Türk Biyokimya Dergisi [Turkish Journal of Biochemistry - Turk J Biochem] 2005; 30 (4); 306-309.

Research Article [Araştırma Makalesi] Yayın tarihi Aralık, 2005 © TurkJBiochem.com

[Published online December, 2005]

Identification and Characterization of Rat GMDS Gene by Using Bioinformatics Tools

[Biyoinformatik Araçlar Kullanarak Sıçan GMDS Geninin Tanımlanması ve Karakterizasyonu]

Lokman Varışlı
Osman Cen

Harran University, Art and Science Faculty, Department of Biology, Sanliurfa, TURKEY
Harran Üniversitesi, Fen Edebiyat Fakültesi, Biyoloji Bölümü, Şanlıurfa, TÜRKİYE

Yazışma Adresi
[Correspondence Address]

Lokman VARIŞLI
Harran Üniversitesi Fen Edebiyat Fakültesi
Biyoloji Bölümü,Osmanbey Kampüsü, Şanlıurfa, Türkiye
Telefon: 0.414.344 00 20
Faks:0.414.344 00 51
E-Mail: vlokman@harran.edu.tr

Kayıt tarihi 29.07.2005; kabul tarihi 01.12.2005
[Received 29.07.2005; accepted 01.12.2005]

ABSTRACT

GDP-mannose 4,6- dehydratese is conserved throughout evolution. This molecule catalyzes conversion of GDP-D-mannose to GDP-4-keto-6-D-deoxymannose by catalyzing the oxidation of the hydroxyl group. This is the first step in conversion of GDP-mannose to GDP-fucose. Defects in GDP-fucose metabolism have been linked to leukocyte adhesion deficiency type II. Here, we identified and characterized the rat GMDS gene using bioinformatics tools. Rat GDP-mannose 4,6- dehydratese gene, consisting of 11 exons, was located within rat genomic contig NW_047491. Rat, human, and mouse GDP-mannose 4,6- dehydratese showed 95-98% total amino acid identity. The conserved Epimerase domain was identified within GDP-mannose 4,6- dehydratese protein. This is the first report on identification and characterization of the rat GMDS gene.

Key Words: GMDS, bioinformatics, rat genome, comparative genomics, GDP-mannose-4,6,dehydratase

ÖZET

GDP - mannoz 4,6- dehidrataz evrimsel süreç boyunca korunmuştur. Bu molekül hidroksil grubunun katalizlenmesi ile GDP-D-mannoz'un GDP-4-keto-6-D-deoxy-mannoza dönüşümünü katalizler. GDP-mannoz'un GDP-fukoz'a dönüşümü için bu ilk adımdır. GDP-fukoz metabolizmasında ortaya çıkabilecek bir defekt tip II lökosit adezyon eksikliğine yol açabilmektedir. Bu çalışmada biyoinformatik araçlar kullanarak sıçan GDP - mannoz 4,6- dehidrataz genini tanımladık ve karakterize ettik. Sıçan GDP - mannoz 4,6- dehidrataz geni 11 eksondan oluşmaktadır ve sıçan genomunda 17. kromozom üzerinde genomik kontig NW_047491 üzerinde yer almaktadır. Sıçan GDP - mannoz 4,6- dehidrataz proteinin fare GDP - mannoz 4,6-dehidrataz proteini ile %98 oranında (365/370) ve insan GDP - mannoz 4,6- dehidrataz proteini ile de %95 oranında (357/370) benzerlik gösterdiği görülmektedir. Sıçan GDP-mannoz 4,6- dehidrataz proteini üzerinde epimeraz domeyni belirlenmiştir. Bu çalışma sıçan GDP - mannoz 4,6- dehidrataz geninin tanımlanması ve karakterizasyonu üzerine yapılmış olan ilk çalışmadır.

Anahtar Kelimeler: GMDS, biyoinformatik, sıçan genomu, karşılaştırmalı genomik, GDP-mannoz 4,6-dehidrataz

INTRODUCTION

GDP-mannose 4,6- dehydratese (GMDS) is conserved throughout evolution (1). GMDS, the first enzyme involved in the de novo biosynthesis pathway of GDP-fucose, has been obtained at different degrees of purity from various sources, both from bacteria and mammals (2). GMDS catalyzes conversion of GDP-D-mannose to GDP-4-keto-6-D-deoxymannose. This is the first step in conversion of GDP-mannose to GDP-fucose (3, 4). Fucose forms a part of a number of glycoconjugates, including the ABO blood groups and the selectin ligand sialyl Lewis X. Defects in GDP-fucose metabolism have been linked to leukocyte adhesion deficiency type II (LADII) abnormalities (5). GMDS is a homodimeric protein with each monomer composed of two domains. The larger N-terminal domain binds the NADP(H) cofactor in a classical Rossmann fold and the C-terminal domain harbors the sugar-nucleotide binding site (5).

Here we identified Rat GMDS gene using bioinformatics tools. In addition to the structure and the chromosomal localization of the rat GMDS gene, the transcribed and translated protein product of the GMDS gene is analyzed *in silico*. This is the first report on identification and characterization of the rat GMDS gene.

MATERIAL AND METHODS

Identification of novel gene fragments in rat genome

Novel rat gene fragments homologous to human GMDS gene was searched for with TblastN (6) program using the homology based gene finding algorithm as described previously (7,8). Amino acid sequence of Human GMDS was used as a query sequence for the TblastN program, in which the query amino acid sequence was compared with the rat genome sequences and rat expressed sequence tags (ESTs) translated in 6 frames.

Structure and chromosomal localization of the novel gene

Exon-intron boundaries around the gene fragments identified by the TblastN program were determined by examining the consensus sequence of exon-intron boundaries ('gt…ag' rule of intronic sequence) and the codon usage within the coding region. Based on the rat chromosomal localization of genome sequences in the NCBI database (http://www.ncbi.nlm.nih.gov), rat chromosomal localization of the novel gene was determined.

Analyses of deduced amino acid sequence

Prediction of coding region as well as translation into amino-acid sequence were performed using ORF Finder program (www.ncbi.nlm.nih.gov). The homolog amino acid sequences of the novel constructed gene product were searched for with the BLASTp program. Amino acid sequences of human, mouse and rat GMDS genes were aligned using the ClustalW program(9). Domain structure of the novel protein was searched for with the pfam program (10).

RESULTS

Identification of rat GMDS gene

Rat genome sequences and ESTs homologous to human GMDS were searched for with the TblastN program using the amino acid sequence of human GMDS as the query sequence. TblastN analysis of human GMDS revealed that GMDS homologous gene fragments were located within rat genome contig NW_047491 on the chromosome 17 and NW_047469 on the chromosome 16. GMDS homologous gene fragments on NW_047469 (chromosome 16) did not seem to code for any protein because there weren't any corresponding ESTs or cDNAs in this region.

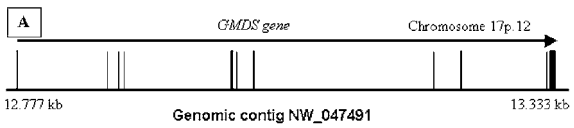
Structure of rat GMDS gene

Rat GMDS gene consists of 11 exons in the rat genome (Figure 1a). Precise exon-intron boundaries of rat GMDS gene were determined based on the contig NW_047491 on the chromosome 17 of rat by examining the consensus sequence of exon-intron boundaries and the codon usage within the coding region (Figure 1b). The rat GMDS gene, consisting of at least 11 exon, was located on the chromosome 17 at the nucleotide position 12.777.888-13.332.937 Therefore, the GMDS gene was found to be a long coding sequence with a size of about 555 kb.

Amino acid sequence of rat GMDS

Nucleotide sequence of rat GMDS mRNA was determined by combining nucleotide sequences of its predicted exons. Rat GMDS mRNA was found to consist of least 28 bp of 5' UTR, 1110 bp of coding region, and at least 332 bp of a 3' UTR region (Figure 2). The predicted mRNA of the GMDS gene encodes a peptide of 370 amino acids.

The Blastp program revealed that rat GMDS has a 98% (365/370) and 95% (357/370) total amino acid identity with GMDS of *Mus musculus*, *H. sapiens*, respectively. ClustalW alignment of GMDS proteins, *Mus musculus* GMDS, and *Homo sapiens* GMDS displayed very high similarity of GMDS proteins in these species (Figure 3). pfam result revealed that rat GMDS protein harbor a big Epimerase domain spanning from amino acid 28 to 358, which also contain a signature of NAD dependent epimerase / dehydratase family proteins (Figure 3).



B

Exon No	Nucleotide sequence around exon-intron boundaries		
	5'-UTR	GGTCGC.....GGCCAG	GTGAGT...
2	..TCCCAG	GATGGT.....TACGAG	GTGAGT...
3	...ACTCAG	GTCCAT.....AAGGAA	GTAAGT...
4	..TAAAAG	ACATGA.....GTCAAG	GTAAGC...
5	..TCCCAG	ATCTCC.....CCTATG	GTAAGG...
6	..CATCAG	GAGCCG.....GAAGAG	GTGAGT...
7	...TTTTAG	GAGCTA.....GACTAT	GTAGAG...
8	..CTCCAG	GCTATG.....CATTGT	GTAAGT...
9	..TTGCAG	GTGGGA.....GAAGTG	GTAAGG...
10	..CCACAG	GACTTT.....TTCGAT	GTAAGT...
11	..TTGCAG	GAGCTG.....TAGCAA	3'-UTR

Figure 1. Structure and chromosomal localization of rat GMDS gene. A, GMDS gene located on rat chromosome 17p.12. B, Exon–Intron boundries of rat GMDS gene. Rat GMDS gene consists of 11 exons, spans nucleotide position 12.777 kb - 13.333 kb of rat chromosome 17 (the Genomic contig NW_047491).

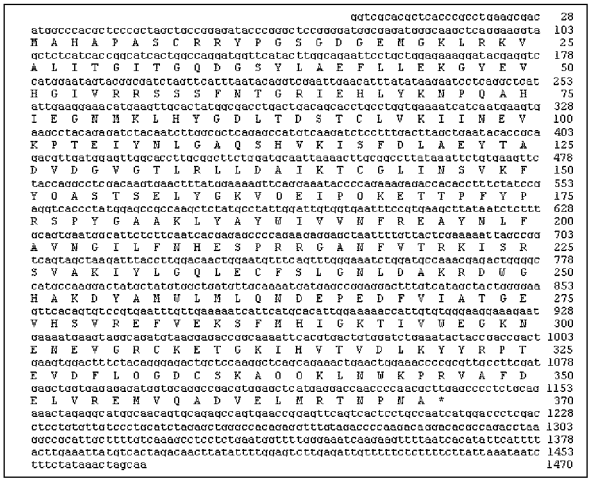


Figure 2. Rat GMDS mRNA and rat GMDS protein. Nucleotide sequence of rat GMDS mRNA and amino acid sequence of rat GMDS protein are shown. Nucleotides and amino acids are numbered on the right.

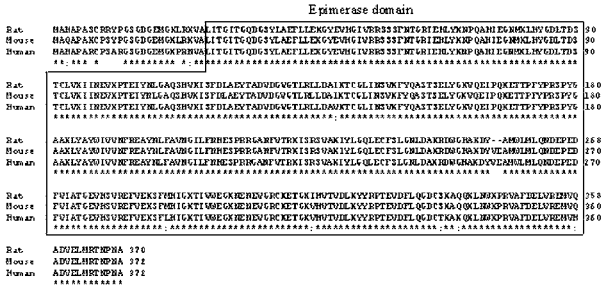


Figure 3. Alignment of rat, human and mouse GMDS protein. Amino acid residues are numbered on the right. Amino acid residues conserved among human, mouse and rat GMSD protein are shown by asterisk below the alignment. Human, mouse and rat GMDS proteins also include epimerase conserved domain structure. This domains are boxed.

DISCUSSION

We identified and characterized the Rat GMDS gene using bioinformatics tools in this study. Complete coding sequence of GMDS gene was determined by assembling the identified 11 exons within the genomic contig NW_047491 of rat chromosome 17p.12 (Figure 1a). Rat GMDS protein (370 aa) was identified based on the nucleotide sequences of the exons (Figure 2). Even though there are some homolog segments on the chromosome 16, these segments did not seem to assemble into a functional gene.

Rat GMDS protein includes an epimerase domain (Figure 3). This domain is found in the NAD dependent epimerase/dehydratase family proteins. Proteins in this family are enzymes utilizing NAD as a cofactor in a variety of biochemical reactions involving nucleotide-sugar molecules as substrates (11).

High homology between the rat, mouse, and human GMDS genes indicate that they may well have been conserved and this conservation may be of paramount importance for its role in the cellular physiology.

Methods of bioinformatics have a role in the confirmation of established laboratory-based work and in *de novo* analysis (12). However results of *in silico* (bioinformatical) studies generally needs confirmation by lab experimentations. Therefore findings of this study need to be confirmed by wet-lab experimentation.

References

1. Becker DJ and Lowe JB (2003) Fucose: Biosynthesis and biological function in mammals. Glycobiology 13(7), 41R – 53R.

2. Tonetti M, Sturla M, Bisso A, Zanardi D, Benatti U, and De Flora A (1998) The metabolism of 6-deoxyhexoses in bacterial and animal cells. Biochimie 80(11), 923-931.

3. Sturla L, Bisso A, Zanardi D, Benatti U, De Flora A, Tonetti M (1997) Expression, purification and characterization of GDP-D-mannose 4,6-dehydratase from Escherichia coli. FEBS Lett. 412(1), 126-130.

4. Wu B, Zhang Y, Wang PG (2001) Identification and characterization of GDP-d-mannose 4,6-dehydratase and GDP-l-fucose snthetase in a GDP-l-fucose biosynthetic gene cluster from Helicobacter pylori. Biochem Biophys Res Commun. 285(2), 364-371.

5. Somoza JR, Menon S, Schmidt H, Joseph-McCarthy D, Dessen A, Stahl ML, Somers WS, Sullivan FX (2000) Structural and kinetic analysis of Escherichia coli GDP-mannose 4,6 dehydratase provides insights into the enzyme's catalytic mechanism and regulation by GDP-fucose. Structure Fold Des. 8(2), 123-135.

6. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol. 215(3), 403-410.

7. Katoh M (2002) Paradigm shift in gene –finding method: From

bench-top approach to desk-top approach. Int.J.Mol.Med. 10. 677-682.

8. Wei L, Liu Y, Dubchak I, Shon J, Park J (2002) Comparative genomics approaches to study organism similarities and differences. J Biomed. Inform. 35(2), 142-150.

9. Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, and Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

10. Bateman A., Birney E., Cerruti L., Durbin R., Etwiller L., Eddy SR., Griffiths-Jones S., Howe K.L., Marshall M. and Sonnhammer EL (2002) The Pfam Protein Families Database. Nucl. Acids Res. 30(1), 276-280.

11. Thoden JB, Hegeman AD, Wesenberg G, Chapeau MC, Frey PA, Holden HM (1997) Structural analysis of UDP-sugar binding to UDP-galactose 4-epimerase from Escherichia coli. Biochemistry 36(21), 6294-6304.

12. Gale CP, and Grant PJ (2004) The characterization and functional analysis of the human glyoxalase-1 gene using methods of bioinformatics. Gene 340, 251–260.