Characterization of Essential Domains in HSD17B13 for Cellular Localization and Enzymatic Activity

Yanling Ma\textsuperscript{1,2}, Suman Karki\textsuperscript{3}, Philip M. Brown\textsuperscript{1,2}, Dennis D. Lin\textsuperscript{1,2}, Maren C. Podszun\textsuperscript{1,2}, Wenchang Zhou\textsuperscript{4}, Olga V. Belyaeva\textsuperscript{3}, Natalia Y. Kedishvili\textsuperscript{3}, and Yaron Rotman\textsuperscript{1,2}

1-Liver and Energy Metabolism Section,  
2-Liver Diseases Branch, NIDDK, NIH, Bethesda, MD  
3-Department of Biochemistry and Molecular Genetics, Schools of Medicine and Dentistry, University of Alabama – Birmingham, Birmingham, AL  
4-Theoretical Molecular Biophysics Laboratory, NHLBI, NIH, Bethesda, MD

Running title: Critical domains for localization and activity of HSD17B13

Correspondence: Yaron Rotman, M.D., M.SC., Liver and Energy Metabolism Section, Liver Diseases Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.  
Address: 10 Center Drive, Building 10, Room 10N248C, MSC1800, Bethesda, MD, 20892-1800  
Phone: 301-451-6553  
Fax: 301-402-0196  
Email: rotman@niddk.nih.gov

Abbreviations: LD (Lipid droplets); HSD17B13 (17-beta hydroxysteroid dehydrogenase 13); PNPLA3 (patatin-like phospholipase domain-containing protein 3); SDRs (Short-chain dehydrogenases/reductases); RDH (retinol dehydrogenase); SNP (single-nucleotide polymorphism); FA (Fatty Acids)
Abstract

Recently, human genetic studies identified an association of single-nucleotide polymorphisms (SNPs) in the *HSD17B13* (17-beta hydroxysteroid dehydrogenase 13) gene with alcoholic and nonalcoholic fatty liver disease (NAFLD) development. Mutant HSD17B13 variants devoid of enzymatic function have been demonstrated to be protective from cirrhosis and liver cancer, supporting the development of HSD17B13 as a promising therapeutic target. Previous studies have demonstrated HSD17B13 is a lipid droplet (LD) associated protein. However, the critical domains that drive LD targeting or determine the enzymatic activity are yet to be defined. Here we used mutagenesis to generate multiple truncated and point-mutated proteins and were able to demonstrate in vitro that the N-terminal hydrophobic domain, PAT-like domain, and a putative α-helix/β-sheet/α-helix domain in HSD17B13 are all critical for LD targeting. Similarly, we characterized the predicted catalytic, substrate-binding, and homodimer interaction sites and found them essential for the enzymatic activity of HSD17B13, in addition to our previous identification of amino acid P260 and cofactor binding site. In conclusion, we identified critical domains and amino acid sites that are essential for the LD localization and protein function of HSD17B13, which may facilitate understanding of its function and targeting of this protein to treat chronic liver diseases.

Keywords:

Non-alcoholic fatty liver disease; Alcoholic liver disease; Retinoids; Lipid droplets; Enzyme regulation; Protein structure; HSD17B13
Introduction

Associated with the global epidemic of lifestyle-associated obesity and metabolic syndrome, non-alcoholic fatty liver disease (NAFLD), has become a major global health burden and one of the leading causes for end-stage liver disease, hepatocellular carcinoma, and liver transplant (1, 2). Limited treatment options are available, stimulating the search for novel molecular targets suitable for therapeutic pharmacological intervention (3). One approach to the discovery of these novel targets, based on the known heritability of NAFLD, is the use of genetic studies (4-6). Several genes related to the incidence of NAFLD and to its progressive form of non-alcoholic steatosis hepatitis (NASH) were identified using this approach, including PNPLA3 (7-12), TM6SF2 (13-15), MBOAT7 (16), and the recently discovered HSD17B13 (17-19). Of those, HSD17B13 (17-beta hydroxysteroid dehydrogenase 13) represents a likely therapeutic target for treatment of chronic liver disease with a potential for pharmaceutical intervention.

HSD17B13 is a hepatic lipid-droplet associated enzyme with steroid substrates, bioactive lipids (19), and retinol (17) suggested as potential enzymatic substrates. Three independent genetic variants in HSD17B13 were found to confer protection from injury in NASH (17, 19, 20), alcoholic liver disease (17, 19), and hepatocellular carcinoma (HCC) (21, 22). We and others identified a splice-site single nucleotide polymorphism (SNP) rs72613567 that leads to the formation of two novel splicing variants (HSD17B13-G insertion and HSD17B13-Exon 6 deletion) (17, 19, 23); the non-synonymous SNP rs62305723 encodes a proline to serine mutation at amino acid position 260 (17); and the rs143404524 SNP leading to premature truncation (24). All three protective variants generate protein products that are devoid, or predicted to be devoid of enzymatic activity, confirming the importance of understanding the enzymatic function of HSD17B13. Collectively, these data also suggest that HSD17B13 activity can be modulated by large domain truncations and deletions, as well as by single or double critical amino acids mutations, implying the possibility to modulate the enzymatic activity therapeutically, either by interfering with gene expression (i.e. by using antisense oligonucleotides) or by inhibiting activity directly using small synthetic compounds.

Excess cellular lipids are esterified to neutral lipids and stored in lipid droplets (LDs), which function as main storage reservoirs of metabolic energy and membrane lipid components (25). LD-associated proteins, which play
pivotal roles in lipid metabolism regulation, are embedded in or adherent to the phospholipid monolayer, which includes phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylinositol, lysoPC, and lysoPE (26). Mass spectrometry proteomic analyses have identified hundreds of LD-associated proteins, among which, a couple of dozen proteins were confirmed to be LD resident proteins (25). Two types of LD targeting signals have been proposed for directing proteins to the LD surface: amphipathic α-helices and hydrophobic hairpins (25, 27). The PAT protein family encompasses the most widely studied LD-associated proteins (Perilipin, ADRP, TIP47, and S3-12) (28), that share conserved PAT domains which are essential for their LD targeting (29, 30).

HSD17B13 is a lipid droplet-associated protein (17, 19, 31, 32) and we identified a PAT-like domain at its N-terminus which is essential for its stability, targeting to LD and enzymatic function (17). However, other domains in HSD17B13 essential for correct trafficking to LDs are not yet fully defined.

Pharmacological efforts to silence HSD17B13 by siRNA have been initiated by pharmaceutical companies (33). However, beyond the naturally occurring mutants and cofactor binding sites we identified (17), other sites or domains critical for its enzymatic activity and targeting are still unknown. In the current study, we aimed to identify critical domains and amino acids that are essential for LD targeting and enzymatic activity of HSD17B13 by studying multiple truncated and point-mutated proteins. Our work may facilitate the design of small molecule inhibitors and lead to better understanding of the protein function which is essential for HSD17B13 to be considered for therapeutic target in chronic liver disease.
Materials and Methods

Hydropathy analysis

Hydropathy analysis of HSD17B13 was performed using the online TMHMM server (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0) (34) and ProtScale (https://web.expasy.org/protscale/) (35) with default settings.

Cell Culture and Stock Solution

HepG2 and HEK293 cells were cultured in DMEM medium (Corning) supplemented with 10% Fetal Bovine Serum (FBS, Sigma Aldrich) under 5% CO₂ at 37 °C. Primary human hepatocytes (Gibco) were plated on poly-lysine (Sigma Aldrich) coated 4-well chamber slide in William’s E medium (Gibco) supplemented with 10% FBS. Oleate and palmitate were solubilized in Phosphate-Buffered Saline (PBS, Corning) solution by heating to 55°C and 65°C respectively. Solubilized fatty acids were conjugated with 10% fatty acid free-bovine serum albumin (BSA) in culture medium to generate fatty acid stock solution. Lipid droplets were induced by adding fatty acid stock solution into culture medium for 48 hrs with a final concentration of 200 µM oleate and 200 µM palmitate (FA) unless otherwise indicated. Transfections of HepG2 cells stably expressed HSD17B13-GFP (17) were carried out using Lipofectamine 3000 (Thermo Fisher). To study homodimerization, cells were grown in 6-well cell culture dishes and transfected using 2 µg of HSD17B13-FLAG plasmid DNA and 10 µl of Lipofectamine 3000 per well. 48 hours after transfection, cells were lysed in 500 µl of lysis buffer (1% Triton X-100, 50mM Tris pH 8, 150mM NaCl) for 10 min on ice and collected after centrifugation at 13000 G for 30 min at 4°C.

Plasmids and Mutagenesis

The full length protein HSD17B13-GFP (RG213132) and the exon-2 deleted variant (Δ71-106, Variant B), HSD17B13-Variant B-GFP (RG227799) were obtained from OriGene (Rockville, USA). HSD17B13-FLAG (VB150430-10020) was designed and constructed by VectorBuilder Inc. (Cyagen Biosciences, Santa Clara, CA). Q5® Site-Directed Mutagenesis Kit (NEB, E0554S) was used to generate mutant HSD17B13 plasmids with designed mutagenesis primers (Supplemental Table S1). HSD17B13-GFP and HSD17B13-FLAG plasmids were

Q5® Site-Directed Mutagenesis Kit (NEB, E0554S) was used to generate mutant HSD17B13 plasmids with designed mutagenesis primers (Supplemental Table S1). HSD17B13-GFP and HSD17B13-FLAG plasmids were
used as mutagenesis templates for the cellular localization study and for the enzymatic assay, respectively, unless otherwise indicated. Mutant plasmids were confirmed by Sanger sequencing.

The selection of sites for point mutation was based on a prediction of sites relevant to activity, using the NCBI conserved domain search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and relying on similarity with other short-chain dehydrogenase/reductases.

**Cellular localization**

Cells were seeded in Nunc™ Lab-Tek™ II Chambered Coverglass (Thermo Fisher Scientific) and transfected with wild type or mutant plasmids using Lipofectamine 3000 Reagent (ThermoFisher Scientific) following kit instruction. To study LD targeting, FA was added 48 hrs after the transfection to induce LD. After 48 hrs of FA treatment cells were fixed in 4% paraformaldehyde (PFA, Electron Microscopy Science) for 10 min and counter stained with Hoechst (1µg/ml, Thermo Fisher Scientific) for nuclei, and LipidTox (1:500, Thermo Fisher Scientific) for LDs.

To study the co-localization of HSD17B13 naturally occurring variant B (HSD17B13-B, Δ71-106) with the endoplasmic reticulum (ER), cells were co-transfected with SEC61β-GFP (ER marker protein, gift from Dr. Alexandre Toulmay) and HSD17B13-B-Flag. Cells were fixed in 4% PFA and permeabilized in 0.3% Triton-X, 3% BSA, and 10% normal goat serum (Vector Laboratories, USA) in PBS. Immunofluorescence staining of HSD17B13-B-FLAG was performed by incubating cells with FLAG M2 antibody (F3165, Sigma) at room temperature for 1 hour. Alexa Fluor 568 goat anti-mouse secondary antibody was used after washing. Hoechst was used to stain nuclei.

Immunofluorescence staining of Apoptosis-inducing factor (AIF), a marker for mitochondria, and HSD17B13 were performed to demonstrate mitochondrial targeting of mutant HSD17B13 after transfection. Antibodies against AIF (5318, Cell Signaling Technology) and FLAG (F3165, Sigma) were used for primary incubation. Alexa Fluor 647 goat anti-rabbit and Alexa Fluor 568 goat anti-mouse secondary antibodies were used to recognize Rabbit anti-AIF and Mouse anti-FLAG antibodies respectively.

Cellular fluorescence images were taken by confocal microscopy (Zeiss LSM 700).
**Retinol Dehydrogenase Activity Assay**

Enzymatic activity was measured using the retinol dehydrogenase (RDH) activity assay, performed as previously described (17, 27). Briefly, HEK293 cells were seeded one day before being transiently transfected in triplicate with HSD17B13, HSD17B13 mutant, or empty vector plasmids. All-trans-retinol (Toronto Research Chemicals, Toronto, Canada) at 2 or 5 µM in ethanol, with a final ethanol concentration ≤ 0.5% (v/v) was added to the culture medium and cells were incubated for 6 or 8 hours. Retinoids were extracted twice by equal volume of ethanol and double volume of hexane and were separated by normal phase HPLC with Spherisorb S3W column (4.6 mm×100 mm) (Waters Corp., Milford, MA). Retinaldehyde and retinoic acid levels were normalized per total protein amount and are shown relative to empty vector. Wild type HSD17B13 was used as positive control, with multiple constructs tested in the same round of experiments sharing the positive control. Aliquots of cell suspensions were taken for protein quantification and western blot analysis.

**Co-immunoprecipitation and Western Blot**

Protein G Dynabeads (Invitrogen Cat# 10007D) were incubated with anti-FLAG antibody (clone F1804) from Sigma Aldrich or anti-turboGFP (clone OTI2H8) antibody from Origene overnight with rotation at 4ºC. For each pulldown, 3 µg of antibody were incubated with 50 µl of bead slurry. Lysates of transfected HepG2 were incubated with antibody conjugated beads with rotation for 1 hour at room temperature. Bound proteins were washed and eluted with SDS containing sample buffer per kit guidelines. Proteins were separated in 4-15% precast PAGE gels (Bio Rad) and transferred to PVDF membranes (Invitrogen). The membranes were blocked with a 5% nonfat milk in Tris-buffered saline containing 0.05% Tween-20 for one hour. Membranes were incubated with polyclonal anti-HSD17B13 antibody (1:2000) from Origene (TA350064) in conjunction with a rabbit horseradish peroxidase (HRP) conjugated secondary antibody (GE Healthcare) and Pierce enhanced chemiluminescent substrate for detection of HRP (Thermo Fisher).

**Statistical analysis**

Differences between wild type HSD17B13 and empty vector or HSD17B13 mutants were tested using Student’s t-test (Prism V.8, GraphPad).
Results

**Essential Domains for Lipid Droplet Targeting of HSD17B13**

Full-length HSD17B13 protein (Variant A) is targeted to LDs when cells are lipid-loaded (Figure 1A). We previously described reduction of protein stability and LD targeting with loss of the PAT-like N-terminus amino acids (AA) 22-28 (Δ22-28) or with loss of AA 71-106 (Variant B, a naturally occurring variant with exon 2 skipping) (17) (Figure 1 and Figure 2A). The N-terminal of HSD17B13 is predicted by hydropathy analysis to be a putative transmembrane domain (Supplemental Figure S1) and thus could serve to anchor the protein to LDs; AA30-300 are likely to reside on the outside membrane surface. We thus extended our study in detail to delineate the characteristics of N-terminal sequences critical for LD targeting. Not surprisingly, a fragment containing only the hydrophobic domain (N1-21) of HSD17B13 was not targeted to LDs in the absence of AA 22-28 and AA 71-106 (Figure 1), suggesting the hydrophobic domain is not sufficient to drive protein LD targeting. Interestingly, addition of the PAT-like domain (AA 22-28) to the hydrophobic domain, generates a peptide fragment (N1-28) of HSD17B13 that localizes to LDs, despite the absence of AA 71-106 (Figure 1). To test whether the hydrophobic domain is necessary for HSD17B13 to target LDs, we further generated an HSD17B13 devoid of the hydrophobic AA 4-16 (Δ4-16) and as expected, found that without this hydrophobic sequence HSD17B13 does not target to LDs (Figure 1A). Interestingly, Δ4-16 has a non-random pattern of cellular distribution and we found out it to be localized in close proximity to mitochondria (Supplemental Figure S2). Thus, our data indicates that at the N-terminus of HSD17B13, both AA4-16 and AA22-28 are necessary for LD targeting.

The naturally-occurring variant B of HSD17B13 is a consequence of exon 2 skipping (without a frame shift), leading to the absence of AA 71-106. Although the structure of HSD17B13 has not been resolved to date, X-ray crystallography data is available for HSD17B11 (PDB ID: 1YB1, https://www.rcsb.org/structure/1YB1), a close paralogue of HSD17B13, with 77% amino acids similarity. Based on the structure of HSD17B11, AA 69-106 in HSD17B13 are predicted to form a α-helix/β-sheet/α-helix structure, which is absent in variant B (Δ71-106). As we previously identified that variant B is not targeted to LDs, despite having the hydrophobic domain (AA4-16) and PAT-like domain (AA22-28), and has no enzymatic function (17), we tested the importance of this region in more detail. To determine which part of the α-helix/β-sheet/α-helix structure is essential for LD targeting, we
generated three mutant proteins of HSD17B13: \(\Delta 69-84\) (first \(\alpha\)-helix deletion), \(\Delta 85-93\) (\(\beta\)-sheet deletion), and \(\Delta 94-106\) (second \(\alpha\)-helix deletion). Loss of any of the three domains was sufficient to impair LD localization (Figure 2), indicating the entire intact structure is required. Without these domains, HSD17B13-variant B is retained in the endoplasmic reticulum (ER), and co-localized with SEC61\(\beta\), an ER marker protein (Supplemental Figure S3). It is possible that the deletion of \(\alpha\)-helix/\(\beta\)-sheet/\(\alpha\)-helix structure leads to misfolding of subsequent sequences in the protein, preventing its transport from ER to LDs. C-terminal deleted variants, either without (N1-70) or with the \(\alpha\)-helix/\(\beta\)-sheet/\(\alpha\)-helix structure (N1-111), localized correctly to LDs (Figure 2). Taken together, we conclude that the N-terminal hydrophobic and PAT-like domains are likely promoting the anchoring of of HSD17B13 to the ER/LD membrane, whereas the \(\alpha\)-helix/\(\beta\)-sheet/\(\alpha\)-helix structure is required for transportation of the full-length protein from ER to LDs. To extend our observations beyond cell lines, we transfected primary human hepatocytes (PHH) with wild type HSD17B13 or mutant proteins with critical LD-targeting domains modified. We confirmed that the same domains are needed in PHH to drive the LD targeting of HSD17B13 (Supplemental Figure S4). To confirm whether HSD17B13 will target LD differently under oleate- or palmitate-induced conditions, we treated PHH with oleate, palmitate, or a combination of both, and found that HSD17B13 targets LD under all conditions (Supplemental Figure S5).

**Essential Domains for Enzymatic Activity of HSD17B13**

Enzymes in the short-chain dehydrogenase/reductase (SDR) superfamily have a conserved Rossmann-fold motif consisting of six parallel \(\alpha\)-helices surrounding the central seven parallel \(\beta\)-sheets (36). This classical Rossmann-fold structure was observed in the predicted 3-dimensional structure of HSD17B13 using HSD17B11 as a template (Figure 3). HSD17B13 is believed to affect chronic liver disease through its enzymatic activity, and enzymatically non-functional mutants are associated with protection from liver injury (17, 19). We therefore sought to identify sites critical for its enzymatic activity. We have previously demonstrated the importance of the conserved cofactor binding site (37) Gly\(^{47}\)/Gly\(^{49}\) for enzymatic activity (17). In a similar manner, we generated plasmids expressing HSD17B13 with mutations in the putative catalytic site, substrate binding site, or homodimer interaction site, and tested their RDH activity.
The conserved Asn^{144}-Ser^{172}-Tyr^{185}-Lys^{189} catalytic tetrad (38) is predicted to be in close proximity to the cofactor, NAD^{+} (Figure 4A). A single amino acid mutation from asparagine to alanine at the predicted catalytic site Asn^{144} (HSD17B13-N144A) abolished the RDH activity of HSD17B13 (Figure 4B) and similar results were obtained when we mutated another putative catalytic site residue at Ser^{172} (Figure 4C). A double mutation of Tyr^{185} and Lys^{189} revealed additional amino acid residues critical for activity (Figure 4D). With the exception of N144A, none of these mutations interfered with protein stability (Supplemental Figure S6) or LD targeting (Supplemental Figure S7).

We also generated mutations in putative substrate binding sites in HSD17B13 and tested the RDH activity of the mutant proteins (Figure 5A). A double mutant at Lys^{153} and Leu^{156} (HSD17B13-K153A/L156A) or at Leu^{199} and Glu^{202} (HSD17B13-L199A/E202A) abolished enzymatic activity of HSD17B13 completely (Figure 5B-C) but was also associated with lower protein levels (Supplemental Figure S6), suggesting decreased stability. A single mutation of Lys^{208} (HSD17B13-K208A) decreased, but did not fully abolish RDH activity (Figure 5D), indicating both sites are critical for the enzymatic activity of HSD17B13. The close paralogue and a major retinol dehydrogenase, RDH10, is functioning either as a homodimer or as an oxidoreductive heterooligometric complex with retinaldehyde reductase 3 (DHRS3), to tightly control the steady-state levels of retinoic acid (39). Given the strong structural similarity, it is plausible that HSD17B13 also requires dimerization for its activity. Indeed, we were able to confirm homodimerization of HSD17B13 in vitro by co-transfecting cells with HSD17B13-GFP and HSD17B13-FLAG and performing coimmunoprecipitation (Supplemental Figure S8). We therefore mutated a putative homodimer interaction site at R97/Y101 and found that mutant protein significantly reduced RDH activity (Figure 5E). A summary of essential sites for HSD17B13 enzymatic activity is shown in Table 1 and Figure 6.
Discussion

Human genetic studies identified an association between the \textit{HSD17B13} gene and NASH (17, 19, 20), alcoholic liver disease (17, 19), and hepatocellular carcinoma (HCC) (21, 22). Enzymatically inactive protein mutants of HSD17B13, generated by genetic variants, confer protection from progression of fatty liver disease (17, 19) confirming the importance of enzymatic activity of HSD17B13, although this protection was not seen in Hsd17b13-knock out mice (40). Pharmaceutical efforts to silence HSD17B13 by siRNA and to inhibit HSD17B13 enzymatic activity are being pursued by pharmaceutical companies (https://investor.regeneron.com/news-releases/news-release-details/regeneron-and-alnylam-pharmaceuticals-announce-collaboration; https://clinicaltrials.gov/ct2/show/NCT04202354). Understanding the critical residues for its enzymatic activity is important to facilitate small molecule inhibitor design and to better explore the protein function. In this study, we focused on residues based on two premises: first, our previous finding that targeting to LDs is required for enzymatic activity (17), and second, \textit{in silico} predictions based on structural similarity to other enzymes. With this approach, we identified critical domains (N-terminal hydrophobic domain, PAT-like domain, and a putative \(\alpha\)-helix/\(\beta\)-sheet/\(\alpha\)-helix domain) for LD targeting and several important residues (predicted catalytic sites, substrates binding sites, and homodimer interaction sites) for the enzymatic activity of HSD17B13.

The N-terminus sequence AA1-35 has been shown to be sufficient in localizing HSD17B13 to LDs (31). Here we scale down the required sequence to the N-terminal AA1-28 and identify it as consisting of two domains with distinct roles - an N-terminal hydrophobic domain predicted to be a transmembrane helix, and an adjacent PAT-like domain (Supplemental Figure S 1), with both domains required for LD targeting. HSD17B13 is structurally related to two families: 17β hydroxysteroid dehydrogenase (HSD) family with steroids, bile acids, or fatty acids as potential substrates (41), and the short-chain dehydrogenase/reductase 16C family (SDR16C), of which some family members have been reported to regulate retinoid metabolism by acting as retinol dehydrogenases (42). Some members of the SDR16C such as HSD17B11, and RDH10 have been also identified as LD-associated proteins (43, 44) and a sequence comparison can help clarify the role of targeting domains. An N-terminus sequence hydrophobic transmembrane domain is known to be required for LD targeting of HSD17B11 and RDH10 (43, 44). The non-canonical PAT motif (PAT-like) has also been described in HSD17B11(44, 45) but is
not present in RDH10, consistent with the known ability of the latter two proteins to localize to other cellular compartments as well.

Interestingly, deletions of the specific HSD17B13 N-terminal domains resulted in different cellular distribution patterns. PAT-like domain deletion (Δ22-28) largely reduced protein stability, whereas deletion of the transmembrane hydrophobic domain (Δ4-16) redistributed HSD17B13 from LDs to mitochondria. Small deletions leading to protein redistribution from LDs to mitochondria have been described previously for ADRP (46), HSD17B11 (44), and RDH10 mutation studies (43), suggesting similarities between LD and mitochondria targeting machinery which may play a role in the close contact and dynamic interaction of these two organelles (47). Notably, the PAT-like domain in HSD17B13 is also a recognition motif for mitochondrial surface import receptor TOM20 (MitoFates - http://mitf.cbrc.jp/MitoFates/cgi-bin/top.cgi) (48). With the loss of the anchoring transmembrane helix, HSD17B13Δ4-16 may be released from the ER to the cytosol, where it can be recognized and imported by the TOM20-TOM40 complex (49). Whether this is has physiological importance or is purely an artefact of laboratory deletions is unknown.

In this manuscript we describe for the first time a unique motif determining the cellular distribution of HSD17B13 in the α-helix/β-sheet/α-helix domain (AA69-106), most of which (AA71-106) is missing in the naturally occurring variant B isoform of HSD17B13. Deletion of this domain led to the protein being retained in the ER and not be targeted to LDs. Interestingly, if the C-terminal domain is deleted, the presence of these domains is no longer critical, evidenced by normal LD targeting of N1-70 and N1-111. One possibility is that the loss of the α-helix/β-sheet/α-helix domain leads to protein misfolding of the C-terminal domain. Instead of exiting ER, misfolded proteins are retained in the ER and subjected to proteasomal degradation through a process termed ER-associated degradation (ERAD) (50). Variant B is less stable than the full length protein (17), hinting that the ERAD may be triggered by HSD17B13-B. Alternatively, this domain may be needed for additional membrane interaction for the full protein. A hairpin α-helix structure is crucial for membrane interaction and targeting from ER to LDs for many proteins (25) and the α-helix/β-sheet/α-helix we report here may play a similar role. The same domain is highly conserved in HSD17B11 and is present in RDH10, though in the latter protein it is disrupted by a 24 AA insertion between the first α-helix and the β-sheet. The high degree of structural
conservation suggest its importance beyond folding only. Furthermore, the conserved Cys$^{80}$ residue contained in this region, was found to be palmitoylated in mouse Hsd17b13 (51), increasing the hydrophobicity and suggesting important membrane interaction.

Beyond localization to LDs, modifying key residues can affect protein function by interfering with sites essential for its enzymatic activity. The genetically-driven splicing variants (HSD17B13-G insertion and HSD17B13-Exon 6 deletion) associated with rs72613567 generate an inactive enzyme, possibly due to protein instability (17, 19, 23). Another inactive variant we previously identified is the P260S mutant, encoded by the non-synonymous genetic variant rs62305723 (17). Although distant from the predicted conserved substrate binding, catalytic, or co-factor binding sites, the highly conserved Pro$^{260}$ is an essential residue for enzymatic activity of HSD17B13 (17), and its mutation to serine may lead to an impaired hinge function and introduce protein conformational and subsequent functional changes (52). An additional variant (rs143404524, encoding A192Lfs) was reported to confer protection against chronic liver disease (24); given the critical role of Pro$^{260}$ (17) and potential protein conformation change, we predict this frame shift protein to be inactive.

NAD(H)/NADP(H) are essential cofactors for SDRs (53) and mutations in conserved cofactor binding glycine residues abolished enzymatic activity of RDH10 (54), and HSD17B13 (17). Despite a relatively low sequence identity, the three-dimensional structures of SDRs display a highly-conserved $\alpha/\beta$ folding pattern with a predicted catalytic tetrat of Asn-Ser-Tyr-Lys (38). Within this catalytic tetrat, TyrXXXLys is the most conserved site across the SDR family (55). We identified HSD17B13 residues Asn$^{144}$, Ser$^{172}$, and Tyr$^{185}$/Lys$^{189}$ as the key components of that tetrat and could abolish enzymatic activity by mutating them. Our findings for the HSD17B13 Y185A/L189A double mutant and the N144A mutant are thus consistent with similar studies in RDH10 and in 3$\beta$-HSD (38, 54). Replacing the Ser$^{172}$ residue in HSD17B13 has also led to loss of function, although replacement of that catalytic serine in other SDRs showed discrepant results; S142A in 17$\beta$-HSD-1 (56) and S138A in 3$\beta$-HSD (38) abolished their activity, while in RDH10 S197A the activity was retained but not in S197T which abolished the enzymatic activity (54). These observations suggest slightly different catalytic structures of each enzyme.

We have also tested the predicted substrate binding site residues Lys$^{153}$, Leu$^{156}$, Leu$^{199}$, Glu$^{202}$, and Lys$^{208}$ found them essential for HSD17B13 activity. The substrate binding loop is the least conserved domain in the SDR
family, which implies a variety of substrate specificities (57, 58), and possibly explains why mutagenetic replacement of Glu282 failed to affect the activity of another enzyme, 17β-HSD-1 (56). Interestingly, some of the essential sites seem to affect protein stability (L199/E202, K153/L156, and N144), a similar phenotype to what has been described in the protective human variants (19).

In conclusion, we identified common and unique domains and amino acid sites that are essential for the LDs targeting (N-terminal hydrophobic domain, PAT-like domain, and a putative α-helix/β-sheet/α-helix domain) and enzymatic activity (catalytic sites, substrates binding sites, and homodimer interaction sites) of HSD17B13 which could facilitate the therapeutic targeting of this protein against chronic liver diseases.

All data are contained within this manuscript and are also available upon request from the corresponding author.

Acknowledgement

The authors would like to thank Dr. Seung Bum Park (NIDDK) for his technical assistance on primary hepatocytes culture. The study was funded by the Intramural Research Program of NIDDK and by National Institute of Alcohol Abuse and Alcoholism (AA012153, to N.Y.K.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
References

1. Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, George J, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Hepatol 2018;15:11-20.

2. Pais R, Barratt AS, Calmus Y, Scatton O, Runge T, Lebray P, Poynard T, et al. NAFLD and liver transplantation: Current burden and expected challenges. J Hepatol 2016;65:1245-1257.

3. Rotman Y, Sanyal AJ. Current and upcoming pharmacotherapy for non-alcoholic fatty liver disease. Gut 2017;66:180-190.

4. Loomba R, Schork N, Chen CH, Bettencourt R, Bhatt A, Ang B, Nguyen P, et al. Heritability of Hepatic Fibrosis and Steatosis Based on a Prospective Twin Study. Gastroenterology 2015;149:1784-1793.

5. Schwimmer JB, Celedon MA, Lavine JE, Salem R, Campbell N, Schork NJ, Shiehmorteza M, et al. Heritability of nonalcoholic fatty liver disease. Gastroenterology 2009;136:1585-1592.

6. Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, Grundy SM, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. Hepatology 2004;40:1387-1395.

7. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, Boerwinkle E, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40:1461-1465.

8. Yuan X, Waterworth D, Perry JR, Lim N, Song K, Chambers JC, Zhang W, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet 2008;83:520-528.

9. Sookoian S, Castano GO, Burgueno AL, Gianotti TF, Rosselli MS, Pirola CJ. A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity. J Lipid Res 2009;50:2111-2116.

10. Rotman Y, Koh C, Zmuda JM, Kleiner DE, Liang TJ, Nash CRN. The association of genetic variability in patatin-like phospholipase domain-containing protein 3 (PNPLA3) with histological severity of nonalcoholic fatty liver disease. Hepatology 2010;52:894-903.

11. Speliotes EK, Butler JL, Palmer CD, Voight BF, Consortium G, Consortium MI, Nash CRN, et al. PNPLA3 variants specifically confer increased risk for histologic nonalcoholic fatty liver disease but not metabolic disease. Hepatology 2010;52:904-912.

12. Valenti L, Al-Serri A, Daly AK, Galmozzi E, Rametta R, Dongiovanni P, Nobili V, et al. Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. Hepatology 2010;51:1209-1217.

13. Holmen OL, Zhang H, Fan Y, Hovelson DH, Schmidt EM, Zhou W, Guo Y, et al. Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk. Nat Genet 2014;46:345-351.

14. Kozlitina J, Smagris E, Stender S, Nordestgaard BG, Zhou HH, Tybjaerg-Hansen A, Vogt TF, et al. Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2014;46:352-356.
15. Mahdessian H, Taxiarchis A, Popov S, Silveira A, Franco-Cereceda A, Hamsten A, Eriksson P, et al. TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content. Proc Natl Acad Sci U S A 2014;111:8913-8918.

16. Buch S, Stickel F, Trepo E, Way M, Herrmann A, Nischalke HD, Brosch M, et al. A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis. Nat Genet 2015;47:1443-1448.

17. Ma Y, Belyaeva OV, Brown PM, Fujita K, Valles K, Karki S, de Boer YS, et al. 17-Beta Hydroxysteroid Dehydrogenase 13 Is a Hepatic Retinol Dehydrogenase Associated With Histological Features of Nonalcoholic Fatty Liver Disease. Hepatology 2019;69:1504-1519.

18. Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, Van der Harst P, Holm H, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 2011;43:1131-1138.

19. Abul-Husn NS, Cheng X, Li AH, Xin Y, Schurmann C, Stevis P, Liu Y, et al. 17- Beta Hydroxysteroid Dehydrogenase 13 Is a Hepatic Retinol Dehydrogenase Associated With Histological Features of Nonalcoholic Fatty Liver Disease. Hepatology 2019;69:1504-1519.

20. Pirola CJ, Garaycoechea M, Flichman D, Arrese M, San Martino J, Gazzi C, Castano GO, et al. Splice variant rs72613567 prevents worst histologic outcomes in patients with nonalcoholic fatty liver disease. J Lipid Res 2019;60:176-185.

21. Stickel F, Lutz P, Buch S, Nischalke HD, Silva I, Rausch V, Fischer J, et al. Genetic variation in HSD17B13 reduces the risk of developing cirrhosis and hepatocellular carcinoma in alcohol misusers. Hepatology 2019.

22. Yang J, Trepo E, Nahon P, Cao Q, Moreno C, Letouze E, Imbeaud S, et al. A 17-Beta-Hydroxysteroid Dehydrogenase 13 Variant Protects From Hepatocellular Carcinoma Development in Alcoholic Liver Disease. Hepatology 2019;70:231-240.

23. Ma Y, Brown PM, Rotman Y. Reply to "Does the HSD17B13 rs72613567 splice variant actually yield a new type of alternative splicing". Hepatology 2019.

24. Kozlitina J, Stender S, Hobbs HH, Cohen JC. HSD17B13 and Chronic Liver Disease in Blacks and Hispanics. N Engl J Med 2018;379:1876-1877.

25. Thiam AR, Farese RV, Jr., Walther TC. The biophysics and cell biology of lipid droplets. Nat Rev Mol Cell Biol 2013;14:775-786.

26. Fujimoto T, Parton RG. Not just fat: the structure and function of the lipid droplet. Cold Spring Harb Perspect Biol 2011;3.

27. Kory N, Farese RV, Jr., Walther TC. Targeting Fat: Mechanisms of Protein Localization to Lipid Droplets. Trends Cell Biol 2016;26:535-546.

28. Bickel PE, Tansey JT, Welte MA. PAT proteins, an ancient family of lipid droplet proteins that regulate cellular lipid stores. Biochim Biophys Acta 2009;1791:419-440.

29. Miura S, Gan JW, Brzostowski J, Parisi MJ, Schultz CJ, Londos C, Oliver B, et al. Functional conservation for lipid storage droplet association among Perilipin, ADRP, and TIP47 (PAT)-related proteins in mammals, Drosophila, and Dictyostelium. J Biol Chem 2002;277:32253-32257.

30. Ohsaki Y, Maeda T, Maeda M, Tauchi-Sato K, Fujimoto T. Recruitment of TIP47 to lipid droplets is controlled by the putative hydrophobic cleft. Biochim Biophys Res Commun 2006;347:279-287.
31. Horiguchi Y, Araki M, Motojima K. 17beta-Hydroxysteroid dehydrogenase type 13 is a liver-specific lipid droplet-associated protein. Biochemical and biophysical research communications 2008;370:235-238.

32. Su W, Wang Y, Jia X, Wu W, Li L, Tian X, Li S, et al. Comparative proteomic study reveals 17beta-HSD13 as a pathogenic protein in nonalcoholic fatty liver disease. Proc Natl Acad Sci U S A 2014;111:11437-11442.

33. Friedman SL, Neuschwander-Tetri BA, Rinella M, Sanyal AJ. Mechanisms of NAFLD development and therapeutic strategies. Nat Med 2018;24:908-922.

34. Sonnhammer EL, von Heijne G, Krogh A. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol 1998;6:175-182.

35. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF. Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 1999;112:531-552.

36. Lukacik P, Kavanagh KL, Oppermann U. Structure and function of human 17beta-hydroxysteroid dehydrogenases. Mol Cell Endocrinol 2006;248:61-71.

37. Duax WL, Pletnev V, Addlagatta A, Bruenn J, Weeks CM. Rational proteomics I. Fingerprint identification and cofactor specificity in the short-chain oxidoreductase (SCOR) enzyme family. Proteins 2003;53:931-943.

38. Filling C, Berndt KD, Benach J, Knapp S, Prozorovski T, Nordling E, Ladenstein R, et al. Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. J Biol Chem 2002;277:25677-25684.

39. Belyaeva OV, Adams MK, Wu L, Kedishvili NY. The antagonistically bifunctional retinoid oxidoreductase complex is required for maintenance of all-trans-retinoic acid homeostasis. J Biol Chem 2017;292:5884-5897.

40. Ma Y, Brown PM, Lin DD, Ma J, Feng D, Belyaeva OV, Podszun MC, et al. Hsd17b13 Deficiency Does not Protect Mice From Obesogenic Diet Injury. Hepatology 2020.

41. Marchais-Oberwinkler S, Henn C, Moller G, Klein T, Negri M, Oster A, Spadaro A, et al. 17beta-Hydroxysteroid dehydrogenases (17beta-HSDs) as therapeutic targets: protein structures, functions, and recent progress in inhibitor development. The Journal of steroid biochemistry and molecular biology 2011;125:66-82.

42. Belyaeva OV, Chang C, Berlett MC, Kedishvili NY. Evolutionary origins of retinoid active short-chain dehydrogenases/reductases of SDR16C family. Chem Biol Interact 2015;234:135-143.

43. Jiang W, Napoli JL. The retinol dehydrogenase Rdh10 localizes to lipid droplets during acyl ester biosynthesis. J Biol Chem 2013;288:589-597.

44. Horiguchi Y, Araki M, Motojima K. Identification and characterization of the ER/lipid droplet-targeting sequence in 17beta-hydroxysteroid dehydrogenase type 11. Arch Biochem Biophys 2008;479:121-130.

45. Tsachaki M, Odermatt A. Subcellular localization and membrane topology of 17beta-hydroxysteroid dehydrogenases. Mol Cell Endocrinol 2019;489:98-106.

46. Nakamura N, Fujimoto T. Adipose differentiation-related protein has two independent domains for targeting to lipid droplets. Biochem Biophys Res Commun 2003;306:333-338.
47. Ingelmo-Torres M, Gonzalez-Moreno E, Kassan A, Hanzal-Bayer M, Tebar F, Herms A, Grewal T, et al. Hydrophobic and basic domains target proteins to lipid droplets. Traffic 2009;10:1785-1801.

48. Fukasawa Y, Tsuji J, Fu SC, Tomii K, Horton P, Imai K. MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. Mol Cell Proteomics 2015;14:1113-1126.

49. Yamamoto H, Itoh N, Kawano S, Yatsukawa Y, Momose T, Makio T, Matsunaga M, et al. Dual role of the receptor Tom20 in specificity and efficiency of protein import into mitochondria. Proc Natl Acad Sci U S A 2011;108:91-96.

50. Smith MH, Ploegh HL, Weissman JS. Road to ruin: targeting proteins for degradation in the endoplasmic reticulum. Science 2011;334:1086-1090.

51. Shen LF, Chen YJ, Liu KM, Haddad ANS, Song IW, Roan HY, Chen LY, et al. Role of S-Palmitoylation by ZDHHC13 in Mitochondrial function and Metabolism in Liver. Sci Rep 2017;7:2182.

52. Park CB, Yi KS, Matsuzaki K, Kim MS, Kim SC. Structure-activity analysis of buforin II, a histone H2A-derived antimicrobial peptide: the proline hinge is responsible for the cell-penetrating ability of buforin II. Proc Natl Acad Sci U S A 2000;97:8245-8250.

53. Marchais-Oberwinkler S, Henn C, Moller G, Klein T, Negri M, Oster A, Spadaro A, et al. 17beta-Hydroxysteroid dehydrogenases (17beta-HSDs) as therapeutic targets: protein structures, functions, and recent progress in inhibitor development. J Steroid Biochem Mol Biol 2011;125:66-82.

54. Takahashi Y, Moiseyev G, Farjo K, Ma JX. Characterization of key residues and membrane association domains in retinol dehydrogenase 10. Biochem J 2009;419:113-122, 111 p following 122.

55. Jornvall H, Persson B, Krook M, Atrian S, Gonzalez-Duarte R, Jeffery J, Ghosh D. Short-chain dehydrogenases/reductases (SDR). Biochemistry 1995;34:6003-6013.

56. Huang YW, Pineau I, Chang HJ, Azzi A, Bellemare V, Laberge S, Lin SX. Critical residues for the specificity of cofactors and substrates in human estrogenic 17beta-hydroxysteroid dehydrogenase 1: variants designed from the three-dimensional structure of the enzyme. Mol Endocrinol 2001;15:2010-2020.

57. Oppermann U, Filling C, Hult M, Shafqat N, Wu X, Lindh M, Shafqat J, et al. Short-chain dehydrogenases/reductases (SDR): the 2002 update. Chem Biol Interact 2003;143-144:247-253.

58. Ghosh D, Vihko P. Molecular mechanisms of estrogen recognition and 17-keto reduction by human 17beta-hydroxysteroid dehydrogenase 1. Chem Biol Interact 2001;130-132:637-650.
Table 1. Critical sites for enzymatic activity of HSD17B13

| Name of protein       | Enzymatic Activity | Possible reason for loss of activity |
|-----------------------|--------------------|--------------------------------------|
| HSD17B13              | Yes                | NA                                   |
| HSD17B13-variant B    | No                 | Non-lipid droplet targeting          |
|                       |                    | Low protein stability                |
| HSD17B13-Δ22-28       | No                 | Non-lipid droplet targeting          |
|                       |                    | Low protein stability                |
| HSD17B13-Δexon 6      | No                 | Not known                            |
| HSD17B13-G insertion  | No                 | Not known                            |
| HSD17B13-P260S        | No                 | Not known                            |
| HSD17B13-G47A/G49A    | No                 | Impaired cofactor binding            |
| HSD17B13-N144A        | No                 | Impaired catalytic site/low protein stability |
| HSD17B13-S172A        | No                 | Impaired catalytic site              |
| HSD17B13-Y185A/K189A  | No                 | Impaired catalytic site              |
| HSD17B13-L199A/E202A  | No                 | Impaired substrate binding/low protein stability |
| HSD17B13-K208A        | No                 | Impaired substrate binding           |
| HSD17B13-K153A/L156A  | No                 | Impaired substrate binding/low protein stability |
| HSD17B13-R97A/Y101A   | No                 | Impaired dimer formation             |
Figure 1  N-terminal domains for LD targeting of HSD17B13. (A) HepG2 cells were transiently transfected with HSD17B13 wild type or mutant plasmids and treated with 200 μM oleate and palmitate to induce LDs. Proteins were C-terminally tagged with GFP, which was used to determine their cellular localization (Green). Nuclei were counter-stained with Hoechst (Blue), and LDs were stained with LipidTox (Red). Images were analyzed by confocal microscopy. Bar indicates 10 μM. (B) Schematic representation of full length HSD17B13 (variant A), and mutant proteins. The two naturally-occurring protective mutants found in humans (G insertion and P260S) are included for comparison. “+” indicates targeting to LDs and “−” indicates no targeting.
Figure 2. Identification of Domains Critical for LD targeting of HSD17B13. (A) HepG2 cells were transiently transfected with HSD17B13 wild type, the naturally-occurring variant B (Δ71-106), or mutant plasmids and treated with fatty acids to induce LDs. Proteins are C-terminally tagged with GFP, which was used to determine their cellular localization (Green). Nuclei are counter stained with Hoechst (Blue), and LDs were stained with LipidTox (Red). Images were analyzed by confocal microscopy. Bar indicates 10 μM. (B) Schematic representation of full length HSD17B13 (variant A), and mutant proteins demonstrated in this figure. “+” indicates targeting to LDs and “−” indicates no targeting.
Figure 3. Predicted 3D structure of HSD17B13 homodimers using HSD17B11 (67% sequence identity on modeled sequence, PDB entry 1YB1) as the template. (A) The structure prediction was performed by SWISS-MODEL homology modeling web-server. The classical Rossmann-fold structure was observed in the predicted 3D structure of HSD17B13. The β-sheets are colored in orange and yellow, α-helices in marine and light blue, loop regions in grey, amino acids in exon 2 (which are missing in HSD17B13-Variant B) in black, catalytic tetrad Asn144-Ser172-Tyr185-Lys189 in cyan, substrate binding sites Leu199, Glu202, Lys208 in magenta. The Pro260 affected by the P260S mutation is marked in red, and the truncated amino acids affected by the rs72613567-driven G-insertion (only 3 amino acids are shown in this model) are marked in pink. (B) Sequence alignment of human HSD17B13 and HSD17B11 showing overall identity of 191/298 (64.1%) and similarity of 232/298 (77.9%) amino acids.
Figure 4. The catalytic tetrad Asn144-Ser172-Tyr185-Lys189 is important for HSD17B13 enzymatic activity. (A) Ribbon diagram for the structure of the catalytic tetrad of HSD17B13. NAD and essential residues are labeled and shown as sticks. The binding pose of NAD and retinol were predicted using the HSD17B13 homology model on SwissDock online web-server. (B-D) Enzymatic activity of mutant HSD17B13. HEK293 cells were seeded one day before and transiently transfected in triplicate with HSD17B13, HSD17B13 mutant, or empty vector plasmids. All-trans-retinol was added to the culture and incubated for 6 or 8 hours. Retinaldehyde and retinoic acid were separated by normal phase HPLC and quantified by retinoids standard. Retinoid levels are normalized to protein concentration and shown as relative value to empty vector controls. Data are presented as Mean±SEM. WT-wild type (variant A).
Figure 5. The amino acid residues in the putative substrate binding sites Lys153, Leu156, Leu199, Glu202, Lys208, and the homodimer binding site Arg97, Tyr101 are important for HSD17B13 enzymatic activity. (A) Ribbon diagram for the predicted structure of the substrate binding sites of HSD17B13 with retinol as substrate. Retinol and essential residues are labeled and shown as sticks. (B-E) Enzymatic activity of mutant HSD17B13. HEK293 cells were seeded one day before and transiently transfected in triplicate with HSD17B13, HSD17B13 mutant, or empty vector plasmids. All-trans-retinol was added to culture and incubated for 6 or 8 hours. Retinaldehyde and retinoic acid were separated by normal phase HPLC and quantified by retinoid standards. Retinoids level has been normalized to protein concentration and shown as relative value to empty vector controls. Data are presented as Mean±SEM. WT – wild type (variant A).
Figure 6 – Essential domains and amino acids in HSD17B13. Domains essential for localization are marked in italics. Amino acids essential for enzymatic activity are marked with predicted site function. The consequence of naturally occurring genetic variants is marked in red. Not to scale.