Isolation and molecular characterization of the shikimate dehydrogenase domain from the Toxoplasma gondii AROM complex

ARTICLE *in* MOLECULAR AND BIOCHEMICAL PARASITOLOGY · APRIL 2014

Impact Factor: 1.79 · DOI: 10.1016/j.molbiopara.2014.04.002 · Source: PubMed

4 AUTHORS, INCLUDING:

Gianni M Castiglione
University of Toronto
2 PUBLICATIONS **1** CITATION

Molecular & Biochemical Parasitology xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Short communication

Isolation and molecular characterization of the shikimate dehydrogenase domain from the *Toxoplasma gondii* AROM complex

James Peek[a], Gianni Castiglione[a], Thomas Shi[a], Dinesh Christendat[a,b,*]

[a] Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada M5S 3B2
[b] Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario, Canada M5S 3B2

ARTICLE INFO

Article history:
Received 25 September 2013
Received in revised form 24 March 2014
Accepted 3 April 2014
Available online xxx

Keywords:
Shikimate dehydrogenase
Toxoplasma gondii
Shikimate pathway
AROM complex
Enzyme kinetics
Enzyme inhibition

ABSTRACT

The apicomplexan parasite *Toxoplasma gondii*, the etiologic agent of toxoplasmosis, is estimated to infect 10–80% of different human populations. *T. gondii* encodes a large pentafunctional polypeptide known as the AROM complex which catalyzes five reactions in the shikimate pathway, a metabolic pathway required for the biosynthesis of the aromatic amino acids and a promising target for anti-parasitic agents. Here, we present the isolation, cloning and kinetic characterization of the shikimate dehydrogenase domain (TgSDH) from the *T. gondii* AROM complex. Recombinant TgSDH catalyzed the $NADP^+$-dependent oxidation of shikimate in the absence of the remaining AROM domains and was sensitive to inhibition by a previously identified SDH inhibitor. Analysis of the TgSDH amino acid sequence revealed a number of novel insertions not found in SDH homologs from other organisms. Nevertheless, a three-dimensional structural model of TgSDH predicts a high level of conservation in the 'core' structure of the enzyme.

The protozoan parasite *Toxoplasma gondii* is the causative agent of toxoplasmosis, a potentially severe disease that can cause encephalitis, chorioretinitis, and lymphandenopathy [1]. Studies have also linked *T. gondii* infection with schizophrenia and suicidal behavior [2,3]. Although the felid family represents its primary host, *T. gondii* can infect a large range of warm-blooded animals and is estimated to infect 10–80% of different human populations [4]. While infection is often asymptomatic, it can have particularly serious consequences for pregnant or immunocompromised individuals.

The identification of the shikimate pathway in members of the phylum Apicomplexa, which includes *T. gondii* and malarial pathogen, *Plasmodium falciparum*, provided a promising new target for anti-parasitic drugs [5,6]. The pathway, which was originally assumed to be present only in plants, fungi and bacteria, is required for the biosynthesis of the aromatic amino acids, folate, ubiquinone, and a large number of additional secondary metabolites [7]. Metazoans lack the pathway, and must therefore obtain the aromatic amino acids and folate from dietary sources. Glyphosate, the active ingredient in the popular herbicide RoundUp™, is a specific inhibitor of 5-enolpyruvyl shikimate 3-phosphate synthase, the enzyme that catalyzes the pathway's sixth reaction [8]. Interestingly, the ability of glyphosate to restrict the *in vitro* growth of *T. gondii* and *P. falciparum* provided early evidence for the existence of the shikimate pathway in these organisms [6]. However, only a single enzyme from the apicomplexan shikimate pathway, *P. falciparum* chorismate synthase, has been recombinantly produced and biochemically characterized [9].

A total of seven enzyme-catalyzed reactions form the shikimate pathway (Fig. S1) [7]. In bacteria, the enzymes of the shikimate pathway are monofunctional. The equivalent plant enzymes are also monofunctional, with the exception of a bifunctional dehydroquinate dehydratase-shikimate dehydrogenase which catalyzes reactions three and four [10]. In fungi, steps two through six of the shikimate pathway are catalyzed by a large pentafunctional polypeptide known as the AROM complex (Fig. S1) [11]. This complex has the following functional domains (from N- to C-terminus): dehydroquinate synthase, 5-enolypyruvylshikimate-3-phosphate synthase, shikimate kinase, dehydroquinate dehydratase, and shikimate dehydrogenase (SDH). These domains catalyze steps 2, 6, 5, 3, and 4 of the shikimate pathway, respectively [7]. The first and last enzymes of the fungal shikimate pathway, 3-deoxy-D-arabinoheptulosonate 7-phosphate synthase and chorismate synthase, are discrete enzymes.

Early analysis of the *T. gondii* genome revealed the presence of a large ~20,000 kb region which shared significant sequence

Abbreviations: SDH, shikimate dehydrogenase; TgSDH, shikimate dehydrogenase domain from the *Toxoplasma gondii* AROM complex; EGCG, epigallocatechin gallate; IC_{50}, half maximal inhibitory concentration.

* Corresponding author at: 25 Willcocks Street, Toronto, Ontario, Canada M5S 3B2. Tel.: +1 416 946 8337; fax: +1 416 978 5878.
E-mail address: dinesh.christendat@utoronto.ca (D. Christendat).

http://dx.doi.org/10.1016/j.molbiopara.2014.04.002
0166-6851/

similarity with the fungal AROM complex [5]. The putative functional domains of the *T. gondii* polypeptide encoded by this region are organized in an identical arrangement to those of the fungal AROM polypeptide (Fig. S1). We identified the SDH domain of the *T. gondii* protein (TgSDH) by aligning the *T. gondii* primary sequence with SDH sequences from other organisms (Fig. S2). In particular, alignment with monofunctional SDH enzymes from bacteria allowed us to delineate the approximate N-terminal limit of the *T. gondii* SDH domain. Based on this information, we designed a series of TgSDH constructs and cloned them into an expression vector with an N-terminal hexahistidine tag (detailed cloning and protein expression protocols utilized in this study are provided as Supplementary material). We were careful to retain a strictly conserved SXS motif found near the N-terminus of the SDH domain as this motif is predicted to be required for binding the substrate, shikimate (Fig. S2) [10]. One of our designed constructs yielded a significant quantity of soluble recombinant protein when expressed in a bacterial host. This construct began on a leucine upstream of the SXS motif and extended to the C-terminal end of the AROM sequence, encompassing a total of 459 amino acids (Fig. S2).

The primary sequence of TgSDH is approximately 200 amino acids longer than typical bacterial SDH sequences. The additional length of the TgSDH sequence is associated with three large amino acid insertions spanning residues 30–93, 212–253, and 382–459 (Fig. S2). These insertions are absent from bacterial SDH sequences, and from representative plant and fungal sequences from *Arabidopsis thaliana* and *Neurospora crassa* (Fig. S2). Amino acid insertions have also been identified in other domains of the *T. gondii* AROM complex [5], and in other apicomplexan proteins including chorismate synthase from *P. falciparum* [9,12,13]. The insertions are relatively hydrophilic in nature and may therefore represent surface exposed loops, although their functional significance is uncertain.

In addition to these insertions, we identified an unusually high number of cysteine residues dispersed throughout the TgSDH primary sequence (Fig. S2). Of the 18 cysteines, 8 are located within or just outside the insertion regions. Interestingly, we also observed a high number of cysteines in the other domains of the *T. gondii* AROM complex. A total of 104 cysteines are found in the complete AROM polypeptide. By comparison, the entire sequence of the *N. crassa* AROM protein contains only 16 cysteines. It is possible that the cysteine residues within the insertion regions of the *T. gondii* AROM sequence are required for assembly of the quaternary structure of the apicomplexan AROM complex, or that they are the site of post-translational modifications such as palmitoylation [14], although these functions have yet to be demonstrated.

In spite of these unique features, the TgSDH primary sequence still shows significant similarity to SDH sequences from other organisms (Fig. S2). We identified a number of conserved active site residues. These residues include Ser16 and Ser18, which form the SXS motif important for substrate binding [10]. The invariant ionizable residues, Lys131 and Asp169, represent key catalytic groups [10]. Our sequence alignment also provided insight into the potential cofactor preference of TgSDH. We observed the conservation of a glycine-rich region (residues 131–134) required for binding the pyrophosphate portion of NAD(P)$^+$ (Fig. S2) [15]. Additionally, a so-called 'basic patch' comprising the motif NRTXXR (residues 260–265) is characteristic of enzymes that favor NADP$^+$ instead of NAD$^+$ [15,16]. The two arginines in this motif form hydrogen bonds with the adenine ribose phosphate of NADP$^+$. Based on the presence of these conserved residues, we predicted that TgSDH would be competent in catalyzing the NADP$^+$-dependent oxidation of shikimate.

We expressed the TgSDH construct in *Escherichia coli* and purified the recombinant protein from the bacterial cells by nickel-nitriloacetic acid affinity chromatography. Due to the cysteine-rich nature of the construct, purification of the protein required dithiothreitol to prevent disulfide bond formation and protein aggregation. Analysis of the purified protein by sodium dodecyl sulfate polyacrylamide gel electrophoresis revealed a distinct band between 50 kDa and 60 kDa, which corresponds to the expected molecular weight of the construct (Fig. 1A).

Table 1
Kinetic properties of TgSDH.[a]

Parameter	SA (nadp)	NADP (sa)	SA (nad)	NAD (sa)
K_{cat} (s^{-1})	2.78 ± 0.04	2.77 ± 0.04	<0.01	<0.01
K_M (μM)	52.7 ± 3.4	9.1 ± 0.6	ND	ND
K_{eff} (10^3 M^{-1} s^{-1})	53	304	ND	ND

[a] Kinetic properties were determined by varying either the substrate or cofactor, while holding the other (shown in parentheses) at saturation. $K_{eff} = K_{cat}/K_M$. SA, shikimate; ND, not determined.

The ability of the isolated SDH domain to retain enzymatic activity was uncertain, as removal of the domain from the AROM complex could alter its tertiary structure. The SDH domain may rely on the adjoining domains to properly fold. Moreover, the effect of the large stretches of amino acid insertions on the activity of the protein was not known, as no kinetic data had been previously published for any domain of the *T. gondii* AROM complex. Nevertheless, the TgSDH construct readily catalyzed the oxidation of shikimate (K_{cat} of $2.78 \pm 0.04\ s^{-1}$), using NADP$^+$ as a cofactor (Table 1 and Fig. 1B and C). We did not detect measurable activity with quinate, a similar compound containing an additional hydroxyl group in the C1-position, demonstrating the high level of substrate specificity of the enzyme. While the measured rate of shikimate turnover is lower than those reported for some bacterial SDH enzymes (*E. coli* SDH, for example, displays a K_{cat} of $237\ s^{-1}$) [15], it is similar to the published rate for the SDH enzyme from *Heliobacter pylori* ($7.7\ s^{-1}$) [17]. It is possible that the turnover rate of the recombinant TgSDH protein is reduced in comparison to the native SDH associated with the *T. gondii* AROM complex as a result of changes in the molecular environment of the protein caused by its isolation. Alternatively, the native SDH domain may also possess a relatively low turnover rate that is compensated for by the proximity of active sites in the AROM complex. Indeed, it has been suggested that the primary benefit of the AROM complex is the potential for the efficient transfer of substrates between the different domains of the protein [11]. Likewise, metabolite channeling between the domains of the bifunctional dehydroquinate dehydratase-shikimate dehydrogenase protein from *A. thaliana* is thought to increase flux through the plant shikimate pathway [10].

We measured a K_M of 52.7 ± 3.4 μM for TgSDH with respect to its substrate shikimate (Table 1 and Fig. 1B). This value is almost identical to those reported for the SDH homologs from *E. coli* (65 μM) [15] and *Mycobacterium tuberculosis* (30 μM) [18]. The similar K_M values of these enzymes are consistent with the conservation of substrate binding residues noted in our sequence analysis (Fig. S2) and may suggest that isolation of the SDH domain from the AROM complex has minimal effects on the overall architecture of the enzyme and its substrate binding site.

With respect to its cofactor, NADP$^+$, TgSDH has a K_M of 9.1 ± 0.6 μM (Table 1 and Fig. 1C). This value is somewhat lower than those reported for a number of bacterial SDH enzymes, including *E. coli* SDH, which exhibits a K_M of 56 μM for the cofactor [15]. We detected only trace TgSDH activity using NAD$^+$ as a cofactor ($K_{cat} < 0.01\ s^{-1}$), demonstrating that TgSDH is an NADP$^+$-specific dehydrogenase. A number of previously characterized SDH enzymes from plants and bacteria similarly display strong preference for NADP$^+$ over NAD$^+$ [10,15,16]. The specificity of TgSDH for NADP$^+$ is consistent with the presence of the NRTXXR motif identified in the primary sequence of the protein (Fig. S2).

Please cite this article in press as: Peek J, et al. Isolation and molecular characterization of the shikimate dehydrogenase domain from the *Toxoplasma gondii* AROM complex. Mol Biochem Parasitol (2014), http://dx.doi.org/10.1016/j.molbiopara.2014.04.002

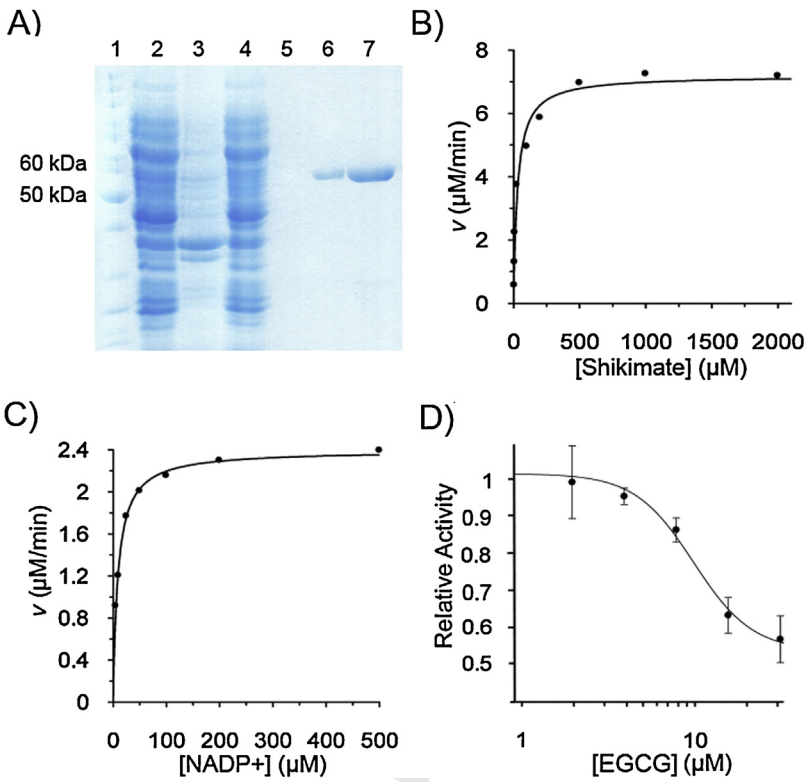

Fig. 1. (A) Sodium dodecyl sulfate polyacrylamide gel electrophoresis analysis of recombinant TgSDH. Lane contents are: (1) molecular weight ladder; (2) crude lysate from *E. coli* cells expressing TgSDH; (3) insoluble fraction from lysate; (4) Ni-NTA column flow-through fraction; (5) column wash fraction; (6) early elution fraction containing purified TgSDH; (7) later elution fraction. (B) and (C) Saturation kinetic profiles for substrate and cofactor usage by TgSDH. Enzyme-catalyzed oxidation of shikimate was monitored by following the reduction of $NADP^+$ ($\varepsilon = 6220\,M^{-1}\,cm^{-1}$) at 340 nm. For substrate-related kinetics, the concentration of shikimate was varied while $NADP^+$ was held at the saturating concentration of 2 mM. For cofactor-related kinetics, the concentration of $NADP^+$ was varied while shikimate was held at 2 mM. (D) Effect of the inhibitor, EGCG, on the activity of TgSDH.

Based on the measurable activity of recombinant TgSDH, we propose that the construct could be used in the future for *in vitro* screening of potential chemotherapeutic agents targeting the apicomplexan shikimate pathway. In the present study, we evaluated the effects of two previously identified SDH inhibitors on the activity of TgSDH. We recently identified the plant polyphenol, epigallocatechin gallate (EGCG), as an inhibitor of the SDH enzyme from the bacterium *Pseudomonas putida* with an IC_{50} of $3.0 \pm 0.2\,\mu M$ [19]. This compound also exhibited potent inhibitory activity against the SDH domain of the dehydroquinate dehydratase-shikimate dehydrogenase protein from *A. thaliana* ($IC_{50} = 2.1 \pm 0.3\,\mu M$). In an earlier study, Han et al. identified the small molecule, curcumin, as an inhibitor of the SDH enzyme from *H. pylori* with an IC_{50} of $15.4 \pm 2.2\,\mu M$ [17]. We tested the ability of both of these compounds to inhibit TgSDH. Although we were unable to detect significant inhibitory activity of curcumin against TgSDH under the assay conditions (pH 8.8), we found that TgSDH could be partially inhibited by EGCG, with an IC_{50} of $9.8 \pm 1.5\,\mu M$ (Fig. 1D). This value is only marginally higher than those calculated for the SDH homologs from *P. putida* and *A. thaliana*.

Given the similar biochemical properties of TgSDH and previously characterized SDH proteins, we predicted that these enzymes might share a conserved three-dimensional structure. To investigate this possibility, we modeled the structure of TgSDH using the I-TASSER protein modeling server (Fig. 2) [20]. The core structure of the TgSDH model (Fig. 2, dark blue region) is characterized by two juxtaposed α/β domains. The N-terminal domain contains four β-strands of strand order β2, β1, β3, β4, flanked by α-helices α1, α4, α5, α6, and α15. The C-terminal domain is composed of β-strands β7, β6, β5, β8, and β9 and α-helices α8, α11, α12, and α13. Two α-helices in the central portion of the protein, α7 and α14, are

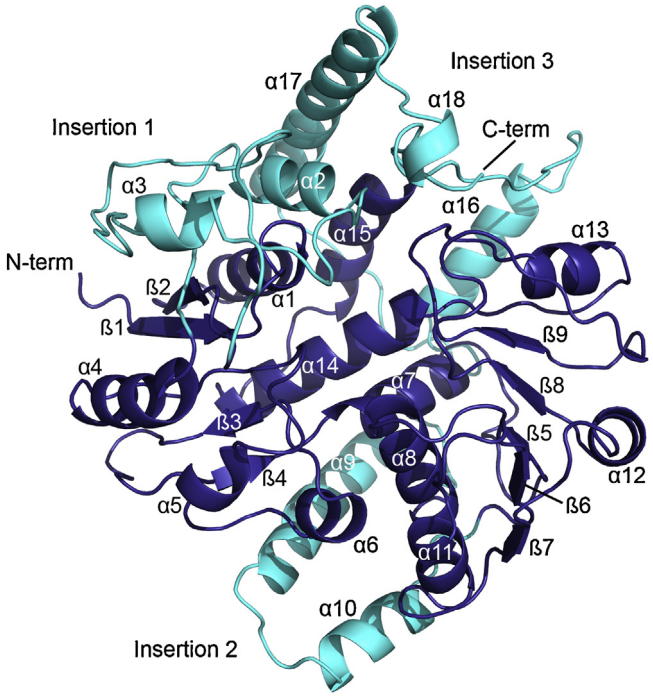

Fig. 2. Structural model of TgSDH. The three large surface-exposed insertions are colored teal, while the 'core' of the protein is colored dark blue. The TgSDH structural model was generated using I-TASSER [20] based on the crystal structure of *Geobacillus kaustophilus* SDH (PDB ID: 2EGG). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.) Q6

Please cite this article in press as: Peek J, et al. Isolation and molecular characterization of the shikimate dehydrogenase domain from the *Toxoplasma gondii* AROM complex. Mol Biochem Parasitol (2014), http://dx.doi.org/10.1016/j.molbiopara.2014.04.002

shared between the N- and C-terminal domains. The large amino acid insertions present in the protein form three distinct α-helical regions in the TgSDH model (Fig. 2, teal regions). The first insertion extends outward from β-strand β1, forms two small α-helices, α2 and α3, and then reconnects with the core structure *via* a-helix α4. The second insertion, which is located on the opposite face of the protein, extends from α-helix a8 and forms a broad loop containing two α-helices, α6 and α10, before rejoining the core structure at β-strand β6. The final insertion attaches to the core structure *via* α-helix α15. The α-helices, α16, α17, and the short helix α18, are formed by this insertion, which extends to the C-terminus of the protein (Fig. 2).

Superimposition of the TgSDH model with the structure of a well-characterized SDH enzyme from the bacterium *Aquifex aeolicus* [16] highlights the conserved core fold of these proteins (Fig. S3). The substrate binding site of the proteins is located in a cleft between their N- and C-terminal domains. Comparison of the TgSDH model with the *A. aeolicus* crystal structure demonstrates the strict conservation and appropriate structural positioning of the important substrate binding residues, Ser16 and Ser18, and the catalytic residues, Lys131 and Asp169 (Fig. S3). These features are predicted to contribute to the ability of TgSDH to catalyze the oxidation of shikimate, as we have observed in our kinetic analysis.

Our structural modeling of TgSDH suggests that the protein's three large amino acid insertions form surface exposed loops with α-helical character. These insertions have not been previously observed in other SDH proteins. Although the functional significance of the insertions is unknown, the unique surfaces they create could be exploited to design novel drugs with high specificity for the apicomplexan enzyme.

Funding

Q2 D.C. receives funding from a discovery grant from the Natural
Q3 Sciences and Engineering Research Council of Canada (NSERC).

Acknowledgement

We thank Dr. Michael Grigg (National Institute of Allergy and Infectious Disease, Bethesda, Maryland) for providing us with *Toxoplasma gondii* mRNA.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.molbiopara.2014.04.002.

References

[1] Montoya JG, Liesenfeld O. Toxoplasmosis. Lancet 2004;363:1965–76.
[2] Ling VJ, Lester D, Mortensen PB, Langenberg PW, Postolache TT. *Toxoplasma gondii* seropositivity and suicide rates in women. J Nerv Ment Dis 2011;199:440–4.
[3] Torrey EF, Bartko JJ, Lun ZR, Yolken RH. Antibodies to *Toxoplasma gondii* in patients with schizophrenia: a meta-analysis. Schizophr Bull 2007;33:729–36.
[4] Pappas G, Roussos N, Falagas ME. Toxoplasmosis snapshots: global status of *Toxoplasma gondii* seroprevalence and implications for pregnancy and congenital toxoplasmosis. Int J Parasitol 2009;39:1385–94.
[5] Campbell SA, Richards TA, Mui EJ, Samuel BU, Coggins JR, McLeod R, et al. A complete shikimate pathway in *Toxoplasma gondii*: an ancient eukaryotic innovation. Int J Parasitol 2004;34:5–13.
[6] Roberts F, et al. Evidence for the shikimate pathway in apicomplexan parasites. Q4 Nature 1998;393:801–5.
[7] Herrmann KM, Weaver LM. The shikimate pathway. Annu Rev Plant Physiol Plant Mol Biol 1999;50:473–503.
[8] Steinrücken HC, Amrhein N. The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid-3-phosphate synthase. Biochem Biophys Res Commun 1980;94:1207–12.
[9] Fitzpatrick T, Ricken S, Lanzer M, Amrhein N, Macheroux P, Kappes B. Subcellular localization and characterization of chorismate synthase in the apicomplexan *Plasmodium falciparum*. Mol Microbiol 2001;40:65–75.
[10] Singh SA, Christendat D. Structure of *Arabidopsis* dehydroquinate dehydratase-shikimate dehydrogenase and implications for metabolic channeling in the shikimate pathway. Biochemistry 2006;45:7787–96.
[11] Hawkins AR, Lamb HK, Moore JD, Charles IG, Roberts CF. The pre-chorismate (shikimate) and quinate pathways in filamentous fungi: theoretical and practical aspects. J Gen Microbiol 1993;139:2891–9.
[12] Dzierszinski F, Popescu O, Toursel C, Slomianny C, Yahiaoui B, Tomavo S. The protozoan parasite *Toxoplasma gondii* expresses two functional plant-like glycolytic enzymes. Implications for evolutionary origin of apicomplexans. J Biol Chem 1999;274:24888–95.
[13] Fox BA, Ristuccia JG, Bzik DJ. Genetic identification of essential indels and domains in carbamoyl phosphate synthetase II of *Toxoplasma gondii*. Int J Parasitol 2009;39:533–9.
[14] Chaudhary K, Donald RG, Nishi M, Carter D, Ullman B, Roos DS. Differential localization of alternatively spliced hypoxanthine-xanthine-guanine phosphoribosyltransferase isoforms in *Toxoplasma gondii*. J Biol Chem 2005;280:22053–9.
[15] Michel G, Roszak AW, Sauve V, Maclean J, Matte A, Coggins JR, et al. Structures of shikimate dehydrogenase AroE and its Paralog YdiB. A common structural framework for different activities. J Biol Chem 2003;278:19463–72.
[16] Gan J, et al. Structural and biochemical analyses of shikimate dehydrogenase AroE from *Aquifex aeolicus*: implications for the catalytic mechanism. Biochemistry 2007;46:9513–22.
[17] Han C, Wang L, Yu K, Chen L, Hu L, Chen K, et al. Biochemical characterization and inhibitor discovery of shikimate dehydrogenase from *Helicobacter pylori*. FEBS J 2006;273:4682–92.
[18] Zhang X, Zhang S, Hao F, Lai X, Yu H, Huang Y, et al. Expression, purification and properties of shikimate dehydrogenase from *Mycobacterium tuberculosis*. J Biochem Mol Biol 2005;38:624–31.
[19] Peek J, Shi T, Christendat D. Identification of novel polyphenolic inhibitors of Q5 shikimate dehydrogenase (AroE). J Biomol Screen 2014.
[20] Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 2010;5:725–38.