b]	24.8	23.8	27.4	21.2	22.1	28.6	23.6

[a] $R_{cryst}=\Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$, F_{obs} and F_{calc} are observed and calculated structure factor amplitudes.
[b] R_{free} as for R_{cryst} using a random subset of the data (5%) excluded from the refinement.

FIGURE LEGENDS

Figure 1. Catalytic reaction mechanism for type II 3-dehydroquinate dehydratase. In a first step of the reaction of Type II DHQase a conserved tyrosine acts to remove the *pro*-S proton of the C2 to form the enolate intermediate. Simultaneously an asparagine holds a conserved water molecule in the correct orientation to stabilize the intermediate and finally a conserved histidine acts as a proton donor to catalyze the elimination of C1 hydroxyl group and liberate a water molecule producing the 3-dehydroshikimate.

Figure 2. MtDQHase ligands: **1** (3-dehydroshikimic acid) the cyclohexene carbon atom numbering progresses counter-clockwise from C1; C1, C3, and C5 of ligand **1** are labeled; **2** (2,3 anhydroquinate); **3** ((1*S*,4*R*,5*R*)-3-(3-benzoylphenyl)-1,4,5-trihydroxycyclohex-2-enecarboxylic acid); **4** (1*R*,4*R,*5*R*)-1,4,5-trihydroxyl-3-(2-phenylcarbamoyl-vinyl)-cyclo-hex-2-enecarboxylic acid); **5** ((1*R*,4*R*,5*R*)-1,4,5-trihydroxy-3-(3-phenoxy-prop-(*E*)-enyl)-cyclohex-2-enecarboxylic acid); **6** ((1*R*,4*R*,5*R*)-3-(tert-butylcarbamoyl)-1,4,5-trihydroxycyclohex-2-enecarboxylic acid); **7** (citrazinic acid represented in both tautomeric forms).

Figure 3. Surface representation of MtDHQase with close-up view of the active site binding of 3-dehydroshikimate and six inhibitors: a) 3-dehydroshikimate (**1**), b) inhibitor **2** and **glycerol**, c) inhibitor **3**, d) inhibitor **4,** e) inhibitor **5,** f) inhibitor **6**, and g) inhibitor **7**. Electron density 2Fo-Fc contour for the inhibitors are shown.

Figure 4. a. Close up comparison of MtDHQase:**1** with the incomplete structure of *apo*-MtDHQase (2DHQ), emphasizing the closure of the active site by flexible loop residues 19-24 upon ligand binding and showing the location of the subpocket located adjacent to the active site observed in the MtDHQase:**2** structure. Subunit G of MtDHQase:**1** (residues 15-28 colored yellow and ligand **1** colored gray) is superimposed individually on the *apo*-MtDHQase subunit (residues 15-28 are colored magenta) using UCSF chimera. The location of the glycerol molecule (cyan) from MtDHQase:**2** is shown based on the individual superposition of MtDHQase:**2** subunit B on MtDHQase:**1** subunit G; all other MtDHQase:**2** atoms were removed for clarity. **Inset.** Surface representation of the MtDHQase homododecamer with a black box encompassing one of the twelve enzyme active sites. Each MtDHQase subunit is individually colored. The boxed active site is represented in the close-up view in Figure 3a. **b.** The 2.52 Å resolution electron density map of MtDHQase:**1** showing ligand **1** (green), loop residues 19-24, and His101 from subunit G, and Asp88* from subunit F (gray). The 2Fo-Fc (1σ; blue) and Fo-Fc (3σ; red) electron density maps were calculated after omitting ligand **1** and residues 19-24, 101, and 88*. The blob feature in XTALVIEW [32] has been applied to limit the electron density display to within 1.5 Å of the subunit residues and the final figure is rendered with Raster3D [36]. **c.** Ligand binding interactions of the MtDHQase:**1** binary complex structure. **1** (green), loop residues 19-24 (yellow), and additional important MtDHQase residues (gray) involved in the binding of **1** are displayed as sticks; hydrogen bond interactions between MtDHQase and **1** are indicated with blue lines and important direct and water mediated intra-protein loop residue [16-21] hydrogen bonds are indicated with black lines.

Figure 5. Residues forming hydrogen bonding and hydrophobic interactions involved in the binding of the different ligands in the active site of the MtDHQase. **a) 1**; **b) 2**; **c) 3**; **d) 4; e) 5.** Different colors for residues are used for different subunits and the compounds are represented in green. The figure for the complexes between MtDHQase and **6** and MtDHQase and **7** are not shown. The complex with MtDHQase and compound **6** has most of the flexible active site loop disordered and the complex with MtDHQase and **7** is very similar those ones observed for the complexes either between MtDHQase and **1** or between MtDHQase and **2**.

Figure 6. a. Glycerol binding site and interactions of MtDHQase:**2**. **Glycerol** and inhibitor **2** (green) and important MtDHQase residues (yellow indicates MtDHQase subunit B, gray indicates neighbouring MtDHQase subunit) involved in the binding of **glycerol** are displayed as sticks; hydrogen bond interactions are indicated with black dashed lines. Electron density 2Fo-Fc contour for **2** and glycerol. The binding of glycerol to this pocket conserves most of the interactions formed by the guanidinium group of Arg19 and makes additional interactions with its three hydroxyl groups. The glycerol molecule bound to the active site flexible loop pocket forms direct hydrogen bonding interactions with the main-chain atoms of Asn12, Arg15, Arg18, and Arg19 and forms water-mediated hydrogen bonds with the side-chain atoms of Asn12 and Asp88* as well as with the main-chain atoms of Gly17, Arg18, Arg19, and Gly25. **b.** Superposition of three different subunits of MtDHQase:**5** (D, E, and G) and the position of glycerol from MtDHQase:2 subunit B (yellow). The figure shows the different ways inhibitor **5** binds in MtDHQase. **c.** Individual superposition of MtDHQase:**4** subunit A (yellow) on MtDHQase:**5** subunit A (blue) with close up view of inhibitors and Tyr24. **d.** Individual superposition of MtDHQase:**6** subunit E (green) on MtDHQase:**5** subunit E (blue) with close up view of inhibitors and key MtDHQase residues involved in inhibitor binding.

Figure 6.a. Individual superposition of MtDHQase:**6** subunit E (purple) on MtDHQase:**2** subunit B (green) with close up view of inhibitors and the position of Tyr24 from the MtDHQase:**2** structure. **b.**

Figure 7 a. Inhibitor binding interactions of the MtDHQase:7 (green) and important MtDHQase residues (gray); hydrogen bond interactions are indicated with blue lines. Loop residues 19-24 are colored yellow. **b.** Individual superposition of MtDHQase:**7** subunit G (cyan) on MtDHQase:**1** subunit G (gray) with close up view of the ligand and active site residue positions/movements. **c.** Individual superposition of MtDHQase:**7** subunit G (cyan) on MtDHQase:**2** subunit G (tan) with close up view of the ligand and active site residue positions/movements.

Figure 1

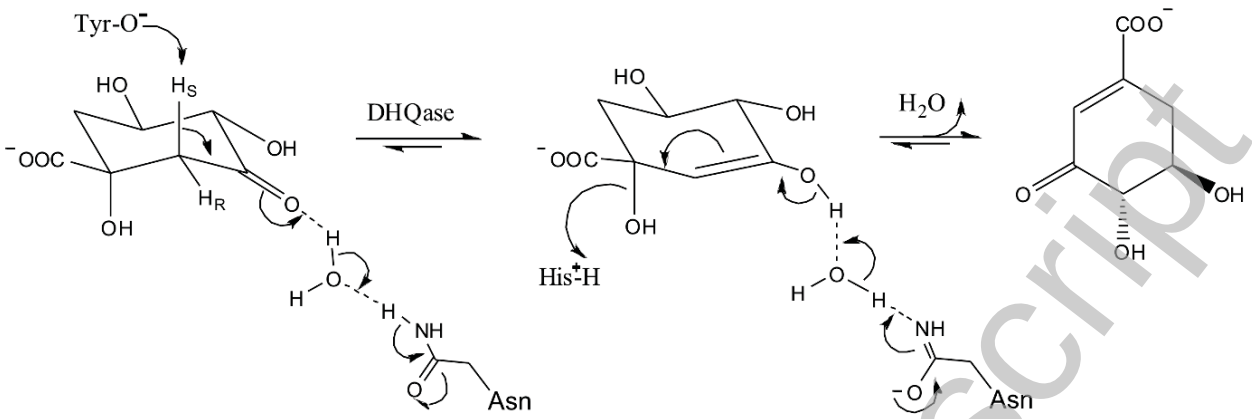


Figure 2

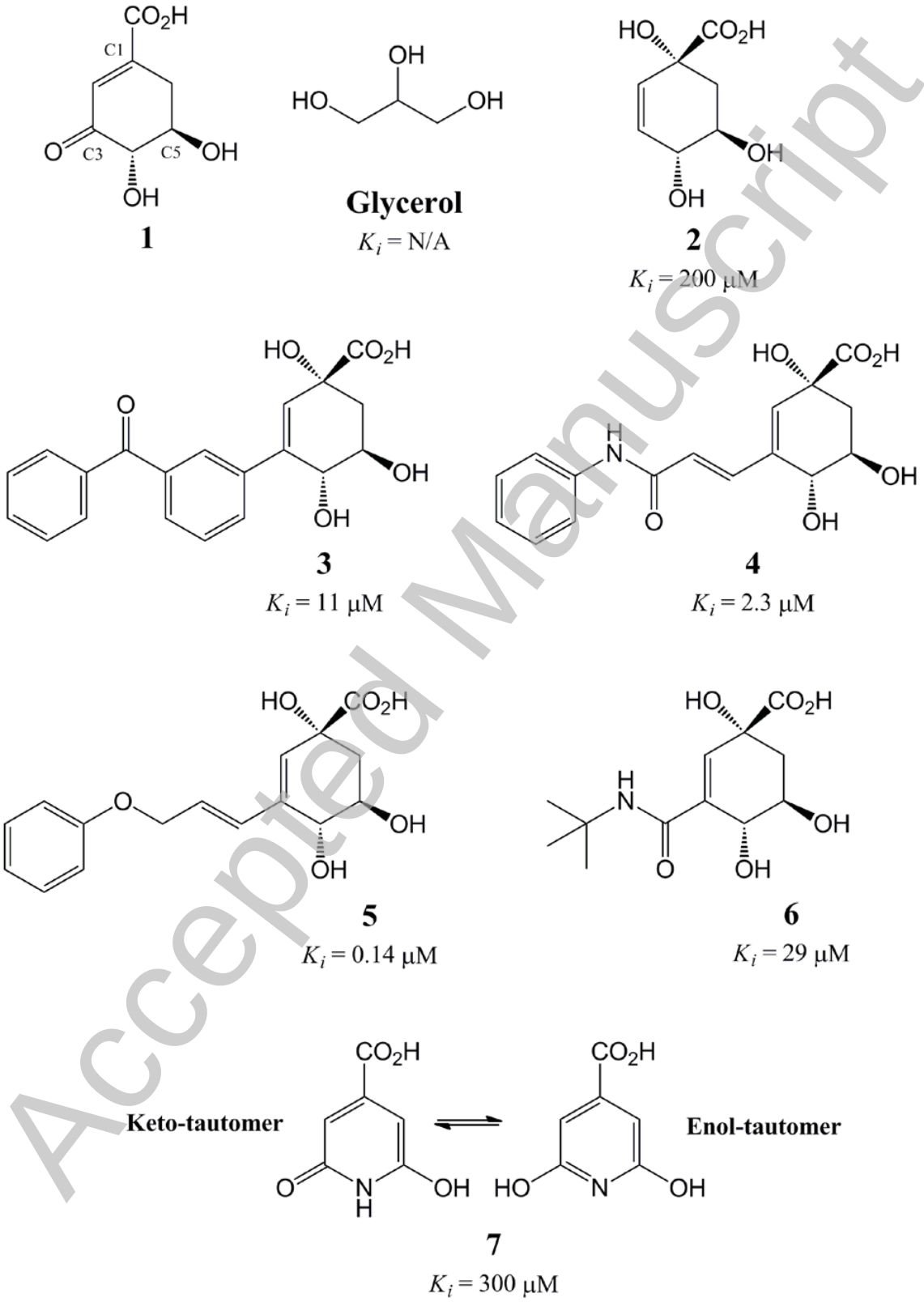
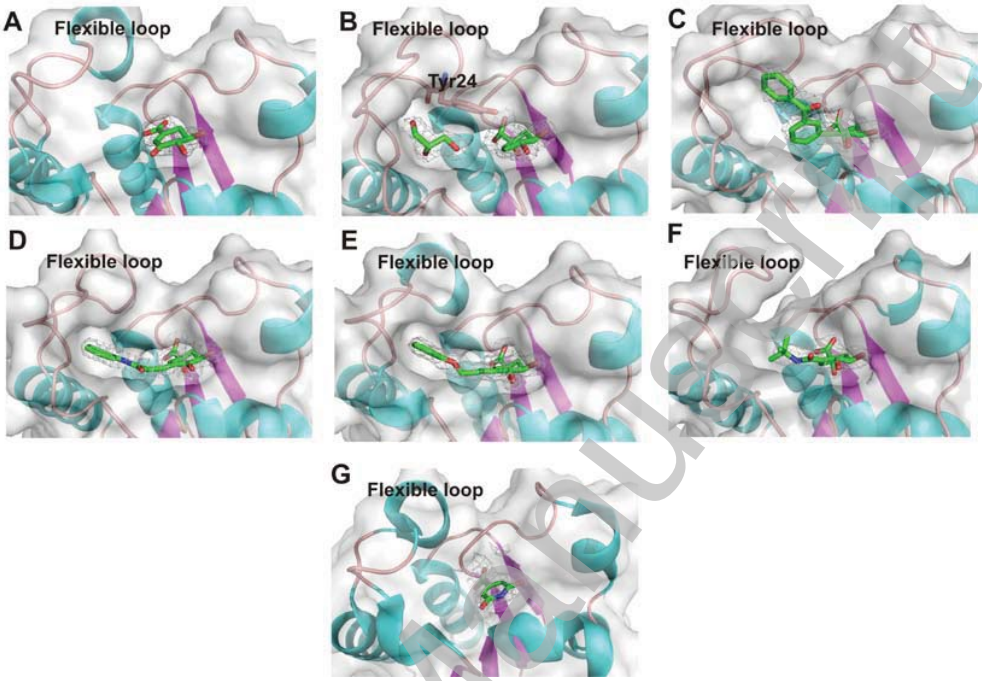


Figure 3



Biochemical Journal Immediate Publication. Published on 17 Mar 2011 as manuscript BJ20110002

Figure 4a

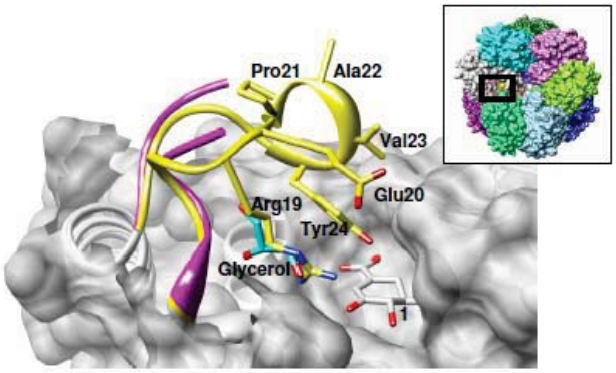


Figure 4b

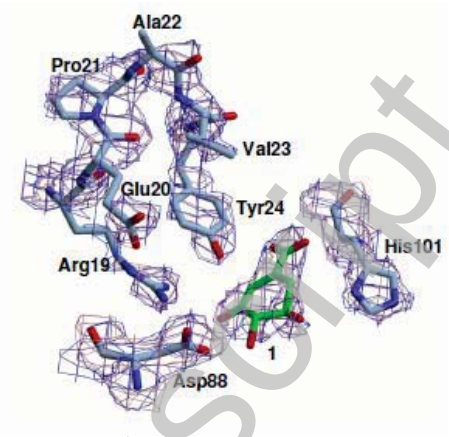


Figure 4c

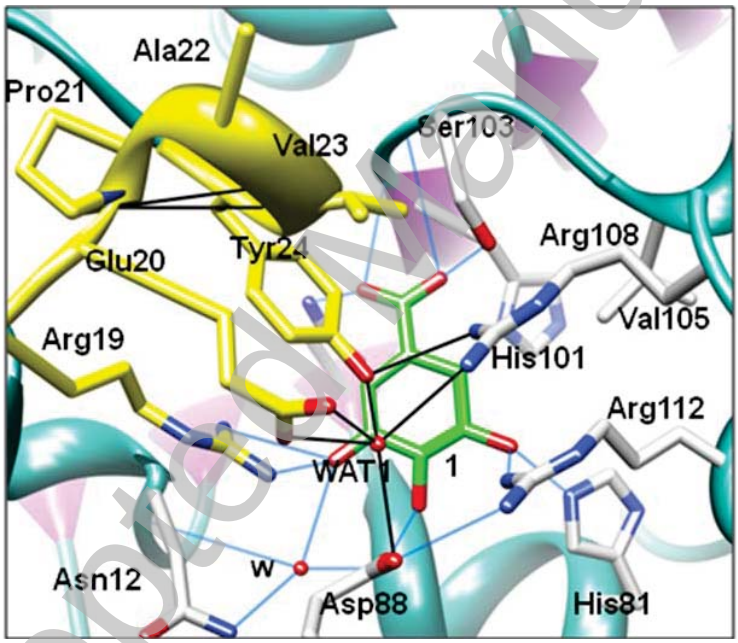


Figure 5

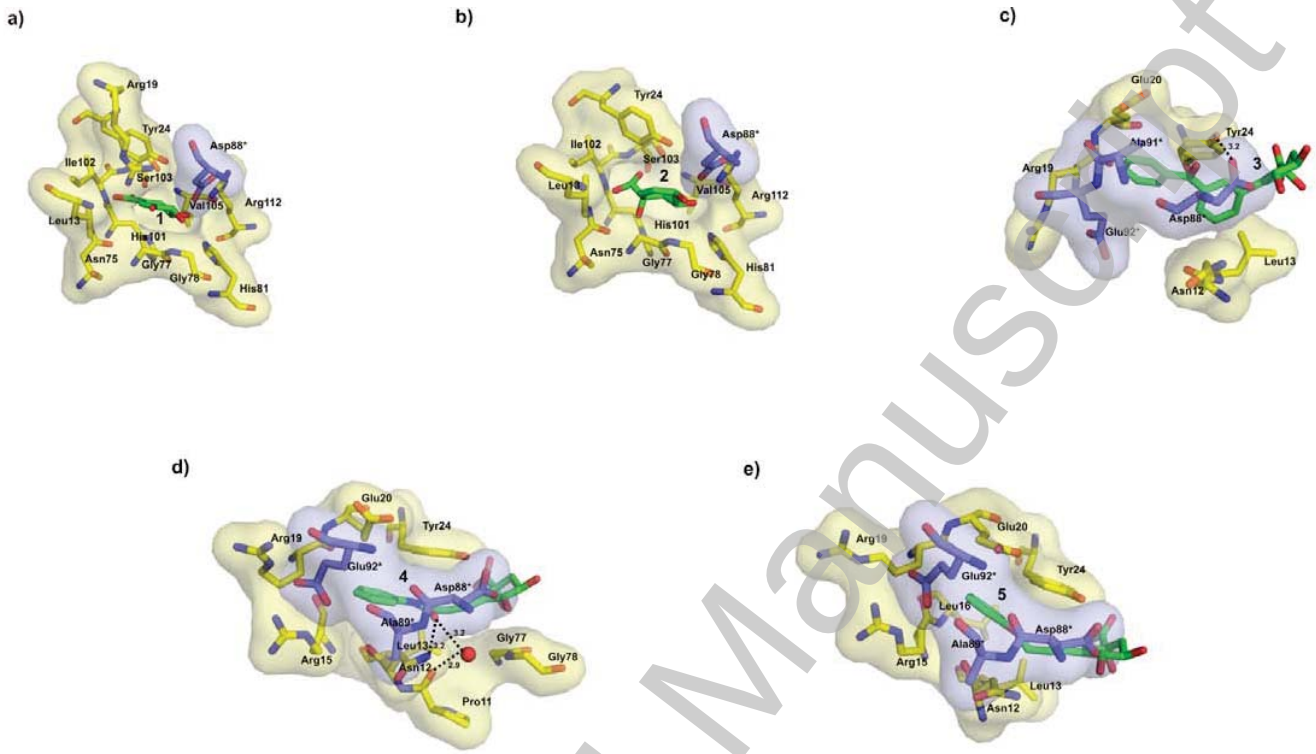


Figure 6

Figure 6a

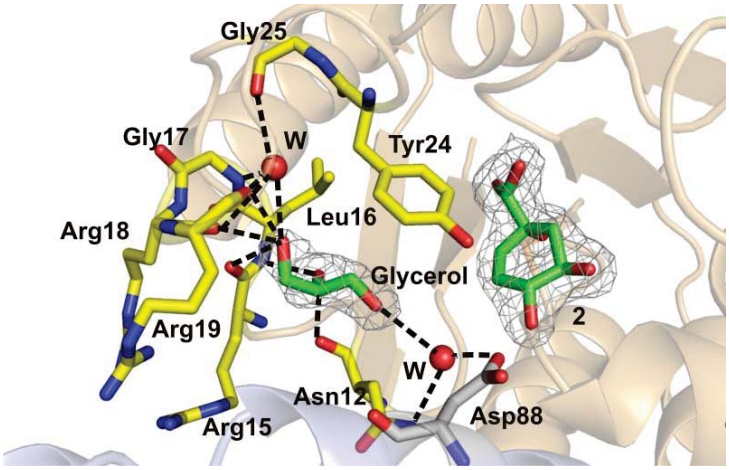


Figure 6b

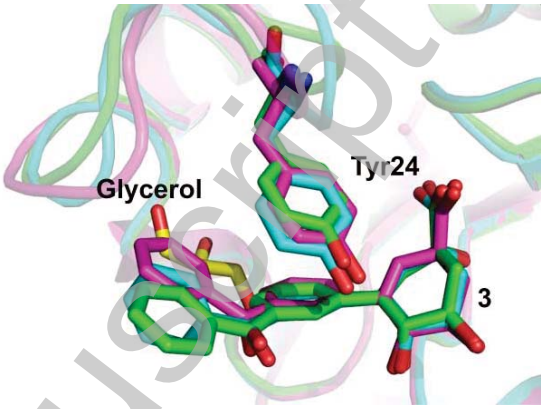


Figure 6c

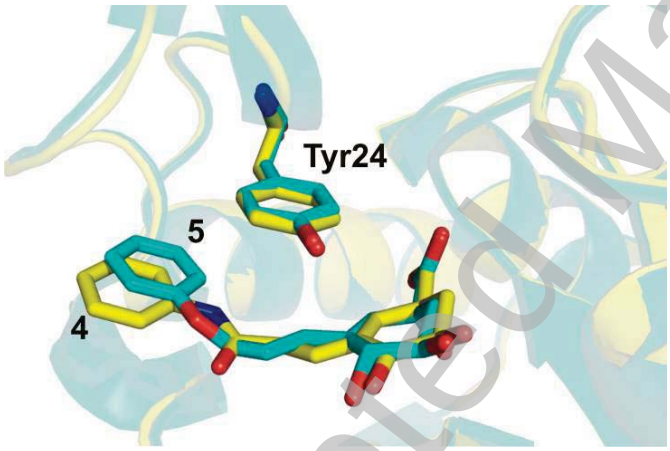


Figure 6d

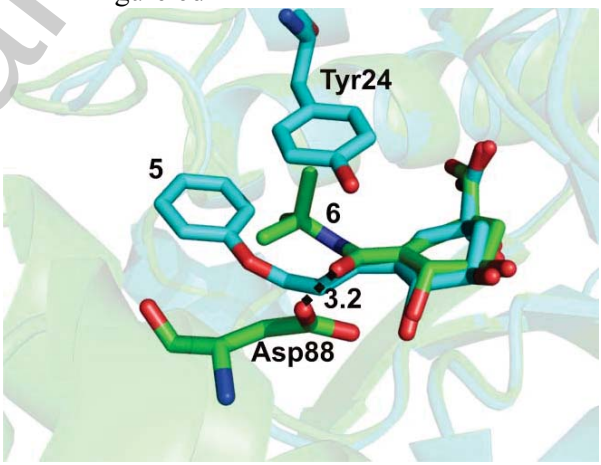


Figure 7

Figure 7a

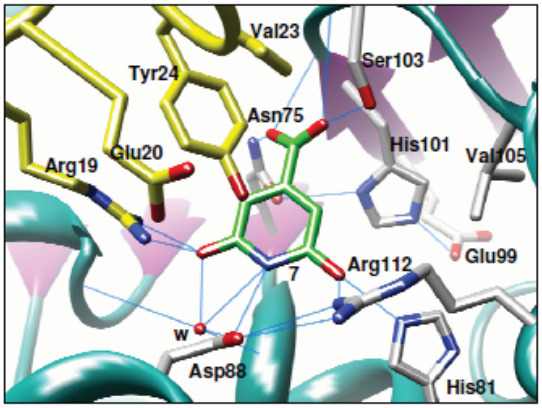


Figure 7b