

# SCIENTIFIC REPORTS

OPEN

## The three-dimensional structure of *Clostridium absonum* 7 $\alpha$ -hydroxysteroid dehydrogenase: new insights into the conserved arginines for NADP(H) recognition

Deshuai Lou<sup>1,2</sup>, Bochu Wang<sup>1</sup>, Jun Tan<sup>1,3</sup>, Liancai Zhu<sup>1</sup>, Xiaoxi Cen<sup>1</sup>, Qingzhi Ji<sup>1</sup> & Yue Wang<sup>1</sup>

7 $\alpha$ -hydroxysteroid dehydrogenase (7 $\alpha$ -HSDH) can catalyse the oxidation of C7  $\alpha$ -OH of the steroid nucleus in the bile acid metabolism. In the paper we determined the crystal structure of 7 $\alpha$ -HSDH from *Clostridium absonum* (CA 7 $\alpha$ -HSDH) complexed with taurochenodeoxycholic acid (TCDCA) and NADP<sup>+</sup> by X-ray diffraction, which, as a tetramer, possesses the typical  $\alpha/\beta$  folding pattern. The four subunits of an asymmetric unit lie in the fact that there are the stable hydrophobic interactions between Q-axis-related subunits. Significantly, we captured an active state of the NADP<sup>+</sup>, confirming that nicotinamide moiety of NADP<sup>+</sup> act as electron carrier in the dehydrogenation. On the basis of crystal structure analysis, site-directed mutagenesis and MD simulation, furthermore, we find that the guanidinium of Arg38 can form the stable cation- $\pi$  interaction with the adenine ring of NADP<sup>+</sup>, and the cation- $\pi$  interaction and hydrogen bonds between Arg38 and NADP<sup>+</sup> have a significant anchor effect on the cofactor binding to CA 7 $\alpha$ -HSDH.

In mammals, the primary bile acids, including cholic acid (CA) and chenodeoxycholic acid (CDCA), are synthesised from cholesterol in hepatocyte, and then intestinal bacteria convert primary bile acids to secondary bile acids such as deoxycholic acid (DA), lithocholic acid (LCA) and ursodesoxycholic acid (UDCA). Bile acids/salts and their conjugations with glycine or taurine<sup>1</sup> achieve an efficient preservation through enterohepatic circulation in the body. In the bile acid metabolism, microbial transformation involves a series of complex processes including epimerization, oxidation, reduction, hydroxylation and dehydroxylation of the steroid nucleus which involves in a variety of microorganisms<sup>2</sup>.

7 $\alpha$ -hydroxysteroid dehydrogenases (7 $\alpha$ -HSDH) is responsible for reversibly catalysing the oxidation of C7  $\alpha$ -oriented hydroxyl of the steroid nucleus in the bile acid metabolism and has been detected in numerous genera of bacteria<sup>3–7</sup> and in mammal<sup>8,9</sup>. To date, three-dimensional structures of two 7 $\alpha$ -HSDHs from *Escherichia coli*<sup>10</sup> and *Brucella Melitensis* have been determined, and the characteristics of folding pattern in both the two 7 $\alpha$ -HSDHs indicate that they belong to the short-chain dehydrogenases/reductases (SDRs), whose members are dimeric or tetrameric and generally share low sequence identity possessing a single-domain globular Rossmann-related fold<sup>11</sup>. Based on the cofactor and active site sequence motifs SDRs can be divided to five subfamilies: “classical”, “extended”, “intermediate”, “divergent” and “complex”<sup>12</sup>. All of the 7 $\alpha$ -HSDHs reported previously belong to the “classical” subfamily with the cofactor binding site sequence motif ‘TA/G/SxxxGIG’ and the active site sequence motif ‘YxxxK’ (x, any amino acid residue)<sup>12</sup>. Three decades ago, scientists began to try to utilise 7 $\alpha$ -HSDH to produce chemical substance with medicinal values<sup>13–15</sup>. Enzyme kinetics data showed 7 $\alpha$ -HSDH had a broad substrate specificity. For example, a great variety of chemicals can be identified by 7 $\alpha$ -HSDH as substrate such as bile acids and their glycine/taurine conjugate<sup>4,6</sup>, benzaldehyde analogues<sup>16</sup> and

<sup>1</sup>Key Laboratory of Biorheological Science and Technology (Chongqing University), Ministry of Education, College of Bioengineering, Chongqing University, Chongqing 400030, China. <sup>2</sup>Postdoctoral research station of biology, Chongqing University, Chongqing 400030, China. <sup>3</sup>School of Biological & Chemical engineering, Chongqing University of Education, Chongqing 400067, China. Correspondence and requests for materials should be addressed to B.W. (email: wangbc2000@126.com) or J.T. (email: tanjunmail@126.com)

aromatic  $\alpha$ -ketoesters<sup>17</sup>. Besides,  $7\alpha$ -HSDH present excellent thermal stability that the activity of *Clostridium absonum*  $7\alpha$ -HSDH (CA  $7\alpha$ -HSDH) has no significant change after incubation at 25 °C for 108 h<sup>18</sup>; *Bacteroides fragilis*  $7\alpha$ -HSDH (BF  $7\alpha$ -HSDH) retains 90% of activity after heat treatment for 1 h at 65 °C<sup>6</sup>.  $7\alpha$ -HSDHs may possess great potential in synthetic application.

All of the dehydrogenases require coenzyme such as NAD(H) or NADP(H) which acts as electron carriers for catalysing the redox reaction, and NADP(H) is much more efficient than NAD(H) in bile acids synthesis process. Therefore, the cofactor specific dehydrogenases attracted a lot of attention. The first crystal structure of NADP(H)-dependent SDR member (mouse lung carbonyl reductase, PDB code: 1CYD) was determined in 1996 and this study<sup>19</sup> drew a conclusion the key determinant for cofactor specificity were the electrostatic environment surrounding the 2'-phosphate or hydroxyl of the adenosine ribose moiety of NADP(H) or NAD(H). The residues Lys17 and Arg39 are the key residues for mouse lung carbonyl reductase since the two residues could strongly interact with the 2'-phosphate of adenine ribose of NADPH through electrostatic attraction<sup>19</sup>, while the residue Thr38 occupies the spatial position that is similar to a residue Asp of the NAD(H)-dependent SDRs interacting with the hydroxyl groups of the adenine ribose<sup>20</sup>. The T38D mutant successfully switched the cofactor specificity of mouse lung carbonyl reductase confirming the vital importance of the negatively-charged residue at the end of the second  $\beta$ -strand for NAD(H)-utilising<sup>20</sup>. Based on the structural information, through primary and tertiary structure comparison, the previous studies presented the basic sequence characteristics of cofactor specificity for SDRs<sup>21,22</sup>. Generally, a negative charged (Asp or Glu) residue at the end of the second  $\beta$  strand indicates that the enzyme is NAD(H)-preferred and this residue forms hydrogen bonds to the 2'- and 3'-hydroxyl group of the adenine ribose moiety<sup>23,24</sup>. NADP(H)-preferred enzymes possess two key basic residues (Arg or Lys). One of them is located at the position just behind the second  $\beta$  strand, and the other is positioned just before the sequence 'GIG' of the cofactor binding sequence motif mentioned above<sup>19,24</sup>. Nevertheless, site-directed mutagenesis performed for altering the cofactor specificity usually cause reduced enzymatic activity<sup>25</sup> indicating that the determinant of cofactor recognition of SDRs may refer to a space structure involving more than one residue. In biological system, redox reaction prefer to utilising NADP(H). However both the two  $7\alpha$ -HSDHs whose three-dimensional structures were determined up to now utilising NAD(H). The three-dimensional structure the  $7\alpha$ -HSDH utilising NADP(H) has not been determined and their mechanism for cofactor specificity is still not clear.

To unravel the NADP(H) recognition in  $7\alpha$ -HSDH, we determined the crystal structure of CA  $7\alpha$ -HSDH complexed with NADP<sup>+</sup> and taurochenodeoxycholic acid (TCDCA) at resolution 2.0 Å. Site-directed mutageneses were performed to investigate the function of the residues involved in cofactor specificity. Furthermore, molecular dynamics (MD) simulation was carried out to study the interactions between  $7\alpha$ -HSDH and NADP(H)/TCDCA.

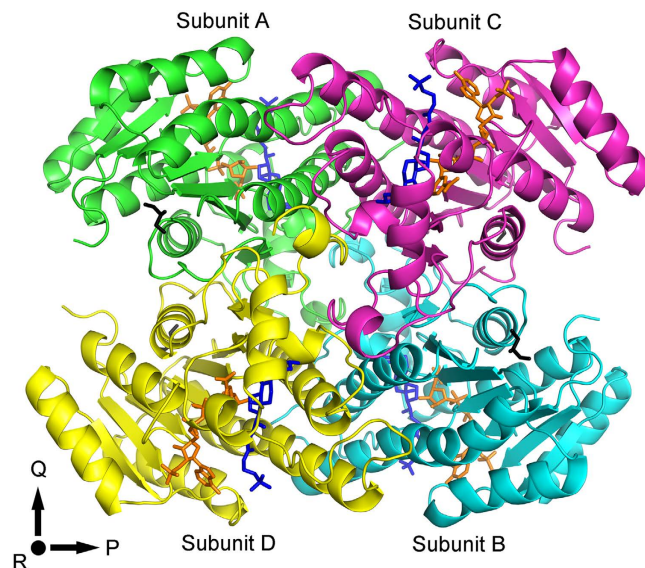
## Results and Discussion

**Enzyme purification, crystallisation, diffraction and subunit association.** The quality of purified enzyme was verified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (see Supplementary Fig. S1) and mass spectrometry (MS) (see Supplementary Fig. S2). Size exclusion chromatography (SEC) was applied to determine the molecular weight distribution and the results (see Supplementary Fig. S3) indicated in the CA  $7\alpha$ -HSDH mainly existed as dimers. However, seemingly a small proportion (arrows) possesses the association pattern of tetramer (see Supplementary Fig. S3B-1). To verify the association of the subunits of CA  $7\alpha$ -HSDH, the dimers from the first SEC were collected and analysed again by SEC (see Supplementary Fig. S3C-1). The tetramers from the two times were determined by MS (see Supplementary Fig. S3B-2 and Fig. S3C-2) and all the molecular weight distributions were corresponding to that of monomer of CA  $7\alpha$ -HSDH. It is proposed that a dynamic equilibrium between dimers and tetramers existed in the aqueous solution. The similar oligomerization state was found in the 6-phosphogluconate dehydrogenase from *Trypanosoma brucei*<sup>26</sup>. The crystal of CA  $7\alpha$ -HSDH is showed in the Fig. S4. The X-ray diffraction of the crystal is presented in Fig. S5.

**Structure quality assessment.** The three-dimensional structure of CA  $7\alpha$ -HSDH complexed with NADP<sup>+</sup> and TCDCA belonged to the  $P2_12_12_1$  space group and was determined at resolution of 2.0 Å by the molecular displacement method and refined to crystallographic  $R$ -factor of 0.182 ( $R_{\text{free}} = 0.222$ ). For the current model, the backbone dihedral angles of 97.7% of the residues are in the most favored region of the Ramachandran plot, and all the non-glycine residues were in the allowed region. More details of data collection and refinement can be found in Supplementary Table S1 and Table S2, respectively.

**Overall structure.** An asymmetric unit of CA  $7\alpha$ -HSDH crystal structure includes four subunits (Fig. 1). There is no major difference for the four subunits through pair-wise comparison (Main-Chain RMSD values ranged from 0.27 to 0.34 Å). Additionally, the glycerols in the crystal are from the cryo-protection solution.

To date, four three-dimensional structures of  $7\alpha$ -HSDHs have been deposited in PDB database. They are *Brucella Melitensis*  $7\alpha$ -HSDH with no ligand (PDB code: 3GAF) at resolution 2.20 Å and *Escherichia coli*  $7\alpha$ -HSDH (EC  $7\alpha$ -HSDH; PDB code: 1FMC, 1AHI and 1AHH). Structure of 1AHH contains ligand NAD<sup>+</sup> at resolution 2.30 Å, while both 1FMC and 1AHI have the ligands NADH and 7-oxo-GCDCA at resolution 1.80 Å and 2.30 Å, respectively. In spite of the low sequence identity (only 15–30%), SDRs generally display the well-conserved Rossmann fold. As one of the SDRs, with the similar structure to the two  $7\alpha$ -HSDHs above, CA  $7\alpha$ -HSDH possesses the typical Rossmann-fold that each monomer consists of seven parallel  $\beta$ -strands as a core ( $\beta$ -sheet core) and  $\alpha$ -helices located between  $\beta$ -strands at the surface of the globular molecules. The three-dimensional structures and folding topologies of CA  $7\alpha$ -HSDH and EC  $7\alpha$ -HSDH are shown as Supplementary Figs S6 and S7, respectively. Although the two conformations are very similar, there are some structural differences in detail. CA  $7\alpha$ -HSDH has longer C-terminal segment including an extra  $\alpha$ -helix ( $\alpha$ I) which involves Gln255-Tyr256-Ser257-Glu258-Tyr259. Differing from that of EC  $7\alpha$ -HSDH, such C-terminal



**Figure 1.** The structure of an asymmetric unit of CA 7 $\alpha$ -HSDH. TCDCA: blue stick; NADP<sup>+</sup>: orange stick; Glycerol: black stick.

structure of CA 7 $\alpha$ -HSDH does not directly surround the substrate-binding pocket but stretch outward. We have previously proved that only 19.1% of catalytic efficiency ( $k_{\text{cat}}/K_m$ ) remained after truncation of 2 amino acids at C-terminal end of CA 7 $\alpha$ -HSDH<sup>18</sup>. Therefore, the C-terminal segment has a great influence on the catalytic activity of CA 7 $\alpha$ -HSDH.

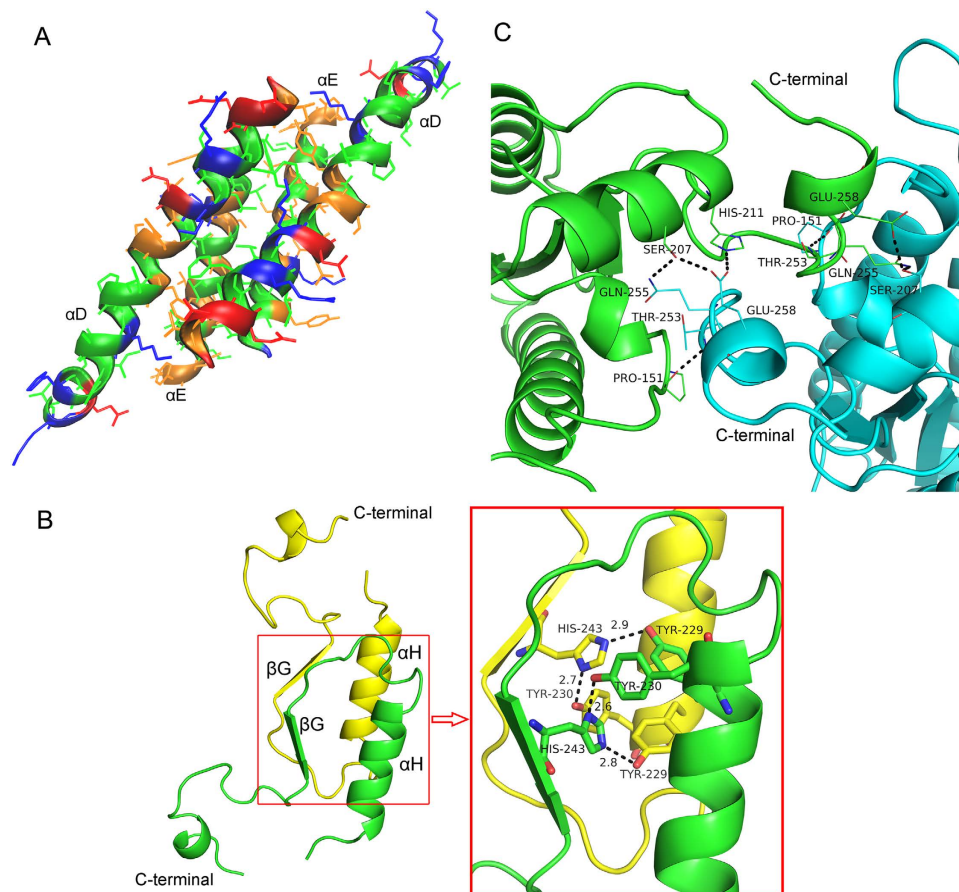
The four subunits in an asymmetric unit of CA 7 $\alpha$ -HSDH are shown in Fig. 1. Following the naming proposed by Ghosh *et al.*<sup>27</sup> the three mutually perpendicular directions are named *P*, *Q* and *R* axes, respectively. Along *Q*-axis, the interactions between subunit A and C (B and D) are mediated by four  $\alpha$ -helices,  $\alpha$ D (residues 107–135) and  $\alpha$ E (residues 156–176), which form  $\alpha$ D- $\alpha$ D and  $\alpha$ E- $\alpha$ E interfaces. As shown in Fig. 2A, the side chains of hydrophobic residues Phe110, Phe111, Phe114, Val118, Val121, Ala125, Ile129, Ile156, Ala157, Ile165, Leu168 and Ile172 are embed in the interspaces among the four helices, resulting in the stable hydrophobic interactions between subunit A and C (B and D). Similar *Q*-axis related interactions exist in the crystal structure of EC 7 $\alpha$ -HSDH<sup>10</sup>. The interactions between *P*-axis related subunits (subunits A and D, or subunits B and C) mainly result from  $\alpha$ H (residues 220–231) and  $\beta$ G (residues 242–245). Figure 2B shows the interaction pattern between subunit A (green) and subunit D (yellow). Phenolic hydroxyls of the residues Tyr230 A and Tyr229 D form hydrogen bonds with the two nitrogen atoms of imidazole ring in the residue His243 A, while phenolic hydroxyls of the residues Tyr230 D and Tyr229 A form hydrogen bonds with the two nitrogen atoms of imidazole ring of His243 D, and vice versa. This particular interaction pattern may play an important role in the subunit association along *P*-axis. In the aqueous solution, solvation effect could weaken the interaction between *P*-axis related subunits, resulting in that a multitude of tetramers dissociated into the dimers. This well explains the dynamic equilibrium between dimers and tetramers of CA 7 $\alpha$ -HSDH in the aqueous solution mentioned above.

In an asymmetric unit, the polar interactions between the two subunits in the diagonal position (subunits A and B, subunits C and D) involves the C-terminal as shown in Fig. 2C. The residues Glu258, Gln255 and Thr253 in one subunit form hydrogen bonds with the residues His211/Ser207, Ser207, Pro151 in the other subunit, respectively. The similar interactions mediated by C-terminal could be found in the asymmetric unit (four subunits) of 3 $\alpha$ , 20 $\beta$ -HSDH from *Streptomyces exfoliatus* (PDB code: 2HSD)<sup>28</sup>.

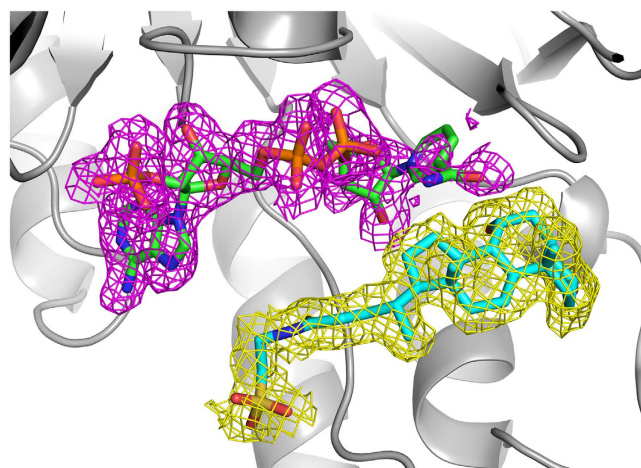
In addition, we found many salt bridges exist in the surface of the monomer such as Asp49-Arg20, Asp85-Arg83, Asp152-Arg155, Asp202-Arg205. The thermostability of CA 7 $\alpha$ -HSDH at 25 °C<sup>18</sup> may be attributed to these salt bridges.

For the molecular NADP<sup>+</sup>, electron density of the nicotinamide moiety of NADP<sup>+</sup> was rather weak (Fig. 3) suggesting this part was in an active state, which confirmed that nicotinamide moiety of NADP<sup>+</sup> act as electron carrier in the dehydrogenation reaction.

**The active site of CA 7 $\alpha$ -HSDH.** The well-conserved active sites with the sequence motif, YxxxK, exist in the three subfamilies (classical, extended and intermediate) of SDRs, and the tyrosine is the most conserved residue which exists in all members of SDRs<sup>12,29</sup>. Generally, a serine located before the tyrosine also play an important role in the catalytic process forming the Ser-Tyr-Lys triad. The active site in EC 7 $\alpha$ -HSDH, consists of Ser146, Tyr159 and Lys163, and the proposed catalysis mechanism suggested the residue Tyr159 might act as a catalytic base through deprotonation of phenolic group and the hydrogen bond with C7 hydroxyl of substrate while the Ser and Lys stabilise the substrate and cofactor respectively through hydrogen bonds<sup>10</sup>. Furthermore, the site-directed mutageneses of these three residues dramatically decrease the enzyme activity especially of the mutant of Tyr159<sup>30</sup>. Sequence alignment<sup>31</sup> suggested the Ser-Tyr-Lys triad is well conserved in five 7 $\alpha$ -HSDHs known to date but not in the CA 7 $\alpha$ -HSDH. However, CA 7 $\alpha$ -HSDH possesses the active site,

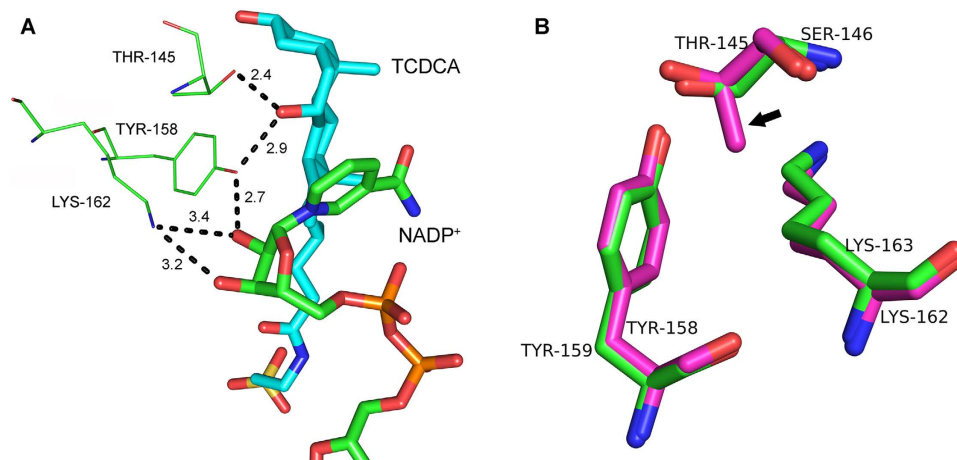


**Figure 2.** The Overall structure and the interactions in the subunits. (A) Helix D and E mediate interactions in Q-axis-related subunits (subunit A and C, B and D). blue: basic residues; red: acidic residues; orange: polar residues; green: nonpolar residues. (B) The C segment mediates interactions in P-axis-related subunits. Green: subunit A; yellow: subunit D. (C) The polar interactions between subunits in the diagonal position of an asymmetric unit structure of CA 7 $\alpha$ -HSDH.



**Figure 3.** Electron density map ( $|F_o| - |F_c|$  omit map) contoured at  $1.0\sigma$ . TCDCA is presented with yellow mesh and NADP<sup>+</sup> with red, respectively.

Thr145-Tyr158-Lys162, rather than Ser-Tyr-Lys triad. Here, the evolutionarily conserved serine was substituted by threonine. Interestingly, such mutation does not exist in the *Clostridium absonum* 7 $\beta$ -HSDH (CA 7 $\beta$ -HSDH). Figure 4A showed the interactions between active sites and NADP<sup>+</sup>/TCDCA and the distances (Å) between



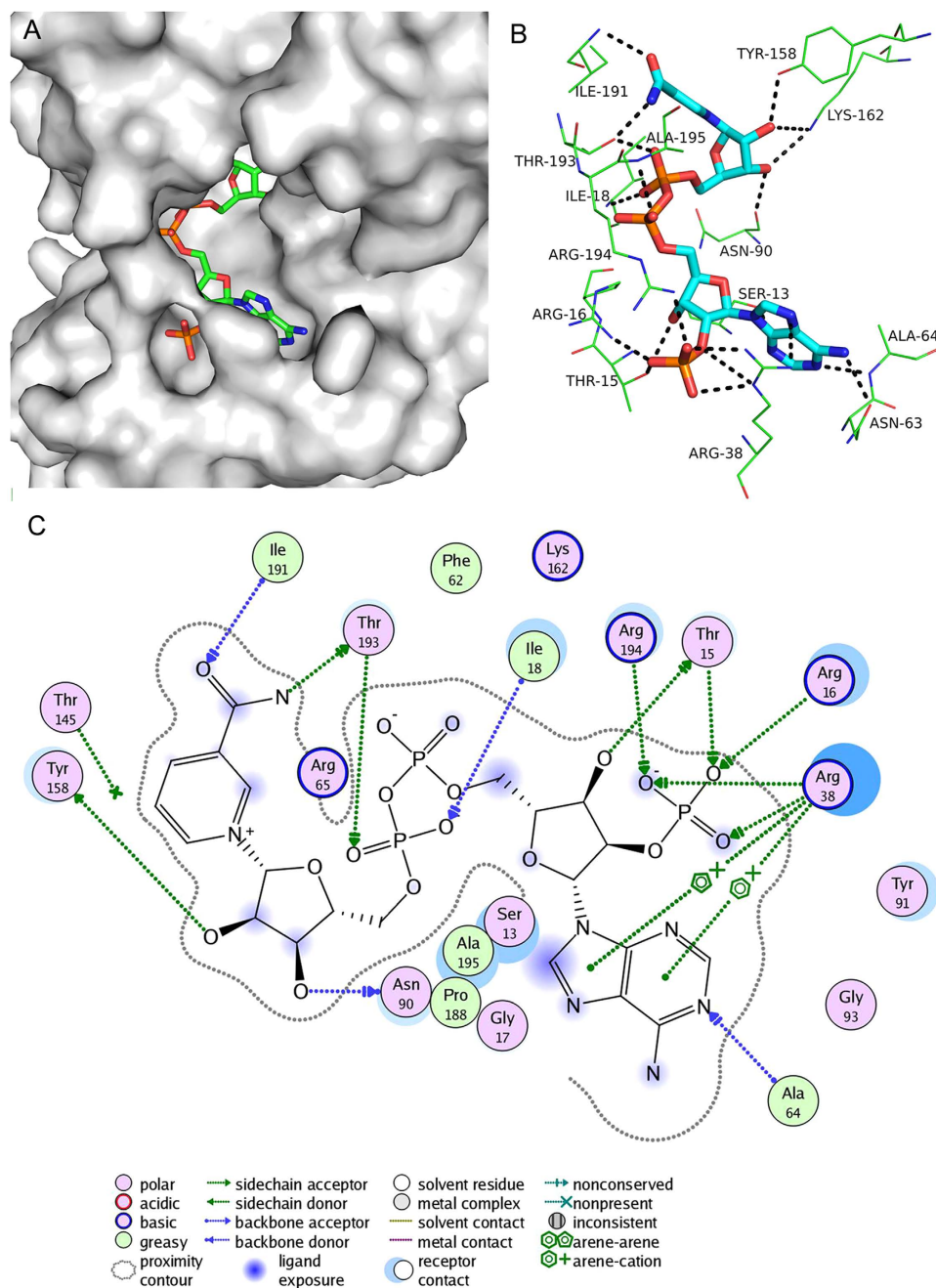
**Figure 4. The active site.** (A) The polar interactions between active sites and NADP<sup>+</sup>/TCDCA in CA 7 $\alpha$ -HSDH. (B) The superposed architecture of active sites of CA 7 $\alpha$ -HSDH (red carbon atoms) and EC 7 $\alpha$ -HSDH (green carbon atoms). An arrow indicates the Thr145 methyl of CA 7 $\alpha$ -HSDH.

relative atoms in CA 7 $\alpha$ -HSDH. The residue Thr145 forms hydrogen bond with C7-OH of TCDCA while the residue Lys162 forms hydrogen bonds with 2'- and 3'-hydroxyl group of the nicotinamide ribose moiety in NADP<sup>+</sup>. The phenolic hydroxyl of Tyr158 form hydrogen bonds with C7-OH of TCDCA and 2'-hydroxyl group of the nicotinamide ribose of NADP<sup>+</sup>. To observe difference of active-site spatial arrangement between CA 7 $\alpha$ -HSDH and EC 7 $\alpha$ -HSDH, we superposed the two catalytic triads (Fig. 4B). Although threonine has an additional methyl comparing with serine, no prominent difference of spatial placement exists for hydroxyl groups of threonine and serine in the active site. Thus, the mutation of serine to threonine during evolution does not cause the complete activity loss of CA 7 $\alpha$ -HSDH.

**Cofactor-binding site.** On account of the conserved hydrogen bonding of cofactor onto protein main chain of SDRs<sup>32</sup>, the conformation and orientation of NADP<sup>+</sup> in CA 7 $\alpha$ -HSDH are similar to that of other NADP(H)-dependent SDRs such as human estrogenic 17 $\beta$ -HSDH<sup>32</sup> (PDB code: 1QYV), trihydroxynaphthalene reductase<sup>33</sup> (PDB code: 1DOH) and tropinone reductase<sup>34</sup> (PDB code: 2AE2). Figure 5 showed the multiple interactions between CA 7 $\alpha$ -HSDH and its cofactor NADP<sup>+</sup> mainly involving Ser13, Thr15, Arg16, Ile18, Arg38, Asn63, Asn90, Tyr158, Lys162, Ile191, Thr193, Arg194, Ala195. Thr15, Arg16, Arg38 and Arg194 formed a positive charge-rich environment surrounding the 2'-phosphate of adenine ribose and these residues determined the cofactor specificity in CA 7 $\alpha$ -HSDH. According to the theory mentioned above, Arg16 and Arg38 are the determinant residues for cofactor specificity. In previous study, we performed the mutation analysis to verified the function of Arg38 of CA 7 $\alpha$ -HSDH in the cofactor recognition and found that the R38D mutant did not display enzyme activity with either NADP<sup>+</sup> or NAD<sup>+</sup><sup>18</sup>. Here, to investigate the function of the residues Arg16, Ala37 and Arg38, nine mutants of CA 7 $\alpha$ -HSDH were prepared as follows: A37D, R16G, R16G/A37D, R16G/R38D, A37D/R38I, A37D/R38Y, A37D/R38V, R16G/A37D/R38I and R16G/A37D/R38V. Interestingly, the results of enzyme assay indicate that only the mutant R16G has enzyme activity with NADP<sup>+</sup> and no activity with NAD<sup>+</sup>. No activity of the other eight mutants was detected with either NADP<sup>+</sup> or NAD<sup>+</sup>. Such results show the alteration of cofactor specificity in CA 7 $\alpha$ -HSDH is more complex and implementation of the residue with negative charges (aspartic acid) at position 37 or 38 could not form proper space orientation for NAD(H) recognition. From a structural point of view, the three arginines in the position 16, 38 and 194 form a positive environment surrounded the 2'-phosphate of adenine ribose of NADP<sup>+</sup> in CA 7 $\alpha$ -HSDH, but not more than two arginines were at the corresponding position in the other 10 SDRs listed as Supplementary Table S3.

Furthermore, the amino groups in Arg16, Arg38 and Arg194 form strong hydrogen bonds with the oxygen atoms in the 2'-phosphate of adenine ribose of NADP<sup>+</sup>. Especially, the nitrogen atom in guanidinium of Arg38 is close (~4 Å) to the adenine ring of NADP<sup>+</sup> forming the cation- $\pi$  stacking. Cation- $\pi$  interactions are ubiquitous in biological systems, which play central roles in molecular recognition, stabilisation of protein structures and nucleic acid structures, and their biological functions. Cheng *et al.*<sup>35</sup> studied the cation- $\pi$  interactions between NH<sub>4</sub><sup>+</sup> and benzene by the density function calculation, finding that the energy difference between the lowest unoccupied molecular orbital (LUMO) of NH<sub>4</sub><sup>+</sup> and the highest occupied molecular orbital (HOMO) of benzene is more than that between LUMO of benzene and HOMO of NH<sub>4</sub><sup>+</sup>. So the electrons preferentially move from the benzene to NH<sub>4</sub><sup>+</sup> combined with the analysis of net charge of atom. We speculate that the cation- $\pi$  interactions might lead to the electrons of the adenine ring of NADP<sup>+</sup> moving to the NH<sub>2</sub><sup>+</sup> of Arg38.

To investigate the effects of this electrostatic environment on NADP<sup>+</sup> binding, the mutants, R16G and R194G, were prepared and the enzyme activities were determined. Table 1 presents the kinetics constants of WT, R16G and R194G toward cofactor NADP<sup>+</sup>. The  $k_{cat}$  and  $K_m$  of R16G mutant increase by more than four times and five times compared with those of WT, respectively, while the catalytic efficiency ( $k_{cat}/K_m$ ) of R16G mutant decrease by 17.26%. The increase in  $K_m$  indicates that affinity of R16G mutant toward NADP<sup>+</sup> become weak, while the cofactor NADP(H) dissociated more easily from the binding site resulting in the increase in  $k_{cat}$ . Moreover, similar

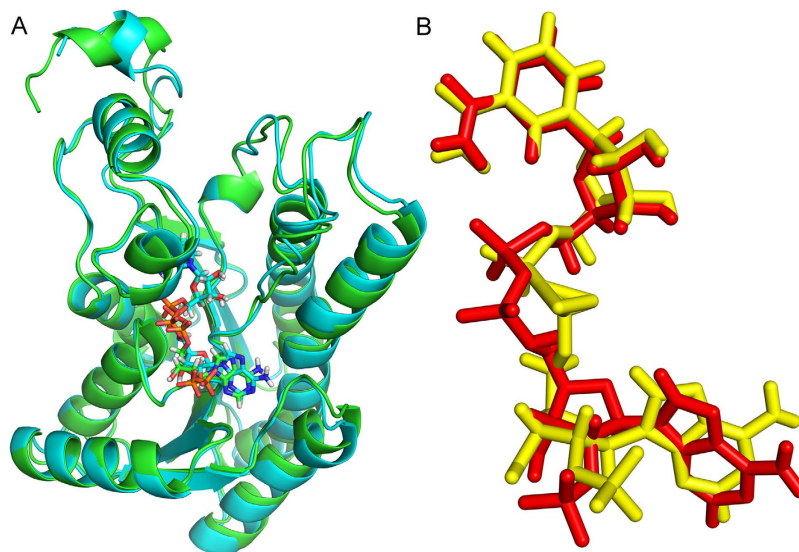


**Figure 5. The cofactor-binding site.** (A) The location and orientation of NADP<sup>+</sup>. (B) The key interactions and H-bond patterns between NADP<sup>+</sup> and CA 7 $\alpha$ -HSDH. (C) The related polar interactions.

Enzyme	$k_{cat}$ (s <sup>-1</sup> )	$K_m$ (mM)	$k_{cat}/K_m$ (s <sup>-1</sup> mM <sup>-1</sup> )
WT	13.76 (0.74)	0.34 (0.01)	40.10
R16G	58.55 (2.32)	1.77 (0.13)	33.18
R194G	28.66 (1.86)	0.72 (0.06)	39.78

**Table 1. Kinetics constants of WT, R16G and R194G toward NADP<sup>+</sup>.** WT: wild type; R16G: the mutant of arginine at position 16 to glycine; R194G: the mutant of arginine at position 194 to glycine; The values of standard deviations (SD) are in parentheses.

change of  $K_m$  value also occurred in the R194G mutants. Hence, the formation of the electrostatic environment which mainly involves the 3 arginines (Arg16, Arg38 and Arg194) mediates the cofactor specificity and recognition in right orientation.



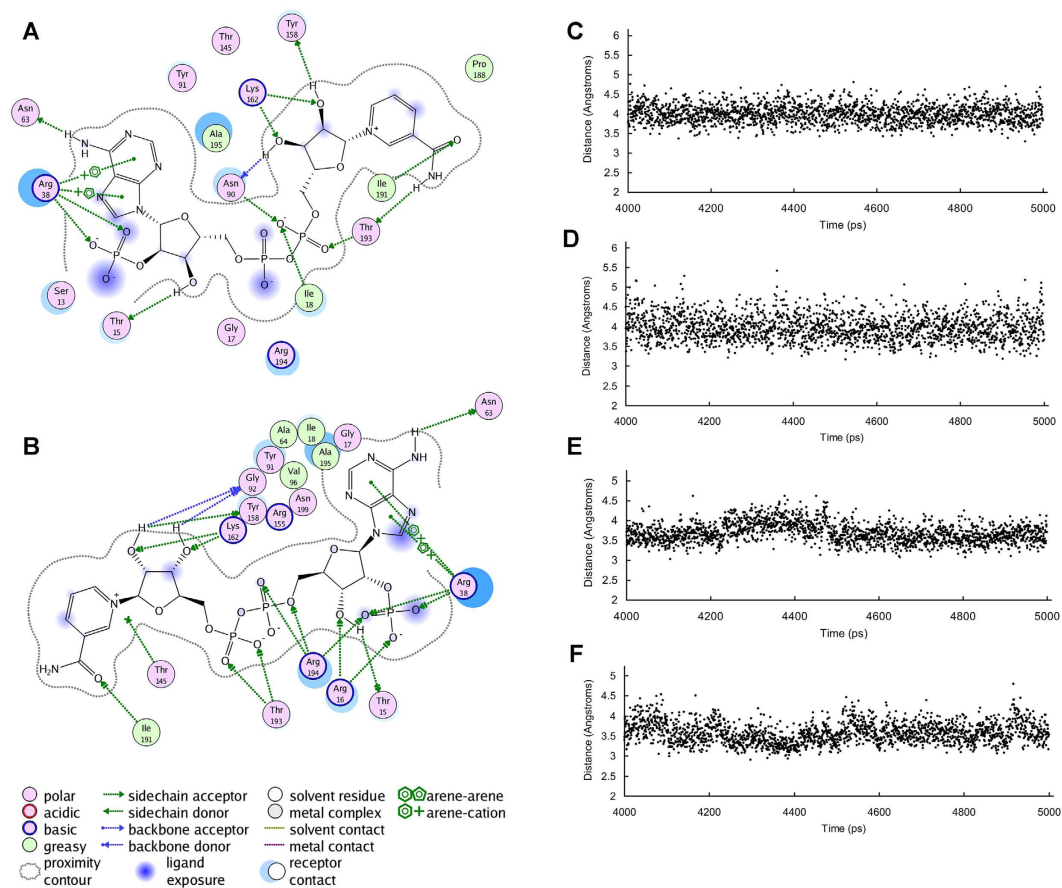
**Figure 6. Structural alignments of WT-ternary and R16G-ternary.** (A) Overall structural alignment (WT: cyan; R16G: green). (B) Cofactor alignment (NADP<sup>+</sup> of WT: red; NADP<sup>+</sup> of R16G: yellow).

**The key role of arginine residue.** To determine the role of the arginine residues Arg16 and Arg38, molecular dynamics (MD) simulations were performed to investigate the dynamics of ternary complexes of wild type, mutant R16G and R38D. The conformational stability was verified by root mean square deviation (RMSD) of the main chain of protein compared with the initial structure of protein (see Supplementary Fig. S8). For ternary complexes WT-TCDCA-NADP<sup>+</sup> (WT-ternary), R16G-TCDCA-NADP<sup>+</sup> (R16G-ternary), R38D-TCDCA-NADP<sup>+</sup> (R38D-ternary), during 5 ns MD simulations the RMSD of the backbone achieved stability quickly and kept fluctuation range from 1.0 to 1.5 Å. Based on MD simulations, Molecular Mechanics Poisson Boltzmann Surface Area (MM/PBSA) calculations were performed to evaluate the binding energies between enzymes and ligands in the three systems mentioned above. The plasticity of WT-ternary, R16G-ternary and R38D-ternary was evaluated through root mean square fluctuations (RMSF) (see Supplementary Fig. S9), and the results indicated that the catalytic site (T145-Y158-K162) was significantly rigid in all the systems. G54 and T103 had higher plasticity in WT-ternary and R16G-ternary respectively, and the same trend of R65 and I153 in the mutant R38D-ternary was observed. Nevertheless, the backbone plasticity was not remarkably influenced by the two mutations R16G and R38D.

Furthermore, the average structures of last 1 ns of the WT-ternary and R16G-ternary systems were calculated and aligned. Figure 6 presented the overall structural superposition suggesting no obvious conformational difference existed between the enzymic average structures of WT and that of the mutant except for the C-terminal. We find that the hydrogen bonds between Arg16, Arg38 and Arg194 and the oxygen atoms in the 2'-phosphate of adenine ribose of NADP<sup>+</sup> in the WT-ternary structure remain stable and the donor-acceptor distances of these H-bonds remain mostly between 2.7 Å and 3.1 Å during 5 ns dynamics (see Supplementary Fig. S10). Furthermore, we find that the distances between =NH<sub>2</sub><sup>+</sup> of guanidinium of Arg38 of WT-ternary and adenine ring of NADP<sup>+</sup> mostly fluctuate between 3.5 Å and 4.5 Å, indicating that the cation- $\pi$  interactions between Arg38 and NADP<sup>+</sup> are also stable (Fig. 7).

After the mutation of arginine at position 16 to glycine, both Gly16 and Arg194 do not form the stable hydrogen bond with NADP<sup>+</sup>, and the pyrophosphate bridge slightly rotation, which is attributed to the change of the surroundings of 2'-phosphate of adenine ribose of NADP<sup>+</sup> in the R16G mutant. Furthermore, we find that the hydrogen bond between Arg38 and 2'-phosphate of adenine ribose of NADP<sup>+</sup> still exists, but the H-bond is very weak because the donor-acceptor distances of the H-bond remain mostly between 3.2 Å and 3.8 Å during 5 ns dynamics (see Supplementary Fig. S11). The distances between =NH<sub>2</sub><sup>+</sup> of guanidinium of Arg38 of R16G-ternary and adenine ring of NADP<sup>+</sup> mostly fluctuate between 3.2 Å and 4.0 Å (Fig. 7). Although the stability of the H-bond between Arg38 and NADP<sup>+</sup> in the R16G-ternary decreased, the cation- $\pi$  interactions enhanced, compared with those of WT-ternary. Fortunately, the nicotinamide moiety keeps in the original position (Fig. 6B), so the R16G mutant still has strong enzyme activity. The MM/PBSA results indicated that the mutation does not influence the binding of TCDCA to enzyme-cofactor complex, while the binding energy of NADP<sup>+</sup> to enzyme-substrate complex significantly increased from -99.84 to -72.99 kcal·mol<sup>-1</sup> after the mutation of arginine at position 16 to glycine (see Supplementary Table S4).

After the mutation of arginine at position 38 to aspartic acid, the hydrogen bonds between Arg16 and NADP<sup>+</sup> remain stable, which is similar to the WT-ternary. In the mutant R38D without any activity, however, the hydrogen bond and cation- $\pi$  interactions between the residue at position 38 and NADP<sup>+</sup> are missing, resulting in the wide fluctuation of the adenine ring of NADP<sup>+</sup>. So the arginine at position 38 in the WT-ternary and the R16G-ternary can anchor the adenine ring of NADP<sup>+</sup> in the appropriate position through the cation- $\pi$  interactions and hydrogen bonds.



**Figure 7. The interactions between NADP<sup>+</sup> and CA 7 $\alpha$ -HSDH.** Residues involving in interactions between NADP<sup>+</sup> and WT-ternary (A) R16G-ternary (B). The distances between NE (C) NH<sub>2</sub> (D) of guanidinium of Arg38 of WT- ternary and adenine ring of NADP<sup>+</sup> in the cation- $\pi$  interactions. The distances between NADP<sup>+</sup> and Arg38 of mutant R16G- ternary and NE (E), NH<sub>2</sub> (F) of Arg 38 and adenine ring of NADP<sup>+</sup> in cation- $\pi$  interactions.

As seen in MD simulation results, it is not hard to get the following points of view. i) The correct binding position of nicotinamide moiety which directly involved in the oxidation-reduction reaction of NAD(P)(H) is the key determinant of the activity of R16G, and the correct spatial orientation ensure that the catalytic sites could form the proper interaction with substrate and cofactor. ii) The rotation of the pyrophosphate moiety of NADP(H) indicates that the flexibility of the pyrophosphate bridge may allow the shaking of nicotinamide moiety for electron transfer. However, molecular modification of CA 7 $\alpha$ -HSDH (such as site-directed mutagenesis of A37D and R38D) may remarkably change the NADP(H) binding pattern that nicotinamide moiety deviates the right position for interactions with substrate-enzyme. iii) Although the side-chain removal of Arg16 changed the electrostatic surroundings constructed by Arg16–Arg38–Arg194, the hydrogen bonds between 2'-phosphate of adenine ribose and Arg38 remained (Fig. 7) in the both WT and R16G systems. Once Arg38 was removed, the mutant R38D would lose enzyme activity. We conclude that the Arg38 is the key determinant of NADP(H) recognition in CA 7 $\alpha$ -HSDH. Such MD simulation result corresponded to that of mutagenesis experiments mentioned above. iv) The anchor effect of Arg38 though cation- $\pi$  interactions with adenine moiety may play an important role in the cofactor binding to CA 7 $\alpha$ -HSDH. Similar interactions could be observed in some of SDRs, such as *Sus scrofa* carbonyl reductase/20 beta-hydroxysteroid dehydrogenase and *Mus musculus* Sepiapterin reductase (see Supplementary Fig. S12).

## Conclusion

We showed the crystal structure of CA 7 $\alpha$ -HSDH including four subunits. CA 7 $\alpha$ -HSDH, as a member of SDRs, possesses the typical  $\alpha/\beta$  folding pattern and the very similar topological structure to EC 7 $\alpha$ -HSDH. Nevertheless, differing from EC 7 $\alpha$ -HSDH, CA 7 $\alpha$ -HSDH is NADP(H)-dependent and has a more complex C-terminal structure. The tetramer of CA 7 $\alpha$ -HSDH lies in the fact that there are the stable hydrophobic interactions between subunit A and C (B and D) and the hydrogen bonds between subunit A and D (B and C). Especially, the latter fixes the four subunits of CA 7 $\alpha$ -HSDH, but in the aqueous solution the hydrogen bonds between subunit A and D (B and C) could be weakened by solvation effect, resulting in the dissociation of the tetramer to the dimer. We further find that the residue Thr145 performs hydrogen abstraction through the hydrogen bond with C7-OH of TCDCA, while the serine executes the same function in the other SDRs. So the active residues related to the



dehydrogenation catalysis involve in Thr-Lys-Tyr combination in CA 7 $\alpha$ -HSDH, differing from the Ser-Lys-Tyr combination in the other SDRs.

Significantly, we captured an active state of the coenzyme NADP<sup>+</sup> with the rather weak electron density of the nicotinamide moiety, which confirmed that nicotinamide moiety of NADP<sup>+</sup> act as electron carrier in the dehydrogenation reaction. In the cofactor-binding site, the three arginines at the position 16, 38 and 194 form a positive environment surrounded the 2'-phosphate of adenine ribose of NADP<sup>+</sup> in CA 7 $\alpha$ -HSDH. The residues Arg16, Arg38 and Arg194 can form strong hydrogen bonds with the oxygen atoms in the 2'-phosphate of adenine ribose of NADP<sup>+</sup>. Especially, the guanidinium of Arg38 with positive charge can form the cation- $\pi$  interaction with the adenine ring of NADP<sup>+</sup>. On the basis of crystal structure analysis, site-directed mutagenesis and MD simulation, we find that the cation- $\pi$  interaction and hydrogen bonds between Arg38 and NADP<sup>+</sup> have a significant anchor effect on the cofactor binding to CA 7 $\alpha$ -HSDH. So we strongly believe that the residue Arg38 play an important role in the dehydrogenation reaction mediated by CA 7 $\alpha$ -HSDH.

## Methods

**Heterologous expression and purification of CA 7 $\alpha$ -HSDH.** CA 7 $\alpha$ -HSDH gene was cloned into plasmid pGEX-6p-1 and the GST-fused enzyme was heterologously expressed in *E. coli* BL21 (DE3) using the previously reported method<sup>18</sup>. The preliminarily purified enzyme was further purified for the protein crystallisation as followings. Enzyme was loaded onto a buffer (50 mM Tris-HCl pH8.0, 50 mM NaCl, 5% Glycerol, 5 mM DTT) pre-equilibrated monoQ 10/100 GL (GE Healthcare) column and 30 column volumes 0–100% buffer (50 mM Tris-HCl pH8.0, 1 M NaCl, 5% Glycerol, 5 mM DTT) was employed to wash and elute the enzyme. Then, the enzyme was loaded onto the buffer (50 mM Tris-HCl pH8.0, 200 mM NaCl) pre-equilibrated superdex 75 column (GE Healthcare). 1.2 column volumes of buffer C were used to wash and elute column and the purified enzyme was collected. Finally, the samples were verified by SDS-PAGE and MS. High purity enzyme was concentrated to a proper concentration for crystallisation experiments. Molecular weight distribution of CA 7 $\alpha$ -HSDH in aqueous solution was analysed through size exclusion chromatography (SEC) with standard protein (GE Healthcare) molecular weight 669, 440, 158, 75, 44, 29 and 13.7 kDa.

**Crystallisation.** Crystal screening using a crystallisation robot was performed at 293 K by sitting-drop vapour-diffusion method. 200 nL protein solution was mixed with 200 nL reservoir solution and equilibrated against 30  $\mu$ L reservoir solution. The concentrations of enzyme, NADP<sup>+</sup> and TCDCA in protein solution are 1.12 mM, 5.6 mM and 5.6 mM, respectively. Commercial crystallisation kits from Hampton Research were used for crystal screening. Initial crystals of complex were obtained in 0.1 M HEPES (pH 7.5) buffer containing 25% PEG-3350.

**Data collection and refinement.** The crystal was transferred to the cryo-protection solution (0.1 M HEPES pH 7.5, 25% PEG-3350, 25% Glycerol). Immediately, we mounted the crystal onto the goniometer which was surrounded by nitrogen at 100 K and all of the data collection works by using FR-E<sup>+</sup> and R-AXIS IV of RIGAKU. The structure of EC 7 $\alpha$ -HSDH (PDB code: 1FMC) was used as model for solving the CA 7 $\alpha$ -HSDH structure by molecular replacement techniques. The software iMosflm<sup>36</sup> used for the data processing and the software Phaser<sup>37</sup> was employed to phasing the crystal structure. The model building and refinement were through software REFMAC<sup>38</sup>, PHENIX<sup>39</sup> and Coot<sup>40</sup>. Software PHENIX<sup>39</sup> and PROCHECK<sup>41</sup> were used for structure validation and reporting MolProbity clashscores and rotamers. PyMOL were used for picture making. The structure of CA 7 $\alpha$ -HSDH has been deposited into the PDB (PDB code: 5EPO).

**Molecular dynamic simulations.** In the asymmetric unit (4 subunits) of CA 7 $\alpha$ -HSDH crystal, Initial structure of subunit A, NADP<sup>+</sup> and TCDCA were extracted for molecular dynamic simulation. Initial coordinates of Arg16Gly mutant was prepared by virtual mutation and structurally minimised using SYBYL program (TRIPOS Inc., St. Louis, MO). Wild-type ternary complex (WT-TCDCA-NADP<sup>+</sup>) or mutant ternary (R16G-TCDCA-NADP<sup>+</sup>) was immersed in TIP3 water box. Na<sup>+</sup> ions were added to neutralise negative charges of the system, then, topology and coordinate files were created by LEaP module. Three rounds of minimisation were carried out: (1) 4,000-step minimisation of solvent and ions; (2) 5,000-step minimisation of solution and side chains; (3) 10,000-step minimisation of whole system. Combination of steepest descent and conjugate gradient method was applied in these minimisations. Temperature of system increased from 0 K to 300 K within 50 ps MD simulation. Then, another 50 ps MD simulation in NPT ensemble and 5 ns final MD simulation were performed at constant temperature (300 K). The Amber Molecular Dynamics Package (AMBER12) was used for all the MD simulation<sup>42</sup>.

**Site-directed mutagenesis.** Table S5 (see Supplementary) summarised the mutagenic forward primers used for site-directed mutagenesis using overlapping PCR for the mutants: A37D, A37D/R38I, A37D/R38Y and A37D/R38V. On the basis of the mutants above, a series of mutants R16G, R16G/R38D, R16G/A37D, R16G/A37D/R38I and R16G/A37D/R38V were performed using conventional PCR with the forward primer 5'-CGGGATCCATGAAACGCCTGGAAGGCAAAGTGGCAATTGTGACCAGCTCTACTGGTGGCATTGGC-3' and the reverse primer 5'-CTTTTGC GGCCGCTTAGCGGGCAGTATTC-3'. These mutants mentioned above were performed by conventional or overlapping PCR with primeSTAR HS DNA polymerase (TaKaRa, Dalian, China) and all the genes were verified by sequencing (TaKaRa, Dalian, China). The gene of mutant R194G was synthesised and verified at Sangon Biotech (Sangon, Shanghai, China).

**Assay of enzyme activity.** Enzyme assay and kinetic analysis of wild-type CA 7 $\alpha$ -HSDH and the mutants were determined by the previously reported methods<sup>18,43</sup>.

## References

- Hofmann, A. F. & Hagey, L. R. Bile acids: chemistry, pathochemistry, biology, pathobiology, and therapeutics. *Cell. Mol. Life Sci.* **65**, 2461–2483 (2008).
- Bortolini, O., Medici, A. & Poli, S. Biotransformations on steroid nucleus of bile acids. *Steroids* **62**, 564–577 (1997).
- Baron, S. F., Franklund, C. V. & Hylemon, P. B. Cloning, sequencing, and expression of the gene coding for bile acid 7 alpha-hydroxysteroid dehydrogenase from *Eubacterium* sp. strain VPI 12708. *J. Bacteriol.* **173**, 4558–4569 (1991).
- Yoshimoto, T. *et al.* Cloning and sequencing of the 7 alpha-hydroxysteroid dehydrogenase gene from *Escherichia coli* HB101 and characterization of the expressed enzyme. *J. Bacteriol.* **173**, 2173–2179 (1991).
- Coleman, J. P., Hudson, L. L. & Adams, M. J. Characterization and regulation of the NADP-linked 7 alpha-hydroxysteroid dehydrogenase gene from *Clostridium sordellii*. *J. Bacteriol.* **176**, 4865–4874 (1994).
- Bennett, M. J., McKnight, S. L. & Coleman, J. P. Cloning and characterization of the NAD-dependent 7 alpha-Hydroxysteroid dehydrogenase from *Bacteroides fragilis*. *Curr. Microbiol.* **47**, 475–484 (2003).
- Ferrandi, E. E. *et al.* In search of sustainable chemical processes: cloning, recombinant expression, and functional characterization of the 7 alpha- and 7 beta-hydroxysteroid dehydrogenases from *Clostridium absonum*. *Appl. Microbiol. Biotechnol.* **95**, 1221–1233 (2011).
- Amuro, Y. *et al.* Partial purification and characterization of 7 alpha-hydroxysteroid dehydrogenase from rat liver microsomes. *Biochim. Biophys. Acta* **917**, 101–107 (1987).
- Shiotsuki, H., Maeda, Y. & Chijiwa, K. Purification and characterization of 7 beta-hydroxysteroid dehydrogenase from rabbit liver microsomes. *J. Steroid Biochem. Mol. Biol.* **91**, 185–190 (2004).
- Tanaka, N. *et al.* Crystal structures of the binary and ternary complexes of 7 alpha-hydroxysteroid dehydrogenase from *Escherichia coli*. *Biochemistry* **35**, 7715–7730 (1996).
- Bennett, M. J. *et al.* Steroid recognition and regulation of hormone action: crystal structure of testosterone and NADP<sup>+</sup> bound to 3 alpha-hydroxysteroid/dihydrodiol dehydrogenase. *Structure* **5**, 799–812 (1997).
- Kavanagh, K. L., Jornvall, H., Persson, B. & Oppermann, U. Medium- and short-chain dehydrogenase/reductase gene and protein families: the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell. Mol. Life Sci.* **65**, 3895–3906 (2008).
- Sutherland, J. D., Macdonald, I. A. & Forrest, T. P. The enzymic and chemical synthesis of ursodeoxycholic and chenodeoxycholic acid from cholic acid. *Prep. Biochem.* **12**, 307–321 (1982).
- Macdonald, I. A., Hutchison, D. M. & Forrest, T. P. Formation of urso- and ursodeoxycholic acids from primary bile acids by *Clostridium absonum*. *J. Lipid Res.* **22**, 458–466 (1981).
- Hirano, S. & Masuda, N. Epimerization of the 7-hydroxy group of bile acids by the combination of two kinds of microorganisms with 7 alpha- and 7 beta-hydroxysteroid dehydrogenase activity, respectively. *J. Lipid Res.* **22**, 1060–1068 (1981).
- Liu, Y. *et al.* The catalytic promiscuity of a microbial 7 alpha-hydroxysteroid dehydrogenase. Reduction of non-steroidal carbonyl compounds. *Steroids* **76**, 1136–1140 (2011).
- Zhu, D., Stearns, J. E., Ramirez, M. & Hua, L. Enzymatic enantioselective reduction of  $\alpha$ -ketoesters by a thermostable 7 $\alpha$ -hydroxysteroid dehydrogenase from *Bacteroides fragilis*. *Tetrahedron* **62**, 4535–4539 (2006).
- Lou, D., Wang, B., Tan, J. & Zhu, L. Carboxyl-terminal and Arg38 are essential for activity of the 7 alpha-hydroxysteroid dehydrogenase from *Clostridium absonum*. *Protein Pept. Lett.* **21**, 894–900 (2014).
- Tanaka, N. *et al.* Crystal structure of the ternary complex of mouse lung carbonyl reductase at 1.8 Å resolution: the structural origin of coenzyme specificity in the short-chain dehydrogenase/reductase family. *Structure* **4**, 33–45 (1996).
- Nakanishi, M. *et al.* Switch of coenzyme specificity of mouse lung carbonyl reductase by substitution of threonine 38 with aspartic acid. *J. Biol. Chem.* **272**, 2218–2222 (1997).
- Duax, W. L., Pletnev, V., Addlagatta, A., Bruenn, J. & Weeks, C. M. Rational proteomics I. Fingerprint identification and cofactor specificity in the short-chain oxidoreductase (SCOR) enzyme family. *Proteins* **53**, 931–943 (2003).
- Kristan, K., Stojan, J., Moller, G., Adamski, J. & Rizner, T. L. Coenzyme specificity in fungal 17 beta-hydroxysteroid dehydrogenase. *Mol. Cell. Endocrinol.* **241**, 80–87 (2005).
- Wierenga, R. K., Terpstra, P. & Hol, W. G. Prediction of the occurrence of the ADP-binding beta alpha beta-fold in proteins, using an amino acid sequence fingerprint. *J. Mol. Biol.* **187**, 101–107 (1986).
- Kallberg, Y., Oppermann, U., Jornvall, H. & Persson, B. Short-chain dehydrogenases/reductases (SDRs) - Coenzyme-based functional assignments in completed genomes. *Eur. J. Biochem.* **269**, 4409–4417 (2002).
- Arnold, P. *et al.* Glutamate-115 renders specificity of human 11 beta-hydroxysteroid dehydrogenase type 2 for the cofactor NAD<sup>+</sup>. *Mol. Cell. Endocrinol.* **201**, 177–187 (2003).
- Hanau, S. *et al.* Evidence for dimer/tetramer equilibrium In *Trypanosoma brucei* 6-phosphogluconate dehydrogenase. *Biochim. Biophys. Acta* **1834**, 2647–2652 (2013).
- Ghosh, D. *et al.* Three-dimensional structure of holo 3 alpha, 20 beta-hydroxysteroid dehydrogenase: a member of a short-chain dehydrogenase family. *Proc. Natl. Acad. Sci. USA* **88**, 10064–10068 (1991).
- Ghosh, D., Wawrzak, Z., Weeks, C. M., Duax, W. L. & Erman, M. The refined three-dimensional structure of 3 $\alpha$ , 20 $\beta$ -hydroxysteroid dehydrogenase and possible roles of the residues conserved in short-chain dehydrogenases. *Structure* **2**, 629–640 (1994).
- Filling, C. *et al.* Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J. Biol. Chem.* **277**, 25677–25684 (2002).
- Tanabe, T. *et al.* Roles of the Ser146, Tyr159, and Lys163 residues in the catalytic action of 7 alpha-hydroxysteroid dehydrogenase from *Escherichia coli*. *J. Biochem.* **124**, 634–641 (1998).
- Ferrandi, E. E. *et al.* In search of sustainable chemical processes: cloning, recombinant expression, and functional characterization of the 7 alpha- and 7 beta-hydroxysteroid dehydrogenases from *Clostridium absonum*. *Appl. Microbiol. Biotechnol.* **95**, 1221–1233 (2012).
- Shi, R. & Lin, S. X. Cofactor hydrogen bonding onto the protein main chain is conserved in the short chain dehydrogenase/reductase family and contributes to nicotinamide orientation. *J. Biol. Chem.* **279**, 16778–16785 (2004).
- Liao, D., Basarab, G. S., Gatenby, A. A., Valent, B. & Jordan, D. B. Structures of trihydroxynaphthalene reductase-fungicide complexes: implications for structure-based design and catalysis. *Structure* **9**, 19–27 (2001).
- Yamashita, A. *et al.* Structure of tropinone reductase-II complexed with NADP<sup>+</sup> and pseudotropine at 1.9 Å resolution: implication for stereospecific substrate binding and catalysis. *Biochemistry* **38**, 7630–7637 (1999).
- Cheng, J. G. *et al.* Research progress in cation- $\pi$  interactions. *Sci. China Ser. B* **51**, 709–717 (2008).
- Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sec. D, Biol. Crystallogr.* **67**, 271–281 (2011).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sec. D, Biol. Crystallogr.* **53**, 240–255 (1997).
- Zwart, P. H. *et al.* Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* **426**, 419–435 (2008).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sec. D, Biol. Crystallogr.* **60**, 2126–2132 (2004).

41. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
42. AMBER core team. *AMBER: A suite of programs for molecular dynamics simulations on biomolecules*. University of California, San Francisco. URL <http://ambermd.org/> (2012)
43. Liu, L., Aigner, A. & Schmid, R. D. Identification, cloning, heterologous expression, and characterization of a NADPH-dependent 7 $\beta$ -hydroxysteroid dehydrogenase from *Collinsella aerofaciens*. *Appl. Microbiol. Biotechnol.* **90**, 127–135 (2011).

### Acknowledgements

This work was supported by National Science and Technology Major Projects for “Major New Drugs Innovation and Development”(2014ZX09301306-007), the Program for Innovation Team Building at Institutions of Higher Education in Chongqing (KJTD201325), the Basic and Frontier Research Project of Chongqing (cstc2015jcyjA10094), the Visiting Scholar Foundation of the Key Laboratory of Biorheological Science and Technology (Chongqing University), the Ministry of Education (CQKLBST-2012-007), and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

### Author Contributions

D.L. solved the crystal structures of 7 $\alpha$ -HSDH, prepared the mutants and performed the kinetic assays. L.Z. and X.C. performed the molecular simulation. Q.J. and Y.W. performed the initial purification of 7 $\alpha$ -HSDH and produced the initial protein crystals. B.W. and J.T. designed experiments and analysed data. All authors reviewed the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Lou, D. *et al.* The three-dimensional structure of *Clostridium absonum* 7 $\alpha$ -hydroxysteroid dehydrogenase: new insights into the conserved arginines for NADP(H) recognition. *Sci. Rep.* **6**, 22885; doi: 10.1038/srep22885 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>