CHEMMEDCHEM

DOI: 10.1002/cmdc.201000281

Understanding the Key Factors that Control the Inhibition of Type II Dehydroquinase by (2*R*)-2-Benzyl-3-dehydroquinic Acids

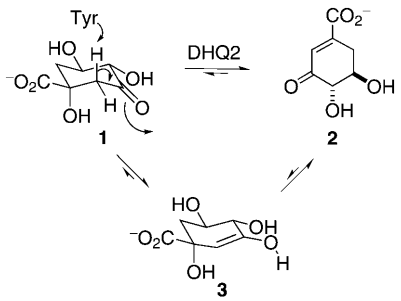
Antonio Peón,[a] José M. Otero,[b, c] Lorena Tizón,[a] Verónica F. V. Prazeres,[a] Antonio L. Llamas-Saiz,[d] Gavin C. Fox,[c] Mark J. van Raaij,[e] Heather Lamb,[f] Alastair R. Hawkins,[f] Federico Gago,[g] Luis Castedo,[a] and Concepción González-Bello*[a]

In memory of Professor Lucas Hernández.

The binding mode of several substrate analogues, (2*R*)-2-benzyl-3-dehydroquinic acids **4**, which are potent reversible competitive inhibitors of type II dehydroquinase (DHQ2), the third enzyme of the shikimic acid pathway, has been investigated by structural and computational studies. The crystal structures of *Mycobacterium tuberculosis* and *Helicobacter pylori* DHQ2 in complex with one of the most potent inhibitor, *p*-methoxybenzyl derivative **4a**, have been solved at 2.40 Å and 2.75 Å, respectively. This has allowed the resolution of the *M. tuberculosis* DHQ2 loop containing residues 20–25 for the first time. These structures show the key interactions of the aromatic ring in the active site of both enzymes and additionally reveal an important change in the conformation and flexibility of the loop that closes over substrate binding. The loop conformation and the binding mode of compounds **4b–d** has been also studied by molecular dynamics simulations, which suggest that the benzyl group of inhibitors **4** prevent appropriate orientation of the catalytic tyrosine of the loop for proton abstraction and disrupts its basicity.

Introduction

Dehydroquinase (3-dehydroquinate dehydratase; EC 4.2.1.10) is the third enzyme in the shikimic acid pathway and catalyzes the reversible dehydration of 3-dehydroquinic acid (**1**) to form 3-dehydroshikimic acid (**2**) (Scheme 1).[1] There are two distinct dehydroquinases, designated as type I (DHQ1) and type II (DHQ2), which have different biochemical and biophysical properties and exhibit little sequence similarity. Both subtypes catalyze the same overall reaction through independent mechanisms and with opposite stereochemistry.[2] The type I enzymes (identified from *Escherichia coli, Salmonella typhi*) use a multistep mechanism that involves covalent imine intermediates between the ketone **1** and a conserved lysine (Lys170 in *E. coli*).[3] These enzymes are only involved in the biosynthesis of shikimate and catalyze an overall *syn* elimination of water



Scheme 1. Proposed E1cB mechanism for the enzymatic conversion of 3-dehydroquinic acid (**1**) to 3-dehydroshikimic acid (**2**) catalyzed by DHQ2. The reaction proceeds via the enol intermediate **3**.

[a] *A. Peón, L. Tizón, V. F. V. Prazeres, Prof. Dr. L. Castedo, Prof. Dr. C. González-Bello*
Departamento de Química Orgánica y Centro Singular de Investigación en Química Biológica y Materiales Moleculares, Universidad de Santiago de Compostela, calle Jenaro de la Fuente s/n, 15782 Santiago de Compostela (Spain)
Fax: (+34) 981-595012
E-mail: concepcion.gonzalez.bello@usc.es

[b] *Dr. J. M. Otero*
Departamento de Bioquímica y Biología Molecular, Facultad de Farmacia, Universidad de Santiago de Compostela, 15782 Santiago de Compostela (Spain)

[c] *Dr. J. M. Otero, Dr. G. C. Fox*
Laboratoire des Proteines Membranaires, Institut de Biologie Structurale J. P. Ebel, 41 Rue Jules Horowitz, Grenoble (France)

[d] *Dr. A. L. Llamas-Saiz*
Unidad de Rayos X, RIAIDT, Edificio CACTUS, Universidad de Santiago de Compostela, 15782 Santiago de Compostela (Spain)

[e] *Dr. M. J. van Raaij*
Instituto de Biología Molecular de Barcelona, Consejo Superior de Investigaciones Científicas (IBMB-CSIC), Parc Científic de Barcelona, Baldiri Reixach 4, 08028 Barcelona (Spain)

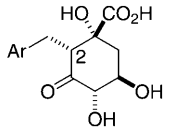
[f] *Dr. H. Lamb, Prof. Dr. A. R. Hawkins*
Institute of Cell and Molecular Biosciences, Medical School, University of Newcastle upon Tyne, Catherine Cookson Building, Framlington Place Newcastle upon Tyne NE2 4HH (UK)

[g] *Prof. Dr. F. Gago*
Departamento de Farmacología
Universidad de Alcalá, 28871 Alcalá de Henares (Spain)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cmdc.201000281.

characterized by the loss of the less acidic *pro*-R hydrogen from C-2 of **1**.[4] In contrast to DHQ1, the type II enzymes catalyze the *anti* elimination of water, which involves the loss of the more acidic *pro*-S hydrogen from C-2 of **1**.[5] DHQ2 enzymes are found on both the shikimate and quinate pathways. Biosynthetic DHQ2 occurs in many bacterial species including *M. tuberculosis*, *Streptomyces coelicolor* and *H. pylori*.[1b] Catabolic DHQ2 have been characterized from *Aspergillus nidulans* and *Neurospora crassa*. DHQ2 from *Amycolatopsis methanolic* operates on both pathways.[1b] The elimination proceeds through a stepwise E1cB mechanism involving an enol intermediate **3** (Scheme 1).[6,7] Initially, an essential tyrosine of the active site removes the *pro*-S hydrogen from C-2 of **1**. The final step is the acid-catalyzed elimination of the C-1 hydroxy group—a reaction mediated by a histidine residue that acts as proton donor. It is believed that the lower pK_a of the tyrosine is the result of a basic environment formed by two conserved arginine residues in close proximity.[7]

In recent years, we have focused on the inhibition of DHQ2 from *M. tuberculosis* (DHQ2-Mt) and *H. pylori* (DHQ2-Hp), essential enzymes in the normal function of these organisms.[8,9] We have shown that substitution of *pro*-R hydrogen of 3-dehydroquinic acid (**1**), the natural substrate of the DHQ2, by benzyl groups generates potent reversible competitive inhibitors (Figure 1).[10] The *p*-methoxybenzyl derivative **4a** and the benzothiophene **4d** proved to be the most potent competitive inhibitors of both DHQ2 enzymes, with K_i values of 26 nM and 28 nM against *M. tuberculosis* and 170 nM and 160 nM against *H. pylori*, respectively. In order to gain further insights into the interaction, we aimed to solve the crystal structures of *M. tuberculosis* and *H. pylori* DHQ2 in complex with one of the most potent inhibitors, *p*-methoxybenzyl derivative **4a**. The binding mode of compounds **4b–d** and the observed conformational changes affecting the loop region that closes over the active site were also studied by molecular dynamics simulations.



4a Ar = (4-MeO)C_6H_4 (0.026 μM; 0.17 μM)
4b Ar = (4-Me)C_6H_4 (0.089 μM; 0.25 μM)
4c Ar = C_6F_5 (0.047 μM; 2.6 μM)
4d Ar = benzo[*b*]thiophen-5-yl (0.028 μM; 0.16 μM)

Figure 1. (2*R*)-2-Benzyl-3-dehydroquinic acids **4** are reversible, competitive inhibitors of DHQ2. The K_i values (μM) for compounds **4** against DHQ2-Mt and DHQ2-Hp are indicated in parentheses in that order.

Results and Discussion

Structural studies

In an effort to obtain structural information on the binding mechanism of (2*R*)-2-benzyl-3-dehydroquinic acids **4**, one of the most potent competitive inhibitors of the series, *p*-methoxybenzyl derivative **4a**, was co-crystallized with DHQ2 from both *M. tuberculosis* and *H. pylori*.

X-ray diffraction data were collected from a cryo-cooled crystal of DHQ2-Mt/**4a** complex using synchrotron radiation and a summary of the statistics following data reduction and processing is given in Table 1. The structure was determined by molecular replacement, using the back-bone chain of DHQ2-Mt bound to 3-hydroxyimino quinic acid (PDB: 1H0S)[11] as a search model (Figure 2), and then refined.

Table 1. Crystallographic data collection and refinement statistics of DHQ2-Mt and DHQ2-Hp complexes with inhibitor **4a**.

Data processing[a]	DHQ2-Mt/**4a**	DHQ2-Hp/**4a**
space group	*F*23	$P4_222$
cell parameters	$a=b=c=126.52$ Å[b]	$a=b=100.56$ Å[b], $c=105.45$ Å
wavelength	0.87260 Å	1.54184 Å
detector	225 mm MarMOSAIC flatpanel	Kappa CCD2000
crystal-to-detector distance	238.9 mm	90.0 mm
resolution range	38.0–2.40 Å (2.53–2.40 Å)	35.0–2.75 Å (2.85–2.75 Å)
observed reflections[c]	6707 (982)[d]	14629 (1446)[d]
wilson B	28.6 $Å^2$	61.5 $Å^2$
multiplicity	10.9 (11.0)	3.6 (3.2)
completeness	1.000 (1.000)	0.999 (0.999)
R_{merge}	0.118 (0.383)	0.044 (0.318)
refinement[e]		
reflections used in refinement[c]	6175 (887)	13844 (1967)
reflections used for R-free	503 (81)	748 (89)
resolution range	20.0–2.40 Å (2.53–2.40 Å)	34.2–2.75 Å (2.90–2.75 Å)
R-factor[f]	0.156 (0.170)	0.185 (0.281)
R-free[g]	0.203 (0.234)	0.250 (0.309)
RMSD (bonds/angles)	0.015 Å/1.5°	0.014 Å/1.4°
protein/inhibitor/Tris/sulfate/citrate/water atoms	1071/22/8/35/-/50	3663/44/-/-/13/21
average B protein/inhibitor/Tris/sulfate/citrate/water	19.1 $Å^2$/18.2 $Å^2$/44.0 $Å^2$/50.2 $Å^2$/-/21.7 $Å^2$	46.1 $Å^2$/40.6 $Å^2$/-/-/41.5 $Å^2$/37.4 $Å^2$
ramachandran statistics[h]	97.1%/99.3%	95.9%/99.8%

[a] Results from SCALA (DHQ2-Mt/**4a**; Reference [12]) or XPREP (DHQ2-Hp/**4a**; Reference [13]). [b] One Angstrom (Å) is 0.1 nm. [c] No sigma cutoff or other restrictions were used for inclusion of reflections. [d] Values in parentheses are for the highest resolution bin, where applicable. [e] Results from REFMAC (Reference [14]). [f] R-factor = $S||Fobs(hkl)|-|Fcalc(hkl)||/S|Fobs(hkl)|$. [g] According to Reference [15]. [h] According to the program MOLPROBITY (Reference [16]). The percentages indicated are for residues in favored and total allowed regions, respectively.

The structure contains a single DHQ2-Mt molecule in the asymmetric unit. Overall, the two structures (PDB: 1H0S and 2BX8) are virtually identical (0.371 Å root mean square difference (RMSD) after superimposition); the only significant differences are noted for Leu3 (the most *N*-terminal resolved residue), Gly17, Arg18, Arg19 and Gly26. Amino acids 20–25, which are located in the loop that closes over the active site following substrate binding, have now been visualized for the first time (Figure 3). It should be noted that in both structures,

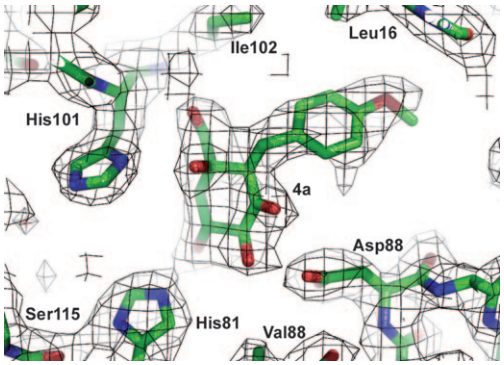


Figure 2. Electron density for inhibitor **4a** in the crystal structure of DHQ2-Mt (PDB: 2XB8). From the model obtained by molecular replacement, amino acids 17–25 were removed and refinement was performed to obtain unbiased density for the inhibitor molecule and other model changes. A maximum-likelihood-weighted $2F_o–F_c$ map[14] contoured at 1σ is shown up to 3 Å around the inhibitor molecule. The final model, as deposited in the database Protein Data Bank and including the inhibitor molecule, is superposed onto the map.

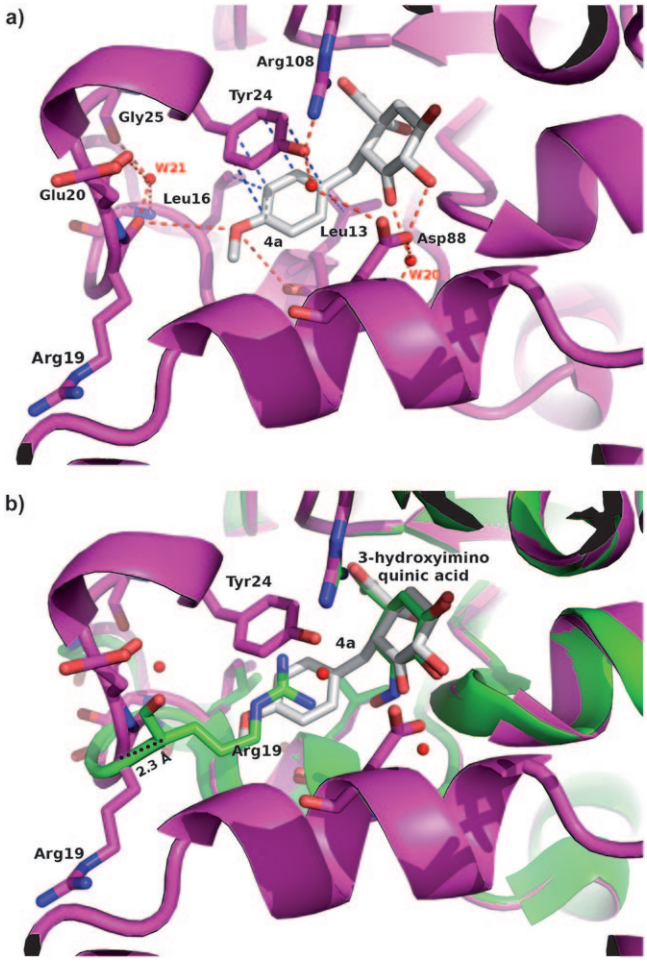


Figure 3. Interactions of inhibitor **4a** with DHQ2-Mt. a) Lipophilic (blue) and hydrophilic (red) interactions between **4a** (grey) and the DHQ2-Mt (purple). b) Superposition of DHQ2-Mt/**4a** complex (PDB: 2XB8; purple/grey) with DHQ2-Mt/3-hydroxyimino quinic acid (PDB: 1H0S;[11] green). Only relevant residues are indicated. The side chain of Arg19 is flipped out of the active site as a result of its displacement by the *p*-methoxybenzyl ring system of **4a**. The position of Arg19 in the DHQ2-Mt/**4a** complex (purple) is displaced by 2.3 Å in relation to its position in DHQ2-Mt/3-hydroxyimino quinic acid complex (green). Arg19 in DHQ2-Mt/**4a** complex (purple) is described as two conformationally possible rotamers.

that is, the structure reported here and PDB entry 1H0S,[11] amino acids belonging to the loop 18–25 have relatively high temperature factors (50 Å^2) compared with the rest of the protein (17.4 Å^2) and that Glu20 of our structure falls outside the expected region of the Ramachandran plot. Therefore, it is likely that this loop is flexible and adopts different conformations, even in cryo-cooled crystals. The cyclohexane ring of inhibitor **4a** occupies approximately the same site as 3-hydroxyimino quinic acid in PDB entry 1H0S, and the carbonyl group of **4a** forms a hydrogen bond with an essential water molecule in the active site. The *p*-methoxybenzyl moiety of **4a** is located in the position occupied by the side chain of Arg19 in PDB entry 1H0S. Therefore, this aromatic group is in close contact with the side chain carbons of Arg15 and Asn12 and with the side chain atoms of Leu16 and Leu13, thus establishing important lipophilic interactions with this part of the active site.

Notably, the side chain of Arg19 is displaced outside of the active site. This amino acid has been identified by chemical modification and site-directed mutagenesis studies as being essential for enzyme activity, and it has been suggested that it plays a role in the stabilization of the enol intermediate.[17,7] Namely, replacement of essential Arg23 of DHQ2 from *S. coelicolor* by lysine, glutamine or alanine causes a reduction in catalytic activity between 3000- and 30000-fold.[2b] Tyr24 in this loop is located close to the aromatic ring of ligand **4a** and interacts with Arg108 and with Asp88 via hydrogen bonding with a bridging water molecule. Additionally, a second water molecule was observed close to the same loop and this interacts through hydrogen bonding with the amide carbonyl groups of Gly25, Glu20 and Arg19 and the amide nitrogen of Gly17. The atoms involved are arranged in a pseudo-square-planar disposition. These attractive interactions facilitate stabilization of the loop in the closed conformation,[18] thus allowing the description of this loop for the first time in DHQ2-Mt.

X-ray diffraction data were also collected for the DHQ2-Hp/**4a** complex at room temperature using a Bruker-Nonius FR591 rotating anode diffractometer. The crystallographic structure was solved by molecular replacement, using the back-bone chain of DHQ2-Hp bound to citric acid (PDB: 2C4V[19]) as a search model (Figure 4), and refined. For data and refinement statistics see Table 1. The DHQ2-Hp/**4a** complex crystallized with three copies of the monomer in the asymmetric unit (designated as chains A, B & C).

The three crystallographically independent copies superimpose well onto each other, with an RMSD value of 0.3 Å when chain A is superimposed onto B, and 0.4 Å when chains A or B are superimposed onto C. PDB entry 2C4V[19] contains a single protein per asymmetric unit; when chains A, B or C are superimposed onto this structure, RMSD values of 0.6, 0.6 and 0.4 Å are observed, respectively; thus, the protein part of the model is virtually identical.

When the electron densities for active sites were observed, chains A and B showed a positive electron density difference,

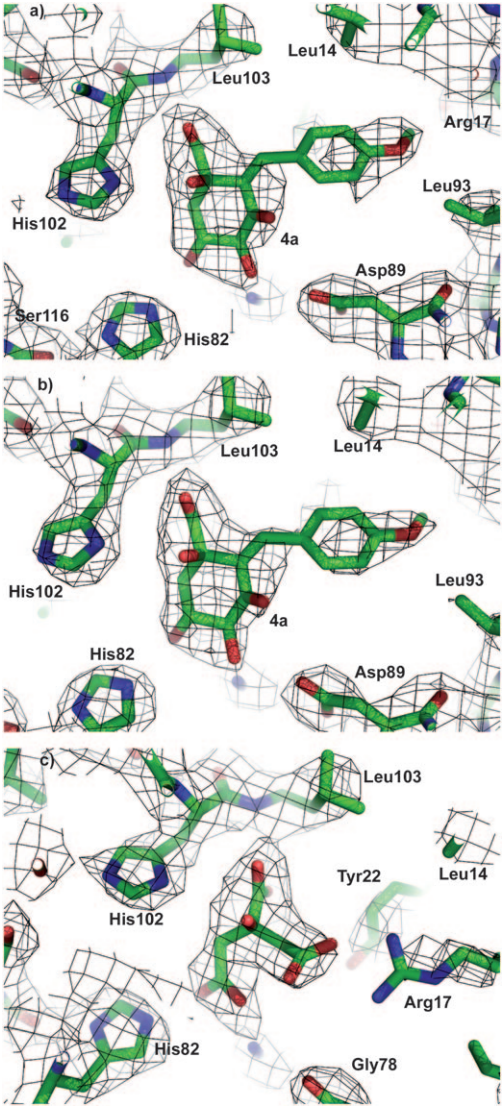


Figure 4. Electron density for inhibitor **4a** in monomers A and B and for citrate in monomer C of the crystal structure of DHQ2-Hp (PDB: 2XB9). From the model obtained by molecular replacement, amino acids 15–23 of each of the three monomers were removed and refinement was performed to obtain unbiased density for the inhibitor and citrate molecules and other model changes. A maximum-likelihood weighted $2F_o–F_c$ map[14] contoured at 1σ is shown up to 3 Å around the inhibitor or citrate molecules. The final model, as deposited into the database and including inhibitor molecules and citrate, is superposed onto the map.

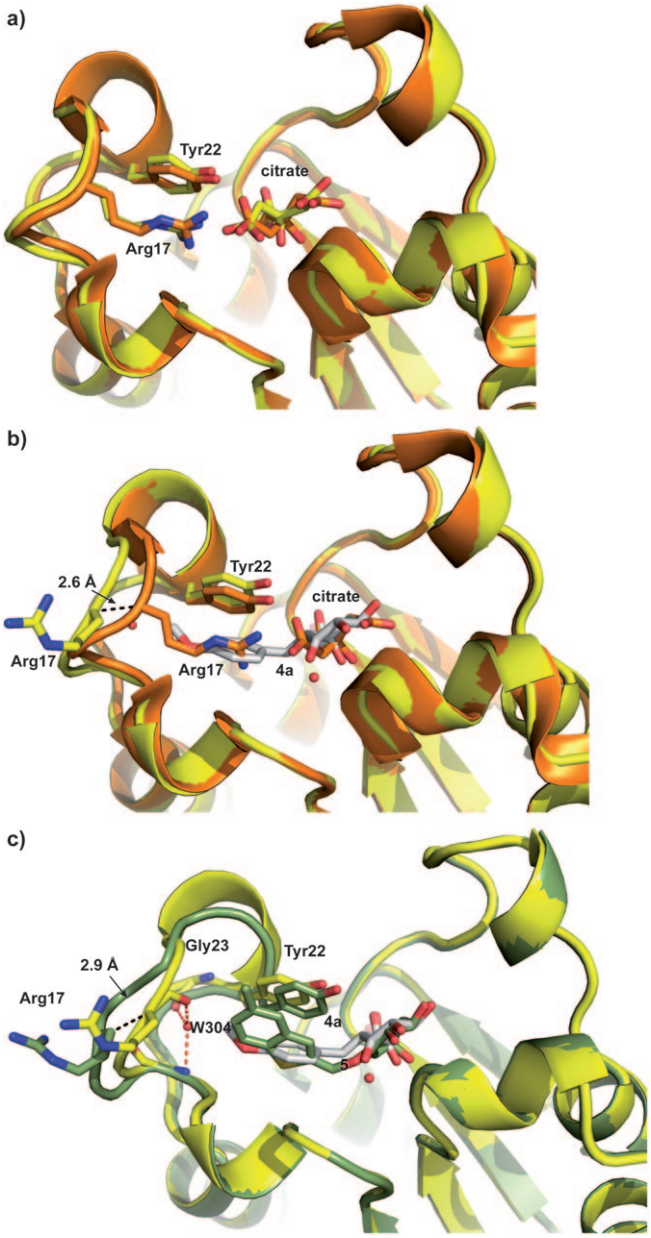


Figure 5. Superposition of: a) Chain C of DHQ2-Hp/**4a** complex (2XB9; yellow) with DHQ2-Hp/citrate (2C4V;[19] orange); b) Chain B of DHQ2-Hp/**4a** complex (yellow with **4a** in grey) with DHQ2-Hp/citrate (PDB: 2C4V;[19] orange); c) Chain B of DHQ2-Hp/**4a** complex (2XB9; yellow with **4a** in grey) with DHQ2-Hp/benzothiophene inhibitor (PDB: 2WKS;[20] pale green). Only relevant residues are indicated. Note how the side chain of Arg17 is flipped out of the active site as a result of its displacement by the aromatic ring system of inhibitors.

indicating that a larger molecule than citrate was present in the active site. However, chain C has citrate originating from the enzyme crystallization buffer (as in PDB entry 2C4V;[19] Figure 4c). Therefore, ligand **4a** was modeled in the active sites of chains A and B and then refined (Figure 4a and b). Temperature factors of the ligand were refined to values similar to those of the protein. In chain C, Arg17 is located close to Tyr22 and it interacts through hydrogen bonding with one of the carboxylate groups of the citrate moiety (Figure 5a) in an analogous way to that found in PDB entry 2C4V.[19] However, in the inhibitor complexes that contain chains A and B, Arg17 is expelled from the active site and faces away from Tyr22 (Figure 5b). In all chains, a water molecule close to the flexible arm helps its organization by hydrogen bonding with the amide carbonyl group of Gly23 and Arg17 and the amide nitrogen of Gly15, forming a pseudo-trigonal-planar disposition. When chains A and B of these structures are compared with the structure recently reported by us for DHQ2-Hp in complex with a benzothiophene derivative ($K_i = 130$ nM; PDB: 2WKS[20]), a small shift of the catalytic Tyr22 and the loop is observed. This shift is caused by the proximity of the benzyl group of in-

hibitor **4a** to the side-chain residues of this part of the active site (Figure 5c). In particular, Tyr22 of chains A and B is shifted about 1.5 Å away from its position in PDB entry 2WKS and, in contrast, Arg17 is displaced 2.9 Å towards the active site.

Molecular dynamics simulations

The solved crystal structures of the DHQ2-Mt/**4a** and DHQ2-Hp/**4a** complexes show not only relevant lipophilic interactions with several residues of the active site, but also an important change in the conformation and flexibility of the loop that closes over the substrate binding site. This partial ordering of the DHQ2-Mt loop appears to be related to the competitive binding of these substrate analogues **4**. In an effort to elucidate whether ligands **4b–d** are likely to have a similar binding mode as the *p*-methoxybenzyl derivative **4a** and how the loop conformation may change as ligands **4** bind to the enzymes, molecular dynamic (MD) simulations were performed.

Firstly, we studied the behavior and stability of the crystalline complexes DHQ2/**4a** in our simulation conditions considering the two possible protonation states of the catalytic tyrosines in both enzymes (Tyr24 and Tyr22 in *M. tuberculosis* and *H. pylori*, respectively). MD studies show that when tyrosine residues are in their acid form, the proposed complex is similar to the crystal structure obtained (Figure 6). However, when MD simulations were performed using the corresponding tyrosinates, the essential arginine of the loop (Arg19 and Arg17 in *M. tuberculosis* and *H. pylori*, respectively) moves into the active site to establish a strong electrostatic interaction with the tyrosine. In addition, comparison of this model with PDB entry 1H0S, which has Arg19 inside the active site, shows that this residue is not as close to Try24 because the benzyl group occupies its position. In both models, as well as in the crystal structures reported here, the catalytic tyrosines are not in the appropriate orientation for proton abstraction (see Supporting Information). This situation has already been demonstrated in other DHQ2 inhibitor complexes, particularly from *S. coelicolor*. For instance, in the 2,3-anhydroquinic acid complex (PDB: 1GU1),[7] in a fluoro derivative complex (PDB: 1V1F),[21] in a biaryl complex (PDB: 2CJF),[22] amongst others.

However, when these calculations were performed using the natural substrate to simulate the enzyme–substrate complex, the essential arginine of the loop is located very close to the catalytic tyrosine (below 3 Å), which is now appropriately orientated for proton abstraction (Figure 7). These results suggests that, on the one hand, Arg109 (Arg108 in DHQ2-Mt) should also be responsible for the stabilization of the catalytic tyrosine in the loop-closed conformation required for the catalysis. On the other hand, the reaction is presumably initiated by the proximity of the essential arginine of the loop, which appropriately orients the tyrosine to initiate the enzymatic reaction, either by a concerted or stepwise process. It has previously been suggested that the proximity of Arg109 (Arg108 in DHQ2-Mt) should lower the pK_a of the catalytic tyrosine and, to a lesser extent, also Arg17 (Arg19 in DHQ2-Mt).[7] Considering the proximity of both residues in the proposed enzyme–substrate complex, the effect of the essential arginine (Arg17 and Arg19) on the pK_a of the tyrosine might be quite significant. In addition, tyrosine pK_a values calculated using the PROPKA[23] program, and the enzyme geometries found in the DHQ2-Hp/**4a** complex and the proposed enzyme–inhibitor complex suggest that this is the case. Namely, the estimated pK_a value of Tyr22 in chains A and B, in which Arg17 is outside the active site, is 8.99. However, for chain C (in complex with citrate) and for the proposed enzyme–substrate complex, in

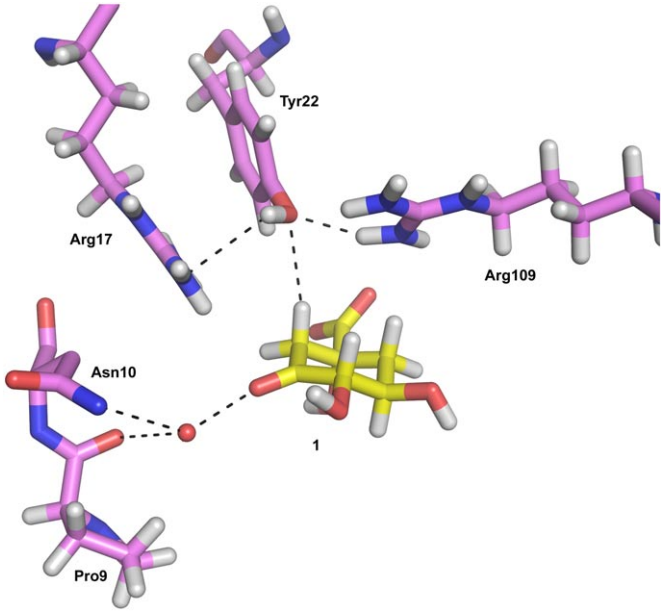


Figure 7. A detailed view of the proposed enzyme–substrate complex as obtained from MD studies with key residues labeled and hydrogen bonds below 3 Å in length shown as dashed lines. The appropriate orientation of Tyr22 for proton abstraction is highlighted.

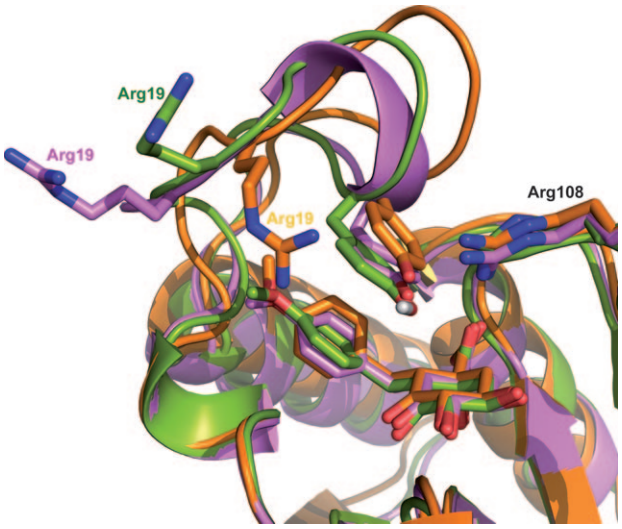


Figure 6. Superposition of DHQ2-Mt/**4a** complex (PDB: 2XB8; violet) and the complex obtained after MD simulations considering essential Tyr22 as its acid form (green) and as tyrosinate (orange). Only relevant residues are indicated.

which Arg17 is inside the active site in close contact to Tyr22, the values are 5.91 and 6.93, respectively.

The binding mode of ligands **4b–d** was then investigated. MD studies show that, as in the crystal structures, the essential arginine of the loop (Arg19 and Arg17 in *M. tuberculosis* and *H. pylori*, respectively) is located outside of the active site. Moreover, the aromatic ring of the inhibitors is located close to the essential tyrosine. For the benzothiophenyl derivative **4d**, one of the most potent inhibitors, the sulfur atom of this benzothiophenyl group is oriented towards the side chain of the essential arginine of the loop (Figure 8).

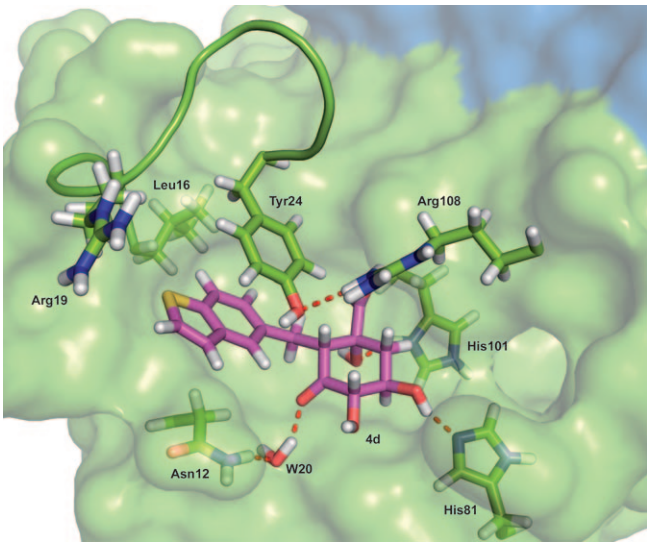


Figure 8. Proposed binding mode of ligand **4d** in DHQ2-Mt. Only relevant residues are indicated.

These structural and computational studies highlight the role of the essential arginine of the loop on the appropriate orientation for proton abstraction and on the reduced pK_a value of the catalytic tyrosine. The proximity to this arginine to the catalytic tyrosine triggers the catalytic process. This fact could explain why (2*R*)-2-benzyl-3-dehydroquinic acids **4**, in which the axial proton on C-2 is free, are not substrates for the enzyme. The benzyl groups of ligands **4** inactivate the enzyme by causing a significant conformational change in the flexible loop—a change that prevents appropriate orientation of the tyrosine for proton abstraction.

Conclusions and Final Remarks

The binding mode of (2*R*)-2-benzyl-3-dehydroquinic acids **4**, which are potent competitive reversible inhibitors of type II dehydroquinase from *M. tuberculosis* and *H. pylori*—the third enzyme of the shikimic acid pathway, has been investigated by structural and computational studies. The crystal structures of *M. tuberculosis* and *H. pylori* type II dehydroquinase in complex with one of the most potent inhibitors, *p*-methoxybenzyl derivative **4a**, have been solved at 2.40 Å and 2.75 Å, respectively, allowing the resolution for the first time of the loop containing residues 20–25. This loop closes over the active site on substrate binding. Both complexes reveal the key interactions that occur between the aromatic ring of the ligand and the catalytic tyrosine present in the active site of both enzymes.

In addition, molecular dynamics studies with ligands **4b–d** predict a similar binding mode as observed for the *p*-methoxybenzyl derivative **4a** in the crystal structures reported here. The results suggest that the benzyl moiety of ligands **4** reduces the loop movement, thus preventing entry of the natural substrate to the active site by changing the appropriate conformation of the catalytic tyrosine for proton abstraction and by altering its basicity. Therefore, compounds capable of altering the conformation and flexibility of the loop that closes over substrate binding should constitute promising candidates for inhibitors of these essential enzymes in *M. tuberculosis* and *H. pylori*, two important bacterial pathogens.

Experimental Section

Type II Dehydroquinase. *H. pylori* and *M. tuberculosis* type II dehydroquinase was expressed and purified as described previously.[24,25]

Crystallization and structure determination. DHQ2-Mt and DHQ2-Hp were concentrated to 20 mg mL^{-1} in 50 mM Tris-HCl (pH 7.5), 1 mM 2-mercaptoethanol, 1 mM EDTA and 200 mM NaCl. (2*R*)-2-*p*-Methoxybenzyl-3-dehydroquinic acid (**4a**) was dissolved at 0.25 M in methanol and added at a ratio of 1:20 (*v*/*v*) to each protein solution to give solutions of approximately 10 equivalents of **4a** per protein monomer.

For DHQ2-Mt, diamond shaped crystals of up to 0.05 mm × 0.05 mm were obtained after two months of vapor diffusion in sitting drops comprised of 2.0 μL of protein/inhibitor solution mixed with 2.0 μL of reservoir solution equilibrated against 0.15 mL of a reservoir solution containing 32% (*v*/*v*) 2-methyl-2,4-pentanediol, 0.3 M ammonium sulfate and 0.1 M 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid sodium salt (HEPES sodium salt), pH 7.5. After harvesting, crystals were frozen by direct immersion in liquid nitrogen. Data were collected on beamline ID23–2 (ESRF; Grenoble, France) from a crystal maintained at 100 K. The data were processed, scaled and analyzed using MOSFLM,[26] SCALA,[12] and other programs within the CCP4 software suite.[27] The structure was solved by molecular replacement, using the program MOLREP,[28] with a search model generated from PDB entry 1H0S.[11] Reflections for the Rfree[15] set were selected randomly. Model building was done with COOT[29] and refinement with REFMAC.[14] Structure validation was performed with MOLPROBITY.[16]

For DHQ2-Hp diamond shaped crystals of up to 0.3 mm × 0.3 mm were obtained after one month of vapor diffusion in sitting drops comprised of 2.0 μL of protein/inhibitor solution mixed with 2.0 μL of reservoir solution equilibrated against 0.15 mL reservoirs containing 31% (*w*/*v*) poly(ethylene glycol) 4000 and 0.1 M sodium citrate (pH 5.0). One crystal was mounted at room temperature in a saturated atmosphere of crystallization solution with a MicroRT Tubing Kit (MiTeGen) and data were collected in a Bruker Nonius FR591 Kappa CCD2000 diffractometer with confocal multilayer monochromatized $Cu_{K\alpha}$ radiation. The COLLECT[30]/HKL2000[31] software was used for data collection and integration. Scaling, merging and absorption correction was performed with the SCALEPACK[31] program. The XPREP[13] program was used for analysis of the three-dimensional diffraction data and processing showing a $P4_22_12$

space group. Finally molecular replacement using the protein chain of PDB entry 2C4V[19] was developed using the program PHASER,[32] which found a correct solution for the crystallographic structure in the $P4_222$ space group with three independent chain proteins per asymmetric crystal unit. Reflections for the Rfree[15] set were selected in thin resolution shells. Model building, refinement and structure validation were performed as described for DHQ2-Mt.

Molecular dynamics simulations. *Ligand minimization*: Ligand geometries were first refined by means of the semi-empirical quantum mechanical program MOPAC[33] using the AM1 Hamiltonian and PRECISE stopping criteria, and further optimized using a restricted Hartree–Fock (RHF) method and a 6–31G(d) basis set, as implemented in the ab initio program Gaussian 09.[34] The resulting wave functions were used to calculate electrostatic potential-derived (ESP) charges employing the restrained electrostatic potential (RESP)[35] methodology, as implemented in the assisted model building with energy refinement (AMBER)[36] suite of programs. The missing bonded and nonbonded parameters were assigned by analogy or through interpolation from those already present in the AMBER database (GAFF).[37]

Generation and minimization of the DHQ2–ligand complexes. Simulations were carried out using the enzyme geometries found in the crystal structures of DHQ2-Mt and DHQ2-Hp in complex with *p*-methoxybenzyl derivative **4a** (PDB: 2XB8 and 2XB9, respectively). Taking into account that unfolding and refolding studies of DHQ2 have shown that the trimer[38] is the minimal biologically active unit of the enzyme, and on the basis of preliminary simulations that showed the monomer to be unstable under our simulation conditions, the trimer was used for these studies. Hydrogen atoms were added to the protein using web-based H++ server,[40] which assigned protonation states to all titratable residues at the chosen pH of 7.0. However, δ and/or ε protonation was manually corrected for His81 (δ) and His101 (dual) of the active site due to the mechanistic considerations and on the basis of results from preliminary MD simulations. Molecular mechanics parameters from the ff03 and GAFF force fields, respectively, were assigned to the protein and the ligands using the LEaP module of AMBER 10.0.[40]

Energy minimization using the implicit solvent GB model was carried out in stages, starting with protein and ligand hydrogen atoms (1000 cycles: half of them steepest descent, the other half conjugate gradient), followed by the 10–34 loop (20000 cycles, idem), protein and ligand hydrogen atoms (5000 cycles, idem), amino acid side chains (5000 cycles, idem) and finally the entire complex (5000 cycles, idem). A positional restraint force constant of 50 kcal mol^{-1} Å^{-2} was applied to those unminimized atoms in each step was applied during all calculations. Thereafter, each refined DHQ2–ligand complex was neutralized by addition of sodium ions[41] and immersed in a truncated octahedron of TIP3P water molecules.[42]

Simulations: MD simulations were performed using the AMBER 10.0 suite of programs and Amber ff03 force field.[41] Periodic boundary conditions were applied and electrostatic interactions were treated using the smooth particle mesh Ewald method (PME)[43] with a grid spacing of 1 Å. The cutoff distance for the nonbonded interactions was 9 Å. The SHAKE algorithm[44] was applied to all bonds containing hydrogen, using a tolerance of 10^{-5} Å and an integration step of 2.0 fs. Minimization was carried out in three steps, starting with the octahedron water hydrogen atoms, followed by solvent molecules and sodium counter ions and finally the entire system. The minimized system was heated at 300 K (1 atm, 25 ps, a positional restraint force constant of 50 kcal mol^{-1} Å^{-2}). These initial harmonic restraints were gradually removed (11 cycles) and the resulting systems were allowed to equilibrate further. MD with constraints of 5 kcal mol^{-1} Å^{-2} were carried out in two stages. Firstly, they were applied to all protein α-carbons for 2 ns and secondly to all protein α-carbons of the two external subunits of the trimer and the β-sheets and α-helix of the central subunit of the trimer for 10 ns. System coordinates were collected every 2 ps for further analysis. A slow-cooling MD simulation was subsequently performed (six steps until 273 K). Minimization of the entire complexes with constraints of 5 kcal mol^{-1} Å^{-2} were applied to protein α-carbons of the two external subunits of the trimer and the β-sheets and α-helix of the central subunit of the trimer.

The data collection, refinement, and model statistics are summarized in Table 1. Coordinates and structure factors are available from the Protein Data Bank (PDB) with accession codes 2XB8 and 2XB9 for *M. tuberculosis* and *H. pylori* and type II dehydroquinases, respectively. All figures were prepared using PyMOL.[45]

Acknowledgements

This work was supported by the Spanish Ministry of Science and Innovation (SAF2007–63533 to CGB; BFU2005–02974/BMC and BFU2008–01588/BMC to MJvR), the CSIC (488IE7 to MJvR) and the Xunta de Galicia (PGIDT07PXIB209080PR and GRC2006/132 to CGB). AP and LT thank the Spanish Ministry of Science and Innovation for FPU fellowships. JMO thanks the Xunta de Galicia and Spanish Ministry of Science and Innovation for "Ángeles Alvariño" and "José Castillejo" fellowships, respectively. VFVP thanks the Portuguese Fundação para a Ciência e a Tecnología for an FCT fellowship. We thank Professors Javier Cañada and Manuel Mosquera for their helpful discussions, Patricia Ferraces-Casais for excellent technical assistance, the ESRF-Grenoble for provision of beam time, and Claire Coderch for technical assistance with the molecular modeling simulations. We are also thankful to the Centro de Supercomputación de Galicia (CESGA) for the use of the Finis Terrae computer.

Keywords: Binding modes · Competitive inhibition · Dehydroquinases · *Helicobacter pylori* · Molecular dynamics · *Mycobacterium tuberculosis*

[1] a) E. Haslam, *The Shikimate Pathway*, Halstead Press, Wiley, New York, **1974**; b) C. Abell in *Comprehensive Natural Products Chemistry*, (Ed.: U. Sankawa), Elsevier Science Ltd., Oxford, **1999**, pp. 573–607.

[2] a) C. K. Kleanthous, R. Deka, K. Davis, S. M. Kelly, A. Cooper, S. E. Harding, N. C. Price, A. R. Hawkins, J. R. Coggins, *Biochem. J.* **1992**, *282*, 687–695; b) D. G. Gourley, A. K. Shrive, I. Polikarpov, T. Krell, J. R. Coggins, A. R. Hawkins, N. W. Isaacs, L. Sawyer, *Nat. Struct. Biol.* **1999**, *6*, 521–525.

[3] a) C. Chaudhuri, K. Duncan, L. D. Graham, J. R. Coggins, *Biochem. J.* **1991**, *275*, 1–6; b) A. Shneier, C. Kleanthous, R. Deka, J. R. Coggins, C. Abell, *J. Am. Chem. Soc.* **1991**, *113*, 9416–9418; c) A. P. Leech, R. James, J. R. Coggins, C. Kleanthous, *J. Biol. Chem.* **1995**, *270*, 25827–25836.

[4] a) K. R. Hanson, I. A. Rose, *Proc. Natl. Acad. Sci. USA* **1963**, *50*, 981–988; b) B. W. Smith, M. J. Turner, E. Haslam, *J. Chem. Soc. Chem. Commun.* **1970**, 842–843; c) E. Haslam, M. J. Turner, D. Sargent, R. S. Thompson, *J. Chem. Soc. C* **1971**, 1489–1495.

[5] a) T. Krell, A. R. Pitt, J. R. Coggins, *FEBS Lett.* **1995**, *360*, 93–96; b) A. Shneier, J. M. Harris, C. Kleanthous, J. R. Coggins, A. R. Hawkins, C. Abell, *Bioorg. Med. Chem. Lett.* **1993**, *3*, 1399–1402; c) J. M. Harris, C. Klean-

thous, J. R. Coggins, A. R. Hawkins, C. Abell, *J. Chem. Soc. Chem. Commun.* **1993**, 1080 – 1081.
[6] J. M. Harris, C. González-Bello, C. Kleanthous, A. R. Hawkins, J. R. Coggins, C. Abell, *Biochem. J.* **1996**, *319*, 333 – 336.
[7] A. W. Roszak, D. A. Robinson, T. Krell, I. S. Hunter, M. Frederickson, C. Abell, J. R. Coggins, A. J. Lapthorn, *Structure* **2002**, *10*, 493 – 503.
[8] a) N. R. Salama, B. Shepherd, S. Falkow, *J. Bacteriol.* **2004**, *186*, 7926 – 7935; b) C. M. Sassetti, D. H. Boyd, E. J. Rubin, *Mol. Microbiol.* **2003**, *48*, 77 – 84.
[9] The Database of Essential Genes (DEG) provides details of key bacterial genes. For details see: a) http://www.essentialgene.org; b) R. Zhang, Y. Lin, *Nucleic Acids Res.* **2009**, *37*, D455 – D458; c) R. Zhang, H. Y. Ou, C. T. Zhang, *Nucleic Acids Res.* **2004**, *32*, D271 – D272.
[10] V. F. V. Prazeres, L. Castedo, H. Lamb, A. R. Hawkins, C. González-Bello, *ChemMedChem* **2009**, *4*, 1980 – 1984.
[11] The X-ray crystal structure is available from the Protein Data Bank (PBD: 1H0S): D. A. Robinson, A. W. Roszak, M. Frederickson, C. Abell, J. R. Coggins, A. J. Lapthorn, *Structural Basis for Specificity of Oxime Based Inhibitors Towards Type II Dehydroquinase from* M. tuberculosis (to be published). Residues 20–25 are not visible, including the essential residue Tyr24.
[12] P. Evans, *Acta Crystallogr. D Biol. Crystallogr.* **2005**, *62*, 72 – 82.
[13] G. Sheldrick, XPREP. Space Group Determination and Reciprocal Space Plots, **1991**. Siemens Analytical X-ray Instruments, Madison, WI, USA.
[14] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. D Biol. Crystallogr.* **1997**, *53*, 240 – 255.
[15] A. T. Brünger, *Methods Enzymol.* **1997**, *277*, 366 – 396.
[16] I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. 3rd Arendall, J. Snoeyink, J. S. Richardson, D. C. Richardson, *Nucl. Acids Res.* **2007**, *35*, W375 – W383.
[17] a) T. Krell, A. R. Pitt, J. R. Coggins, *FEBS Lett.* **1995**, *360*, 93 – 96; b) T. Krell, M. J. Horsburgh, A. Cooper, S. M. Kelly, J. R. Coggins, *J. Biol. Chem.* **1996**, *271*, 24492 – 24497.
[18] DHQ2 can exist in an open or closed conformation depending on the position adopted by the loop that covers its active site: D. Maes, L. A. González-Ramírez, J. López-Jaramillo, B. Yu, H. De Bondt, I. Zegers, E. Afonina, J. M. García-Ruíz, S. Gulnik, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 463 – 471.
[19] D. A. Robinson, K. A. Stewart, N. C. Price, P. A. Chalk, J. R. Coggins, A. J. Lapthorn, *J. Med. Chem.* **2006**, *49*, 1282 – 1290.
[20] V. F. V. Prazeres, L. Tizón, J. M. Otero, P. Guardado-Calvo, A. L. Llamas-Saiz, M. J. van Raaij, L. Castedo, H. Lamb, A. R. Hawkins, C. González-Bello, *J. Med. Chem.* **2010**, *53*, 191 – 200.
[21] M. Frederickson, A. W. Roszak, J. R. Coggins, A. J. Lapthorn, C. Abell, *Org. Biomol. Chem.* **2004**, *2*, 1592 – 1596.
[22] R. J. Payne, A. Riboldi-Tunnicliffe, O. Kerbarh, A. D. Abell, A. J. Lapthorn, C. Abell, *ChemMedChem* **2007**, *2*, 1010 – 1013.
[23] a) H. Li, A. D. Robertson, J. H. Jensen, *Proteins Struct. Funct. Bioinf.* **2005**, *61*, 704 – 721; b) D. C. Bas, D. M. Rogers, J. H. Jensen, *Proteins Struct. Funct. Bioinf.* **2008**, *73*, 765 – 783.
[24] D. G. Gourley, J. R. Coggins, N. W. Isaacs, J. D. Moore, I. G. Charles, A. R. Hawkins, *J. Mol. Biol.* **1994**, *241*, 488 – 491.
[25] C. Sánchez-Sixto, V. F. V. Prazeres, L. Castedo, S. W. Suh, H. Lamb, A. R. Hawkins, F. J. Cañada, J. Jiménez-Barbero, C. González-Bello, *ChemMedChem* **2008**, *3*, 756 – 770.
[26] A. G. Leslie, *Acta Crystallogr. D Biol. Crystallogr.* **2005**, *62*, 48 – 57.
[27] M. D. Winn, *J. Synchrotron Radiat.* **2002**, *10*, 23 – 25.
[28] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022 – 1025.
[29] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. A Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126 – 2132.
[30] COLLECT, Nonius BV, R. W. Hooft, Delft, **1998**.
[31] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307 – 326.
[32] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658 – 674.
[33] J. Stewart, *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1 – 45.
[34] Gaussian 09, Revision A.2: M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford, CT (USA), **2009**.
[35] a) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179 – 5197; b) http://q4md-forcefieldtools.org/RED/ (Last accessed: August 9, 2010).
[36] http://amber.scripps.edu/ (Last accessed: August 9, 2010).
[37] a) J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157 – 1174; b) J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graphics Modell.* **2006**, *25*, 247 – 260.
[38] N. C. Price, D. J. Boam, S. M. Kelly, D. Duncan, T. Krell, D. G. Gourley, J. R. Coggins, R. Virden, A. R. Hawkins, *Biochem. J.* **1999**, *338*, 195 – 202.
[39] a) http://biophysics.cs.vt.edu/H++; b) J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath, A. Onufriev, *Nucleic Acids Res.* **2005**, *33*, W368 – W371; c) R. Anandakrishnan, A. Onufriev, *J. Comput. Biol.* **2008**, *15*, 165 – 184.
[40] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, O. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668 – 1688.
[41] J. Aaqvist, *J. Phys. Chem.* **1990**, *94*, 8021 – 8024.
[42] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, *J. Chem. Phys.* **1983**, *79*, 926 – 935.
[43] T. A. Darden, D. York, L. G. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089 – 10092.
[44] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327 – 341.
[45] W. L. DeLano, *The PyMOL Molecular Graphics System*, **2008**, DeLano Scientific LLC, Palo Alto, CA (USA): http://www.pymol.org (Last accessed: August 9, 2010).

Received: July 9, 2010
Published online on September 2, 2010

