

# Mechanistic Insight Into the Reaction Catalysed by Bacterial Type II Dehydroquinases‡

Claire CODERCH,<sup>φ,1</sup> Emilio LENCE,<sup>§,1</sup> Antonio PEÓN,<sup>§</sup> Heather LAMB,<sup>¶</sup> Alastair R. HAWKINS,<sup>¶</sup> Federico GAGO<sup>φ,\*</sup> and Concepción GONZÁLEZ-BELLO<sup>§,\*</sup>

<sup>φ</sup>Departamento de Ciencias Biomédicas, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain.

<sup>§</sup>Centro Singular de Investigación en Química Biológica y Materiales Moleculares (CIQUS), Universidad de Santiago de Compostela, 15782 Santiago de Compostela, Spain.

<sup>¶</sup>Institute of Cell and Molecular Biosciences, Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, UK

**ABSTRACT:** The type II dehydroquinase, which is an essential enzyme in *Helicobacter pylori* and *Mycobacterium tuberculosis* and does not have any counterpart in human cells, is recognized to be an attractive target for the development of new antibacterial agents. Computational and biochemical studies that help understand in atomic detail the catalytic mechanism these bacterial enzymes are reported. A previously unknown key role of certain conserved residues of these enzymes, as well as the structural changes responsible for triggering the release of the product from the active site were identified. Asp89\*/Asp88\* from a symmetry-related neighboring enzyme subunit proved to be the residue responsible for the deprotonation of the essential tyrosine to afford the catalytic tyrosinate, which triggers the enzymatic process. The essentiality of this residue is supported by results from site-directed mutagenesis. For *Helicobacter pylori* DHQ2, this reaction takes place through the assistance of a water molecule, while for *Mycobacterium tuberculosis* DHQ2, the tyrosine is directly deprotonated by the aspartate residue. The participation of a water molecule in this deprotonation reaction is supported by solvent isotope effects and proton inventory studies. Molecular dynamics simulation studies provide details of the required motions for the catalytic turnover, which provides a complete overview of the catalytic cycle. The product is expelled from the active site by the essential arginine and after a large conformational change of a loop containing two conserved arginines (Arg109/Arg108 and Arg113/Arg112), which reveals a previously unknown key role of these residues. These studies highlight the key role of the aspartate residue whose blockage could be useful in the rational design of inhibitors and the mechanistic differences between both enzymes.

## Corresponding Author Information

\*Prof. Dr. Federico Gago, Departamento de Ciencias Biomédicas, Universidad de Alcalá, Madrid, Spain.

FAX: +34 918854591; e-mail: [federico.gago@uah.es](mailto:federico.gago@uah.es)

\*Prof. Dr. Concepción González-Bello, Centro Singular de Investigación en Química Biológica y Materiales Moleculares (CIQUS), Universidad de Santiago de Compostela, calle Jenaro de la Fuente s/n, 15782

Santiago de Compostela, Spain. Tel: +34 881815726; FAX: +34 881815704; e-mail:

[concepcion.gonzalez.bello@usc.es](mailto:concepcion.gonzalez.bello@usc.es)

<sup>1</sup>These authors contributed equally to this work.

<sup>‡</sup>Dedicated to Prof. Gregorio Asensio on the occasion of his 65<sup>th</sup> birthday

Key words: enzymatic mechanism, dehydroquinase, QM/MM, MD simulations, *Mycobacterium tuberculosis*, *Helicobacter pylori*, essential residues, solvent isotope effects

Abbreviations used: DHQ2, Type II Dehydroquinase; *Hp*-DHQ2, Type II Dehydroquinase from *Helicobacter pylori*; *Mt*-DHQ2, Type II Dehydroquinase from *Mycobacterium tuberculosis*; *Sc*-DHQ2, Type II Dehydroquinase from *Streptomyces coelicolor*; *An*-DHQ2, Type II Dehydroquinase from *Aspergillus nidulans*; QM, quantum mechanics; MD, molecular dynamics; MM, molecular mechanics; DFT, density functional theory; PME, particle mesh Ewald.

## INTRODUCTION

The decreased effectiveness of antibacterial agents against infectious diseases caused by the development of drug resistance has become one of the most important public health issues of the early 21st century. Only 70 years after the introduction of penicillin, we face the possibility of a future without effective antibiotics against several types of bacteria that cause infections in humans. Antibacterial drug resistance also has a huge economic and social impact [1,2]. Thus, the European Centre for Disease Prevention and Control (ECDC) estimates that antimicrobial resistance results each year in 25,000 deaths and related costs of over €1.5 billion in healthcare expenses and productivity losses. Despite the recognized need for new antibiotics, only two new classes of antibiotics have been brought to the market in the last three decades and there are only a few antibacterial agents under development to face the challenge of multidrug resistance [3-5]. Therefore, the discovery of novel drugs and therapies to treat antibiotic-resistant infections and, particularly of drugs with new mechanisms of action is needed. In this context, there is a great and increasing interest in the detailed knowledge of the catalytic mechanism of selected enzymes for the rational design and development of new inhibitors that can be used as drugs. However, enzymes are very complex entities, and in many cases it is difficult to establish precisely how they perform only on the basis of biochemical and structural data. In recent years, computational methods have consolidated their value as important complementary tools that can assist in the elucidation of how these reactions are catalyzed [6-8]. Here we report results from computational and biochemical studies that help understand in atomic detail the catalytic mechanism of two essential bacterial enzymes that are recognized to be attractive targets for the development of new antibacterial agents [9]. These studies also allowed us to identify a previously unknown key role of certain conserved residues as well as the structural changes responsible for triggering the release of the product from the active site. This information could be very useful for structure-based inhibitor design.

During the past few years, a great deal of effort has been focused on the development of new antibacterial agents that inhibit the enzymes involved in the shikimic acid pathway, in which chorismic acid is synthesized. This compound is the precursor of aromatic amino acids and other metabolites, including folates, ubiquinone, and vitamins E and K [10]. This biosynthetic route is present in bacteria, fungi, higher plants and certain apicomplexan parasites, which are the cause of malaria and toxoplasmosis [11]. The fact that it does not have any counterpart in human cells makes this pathway an attractive target for the development of herbicides and antimicrobial agents. We have focused on the inhibition of the third enzyme of the pathway, the type II dehydroquinase (3-dehydroquinone dehydratase, EC 4.2.1.10, DHQ2), one of the two known DHQ enzymes, because it is essential in important pathogenic bacteria such as *Mycobacterium tuberculosis* (*aroD* gene, *Mt*-DHQ2) [12], responsible for causing tuberculosis, and *Helicobacter pylori* (*aroD/aroQ* gene, *Hp*-DHQ2) [12], the causative agent of gastric and duodenal ulcers, which has also been classified as a type I carcinogen.

DHQ2 catalyzes the reversible dehydration of 3-dehydroquinic acid (**1**) to form 3-dehydroshikimic acid (**2**) by the *anti* elimination of water involving the loss of the more acidic pro-*S* hydrogen from C2 of **1** (Scheme 1). DHQ2 is a dodecamer formed from a tetramer of trimers, a trimer being the minimum catalytic unit of the enzyme [13]. After substrate binding, a flexible loop forms a lid that closes the active site. Two residues, an arginine (Arg17/Arg19 in *Hp*-DHQ2 and *Mt*-DHQ2, respectively) and a tyrosine (Tyr22/Tyr24 in *Hp*-DHQ2 and *Mt*-DHQ2, respectively), both located in the flexible loop, were identified by chemical modification and site-directed mutagenesis studies as essential for enzyme activity [14,15]. In the absence of structural data, earlier kinetic isotope effects and pH profile studies carried out on *Mt*-DHQ2 and *Aspergillus nidulans* DHQ2 (*An*-DHQ2) suggested that the elimination proceeds by a stepwise E<sub>1</sub>CB mechanism involving an enolate intermediate **3** (Scheme 1) [16]. Subsequent crystallization and resolution of the DHQ2 from *Streptomyces coelicolor* (*Sc*-DHQ2) by Laphorn *et al.* [17] allowed a detailed description of the active site and the assignment of specific functions to individual amino acid residues. In particular, an important part of the substrate recognition site is the carboxylate binding pocket that gives rise to four strong hydrogen bonding interactions involving the main chain amide NH groups of Leu102 and Thr102, the side chain OH group of Thr102 and the side chain amide of the conserved Asn76 [18]. In addition, the latter residue also acts in correctly positioning the C1 hydroxyl group to accept a proton from the conserved histidine (His102/His101 in *Hp*-DHQ2 and *Mt*-DHQ2, respectively). Laphorn *et al.* [17] proposed that the reaction is

initiated by the essential tyrosine of the active site removing the pro-*S* hydrogen from C2 of **1** whereas the final step is the acid-catalyzed elimination of the C1 hydroxyl group – a reaction mediated by the conserved histidine acting as a proton donor. The reduced pKa of the essential tyrosine arises from the basic environment formed by the conserved arginine of the active site (Arg109/Arg108 in *Hp*-DHQ2 and *Mt*-DHQ2, respectively). Molecular dynamics (MD) simulation studies suggested that the essential arginine side chain located on the loop must be close to the essential tyrosine side chain for catalysis [19]. Under this arrangement, the arginine controls the appropriate orientation of the tyrosine roughly perpendicular to the cyclohexane ring of the natural substrate by a cation- $\pi$  interaction and contributes to decreasing the pKa of the phenol ring to about 7 [19].

Due to the absence of any suitable residue that could stabilize an enolate intermediate and the presence of a conserved active site water molecule (henceforth WAT1) in *Sc*-DHQ2 crystal structures, Laphorn *et al.* [17] suggested the formation of the enol intermediate **4** rather than the enolate intermediate **3** (Scheme 1). This water molecule interacts through hydrogen-bonding with a conserved asparagine (Asn10/Asn12 in *Hp*-DHQ2 and *Mt*-DHQ2, respectively), the carbonyl group of a conserved proline (Pro9/Pro11 in *Hp*-DHQ2 and *Mt*-DHQ2, respectively), and the main chain amide of a glycine or an alanine (Ala79/Gly78 in *Hp*-DHQ2 and *Mt*-DHQ2, respectively). On the basis of these structural considerations, the authors proposed that the enol intermediate **4** is formed upon proton transfer from WAT1 assuming that the side chain of a conserved asparagine (Asn10/Asn12 in *Hp*-DHQ2 and *Mt*-DHQ2, respectively) would be deprotonated by this water molecule. Later on, Blomberg *et al.* [20], considering the high energy required to deprotonate the side chain of an asparagine and on the basis of results from quantum mechanical (QM) calculations using different Hamiltonians on *Sc*-DHQ2, suggested the formation of enolate intermediate **3** because of its significantly lower energy compared to that of enol intermediate **4**. Even though the reaction mechanism is generally well understood, there are still some unsolved issues, mainly regarding how the catalytic form of the essential tyrosine (tyrosinate) is formed and the discrepancy in the mechanisms comparing DHQ2 from different sources to explain previously reported solvent isotope effects (mainly with *Mt*-DHQ2 and *An*-DHQ2) [16]. In addition, a more detailed knowledge of the role(s) of the essential arginine also located in the flexible loop is required.

To gain further insight into these unanswered questions and to improve our understanding of the catalytic mechanism of both *Hp*-DHQ2 and *Mt*-DHQ2 in atomic detail, computational and biochemical studies were undertaken including site-directed mutagenesis, solvent isotope effects and proton inventories. For the computational studies, we used a combination of standard MD simulations on a catalytically active trimer in aqueous solution, using the molecular mechanics force field AMBER [21], and hybrid QM/MM calculations inside the active site, using density functional theory (DFT). The main advantage of this latter approach, which is implemented in the AMBER 12 release, over previous QM studies that focused on simplified model systems is that bond making and breaking are possible in a dynamic context and also that reorganization of active site residues, substrate, and water molecules is realistically allowed during the whole reaction coordinate. The studies described here identified the residue that deprotonates the essential tyrosine that triggers the DHQ2 catalytic process, whose essentiality is supported by results from site-directed mutagenesis. This allowed us to report a previously unnoticed key role for the conserved residue Asp89\*/Asp88\* (in *Hp*-DHQ2 and *Mt*-DHQ2, respectively) from a symmetry-related neighboring enzyme subunit (residues from this subunit will be marked with an asterisk). This residue was previously thought to play only a structural role because it is involved in the formation of a salt bridge with the guanidinium of the conserved Arg113/Arg112 (in *Hp*-DHQ2 and *Mt*-DHQ2, respectively) and of a hydrogen bond with the C4 hydroxyl group of the substrate [17]. In addition, the studies reported here show the required enzymatic motions for catalytic turnover and also identify a key role of the essential arginines Arg17/Arg19 and Arg109/Arg108 in product release.

## EXPERIMENTAL

### Generation and minimization of the Michaelis complex

a) *Helicobacter pylori* DHQ2 model: These studies were carried out using the enzyme coordinates found in the crystal structure of *Hp*-DHQ2 in complex with an analog of the natural substrate, (2*R*)-2-(4-

methoxybenzyl)-3-dehydroquinic acid (PDB code 2XB9) [19] that co-crystallized with three copies of the monomer in the asymmetric unit (designated as chains A, B and C). Chains A and B contain the inhibitor and chain C has citrate originating from the enzyme crystallization buffer as in PDB entry 2C4V [22]. The citrate-bound chain C was used to create a trimer after superimposition to chains A, F and G of *Mt*-DHQ2 in complex with its reaction product (PDB code 3N59) [23].

a) *Mycobacterium tuberculosis* DHQ2 model: Simulations were carried out using the enzyme coordinates found in the crystal structures of *Mt*-DHQ2 in complex with either (2*S*)-2-(perfluorobenzyl)-3-dehydroquinic acid (PDB code 4B6P, model 1) [24] or 3-dehydroshikimic acid (**2**) (PDB code 3N59 [23], model 2). These binary complexes contain a single *Mt*-DHQ2 molecule in the crystallographic asymmetric unit. For model 1, the trimer was created from this monomer after superimposition to chains B, C and E of *Mt*-DHQ2 in complex with 3-dehydroshikimic acid (**2**) (PDB code 3N59 [23]). Both models provided similar results for the first reaction coordinate but model 2 gave results for the dehydration step that were in better agreement with the experimental data.

For all the binary complexes, the substrate was manually docked into the active site. No crystallographic water molecules were maintained except for the conserved water molecule (WAT1). Addition of missing hydrogen atoms and computation of the protonation state of titratable groups at pH 7.0 were carried out using the H++ Web server [25,26], which relies on AMBER [21] force-field parameters and finite difference solutions to the Poisson-Boltzmann equation [27]  $\delta$  and/or  $\epsilon$  protonation was manually corrected for His102/His101 (dual) residues due to mechanistic considerations and on the basis of results from preliminary MD simulations. The charge distribution for each ligand studied was obtained by fitting the quantum mechanically calculated (RHF/6-31G\*\*/B3LYP/cc-pVTZ and IEF-PCM as solvation model) molecular electrostatic potential (MEP) of the geometry-optimized molecule to a point charge model, as implemented in Gaussian 03 and 09 [28]. The general AMBER force field, GAFF [29], was used to assign bonded and nonbonded parameters (parm03) to 3-dehydroquinic acid (**1**) atoms.

### Molecular dynamics simulations

Each molecular system was immersed in a truncated octahedron containing TIP3P water molecules [30] and Na<sup>+</sup> ions [31] to achieve electroneutrality (~30,000 water molecules and 16 Na<sup>+</sup> ions for *Hp*-DHQ2 model, ~14,200 water molecules and 15 Na<sup>+</sup> ions for model 1 of *Mt*-DHQ2 and ~20,000 water molecules and 21 Na<sup>+</sup> ions for model 2 of *Mt*-DHQ2). The *sander* and *pmemd* modules from the AMBER 12 [32] suite were used for the restrained and unrestrained MD simulations, respectively. Periodic boundary conditions were applied and electrostatic interactions were treated using the smooth particle mesh Ewald (PME) method [33] with a grid spacing of 1 Å. The cutoff distance for the nonbonded interactions was 9 Å, the SHAKE [34] algorithm was applied to all bonds and an integration step of 2.0 fs was used throughout. After an initial energy minimization of the water molecules and counterions, the system was heated to 300 K in 25 ps. Then it was equilibrated by first allowing the solvent to redistribute around the positionally restrained solute for 220 ps; after this time, the system was allowed to evolve freely without restraints so as to explore the mutual adaptation between ligand and enzymes. Snapshots from the 10 ns MD trajectory were collected every 20 ps.

### Hybrid QM/MM calculations

Classical molecular mechanics (MM) methods lack the ability to treat fundamentally electronic processes such as bond breaking/forming and charge fluctuations as a function of geometry [35] but it is possible to treat a subsection of the system quantum mechanically (QM) and use a coupling potential to connect the MM and QM regions. This hybrid QM/MM approach is seamlessly integrated in the *sander* module of AMBER 12 [32] and also includes a complete treatment of long-range electrostatics by using a QM/MM modified PME method. The QM region, defined by the *iqmatoms* keyword, encompassed the active site region where bonds are broken and formed, and the MM region included all the remaining protein, solvent atoms and counterions. Care was taken not to cut any polarizable bonds when defining the QM/MM boundary. Thus, the QM region contained the whole of the natural substrate and the following side chains residues: a) for the *Hp*-DHQ2 enzyme, the side chains of Asn10, Tyr22, Asn76, His82, His102, Thr104 (starting at C $\beta$ ), Arg17, Arg109 and Arg113 (starting at C $\delta$ ) of the central subunit, and Asp89\* of the neighboring subunit, and the structural (WAT1) and catalytic (WAT2) water molecules. b) for the *Mt*-DHQ2

enzyme, the side chains of Asn12, Tyr24, Asn75, His81, His101, Ser103 (starting at C $\beta$ ), Arg19, Arg108 and Arg112 (starting at C $\delta$ ) of the central subunit, and Asp88\* of the neighboring subunit and the structural water molecule WAT1. We selected the Density Functional Tight-binding method (DFTB) [36] and full electrostatic interactions between the QM charge density and the point charges of the MM atoms. To proceed with abstraction of the C2 axial hydrogen of **1**, the active site region of each solvated complex was subjected to a QM treatment during the course of a steered 10-ps MD simulation [37]. For *Hp*-DHQ2, three double reaction coordinates were defined for simultaneously shortening the (WAT2)-H-(Asp89\*)-O, (WAT2)-O-(Tyr22)-H and (Tyr22)-O-(**1**)-H<sub>2ax</sub> distances while stretching the (WAT2)-H-(WAT2)-O, (Tyr22)-H-(Tyr22)-O and (**1**)-H<sub>2ax</sub>-(**1**)-C2 bonds. For *Mt*-DHQ2, two double reaction coordinates were defined for simultaneously shortening the (Asp88\*)-O-(Tyr24)-H and (Tyr24)-O-(**1**)-H<sub>2ax</sub> distances while stretching the (Tyr24)-H-(Tyr24)-O and (**1**)-H<sub>2ax</sub>-(**1**)-C2 bonds (see Supplementary Online Data). Dehydration of **3** through His102<sup>+</sup> deprotonation was achieved by elongating the covalent bond between C1 and the OH oxygen during the course of a 5-ps MD simulation. The recommended force constant of 1000 kcal mol<sup>-1</sup> Å<sup>-2</sup> was used in all the reaction coordinates. During the QM/MM part of the MD simulation SHAKE for the QM region (keyword *qmshake*) was turned off and the integration step was reduced to 0.5 fs. The molecular graphics program PyMOL [38] was employed for visualization and model building. Interatomic distances were monitored by using the *ptraj* module in AMBER.

### Site-directed mutagenesis

The DNA sequences encoding variant forms of the *Mt*-DHQ2 in which Asp88\* was replaced by Asn and Asn12 was replaced by Ser were synthesized by the company GeneArt, and subcloned into the *E. coli* expression vector pET21a to yield plasmids pMUT308 and pMUT309, respectively.

### Protein production and purification of variant proteins

*E. coli* strain BL21 DE3 was transformed singly each plasmid and transformants were selected on 100  $\mu$ g mL<sup>-1</sup> ampicillin. The transformed strains were grown in drug supplemented Luria Broth at 37 °C until an attenuation of 0.6. Upon addition of 0.2 mg mL<sup>-1</sup> IPTG, incubation continued for a further 5 h. After harvesting by centrifugation, 50g cells producing D88N or N12S were disrupted by sonication in 50 mM potassium phosphate buffer pH 7.2, 1 mM DTT, 1 mM benzamidine (buffer 1) and centrifuged at 10,000xg for 40 min. The supernatants were applied to Q Sepharose columns (5 $\times$ 15 cm) and were washed with 1.0 L of 50 mM potassium phosphate buffer pH 7.2, 1 mM DTT (buffer 2). The columns were then eluted with a NaCl gradient (1 L, 0 to 1 M) in buffer 2 collecting 10 mL fractions. DHQ2-encoding fractions were identified by direct assay and SDS PAGE and pooled appropriately. After dialysis into 2 $\times$ 5 L changes of 50 mM potassium phosphate buffer pH 6.6, 1 mM DTT (buffer 3), the D88N and N12S variant proteins were loaded onto hydroxyapatite (5 $\times$ 7.5 cm) columns and washed with 500 mL of buffer 4. The columns were then eluted with a gradient of potassium phosphate buffer pH 6.6 containing 1 mM DTT (1 L, 50 to 400 mM). For N12S protein, DHQ2-encoding fractions were identified by direct enzyme assay and SDS PAGE and pooled appropriately, yielding approximately 300 mg of N12S variant protein of at least 95% purity. For the D88N protein, DHQ2-encoding fractions were identified by SDS PAGE and pooled appropriately. D88N protein was then dialysed against 2 $\times$ 5 L changes of 50 mM potassium phosphate buffer pH 6.0, 1 mM DTT (buffer 4) and chromatographed on a MONO Q 10/100 GL FPLC column. Following a 3 column volume wash, protein was eluted by a 20 column volume 0–1.0 M NaCl gradient in buffer 4. The flow rate was 0.5 mL min<sup>-1</sup> and the fraction size was 0.5 mL. This procedure yielded approximately 60 mg of D88N protein at greater than 95% purity.

### Enzyme assay conditions

Wild-type *Hp*-DHQ2, *Sc*-DHQ2 and *Mt*-DHQ2 enzymes were purified as described previously [39,40]. Concentrated solutions of *Hp*-DHQ2 (6.4 mg mL<sup>-1</sup>) were stored in potassium phosphate buffer (50 mM, pH 7.2), DTT (1 mM) and NaCl (150 mM) and *Sc*-DHQ2 (5 mg mL<sup>-1</sup>) in Tris HCl (20 mM, pH 7.5) and DTT (0.5 mM), respectively. When required for assays, aliquots of the enzyme stocks were diluted in water and buffer and stored on ice. DHQ2 was assayed in the forward direction by monitoring the increase in absorbance at 234 nm in the UV spectrum due to the absorbance of the enone-carboxylate chromophore of 3-dehydroshikimic acid (**2**) ( $\epsilon$ /M<sup>-1</sup> cm<sup>-1</sup> 12 000). Standard assay conditions were pH 7.0 at 25 °C in Tris HCl

or Tris HOAc (50 mM). Each assay was initiated by addition of the substrate. Solutions of 3-dehydroquinic acid (**1**) were calibrated by equilibration with DHQ2 and measurement of the change in the UV absorbance at 234 nm due to the formation of the enone-carboxylate chromophore of 3-dehydroshikimic acid (**2**). The initial rates at fixed enzyme and substrate concentrations (0.2–5  $K_m$ ) were measured. Kinetic constants were evaluated using the GraFit 5 program (Erithacus Software Ltd.).

### Solvent isotope effects

The solvent isotope effects on  $V_{max}$  and  $V_{max}/K_m$  were determined in  $^1\text{H}$  and  $^2\text{H}$  Tris-chloride buffer solutions at pH or  $p^2\text{H}$  7.0 over a range of substrate concentrations (0.2–5  $K_m$ ). Deuterated Tris-chloride buffer was prepared from commercially available 1 M Tris-d11 solution in  $^2\text{H}_2\text{O}$  with 98% deuterium content,  $^2\text{H}_2\text{O}$  with 99.9% deuterium content and deuterium chloride solution 35 wt.% in  $^2\text{H}_2\text{O}$  with 99% deuterium content, all purchased from Sigma-Aldrich. The  $p^2\text{H}$  was calculated by adding 0.4 to the electrode reading.

### Proton inventory

The proton inventory studies were performed in isotopic buffers at a range of molar fraction of deuterium ( $n$ ) from 0–1 under  $V_{max}$  conditions (see Supplementary Online Data) [41]. Buffers in isotopic mixtures of water were prepared by volumetric mixing. The difference in the density between the two buffer stock solutions in  $^1\text{H}_2\text{O}$  and  $^2\text{H}_2\text{O}$  and the isotopic dilution of the  $^2\text{H}_2\text{O}$  solutions was not considered when calculating the molar fraction of deuterium present.

## RESULTS

### Computational Studies

#### Michaelis complexes

In the *Hp*-DHQ2 Michaelis complex, after system equilibration a water molecule from the bulk solvent (henceforth WAT2) entered the active site and remained fixed between the side chain of Asp89\* and Tyr22 (Figure 1), with one of its protons engaged in a hydrogen bond with a carboxylate oxygen of Asp89\* and one of its oxygen's lone pairs accepting a hydrogen bond from the guanidinium group of Arg109. As a result, the orientation of the other lone pair was locked towards the phenolic proton of Tyr22 whereas the phenolic oxygen was placed at an average distance of 2.0 Å from the axial hydrogen at C2 of **1**. Besides, a number of strong and highly directional electrostatic interactions were established (Figure 2 and Supplementary Online Data). Most importantly, the Asp89\* carboxylate oxygen binding WAT2 was firmly held in position by two hydrogen bonds established with the guanidinium group of Arg113 whereas the other one was hydrogen bonded to the guanidinium group of the essential Arg17 and the amide of Asn10, which in turn was held in place by WAT1. We considered this architecture of the *Hp*-DHQ2 Michaelis complex, which remained stable during the MD simulation, as ideal for initiation of the reaction coordinate using the hybrid QM/MM method.

In contrast, in the simulation of the Michaelis *Mt*-DHQ2 complex, a fixed water molecule between Asp88\* and Tyr24 was not observed during the simulation. Instead, Tyr24 was found to be sufficiently close to Asp88\* to effect a direct deprotonation. We believe that these differences arise from the distinct flexibilities and motions of the two Tyr-containing loops, which we have studied in detail in the two enzymes over the years [24]. Compared to *Hp*-DHQ2, this system proved to be somewhat less stable. In particular, significant changes in the relative distances between the phenolic oxygen of Tyr24 and the C2 axial hydrogen atom of **1** and between the Asp88\* carboxylate and the guanidinium group of Arg19 were observed. This finding is not surprising considering that this carboxylate is held in place by the arginine but at the same time needs to get close enough to the tyrosine to abstract the proton so that the ensuing tyrosinate will remove the C2 axial hydrogen of **1** (Figure 2). Apart from these differences, the substrate interacts with the enzyme active site making use of the same type of strong hydrogen bonding interactions with the key residues of the active site (see Supplementary Online Data). Several snapshots from the simulation were chosen as starting points to study the reaction mechanism using the hybrid QM/MM method, namely those with the shortest distances between the phenolic oxygen of Tyr24 and the C2 axial hydrogen and between the carboxylate oxygen of Asp88\* and the phenolic hydrogen of Tyr24.

## Enolate formation

For *Hp*-DHQ2, the first reaction coordinate led to a neutral carboxylic group in the side chain of Asp89\* upon capture of one of the WAT2 protons and the regeneration of WAT2 with the proton transferred from Tyr22, which in turn abstracted the C2 axial hydrogen of **1** and maintained the  $\pi$ -cation interaction with the planar guanidinium moiety of Arg17 (Figures 1 and 2). This positively charged group also stabilized the enolate intermediate **3** through electrostatic interactions. In the resulting complex, WAT2 accepted a hydrogen bond from Arg109 and the protonated Asp89\* had its orientation stabilized by Arg17 as well. The enolate intermediate **3** displayed a C-OH bond at C1 that was  $\sim 0.1$  Å longer than the rest of C-OH bonds present in the substrate, and one of the lone pairs of the oxygen in the C1 hydroxyl group was oriented towards the N $\delta$  proton of His102. The electrostatic interaction between the hydrogen of the leaving OH and the side chain of Asn76 also favors this geometry as it locks the orientation of the oxygen lone pair.

For *Mt*-DHQ2, direct removal of the Tyr24 proton by Asp88\* followed by abstraction of the C2 axial hydrogen of **1** by the tyrosinate led to enolate **3** (Figures 2 and 3). For both enzymes, the enolate was engaged in a strong electrostatic interaction with the guanidinium group of Arg19/Arg17 and also in a good hydrogen bond with WAT1 (see Supplementary Online Data). Moreover, when the DHQ2/**3** binary complex was subjected to 10 ns of MD simulation, those interactions remain stable during the whole simulation and in addition, after 0.6 ns of MD simulation, the neutral Asp89\*/Asp88\* rotates to establish a hydrogen bond interaction with the enolate **3** (1.5–2.0 Å) that remains stable after that (see Supplementary Online Data). However, under similar conditions, the DHQ2/**4** binary complex in which the enol intermediate **4** is formed after deprotonation of the neutral Asp89\*/Asp88\*, the guanidinium group of Arg19/Arg17 and Asp89\*/Asp88\* rapidly moves away from enol **4** ( $> 3$  Å, see Supplementary Online Data).

In line with previous proposals [16,20], our studies also suggested that enolate formation is the rate-determining step as the calculated energy barrier obtained for this reaction was 13.9 and 19.7 kcal mol<sup>-1</sup>, which compares well with the experimentally measured and rate-determining barrier of 17.4 and 16.5 kcal mol<sup>-1</sup> [42], for *Hp*-DHQ2 and *Mt*-DHQ2, respectively.

## Enolate dehydration and product formation

Upon elongation of the covalent bond between C1 and the OH oxygen, the natural evolution of both systems was the spontaneous abstraction of the His102/His101 N $\delta$  proton and formation of the leaving water molecule. As a result, 3-dehydroshikimic acid (**2**) was formed and the carboxylate binding site geometry was structurally reorganized because the C1 hybridization changed from tetrahedral sp<sup>3</sup> to planar sp<sup>2</sup>. As will be discussed below, this change does not appear to be the driving force for product release, contrary to earlier suggestions [17]. The calculated energy barriers for the dehydration step were 7.7 and 10.3 kcal mol<sup>-1</sup> for *Hp*-DHQ2 and *Mt*-DHQ2, respectively.

## Product release

In order to complete the catalytic cycle, two possibilities were then considered: (1) proton transfer from the neutral carboxylic acid of Asp89\*/Asp88\* to a water molecule (WAT2 in *Hp*-DHQ2), followed by product release, or (2) the reverse order. However, when the resulting DHQ2/**2** binary complexes containing Asp89\*/Asp88\* as a carboxylate were subjected to 10 ns of standard MD simulation, no significant motions of the reaction products were observed (see Supplementary Online Data).

Despite the aforementioned change of hybridization at the C1 position in **2**, the carboxylate group maintained four strong hydrogen bonds with the enzyme, namely, with the main chain NH groups of Thr104 and Leu103 and the side chains of Asn76 and Thr104. Furthermore, the enzyme snugly accommodated a water molecule, which was exchanged with other water molecules from the bulk solvent, in this part of the active site by bridging His101 and Asn76 residues (see Supplementary Online Data). Both facts provided a very stable *Hp*-DHQ2/**2** binary complex during the 10 ns of MD simulation. However, under the same conditions, in the *Mt*-DHQ2/**2** complex containing the neutral Asp88\* side chain, the guanidinium moiety of Arg112 was expelled from the active site (up to  $\sim 6.4$  Å) and that of the essential Arg19 moved toward the C1 carboxylate (Figure 4b). In order to evaluate whether these structural changes could be responsible for triggering product release from the active site and to get further details of the process, this complex was further subjected to 40 ns of MD simulation at 77 °C (Figure 4).

The results obtained showed that the large displacement of residues 101–112 and 126–130 (by up to 6.4 and 8.3 Å, respectively) containing H2 and H3 3<sub>10</sub> helix and the conserved Arg108, Glu109 and Arg112 residues and the electrostatic interaction of the guanidinium group of Arg19 by pulling on the carboxylate group in **2**, were responsible for breaking the favorable interactions of **2** with the active site residues, particularly the stabilizing hydrogen bonds with Asn76, Thr104 and Leu103 residues. As a result, the product would then be released from the active site through the cavity lined by the side chains of Leu16 and Ile102. With the product on the way out of the active site, the large loop containing H2 and H3 3<sub>10</sub> helix recovers its position. Moreover, the proton transfer reaction from the neutral carboxylic acid of Asp89\*/Asp88\* to a water molecule might take place after disruption of the product interactions with the C1 carboxylate binding site.

### Biochemical Studies

To get further evidence supporting the key role of Asp89\*/Asp88\* in the generation of the catalytic tyrosinate and the conserved asparagine, Asn10/Asn12, in the stabilization of the protonated Asp89\*/Asp88\* during the enolate generation step rather than in engagement in a hydrogen bonding interaction with the conserved water molecule WAT1, site-directed mutagenesis experiments were carried out. Moreover, to demonstrate that deprotonation of the essential tyrosine in *Hp*-DHQ2 is assisted by a water molecule, solvent isotope effects and proton inventory studies were performed.

Asn12 and Asp88\* of *Mt*-DHQ2 were replaced by serine and asparagine residues, respectively. The kinetic parameters of both variants, N12S and D88N, were determined and compared to those measured for the wild-type enzyme (Table 1). When the kinetic parameters of both enzyme variants were determined at pH 7 and compared to those measured for the wild-type enzyme, the N12S variant proved to have a similar  $K_m$  value but a much smaller  $k_{cat}$  whereas the activity of the D88N variant was dramatically reduced to less than 0.05% that of the wild-type protein. Evaluation of the D88N variant in a pH range between 7.0 and 9.0 and using large amounts of the variant protein (20 units) showed some residual enzymatic activity at pH of 7.8–8.2 whereas no uncatalyzed reaction was observed under these conditions. At pH 8.0, a  $K_m$  more than one order of magnitude larger than the wild-type enzyme and a severely decreased  $k_{cat}$  value were found (Table 1). In addition, in order to discard that the inactivity of D88N variant was due to a destabilization of its quaternary structure, this protein was analyzed by dynamic light scattering and by size exclusion chromatography and compared with the N12S variant (see Supplementary Online Data). The results showed that the N12S variant is a dodecamer whereas D88N is a trimer.

Solvent isotope effects were determined on  $V_{max}$  and  $V_{max}/K_m$  for *Hp*-DHQ2 and *Sc*-DHQ2 (Table 2). These enzymes proved to have a large solvent isotope effect on  $V_{max}$  and  $V_{max}/K_m$ , being the latter smaller than for *An*-DHQ2. Moreover, in order to get more information regarding the number of hydrogenic sites contributing to the observed solvent isotope effects on  $V_{max}$  proton inventory studies were carried out (see Supplementary Online Data). The results showed that, for *Hp*-DHQ2, there are at least two protons involved in the reaction, one associated with the elimination step and another/others involved in tyrosine deprotonation, in agreement with the QM/MM proposal that a water molecule assists in Tyr deprotonation. On the contrary, for *Mt*-DHQ2, there is a single proton contributing to the observed solvent [13], showing that this reaction is not assisted by a water molecule.

### DISCUSSION

The catalytic mechanism of *H. pylori* and *M. tuberculosis* DHQ2, the third enzyme of the shikimic acid pathway and an essential enzyme in these pathogenic bacteria, has been studied in atomic detail by means of a hybrid QM/MM approach. A combination of standard MD on a catalytically active trimer in aqueous solution, using the MM force field AMBER in a box of water molecules and counterions, and QM DFT inside the active site during the course of several steered MD trajectories has allowed us to gain unique insight into the reaction mechanism.

Given that the minimum catalytic unit of DHQ2 was experimentally determined to be a trimer [13], the latter was used for modeling and simulating first the Michaelis complexes with 3-dehydroquinic acid (**1**) and then the reaction to form 3-dehydroshikimic acid (**2**). In addition, previous MD simulation studies suggested that the essential arginine side chain must be close to the essential tyrosine side chain for catalysis because under this arrangement, it controls the appropriate orientation of the tyrosine roughly perpendicular to the cyclohexane ring of the natural substrate by a cation- $\pi$  interaction and contributes to decreasing the pKa of



the phenol ring [19]. Taking this into account, crystal structures of both enzymes with the two essential residues, Arg17/Arg19 and Tyr22/Tyr24, together pointing towards the active site were chosen. In particular, the *Hp*-DHQ2 Michaelis complex was generated using the enzyme geometries found in chain C of PDB 2XB9 [19]. For the *Mt*-DHQ2 Michaelis complex, two models were evaluated using the enzyme coordinates found in the crystal structures of *Mt*-DHQ2 in complex with either (2*S*)-2-(perfluorobenzyl)-3-dehydroquinic acid (PDB code 4B6P, model 1) [24] and 3-dehydroshikimic acid (**2**) (PDB code 3N59 [23], model 2). Both models provided similar results for the first reaction coordinate but model 2 gave results for the dehydration step that were in better agreement with the experimental data.

For the deprotonation of the essential tyrosinate and on the basis of the MD simulation studies carried out with the *Hp*-DHQ2 and *Mt*-DHQ2 Michaelis complexes, Asp89\*/Asp88\* and the essential Arg17/Arg19 were considered as the general base involved in the enzymatic mechanism. However, the latter possibility was discarded based on the extremely high calculated energy barriers required (over 40 kcal mol<sup>-1</sup>, see Supplementary Online Data) compared with the obtained with Asp89\*/Asp88\* as the general base (see Results section). Furthermore, for the generation of the reaction intermediate, two possible mechanisms were assessed, thus, a two-step mechanism involving the generation of the catalytic tyrosinate followed by abstraction of the C2 axial hydrogen of **1** by the tyrosinate and a single-step mechanism. The lower calculated energy barrier obtained for the latter (13.9 and 19.7 kcal mol<sup>-1</sup>), which compares well with the experimentally measured and rate-determining barrier of 17.4 and 16.5 kcal mol<sup>-1</sup> [42], for *Hp*-DHQ2 and *Mt*-DHQ2, respectively, suggested a concerted mechanism for the generation of the reaction intermediate. For both enzymes, the enolate **3** was engaged in a strong electrostatic interaction with the guanidinium group of Arg19/Arg17 and also in a good hydrogen bond with WAT1 and the neutral Asp89\*/Asp88\*. The results from the MD simulation studies carried out with the DHQ2/3 and DHQ2/4 binary complexes suggested that the strong and efficient electrostatic interaction between enolate **3** with the guanidinium group of Arg19/Arg17 is key in the stabilization of this charged intermediate.

The alternative mechanism in which the neutral Asp88\*/Asp89\* could protonate the enolate intermediate **3** to afford enol intermediate **4** was also studied. However, the energy barrier of the subsequent dehydration step was calculated to be much higher. The formation of an enol **4** vs enolate **3** was also previously studied by Blomberg *et al.* [20]. In this case, formation of **4** was studied upon proton transfer from WAT1 assuming that the side chain of a conserved asparagine (Asn10/Asn12) would be deprotonated by this water molecule as previously suggested by Laphorn *et al.* [17]. The authors considering the high energy required to deprotonate the side chain of an asparagine and on the basis of results from QM calculations, suggested the formation of enolate intermediate **3** because of its significantly lower energy compared to that of **4**. We have found the same results as Blomberg *et al.* [20] considering a proton transfer from neutral Asp88\*/Asp89\*.

Due to the limited repertoire of titratable groups, proteins can activate these residues as nucleophiles, electrophiles, general-base or general-acid catalysts as a direct result of the microenvironment created around the functional group [43]. There are numerous examples of ionizable groups with highly perturbed pKa values [44]. For instance, a good example of an enzyme with a perturbed Tyr and Asp in the active site, as might occur for the DHQ2 enzyme, is ketosteroid isomerase, in which an active site aspartate (Asp38, pKa 9.5) acts as a general base and removes the ketone  $\alpha$ -hydrogen atom of the substrate (pKa 12.7) [45-47]. The reaction catalysed by this enzyme also involves a tyrosine with a highly perturbed pKa value [43]. In the DHQ2 enzymes, the proximity of the conserved Arg108/Arg109 residue and the cation- $\pi$  interaction with the essential Arg19/Arg17 would be expected to have a significant effect on the pKa of Tyr22/Tyr24, as was previously suggested by Laphorn *et al.* [17]. This fact was the basis of several substrate analogs developed in our research group that still have the axial hydrogen to be removed by the enzyme. These (2*R*)-2-benzyl-3-dehydroquinic acids proved to be nanomolar reversible competitive DHQ2 inhibitors [19, 24]. The benzyl group of those inhibitors prevent the appropriate orientation of the catalytic tyrosine of the loop and disrupts its basicity by expelling the Arg19/Arg17 side-chain from the active site. These results were supported by the resolution of the crystal structure of several binary complexes (PDB 2XB8, 2XB9, etc.). These results allowed us to solve the structure of the loop in *Mt*-DHQ2 for the first time because its flexibility is dramatically reduced by the inhibitors.

To obtain further evidence supporting the role of Asp88\*/Asp89\* as a general base in the enzymatic mechanism reported here, the catalytic activity of D88N variant was studied in a pH range of 7.0–9.0 to show that the dehydration reaction takes place if the catalytic tyrosinate is generated, by the buffer in this case. It is important to highlight that DHQ2 enzymes show a pH-dependence for  $K_m$  and  $k_{cat}$ , which increases sharply above pH 9 [13]. The maximum range of activity was found between pH 7.8–8.2. Using large large quantities of D88N variant and at pH 8.0, the  $K_m$  and  $k_{cat}$  values were determined and the enzymatic conversion was also monitored by  $^1\text{H}$  NMR (see Supplementary Online Data). The significant reduced affinity for the substrate is not surprising considering that Asp88\* is also involved in the recognition of the C4 hydroxyl group of the substrate. In addition, the fact that the D88N variant is a trimer rules out the possibility of a major structural change at the monomer interfaces that we have studied. Moreover, considering that 0.5 M guanidinium chloride dissociates wild-type *Mt*-DHQ2 into trimers that retain their enzymatic activity [13], it is unlikely that the observed loss of activity in the D88N variant is a consequence of its trimeric state. Therefore, we think that the site-directed mutagenesis studies and the pH effect studies with D88N variant demonstrate the essentially of Asp88\* as a general base in the enzymatic mechanism, in line with the results from our QM/MM studies. In addition, the decreased catalytic activity of the N12S variant enzyme also confirms that the conserved Asn12 is likely to have a role beyond establishing a hydrogen-bonding interaction with the conserved water molecule.

On the other hand, similar substrate isotope effects on  $V_{max}$  and  $V_{max}/K_m$  reported earlier for *Mt*-DHQ2 and *An*-DHQ2 suggested that hydrogen abstraction is, at least partially, the rate-determining step of the reaction and that it occurs at or before the highest energy barrier leading to the transition state(s) [16]. However, important differences between both enzymes were found in the solvent isotope and proton inventory studies reported. For *Mt*-DHQ2, no solvent isotope effect on  $V_{max}/K_m$  was obtained and a single exchangeable proton in the reaction was found that is associated with the elimination step. On the contrary, for *An*-DHQ2, a solvent isotope effect on  $V_{max}/K_m$  was observed and at least two protons transferred from water or two water-exchangeable groups were found to be involved in the reaction. These previously reported experimental results were in nice agreement with our theoretical findings and we thought they might well relate to the catalytic water molecule involved in the reaction. The results from solvent isotope effects and proton inventory studies reported here showed that, for *Hp*-DHQ2, there are at least two protons involved in the reaction, one associated with the elimination step and another/others involved in tyrosine deprotonation, in agreement with the QM/MM proposal that a water molecule assists in Tyr deprotonation. For *Mt*-DHQ2, on the contrary, a single proton contributing to the observed effect confirms that the reaction catalyzed by this enzyme is not assisted by a water molecule.

In conclusion, the results reported here strongly support a catalytic mechanism consisting of abstraction of the axial hydrogen of the substrate by the essential Tyr22/Tyr24, enolate formation that is stabilized by the positively charged side chain of the essential Arg17/Arg19, and  $\beta$ -dehydration mediated by the conserved His102/His101 acting as a proton donor. No evidence is found for an earlier suggestion of proton transfer from the side chain of Asn10/Asn12 to a water molecule but, instead, a previously unnoticed key role for Asp89\*/Asp88\* from a symmetry-related neighboring enzyme subunit in the reaction is proposed. The essentiality of this residue is supported by results from site-directed mutagenesis. Thus, Asp89\*/Asp88\* is not only involved in the formation of intersubunit electrostatic interactions with the conserved Arg113/Arg112 and also with the C4 hydroxyl group of the substrate, as previously suggested. Asp89\*/Asp88\* turns out to be the residue responsible for the deprotonation of the essential Tyr22/Tyr24 to afford the catalytic tyrosinate, which triggers the enzymatic process. For the *Hp*-DHQ2 enzyme, this deprotonation takes place through the assistance of a water molecule, while for *Mt*-DHQ2, Tyr24 is directly deprotonated by Asp88\*. The participation of a water molecule in this deprotonation reaction is supported by solvent isotope effects and proton inventory studies. Site-directed mutagenesis studies also confirmed the important role of the conserved Asn10/Asn12 in stabilizing the neutral form of Asp89\*/Asp88\* in the enolate generation step as the QM/MM studies suggested.

Our simulations also pinpoint the essential Arg17/Arg19 as the residue that most contributes to expel the product from the active site following a large conformational change involving residues 101–112 and 126–130 which contain the conserved Arg108, Glu109 and Arg112 (for *Mt*-DHQ2). These findings reveal a previously unknown key role of these essential and conserved residues and provide details of the required

motions for the catalytic turnover. Moreover, the key catalytic role played by Asp89\*/Asp88\* suggests that blocking this residue could be a useful strategy in the structure-based design of novel inhibitors of this attractive bacterial target.

## AUTHOR CONTRIBUTION

Claire Coderch and Federico Gago designed and performed the initial QM/MM studies on the *Hp*-DHQ2 enzyme. Emilio Lence performed the QM/MM studies on the *Hp*-DHQ2 enzyme and studied the product release. Antonio Peón did the QM/MM studies on the *Mt*-DHQ2 enzyme. Heather Lamb and Alastair R. Hawkins performed the protein expression, purification and biophysical studies of the *Mt*-DHQ2 N12S and D88N variant proteins. Concepción González-Bello carried out the enzymatic assays, solvent isotope effects, proton inventory and NMR studies. Concepción González-Bello and Federico Gago supervised the whole work and wrote the paper.

## ACKNOWLEDGMENTS

We are grateful to the Centro de Supercomputación de Galicia (CESGA) and Centro Nacional de Supercomputación (BSC-CNS) for generous allowances on Finis Terrae and Mare Nostrum computers. We thank Prof. A. J. Laphorn for providing the *Sc*-DHQ2 enzyme for the solvent isotope studies and Dr. Pedro A. Sánchez-Murcia for help and many fruitful discussions.

## FUNDING

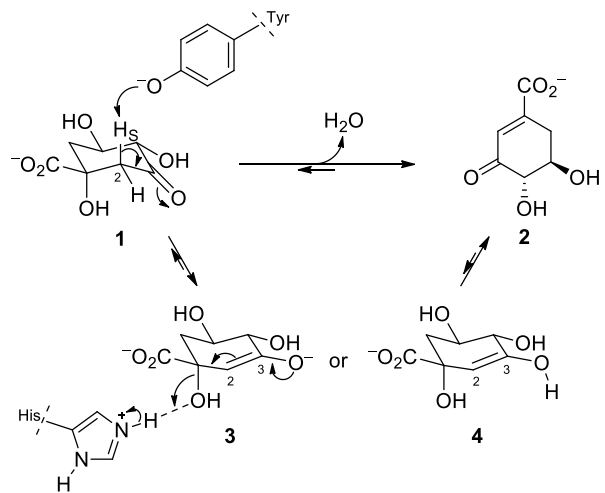
Financial support from the Comunidad de Madrid (S2010-BMD-2457 to F.G.), Xunta de Galicia (10PXIB2200122PR and GRC2010/12 to C.G.-B.) and the Spanish Ministry of Science and Innovation (SAF2009-13914-C02-02 to F.G. and SAF2010-15076 to C.G.-B.) is gratefully acknowledged. C.C. and A.P. thank the Spanish Ministry of Science and Innovation for their respective FPU fellowships.

## REFERENCES

- 1 Cosgrove, S. E., and Carmeli, Y. (2003) The impact of antimicrobial resistance on health and economic outcomes. *Clin. Infect. Dis.* **36**, 1433–1437.
- 2 Zaoutis, T. E. (2009) Antibiotic resistance: who will pay the bills. *Clin. Infect. Dis.* **49**, 1185–1186.
- 3 Fischbach, M. A., and Walsh, C. T. (2009) Antibiotics for emerging pathogens. *Science* **325**, 1089–1093.
- 4 Lewis, K. (2013) Platforms for antibiotic discovery. *Nat. Rev. Drug Discovery* **12**, 371–387.
- 5 Busch, K. *et al.* (2011) Tackling antibiotic resistance. *Nature Rev. Microbiol.* **9**, 894–896.
- 6 Lonsdale, R., Harvey, J. N., and Mulholland, A. J. (2012) A practical guide to modelling enzyme-catalysed reactions. *Chem. Soc. Rev.* **41**, 3025–3038.
- 7 van der Kamp, M. W., and Mulholland, A. J. (2013) Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* **52**, 2708–2728.
- 8 van der Kamp, M. W., Chaudret, R., and Mulholland, A. J. (2013) QM/MM modelling of ketosteroid isomerase reactivity indicates that active site closure is integral to catalysis. *FEBS J.* **280**, 3120–3131.
- 9 Lamichhane, G., Freundlich, J. S., Ekins, S., Wickramaratne, N., Nolan, S. T., and Bisha, W. R. (2011) Essential metabolites of *Mycobacterium tuberculosis* and their mimics. *MBio* **2**, e00301-10.
- 10 Abell, C. *Enzymology and Molecular Biology of the Shikimate Pathway*, in *Comprehensive Natural Products Chemistry*, Ed. U. Sankawa, Pergamon, Elsevier Science Ltd., Oxford, 1999, p 573.
- 11 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.
- 12 Data base for essential genes in bacteria see [www.essentialgene.org](http://www.essentialgene.org).
- 13 Price, N. C., Boam, D. J., Kelly, S. M., Duncan, D., Krell, T., Gourley, D. G., Coggins, J. R., Virden, V., and Hawkins, A. R. (1999) The folding and assembly of the dodecameric type II dehydroquinases. *Biochem. J.* **338**, 195–202.

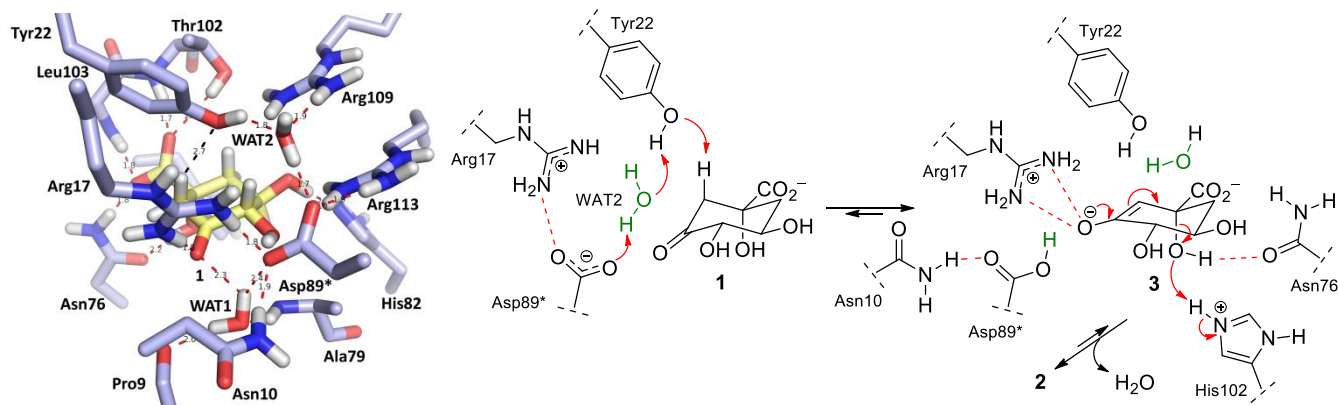
- 14 Krell, T., Pitt, A. R., and Coggins, J. R. (1995) The use of electrospray mass spectrometry to identify an essential arginine residue in type II dehydroquinases. *FEBS Lett.* **360**, 93–96.
- 15 Krell, T., Horsburgh, M. J., Cooper, A., Kelly, S. M., and Coggins, J. R. (1996) Localization of the active site of type II dehydroquinase. Identification of a common arginine-containing motif in the two classes of dehydroquinases. *J. Biol. Chem.* **271**, 24492–24497.
- 16 Harris, J., González-Bello, C., Kleanthous, C., Coggins, J. R., Hawkins, A. R., and Abell, C. (1996) Evidence from kinetic isotope studies for an enolate intermediate in the mechanism of type II dehydroquinases. *Biochem. J.* **319**, 333–336.
- 17 Roszak, A. W., Robinson, D. A., Krell, T., Hunter, I. S., Frederickson, M., Abell, C., Coggins, J. R., and Laphorn, A. J. (2002) The structure and mechanism of the type II dehydroquinase from *Streptomyces coelicolor*. *Structure* **10**, 493–503.
- 18 Numbering corresponds to DHQ2 from *H. pylori*.
- 19 Peón, A., Otero, J. M., Tizón, L., Prazeres, V. F. V., Llamas-Saiz, A. L., Fox, G. C., van Raaij, M. J., Lamb, H., Hawkins, A. R., Gago F., Castedo, L., and González-Bello, C. (2010) Understanding the key factors that control the inhibition of type II dehydroquinase by (2*R*)-2-benzyl-3-dehydroquinic acids. *ChemMedChem* **5**, 1726–1733.
- 20 Blomberg, L. M., Mangold, M., Mitchell, J. B. O., and Blumberger, J. (2009) Theoretical study of the reaction mechanism of *Streptomyces coelicolor* type II dehydroquinase. *J. Chem. Theory Comput.* **5**, 1284–1294.
- 21 Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
- 22 Robinson, D. A., Stewart, K. A., Price, N. C., Chalk, P. A., Coggins, J. R., and Laphorn, A. J. (2006) Crystal structures of *Helicobacter pylori* type II dehydroquinase inhibitor complexes: new directions for inhibitor design. *J. Med. Chem.* **49**, 1282–1290.
- 23 Dias, M. V. B., Snee, W. C., Bromfield, K. M., Payne, R. J., Palaninathan, S. K., Ciulli, A., Howard, N. I., Abell, C., Sacchettini, J. C., and Blundell, T. L. (2011) Structural investigation of inhibitor designs targeting 3-dehydroquinase dehydratase from the shikimate pathway of *Mycobacterium tuberculosis*. *Biochem. J.* **436**, 729–739.
- 24 Lence, E., Tizón, L., Otero, J. M., Peón, A., Prazeres, V. F. V., Llamas-Saiz, A. L., Fox, G. C., van Raaij, M. J., Lamb, H., Hawkins, A. R., and González-Bello, C. (2013) Mechanistic basis of the inhibition of type II dehydroquinase by (2*S*)- and (2*R*)-2-benzyl-3-dehydroquinic acids. *ACS Chem. Biol.* **8**, 568–577.
- 25 Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S., and Onufriev, A. (2005) H<sup>++</sup>: a server for estimating pK<sub>a</sub>s and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **33** (Web Server issue), W368–371.
- 26 <http://biophysics.cs.vt.edu/H++>
- 27 Bashford, D., and Karplus, M. (1990) pK<sub>a</sub>s of ionizable groups in proteins: atomic detail from a continuum electrostatic mode. *Biochemistry* **29**, 10219–10225.
- 28 Gaussian 09, Revision A.2, Frisch, M. J. *et al.* Gaussian, Inc., Wallingford CT, 2009.
- 29 Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general Amber force field. *J. Comput. Chem.* **25**, 1157–1174.
- 30 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
- 31 Aqvist, J. (1990) Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.* **94**, 8021–802.
- 32 <http://ambermd.org/>
- 33 Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An *N*·log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.

- 34 Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (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.
- 35 Seabra, G. M., Walker, R. C., Elstner, M., Case, D. A., and Roitberg, A. E. (2007) Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package. *J. Phys. Chem. A* **111**, 5655–5664
- 36 Cui, Q., Elstner, M., Kaxiras, E., Frauenheim, T., and Karplus, M. A. (2000) A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *J. Phys. Chem. B* **105**, 569–585.
- 37 Crespo, A., Marti, M. A., Estrin, D. A., and Roitberg, A. E. (2005) Multiple-steering QM–MM calculation of the free energy profile in chorismate Mutase. *J. Am. Chem. Soc.* **127**, 6940–6941.
- 38 DeLano, W. L. The PyMOL Molecular Graphics System. (2008) DeLano Scientific LLC, Palo Alto, CA, USA. <http://www.pymol.org>
- 39 Gourley, D. G., Coggins, J. R., Isaacs, N. W., Moore, J. D., Charles, I. G., and Hawkins, A. R. (1994) Crystallization of a type II dehydroquinase from *Mycobacterium tuberculosis*. *J. Mol. Biol.* **241**, 488–491.
- 40 Prazeres, V. F. V., Sánchez-Sixto, C., Castedo, L., Shuh, S. W., Lamb, H., Hawkins, A. R., Cañada, F. J., Jiménez-Barbero, J., and González-Bello, C. (2008) Competitive inhibitors of *Helicobacter pylori* type II dehydroquinase: synthesis, biological evaluation, and NMR studies. *ChemMedChem* **3**, 756–770.
- 41 Schown, K. B., and Schowen, R. L. (1982) Solvent isotope effects of enzyme systems. *Methods Enzymol.* **87**, 551–606.
- 42 Calculated from  $k_{cat}$  values of  $0.9\text{ s}^{-1}$  and  $5.2\text{ s}^{-1}$  for *Hp*-DHQ2 and *Mt*-DHQ2, respectively, using transition-state theory. Data from: Evans, L. D. B., Roszak, A. W., Noble, L. J., Robinson, D. A., Chalk, P. A., Matthews, J. L., Coggins, J. R., Price, N. C., Laphorn, A. J. (2002) Specificity of substrate recognition by type II dehydroquinases as revealed by binding of polyanions. *FEBS Lett.* **530**, 24–30.
- 43 Harris, T. K. and Turner, G. J. (2002) Structural basis of perturbed pKa values of catalytic groups in enzyme active sites. *IUBMB Life* **53**, 85–98.
- 44 Schwans, J. P., Sunden, F., Gonzalez, A., Tsai, Y., and Herschlag, D. (2013) Uncovering the determinants of highly perturbed tyrosine pKa in the active site of ketosteroid isomerase. *Biochemistry* **52**, 7840–7855.
- 45 Wu, Z. R., Ebrahimian, S., Zawrotny, M. E., Thornburg, L. D., Perez-Alvarado, G. C., Brothers, P., Pollack, R. M., and Summers, M. F. (1997) Solution structure of 3-oxo- $\Delta^5$ -steroid isomerase. *Science* **276**, 415–418.
- 46 Schwans, J. P., Sunden, F., Lassila, J. K., Gonzalez, A., Tsai, Y., and Herschlag, D. (2013) Use of anion-aromatic interactions to position the general base in the ketosteroid isomerase active site. *PNAS* **110**, 11308–11313.
- 47 Cho, H.-S., Choi, G., Choi, K. Y., and Oh, B.-H. (1998) Crystal structure and enzyme mechanism of  $\Delta^5$ -ketosteroid isomerase pseudomonas testosterone. *Biochemistry* **37**, 8325–8330.



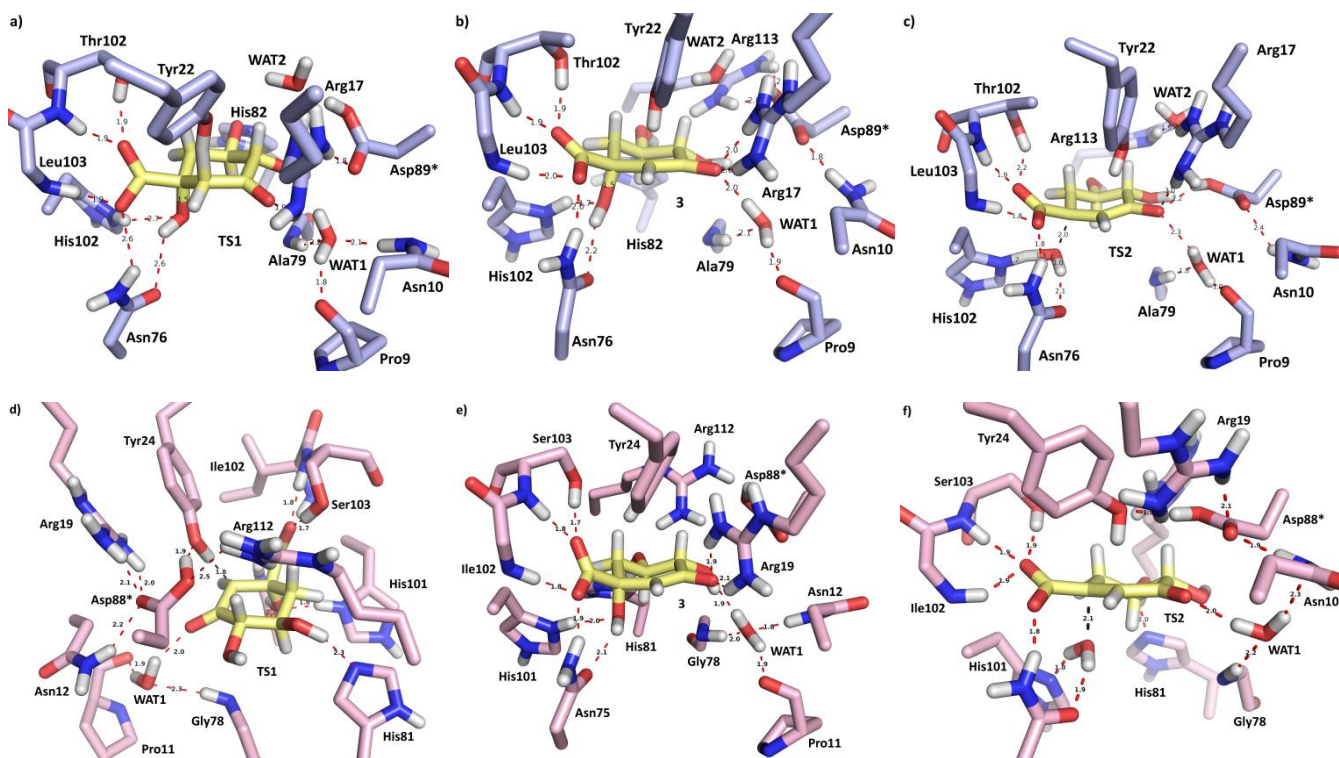
**Scheme 1 Enzymatic conversion of 3-dehydroquinic acid (1) to 3-dehydroshikimic acid (2)**

Previously proposed reaction mechanism of the reaction catalyzed by DHQ2. Proposed reaction intermediates are enolate **3** and enol **4**.



**Figure 1 *Hp*-DHQ2: Michaelis complex and proposed reaction mechanism**

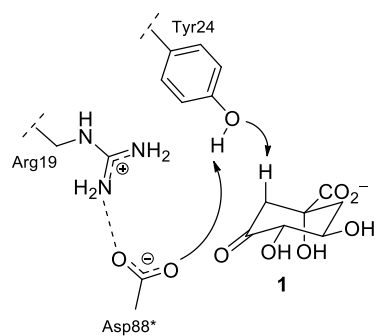
Relevant residues are indicated and labeled. Hydrogen bonds are shown as dashed red lines. Note how Asp89\* from a symmetry-related neighboring enzyme subunit is responsible, assisted by a water molecule (WAT2), for the generation of the catalytic form of the essential tyrosine (tyrosinate).



**Figure 2 Transition states TS1 and TS2 and enolate intermediate 3 geometries**

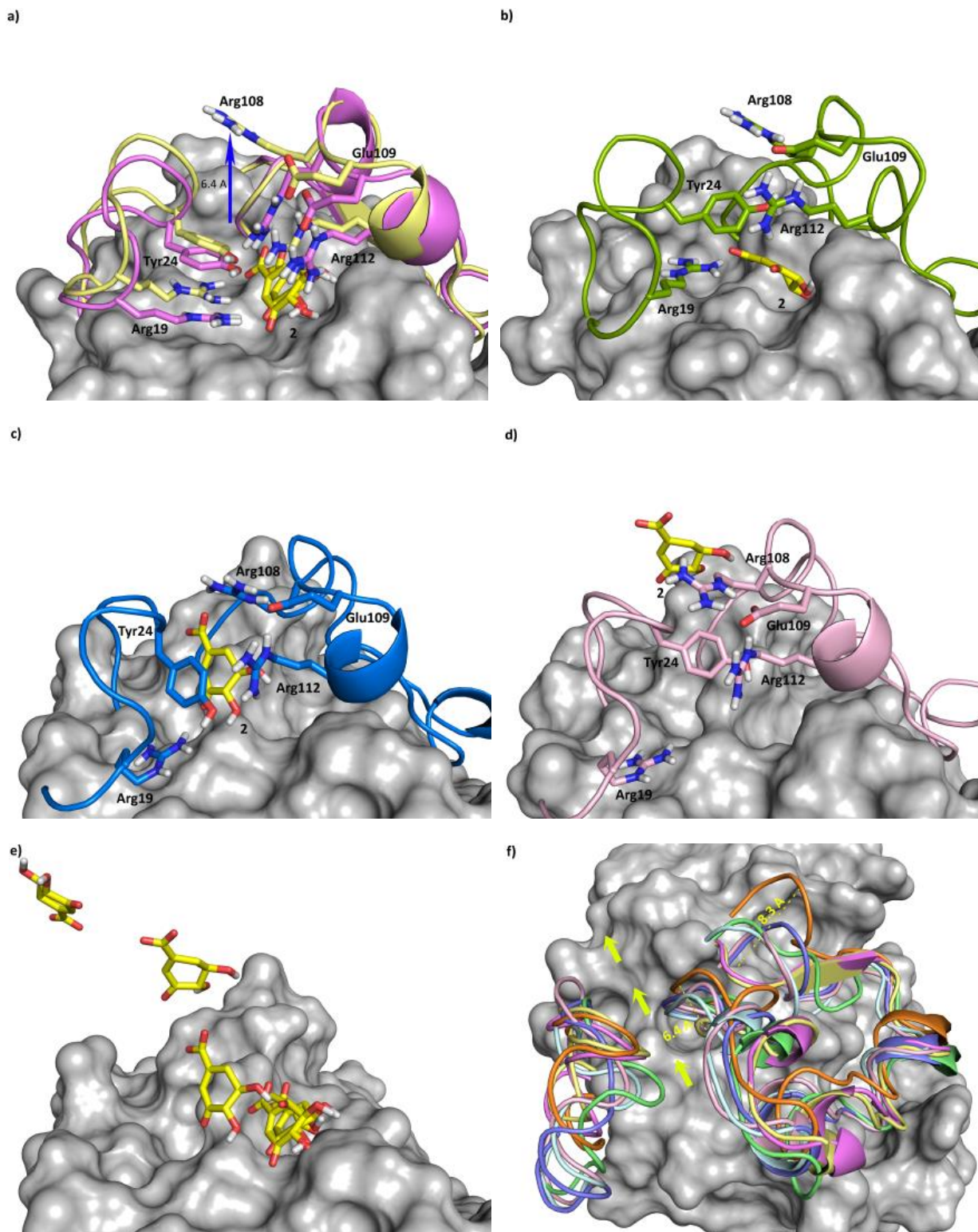
Representative snapshots of transition states TS1 (a,d) and TS2 (c,f), and enolate intermediate **3** (b,e) geometries along the reaction pathway of *Hp*-DHQ2 (blue) and *Mt*-DHQ2 (pink) enzymes. Relevant residues and water molecules are indicated and labeled. Hydrogen bonds (red) and key distances (black) are shown.





**Figure 3 Generation of the catalytic tyrosinate in *Mt*-DHQ2**

Proposed mechanism for the deprotonation of the essential tyrosine (Tyr24) by Asp88\* from a symmetry-related neighbor subunit in the reaction catalyzed by *Mt*-DHQ2 enzyme.



#### Figure 4 Product release

Snapshots of 3-dehydroshikimic acid (**2**) (yellow) release from the active site of *Mt*-DHQ2 obtained from the MD simulations: a) superposition of results of the second reaction coordinate and previous MD simulation (purple) and after 10 ns of simulation at 25 °C (yellow); b) after 10 ns at 25 °C and 20 ns at 77 °C (green); c) after 10 ns at 25 °C and 36 ns at 77 °C (blue); d) after 10 ns at 25 °C and 37 ns at 77 °C; e) a general view of how the product exits the active site; f) superposition of 10 ns at 25 °C and 40 ns at 77 °C. Note the large conformational changes of residues 101–112 and 126–130 during product release. Relevant side chain residues are shown and labeled.

**Table 1.** Kinetic parameters of wild-type and variant *Mt*-DHQ2 enzymes

Enzyme <sup>c</sup>	$k_{cat}$ (ms <sup>-1</sup> )	$K_m$ (μM)	$k_{cat}/K_m$ (s <sup>-1</sup> M <sup>-1</sup> )
wild-type <sup>a</sup>	154.2 ± 6.7	14.6 ± 1.8	10.5
N12S <sup>a</sup>	6.5 ± 0.2	11.4 ± 1.4	0.6
D88N <sup>a</sup>	< 0.05%	--	--
D88N <sup>b</sup>	0.60 ± 0.03	177.6 ± 34.6	0.03

<sup>a</sup>Assay conditions: Tris HOAc (50 mM, pH 7.0), 25 °C. <sup>b</sup>Assay conditions: Tris HOAc (50 mM, pH 8.0), 25 °C. <sup>c</sup>enzyme units: 0.2 (wild-type), 3 (N12S) and 20 (D88N).

**Table 2.** Summary of observed solvent isotope effects on several DHQ2 enzymes<sup>a</sup>

Enzyme	$(V_{max})_{1H_2O} / (V_{max})_{2H_2O}$	$(V_{max}/K_m)_{1H_2O} / (V_{max}/K_m)_{2H_2O}$
<i>Mt</i> -DHQ2	$2.48 \pm 0.05^{[13]}$	$1.0 \pm 0.1^{[13]}$
<i>An</i> -DHQ2	$4.1 \pm 0.2^{[13]}$	$2.5 \pm 0.5^{[13]}$
<i>Sc</i> -DHQ2 <sup>a</sup>	$3.97 \pm 0.04$	$1.32 \pm 0.11$
<i>Hp</i> -DHQ2 <sup>a</sup>	$4.12 \pm 0.05$	$1.49 \pm 0.14$

<sup>a</sup>Assay conditions: pH(<sup>2</sup>H) 7.0, 25 °C in Tris.HCl (50 mM) or Tris-*d*<sub>11</sub>. <sup>2</sup>HCl (50 mM).