Tuberculosis 108 (2018) 143–150

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: www.elsevier.com/locate/tube

Activity loss by H46A mutation in *Mycobacterium tuberculosis* isocitrate lyase is due to decrease in structural plasticity and collective motions of the active site

Rohit Shukla[1], Harish Shukla[1], Timir Tripathi*

Molecular and Structural Biophysics Laboratory, Department of Biochemistry, North-Eastern Hill University, Shillong 793022, India

ARTICLE INFO

Keywords:
Mycobacterium tuberculosis
Molecular dynamic simulation
Principal component analysis
Active site
Collective motions
Residue interaction network

ABSTRACT

Mycobacterium tuberculosis isocitrate lyase (MtbICL) is a crucial enzyme of the glyoxylate cycle and is a validated anti-tuberculosis drug target. Structurally distant, non-active site mutation (H46A) in MtbICL has been found to cause loss of enzyme activity. The aim of the present work was to explore the structural alterations induced by H46A mutation that caused the loss of enzyme activity. The structural and dynamic consequences of H46A mutation were studied using multiple computational methods such as docking, molecular dynamics simulation and residue interaction network analysis (RIN). Principal component analysis and cross correlation analysis revealed the difference in conformational flexibility and collective modes of motions between the wild-type and mutant enzyme, particularly in the active site region. RIN analysis revealed that the active site geometry was disturbed in the mutant enzyme. Thus, the dynamic perturbation of the active site led to enzyme transition from its active form to inactive form upon mutation. The computational analyses elucidated the mutant-specific conformational alterations, differential dominant motions, and anomalous residue level interactions that contributed to the abrogated function of mutant MtbICL. An understanding of interactions of mutant enzymes may help in modifying the existing drugs and designing improved drugs for successful control of tuberculosis.

1. Introduction

Tuberculosis (TB), one of the leading causes of death by infectious diseases worldwide, is caused due to *Mycobacterium tuberculosis* infection [1]. The prevalence of *Mycobacterium* infection is in part owed to its ability to adapt and persist in changing environments. Most of the infected individuals show no signs or symptoms as the infecting bacteria establishes a latent infection within the host that is asymptomatic and persists for a longer time, sometimes even for several decades [2–4]. It has been estimated that ~35% of the world population acts as a reservoir for the latent infection of TB. In the current scenario, new drugs are required due to the emergence of multidrug resistant (MDR) and extreme drug resistant (XDR) strains of *M. tuberculosis*. The current drugs mainly act against the growing bacteria but are not good enough for persistent latent *Mycobacterium*. The drug-resistant form of TB is caused by the development of different mutations in the target and has been the chief cause of the current spread of TB [5–8]. Thus, drug resistance poses a challenge that requires improvement in diagnostics, surveillance, and therapeutic methods. Global reports of increasing MDR TB cases have underlined the need for research into the mechanism of drug resistance [9].

The enzymes of glyoxylate shunt play an important role during the latent infection of *M. tuberculosis*. When the fatty acids or two-carbon compounds become the sole carbon source of *M. tuberculosis* growth, the glyoxylate shunt is highly activated [10,11]. The first enzyme of the glyoxylate shunt isocitrate lyase (ICL) catalyzes the formation of glyoxylate and succinate from isocitrate. In the next reaction, conversion of glyoxylate to malate, proceeds via malate synthase (MS). In the light of the importance of enzymes of glyoxylate shunt, several studies have targeted ICL and MS for drug development. Several molecules have been identified against the *M. tuberculosis* isocitrate lyase (MtbICL) [11–15], but till date, no drug has been developed [16–18]. The catalytic mechanism, including the various steps involved in the reaction and the putative catalytic residues, has been proposed [11,19–21]. Structurally, MtbICL is organized in a tetrameric conformation where each monomer consists of 14 α helices and 14 β sheets forming α/β barrel structure (Fig. 1A). α12, α13 along with α14 are involved in the tetramerization through domain swapping with the neighboring subunit [11]. Till date, a total of four MtbICL crystal structures have been reported (PDB ID: 1F61; 1F8I; 1F8M; 5DQL). The structures of MtbICL

* Corresponding author.
E-mail addresses: timir.tripathi@gmail.com, ttripathi@nehu.ac.in (T. Tripathi).
[1] Both authors contributed equally to the work.

https://doi.org/10.1016/j.tube.2017.11.013
Received 3 October 2017; Received in revised form 25 November 2017; Accepted 27 November 2017

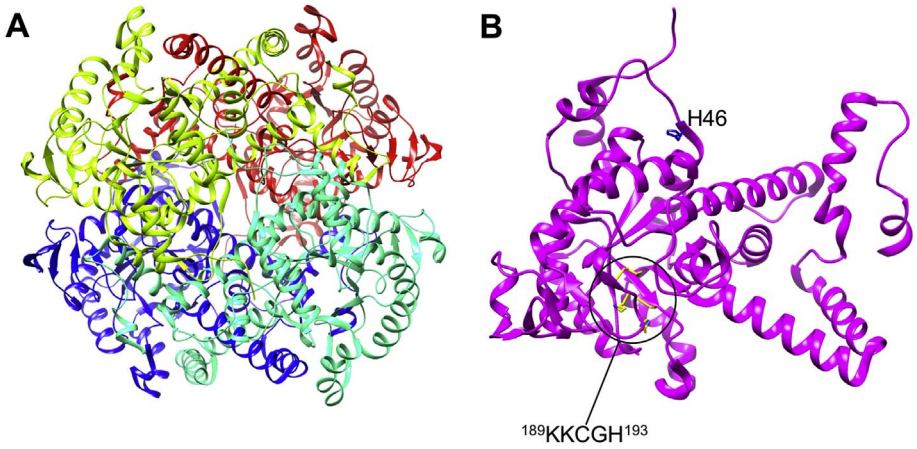

Fig. 1. Structure of MtbICL. (A) Ribbon structure of the tetrameric MtbICL (PDB ID: 1F8I). Each subunit is shown in different color. (B) Chain A of MtbICL tetramer showing the position of His46 in blue stick and active site signature sequence (^{189}KKCGH^{193}) in yellow sticks. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

have been elucidated in both apo-form and inhibitor bound form [11,22]. The conformation of apo-MtbICL differs from its substrate analog bound forms, especially in regions that control access to the active site. Structurally, they show that there are two active site regions in MtbICL; the first region contains the ICL signature sequence (^{189}KKCGH^{193}), and the second region consists of the last 18 residues (^{411}PNSSTTALTGSTEEGQFH^{428}) at the C-terminal end of the adjacent subunit (lid portion). Substrate binding at the active site loop leads to a movement of 10–15 Å to attain a closed conformation. The closed conformation brings the catalytic Cys191 of ICL signature sequence next to the substrate and completely closes off the active site. Closure of the active site loop blocks the accessibility to the catalytic residues that result in the movement of C-terminus lid (residues 411 to 428) of the adjacent subunit completing the catalytic conformation [11,23,24]. It is reported that MtbICL is inhibited by quercetin with an IC_{50} of 3.57 μM [12]. Quercetin is a natural flavanoid that inhibit mycobacterial growth *in vivo* [25,26]. A point mutation at H46 has been shown to develop resistance to quercetin [12].

Several research groups including ours have delineated the roles of specific amino acid residues in maintaining the structure and function of MtbICL. We recently showed a number of mutations far away from the active site region that modulated the environment at the substrate binding site, leading to global structural changes and complete loss of activity [12,23,24,27]. It has been reported that a distant H46A mutation leads to complete loss if MtbICL activity [12]. To understand how H46A mutation in MtbICL (position shown in Fig. 1B) leads to loss of activity, we performed extensive molecular dynamics simulations for wild-type and mutant MtbICL ($MtbICL_{H46A}$). Our result suggested that the activity of the enzyme is lost due to a decrease in structural plasticity and collective motions of the active site in $MtbICL_{H46A}$.

2. Methods

2.1. Molecular dynamics simulation (MDS)

Crystal structure of tetrameric MtbICL (PDB ID: 1F8I) was used as the starting conformation for computational analysis. MDS was carried out with Amberff99SB force field [28] in the GROMACS 4.6.5 [29–31] on an in-house supercomputer as earlier [13,23,27,32]. MDS was performed to sample alternative conformations of static X-ray structure and create receptor ensemble for subsequent ensemble docking. Wild-type and mutant MtbICL were solvated in a rectangular TIP3P water box. 72 Na^+ ions were added for neutralizing the systems using the genion tool. All systems were subjected to steepest energy minimization to give the maximum force below 1000 kJ mol^{-1} nm^{-1} to remove steric clashes of the systems. Long-range electrostatic forces were calculated using the Particle Mesh Ewald method [33]. A 1.0 nm radius cut-off was used for the computation of Lennard-Jones and Coulomb interactions. Bond lengths were constrained using the LINCS algorithm [34]. SHAKE algorithm was used to constrain all water bonds. After energy minimization, the systems were equilibrated. Then the position restraint simulation of 1 ns was carried out under NVT (the constant Number of particles, Volume and Temperature) and NPT (the constant Number of particles, Pressure and Temperature) conditions. Finally, both the systems were submitted to 50 ns MDS. A 2 fs interval was given for saving the coordinates. The root mean square deviation (RMSD) and root mean square fluctuation (RMSF) were calculated using g_rms and g_rmsf tools respectively. Time dependent secondary structures were calculated using do_dssp tool. Finally, the trajectories were analyzed by Visual Molecular Dynamics [35] and Chimera [36]. GRACE and Origin softwares were used for generating and visualizing the plots (http://plasma-gate.weizmann.ac.il/Grace).

2.2. Principal component analysis (PCA)

In order to explore the influence of H46A mutation on protein dynamics, the collective motions of wild-type and mutant MtbICL were investigated using PCA. PCA was used to calculate eigenvectors, eigenvalues and their projection along the first two principal components (PCs) of both wild-type and mutant MtbICL. The PCA method was present in the GROMACS software package. The concerted motions of macromolecules were extracted during MDS by PCA that are important for biological function [37,38]. Rotational and translational movements were removed from the trajectory used in the construction of variance/co-variance matrix. The positional covariance matrix C of atomic coordinates and its eigenvectors were used. The elements of the positional covariance matrix C were calculated by the following equation:

$$C_i = \langle (q_i - \langle q_i \rangle)(q_j - \langle q_j \rangle) \rangle \, (i, j = 1,2,...,3N) \quad (1)$$

where q_i is the Cartesian coordinate of the ith Cα atom and N is the number of Cα atoms in wild-type and mutant MtbICL. The average was calculated over the equilibrated trajectory after superimposition on a reference structure to remove overall translations and rotations by using a least-square fit method. The matrix C was symmetric and was diagonalized by an orthogonal coordinate transformation diagonal matrix Λ, for predicting the set of eigenvectors and eigenvalues λ_i:

$$\Lambda = T^T C_{ij} T \quad (2)$$

where the columns are the eigenvectors indicating to the direction of motion relative to $\langle q_i \rangle$ and each eigenvector associated with an eigenvalue that signified the total mean-square fluctuation of the system along the corresponding eigenvector. The last 20 ns production runs were used to perform the analysis. The amplitude of eigenvector was represented by eigenvalues in the multidimensional space. The movement of atoms along each eigenvector indicated the protein's concerted motions along each direction. The Cartesian trajectory coordinates