



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

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

CrossMark

Data Article

Biochemical and computational analyses of two phenotypically related GALT mutations (S222N and S135L) that lead to atypical galactosemia

Benjamin Cocanougher^{a,*}, Umut Aypar^b, Amber McDonald^b,
Linda Hasadsri^b, Michael J. Bennett^{c,d}, W. Edward Highsmith^b,
Kristin D'Aco^a

^a University of Rochester School of Medicine and Dentistry, Department of Pediatrics, Division of Genetics, Rochester, NY, USA

^b Molecular Genetics Laboratory, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

^c Metabolic Disease Laboratory, Children's Hospital of Philadelphia, USA

^d Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA

ARTICLE INFO

Article history:

Received 5 January 2015

Accepted 5 January 2015

Available online 7 February 2015

ABSTRACT

Galactosemia is a metabolic disorder caused by mutations in the *GALT* gene [1,2]. We encountered a patient heterozygous for a known pathogenic H132Q mutation and a novel S222N variant of unknown significance [3]. Reminiscent of patients with the S135L mutation, our patient had loss of *GALT* enzyme activity in erythrocytes but a very mild clinical phenotype [3–8]. We performed splicing experiments and computational structural analyses to investigate the role of the novel S222N variant. Alamut software data predicted loss of splicing enhancers for the S222N and S135L mutations [9,10]. A cDNA library was generated from our patient's RNA to investigate for splicing errors, but no change in transcript length was seen [3]. *In silico* structural analysis was performed to investigate enzyme stability and attempt to understand the mechanism of the atypical galactosemia phenotype. Stability results are publicly available in the *GALT* Protein Database 2.0 [11–14]. Animations were created to give the reader a dynamic view of

DOI of original article: <http://dx.doi.org/10.1016/j.yimgmr.2014.12.004>

* Corresponding author.

E-mail address: benjamin_cocanougher@urmc.rochester.edu (B. Cocanougher).

<http://dx.doi.org/10.1016/j.dib.2015.01.001>

2352-3409/© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

the enzyme structure and mutation locations. Protein database files and python scripts are included for further investigation.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications table

Subject area	Medicine
More specific subject area	Medical genetics, galactosemia
Type of data	Graph, protein structure animation, pdb files, python scripts
How data was acquired	<i>In silico</i> structural analyses performed with Chimera (UCSF) and modeller programs. Splicing studies performed with Alamut Splicing Predictor (Cold Spring Harbor Laboratories). RNA studies performed on whole blood sample from patient.
Data format	Analyzed molecular data, Analyzed computational data, and raw.pdb files for structural analysis
Experimental factors	Computational analyses were performed using Alamut software, Modeller, and UCSF Chimera.
Experimental features	Splicing enhancer studies performed using Alamut software. Animations and structural analysis performed with Modeller and UCSF Chimera software.
Data source location	Not applicable
Data accessibility	Structural data available in GALT protein database (GALT Protein Database 2.0). Biochemical data available in article.

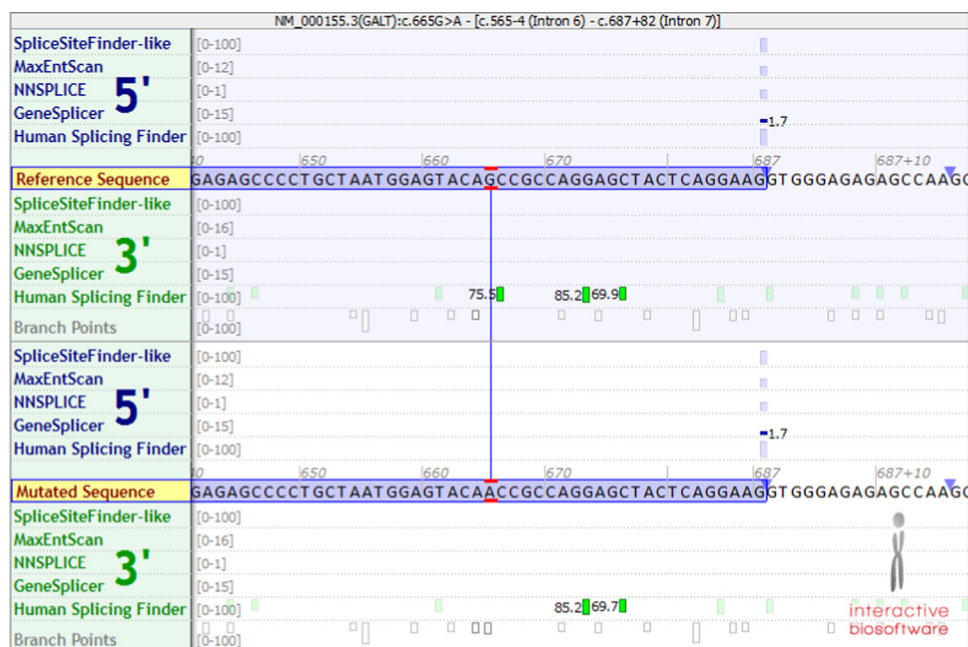


Fig. 1. Screenshot from the Alamut (Cold Spring Harbor Laboratories) Splicing Effects window around the GALT c665G > A variant [9,10]. The top box represents the wild type sequence, with a G at position c.665, and the bottom panel representing the mutant sequences, with an A at c.665. In both panels the nucleotides boxed in lavender represent the 3' end of exon 7. The mutant sequence is not predicted to create or destroy a predicted splice site.

Value of the data

- Heterozygosity for novel S222N GALT mutation leads to atypical galactosemia similar to known S135L cases.
- **Splicing enhancers are predicted to be lost for both S222N and S135L mutations.**
- *In silico* analyses predict stability/instability of dimers present in our patient and allow for visualization of mutation types.

1. Data, experimental design, materials and methods

1.1. Case report

Our data is based on a single case of atypical galactosemia caused by compound heterozygosity for a known pathogenic H132Q mutation and a novel S222N variant of unknown significance [3]. Briefly, our patient had a positive newborn screen for galactosemia with subsequent biochemical

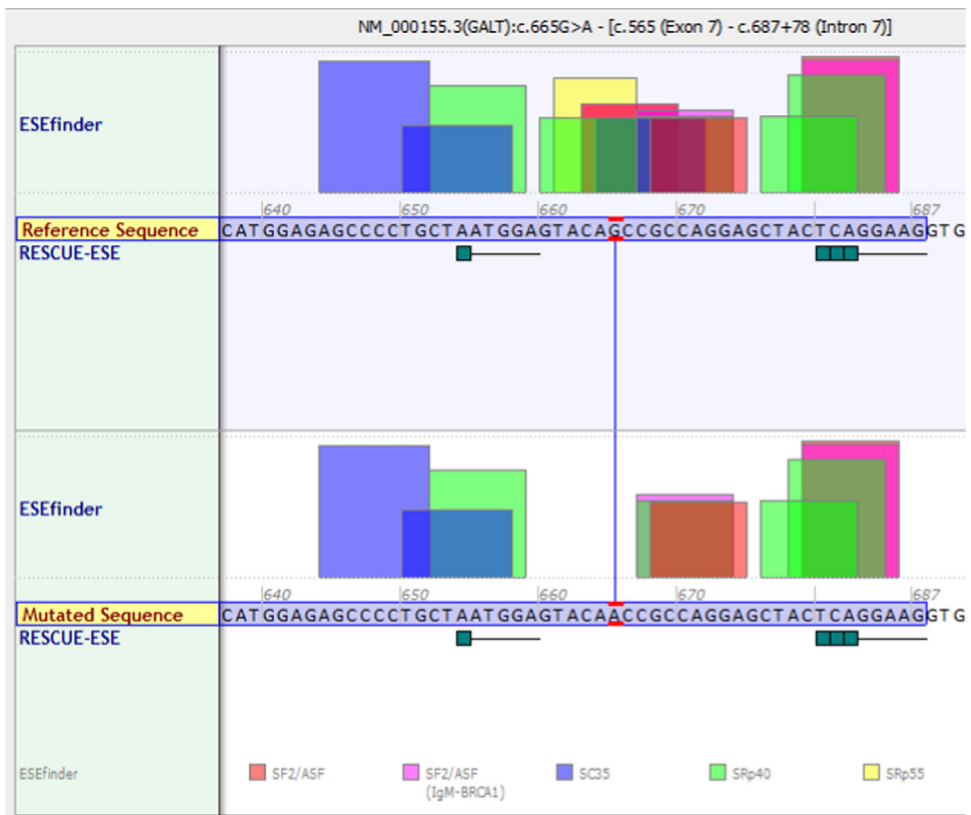


Fig. 2. Screenshot from the Alamut (Cold Spring Harbor Laboratories) Splicing Predictor/ESE Predictions window around the GALT c.665G > A variant [9,10]. The top box represents the wild type sequence, with a G at position c.665, and the bottom panel representing the mutant sequences, with an A at c.665. In both panels the nucleotides boxed in lavender represent the 3' end of exon 7. The mutant sequence is predicted to lose binding sites for SRp40 (green box), SRp55 (yellow box), and SF2/ASF (pink/red box).

confirmation of absent GALT enzyme activity, was treated with a galactose-restricted diet until the age of 3 with no symptoms, and was later lost to follow-up. The patient presented to our clinic at age 17 with no stigmata of galactosemia (apart from mild anxiety) to determine his disease status [15]. Testing revealed nearly complete absence of GALT enzyme activity in erythrocytes without any increase in the clinical biomarkers of galactosemia, galactose-1-phosphate and galactitol [16,17]. Sequencing revealed compound heterozygosity for H132Q and S222N mutations.

Given the ability of our patient to oxidize galactose in tissues other than red blood cells, as evidenced by galactose-1-phosphate and galactitol levels in the normal range, we suspected a splicing defect. Differential splicing seemed a reasonable explanation for differential tissue activity of the GALT enzyme.

1.2. Splicing experiments

We first examined the S222N mutation for changes in splice sites due to genomic c.665G > A. Fig. 1 is a screenshot of the Alamut software program showing that this mutation does not create or destroy a predicted splice site [9,10].

Though this mutation does not directly lead to changes in a splice site, further investigation revealed a predicted loss of a cluster of exonic splicing enhancer (ESE) sites due to the c.665G > A mutation (Fig. 2). This could potentially explain the loss of activity only in erythrocytes if these enhancers were specific to red blood cells and unused in other tissues. Interestingly, the c.404C > T mutation, which leads to the S135L protein, also results in a predicted loss of ESE clusters (Fig. 3).

In order to further investigate this hypothesis, RNA was isolated from the whole blood sample, a cDNA library was created, and PCR was used to amplify the entire gene in several overlapping pieces. No shift in cDNA size was seen, which would be expected in the case of interrupted or aberrant splicing

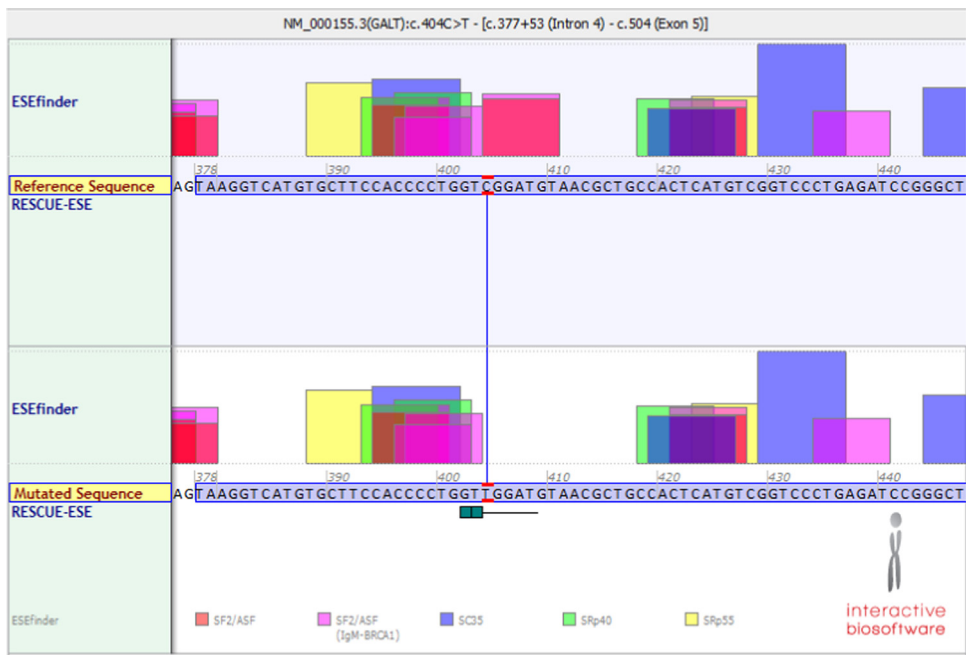


Fig. 3. Screenshot from the Alamut (Cold Spring Harbor Laboratories) Splicing Predictor/ESE Predictions window around the GALT c.404 C > T variant responsible for the S135L mutation [8,9]. The top box represents the wild type sequence, with a G at position c.665, and the bottom panel representing the mutant sequences, with an A at c.665. In both panels the nucleotides boxed in lavender represent the 3' end of exon 7. The mutant sequence is predicted to lose binding sites for SF2/ASF (pink/red box).

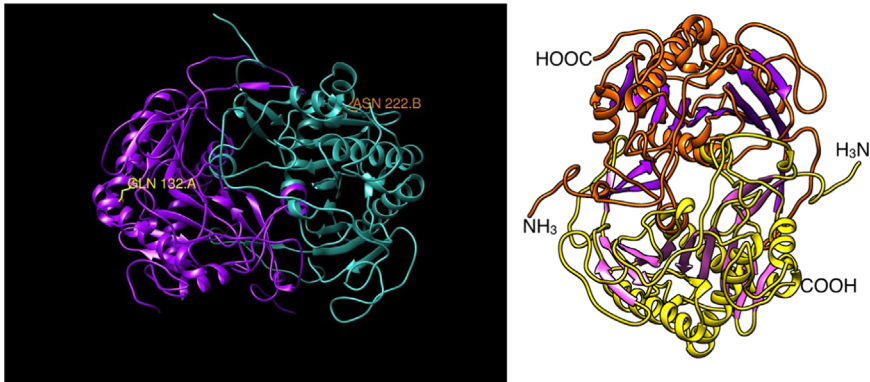


Fig. 4. Video stills of two movies generated to examine the structure of GALT enzyme variants. Supplementary Movie 1 examines the location and neighboring residues for the H132Q/S222N heterodimer. Supplementary Movie 2 examines the location of the S222N and S135L mutation residues.

(Supplementary Fig. 1) [3]. The cDNA was sequenced to look for splicing defects that may not be apparent on gel electrophoresis, but there were no changes in exon structure from the reference or the father's cDNA sequence (Supplementary Fig. 2) [3]. The only base pair change identified was the c.665G > A variant itself.

1.3. Structure analysis

Protein modeling was performed to evaluate the stability of the three types of dimers produced by our patient (H132Q homodimer, H132Q/S222N heterodimer, and S222N homodimer) and the phenotypically similar S135L homodimer (Supplementary Files 1–4). We submitted our mutation to the GALT protein database run by Dr. Anna Marabotti and the Laboratory of Bioinformatics at the Institute of Food Science. The models were mutated using the “mutate model” function in Modeller. Data on stability of each of the various GALT dimers are accessible in the GALT Protein Database 2.0 [11,12].

Using the UCSF Chimera software, we investigated the residues involved in our four dimeric structures. As evidenced in the included animations, the mutations in the H132Q and S222N heterodimer are too far from one another in the structure to interact directly (Supplementary Movie 1) (Fig. 4).

A separate analysis was performed to visualize the S222N and S135L residue locations. The two are located on separate ends of the dimer, but both appear to be peripherally accessible for phosphorylation in the non-mutated state (Supplementary Movie 2). This animation was colored to match the original publication of the GALT enzyme structure [18]. Python scripts created to generate both of the animations above have been included (Supplementary Files 5,6). They can be opened and modified using the UCSF Chimera software [14].

Appendix A. Supplementary information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2015.01.001>.

References

- [1] J.L. Fridovich-Keil, J.H. Walter, Galactosaemia, in: D. Valle, B. Vogelstein, K.W. Kinzler, S.E. Antonarakis, A. Ballabio, et al., (Eds.), *The Online Metabolic and Molecular Bases of Inherited Disease*, OMMBID, Part 7: Carbohydrates, McGraw Hill, New York 2008, p. 2008. (Chapter 72).

- [2] A.M. Bosch, L. Ijlst, W. Oostheim, J. Mulders, H.D. Bakker, F.A. Wijburg, R.J. Wanders, H.R. Waterham, Identification of novel mutations in classical galactosemia, *Hum. Mutat.* 25 (5) (2005) 502.
- [3] B. Cocanougher, U. Aypar, A. McDonald, L. Hasadsri, M.J. Bennett, W.E. Highsmith, K. D'Aco, Compound heterozygosity with a novel S222N GALT mutation leads to atypical galactosemia with loss of GALT activity in erythrocytes but little evidence of clinical disease, *Mol. Genet. Metab. Rep.* (2014).
- [4] L. Baker, W.J. Mellman, T.A. Tedesco, S. Segal, Galactosemia: symptomatic and asymptomatic homozygotes in one Negro sibship, *J. Pediatr.* 68 (1966) 551–558.
- [5] S. Segal, P. Cuatrecasas, The oxidation of C 14 galactose by patients with congenital galactosemia, *Am. J. Med.* 44 (1968) 340–347.
- [6] S. Segal, S. Rogers, P.G. Holtzapple, Liver galactose-1-phosphate uridylyl transferase: activity in normal and galactosemic subjects, *J. Clin. Invest.* 50 (1971) 500–506.
- [7] K. Lai, S.D. Langley, R.H. Singh, P.P. Dembure, L.N. Hjelm, L.J. Elsas, A prevalent mutation for galactosemia among black Americans, *J. Pediatr.* 128 (1) (1996) 89–95.
- [8] N. Manga, T. Jenking, H. Jackson, D.A. Whittaker, A.B. Lane, The molecular basis of transferase galactosaemia in South African negroids, *J. Inherit. Metab. Dis.* 22 (1) (1999) 37–42.
- [9] P.J. Smith, C. Zhang, J. Wang, S.L. Chew, M.Q. Zhang, A.R. Krainer, An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers, *Hum. Mol. Genet.* 15 (16) (2006) 2490–2508.
- [10] L. Cartegni, J. Wang, Z. Zhu, M.Q. Zhang, A.R. Krainer, ESEfinder: a web resource to identify exonic splicing enhancers, *Nucl. Acid Res.* 31 (13) (2003) 3568–3571.
- [11] A. Marabotti, A.M. Facchiano, Homology modeling studies on human galactose-1-phosphate uridylyltransferase and on its galactosemia-related mutant Q188R provide an explanation of molecular effects of the mutation on homo- and heterodimers, *J. Med. Chem.* 48 (2005) 773–779.
- [12] A. d'Acerno, A. Facchiano, A. Marabotti, GALT protein database: querying structural and functional features of GALT enzyme, *Hum. Mutat.* 35 (2014) 1060–1067, <http://dx.doi.org/10.1002/humu.22613>.
- [13] N. Eswar, M.A. Marti-Renom, B. Webb, M.S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali, Comparative protein structure modeling with MODELLER, *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc., 2006 (Supplement 15, 5.6.1–5.6.30).
- [14] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (13) (2004) 1605–1612.
- [15] S.E. Waisbren, N.L. Potter, C.M. Gordon, R.C. Green, P. Greenstein, C.S. Gubbels, E. Rubio-Gozalbo, D. Schomer, C. Welt, V. Anatasoia, The adult galactosemic phenotype, *J. Inherit. Metab. Dis.* 35 (2) (2012) 279–286.
- [16] G.T. Berry, I. Nissim, L. Zhiping, A.T. Mazur, J.B. Gibson, S. Segal, Endogenous synthesis of galactose in normal men and patients with hereditary galactosaemia, *Lancet* 346 (1995) 1073–1074.
- [17] C. Jakobs, S. Schweitzer, B. Dorland, Galactitol in galactosemia, *Eur. J. Pediatr.* 154 (Suppl. 2) (1995) S50–S52.
- [18] J.E. Wedekind, P.A. Frey, I. Rayment, Three-dimensional structure of galactose-1-phosphate uridylyltransferase from *Escherichia coli* at 1.8 Å resolution, *Biochemistry* 34 (1995) 11049–11061.