Gene 526 (2013) 318–324

Contents lists available at SciVerse ScienceDirect

Gene

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

Comparison of dynamics of wildtype and V94M human UDP-galactose 4-epimerase—A computational perspective on severe epimerase-deficiency galactosemia

David J. Timson [a], Steffen Lindert [b,c,*]

[a] School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, UK
[b] Department of Pharmacology, University of California San Diego, La Jolla, CA 92093, USA
[c] Center for Theoretical Biological Physics, La Jolla, CA, USA

ARTICLE INFO

Article history:
Accepted 21 May 2013
Available online 31 May 2013

Keywords:
Type III galactosemia
Molecular dynamics
UDP-galactose 4′-epimerase

ABSTRACT

UDP-galactose 4′-epimerase (GALE) catalyzes the interconversion of UDP-galactose and UDP-glucose, an important step in galactose catabolism. Type III galactosemia, an inherited metabolic disease, is associated with mutations in human GALE. The V94M mutation has been associated with a very severe form of type III galactosemia. While a variety of structural and biochemical studies have been reported that elucidate differences between the wildtype and this mutant form of human GALE, little is known about the dynamics of the protein and how mutations influence structure and function. We performed molecular dynamics simulations on the wildtype and V94M enzyme in different states of substrate and cofactor binding. In the mutant, the average distance between the substrate and both a key catalytic residue (Tyr157) and the enzyme-bound NAD^+ cofactor and the active site dynamics are altered making substrate binding slightly less stable. However, overall stability or dynamics of the protein is not altered. This is consistent with experimental findings that the impact is largely on the turnover number (k_{cat}), with less substantial effects on K_m. Active site fluctuations were found to be correlated in enzyme with substrate bound to just one of the subunits in the homodimer suggesting inter-subunit communication. Greater active site loop mobility in human GALE compared to the equivalent loop in *Escherichia coli* GALE explains why the former can catalyze the interconversion of UDP-N-acetylgalactosamine and UDP-N-acetylglucosamine while the bacterial enzyme cannot. This work illuminates molecular mechanisms of disease and may inform the design of small molecule therapies for type III galactosemia.

1. Introduction

UDP-galactose 4′-epimerase (GALE, EC 5.1.3.2) catalyzes the reversible interconversion of UDP-galactose and UDP-glucose (Timson, 2006). In many organisms, including humans, the enzyme can also catalyze the interconversion of the N-acetylated forms of these UDP-sugars (McCorvie and Timson, 2013a). The reaction with UDP-galactose is important because it plays a critical role in the Leloir pathway of galactose catabolism in which galactose is converted to the glycolytic intermediate glucose 6-phosphate (Frey, 1996; Leloir, 1951). Both UDP-sugar interconversions are important in the synthesis of glycoproteins and glycolipids. The majority of GALE enzymes characterized to date are homodimeric with each subunit containing one active site. These active sites contain an NAD^+ molecule whose function is to transiently oxidize the UDP-sugar substrate at the C4-OH position. This oxidation is followed by reorientation of the substrate in the active site and re-reduction of the ketone to an alcohol (Maitra and Ankel, 1971; Wong and Frey, 1977). The structure of human GALE (HsGALE) revealed that a key tyrosine residue (Tyr157) assists in this oxidation/re-reduction reaction cycle (Thoden et al., 2000).

Mutations in enzymes of the Leloir pathway are associated with the inherited metabolic disease, galactosemia. This disease has highly varied symptoms which range from altered blood chemistry to death in childhood. The most commonly detected form, type I galactosemia (galactose 1-phosphate uridylyltransferase deficiency, OMIM #230400) generally has serious clinical manifestations which get progressively worse through childhood (McCorvie and Timson, 2011a,b). Currently the only treatment is the restriction of dietary galactose (and its precursors such as lactose). While this slows the appearance of symptoms and can lessen their severity, it is rarely completely effective and it is often difficult to impose a strict diet onto young children (Fridovich-Keil, 2006). In contrast, galactokinase deficiency which results in type II galactosemia (OMIM #230200) is a

Abbreviations: GALE, UDP-galactose 4′-epimerase; HsGALE, human UDP-galactose 4′-epimerase; k_{cat}, turnover number; OMIM #230400, type I galactosemia, galactose 1-phosphate uridylyltransferase deficiency; OMIM #230200, type II galactosemia; OMIM #230350, type III galactosemia; MD, molecular dynamics; UDP-glc, uridine-5′-diphosphate glucose; UDP-gal, uridine-5′-diphosphate galactose; RMSF, root mean square fluctuations; DCCM, dynamical cross-correlation matrix; RMSD, root mean square distance.

* Corresponding author at: Department of Chemistry & Biochemistry, University of California San Diego, 9500 Gilman Drive, Mail Code 0365, La Jolla, CA 92093-0365, USA. Tel.: +1 858 534 2913(office); fax: +1 858 534 4974.
E-mail address: slindert@ucsd.edu (S. Lindert).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.05.027

relatively mild disease in which the only verified symptom is early onset cataracts (Bosch et al., 2002). Mutations in the gene encoding human GALE (HsGALE) are associated with type III galactosemia (OMIM #230350) (Fridovich-Keil et al., 1993; Holton et al., 1981; Openo et al., 2006). Historically, this disease was divided into two forms. The more severe, or "generalized" form was associated with severe symptoms similar to those of type I galactosemia and the recommended treatment is similar. The milder, or "peripheral", form was associated with increased levels of galactose metabolites in the blood and no intervention was considered necessary. However, it has now been conclusively shown that this is an over-simplification. The disease presents a continuum of severities of symptoms which are influenced by the mutations(s) present in the patient's *GALE* genes and on their environment (in which diet is a particularly important factor) (Chhay et al., 2008; Openo et al., 2006). A summary of mutations of hGALE and their effect on the protein can be found in McCorvie and Timson (2013b). While the focus of this work is V94M, a substrate binding site mutation that causes the most severe form of type III galactosemia, there are other known mutations with severe impact on the enzyme. The G90E mutant is very severely impaired kinetically and, along with L183P, also is highly destabilized compared to wildtype (Timson, 2005). K161N is also very severely impaired kinetically, with the likely loss of the NAD cofactor. Interestingly, this mutant is thermally more stable than wildtype (McCorvie et al., 2012).

It is not clear why aberrant galactose metabolism can result in such severe consequences. Several studies have suggested that a build-up of galactose 1-phosphate is toxic to eukaryotic cells, but the exact, molecular cause of this toxicity is unknown (Lai et al., 2009). Altered UDP-sugar metabolism in type III galactosemia affects the pools of these precursors for glycoprotein and glycolipid synthesis (Daenzer et al., 2012; Kingsley et al., 1986). In addition, aggregation of HsGALE has been observed in cells expressing disease-associated variants (Bang et al., 2009). As yet, there are no small molecule therapies for any form of galactosemia, although it has been suggested that a "small molecule chaperone" approach may be viable since many of the mutations result in decreased protein stability (Chhay et al., 2008; McCorvie et al., 2011, 2012; Timson, 2005, 2006; Wohlers and Fridovich-Keil, 2000; Wohlers et al., 1999). In order to implement such therapies, however, it is important to understand more about not just the structures of the proteins and the disease-associated variants but also the flexibility and dynamics of these structures. Molecular dynamics (MD) has been established as one of the prime methods to computationally probe the dynamics of biomolecular systems (Adcock and McCammon, 2006). It is thus ideally suited to compare the dynamics of different mutant forms of the same protein. In the realm of computer-aided drug discovery, MD can be used to account for receptor flexibility (Sinko et al., 2013). Therefore, we undertook a detailed molecular dynamics investigation of HsGALE and the V94M variant which is the one most commonly associated with the most severe form of type III galactosemia (Wohlers et al., 1999).

2. Material and methods

2.1. System preparation

Systems based on two different crystal structures were prepared for simulations: PDB entries 1EK6 (Thoden et al., 2000) and 1I3L (Thoden et al., 2001a) for wildtype and V94M human UDP-galactose 4-epimerase respectively. These structures contain the substrate uridine-5′-diphosphate glucose (UDP-glc) and uridine-5′-diphosphate galactose (UDP-gal) respectively and were modeled as dimers. The cofactor was modeled as NAD^+ in all structures. Missing residues (347,348 in 1EK6A, 1,347,348 in 1EK6B, 1 in 1I3LA and 1,347,348 in 1I3LB) were built in using Prime (Jacobson et al., 2002, 2004). For each of the starting structures, four different systems were built: apo (no substrates, no NAD^+ cofactors), NAD^+-bound (no substrates, with two NAD^+ cofactors), single-substrate–NAD^+-bound (with one substrate and two NAD^+ cofactors) and substrate–NAD^+-bound (with two substrates and two NAD^+ cofactors). Tleap (Case et al., 2005) was used to neutralize the systems by adding Na^+ counter ions (6, 8, 10, and 12 Na^+ for the apo, NAD^+-bound, single-substrate–NAD^+-bound and substrate–NAD^+-bound systems respectively) and solvating using a TIP3P water box. The fully solvated substrate–NAD^+-bound systems contained 109,037 (1EK6), and 115,573 (1I3L) atoms, respectively. Simulations were performed on each of the eight different systems. Minimization using SANDER (Case et al., 2005) was carried out in two stages: 1000 steps of minimization of solvent and ions with the protein, substrate, and cofactor restrained using a force constant of 500 kcal/mol/$Å^2$, followed by a 2500 step minimization of the entire system. A short initial 20 ps MD simulation with weak restraints (10 kcal/mol/$Å^2$) on the protein, substrate, and cofactor residues was used to heat the system to a temperature of 300 K. Subsequently, 150 ns of MD simulations were performed on each of the eight systems under investigation.

2.2. Molecular dynamics simulations

All MD simulations were performed under the NPT ensemble at 300 K using AMBER (Case et al., 2005) and the ff99SBildn force field (Hornak et al., 2006; Lindorff-Larsen et al., 2010). Periodic boundary conditions were used, along with a non-bonded interaction cutoff of 10 Å. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm (Ryckaert et al., 1977), allowing for a time step of 2 fs. For each system, 150 ns MD trajectories were generated totaling a simulation time of 1.2 μs.

2.3. Distance measurements

Assuming the proposed catalysis mechanism is correct, NAD^+ needs to oxidize the sugar at the C4-OH group of the substrate. Similarly the Tyrosine 157 residue is required for the mechanism. The proximity of the nicotinamide group, the substrate C4-OH, as well as the Tyr157 were monitored for wildtype and mutant substrate–NAD^+-bound simulations. For this the pairwise distances between the C13 substrate atom, the C18 atom on the cofactor nicotinamide group and the OH on Tyr157 were calculated.

Additionally, stabilizing hydrogen bonds between protein active site residues and the substrate or cofactor have been monitored for wildtype and mutant substrate–NAD^+-bound simulations. This way changes in dynamics caused by the mutation that might affect cofactor or substrate binding stability can be investigated. A total of 12 H-bonds that are present both in the wildtype and mutant crystal structures were investigated: N224 O—substrate H1, F226 H—substrate O1, R300 2HH1—substrate O8, R300 2HH2—substrate O8, N187 1HD2—substrate O11, R239 HE—substrate O11, D66 OD1—cofactor H11, N34 H—cofactor N5, N37 H—cofactor O6, D33 OD1—cofactor H5, K92 HZ1—cofactor O8, and I14 H—cofactor O9. Distances between 2.0 and 3.5 Å are representative of hydrogen bonds (Wallwork, 1962). All distance measurements were carried out at 200 ps intervals.

2.4. Secondary structure, RMSF, and cross correlation analysis

In order to investigate a possible change in stability upon the V94M mutation, the secondary structure content over the course of the simulations were calculated for all eight systems. Frames every 20 ps were extracted from the trajectories. The STRIDE algorithm (Frishman and Argos, 1995) was used to calculate helical and strand content for each of the frames. Average secondary structure contents for the entire simulation were also calculated. The Root Mean Square Fluctuations (RMSF) were calculated using ptraj (AmberTools 12). For the comparison between human and *Escherichia coli* loop fluctuations, the average RMSF of residues 298 through 311 (human) and their equivalent counterparts