BBA - General Subjects 1862 (2018) 866–876

Contents lists available at ScienceDirect

BBA - General Subjects

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

Functional characterization of naturally occurring genetic variations of the human guanine-rich RNA sequence binding factor 1 (GRSF1)

Sajad Sofi[a], Julia C. Fitzgerald[b], Désirée Jähn[a], Bernhard Dumoulin[a], Sabine Stehling[a], Hartmut Kuhn[a], Christoph Ufer[a,*]

[a] Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health, Institute of Biochemistry, Chariteplatz 1, CCO-Building, Virchowweg 6, D-10117 Berlin, Germany

[b] Functional Neurogenomics Laboratory, Department of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, Centre for Integrative Neuroscience (CIN), the German Centre For Neurodegenerative Diseases (DZNE), University of Tuebingen, Otfried-Mueller-Strasse 27, D-72075 Tuebingen, Germany

ARTICLE INFO

Keywords:
hnRNP F/H
RNA-binding
RNA recognition motif
Thermostability
Genetic variation

ABSTRACT

The guanine-rich RNA sequence binding factor 1 (GRSF1) constitutes an ubiquitously occurring RNA-binding protein (RBP), which belongs to the family of heterogeneous nuclear ribonucleoprotein F/H (hnRNP F/H). It has been implicated in nuclear, cytosolic and mitochondrial RNA metabolism. Although the crystal structures of GRSF1 orthologs have not been solved, amino acid alignments with similar RNA-binding proteins suggested the existence of three RNA-binding domains designated quasi-RNA recognition motifs (qRRMs). Here we established 3D–models for the three qRRMs of human GRSF1 on the basis of the NMR structure of hnRNP F and identified the putative RNA interacting amino acids. Next, we explored the genetic variability of the three qRRMs of human GRSF1 by searching genomic databases and tested the functional consequences of naturally occurring mutants. For this purpose the RNA-binding capacity of wild-type and mutant recombinant GRSF1 protein species was assessed by quantitative RNA electrophoretic mobility shift assays. We found that some of the naturally occurring GRSF1 mutants exhibited a strongly reduced RNA-binding activity although the general protein structure was hardly affected. These data suggested that homozygous allele carriers of these particular mutants express dysfunctional GRSF1 and thus may show defective GRSF1 signaling.

1. Introduction

The guanine-rich RNA sequence binding factor 1 (GRSF1) is a member of the heterogeneous nuclear ribonucleoprotein (hnRNP) F/H family. The members of the family are dominantly nuclear proteins, which have been implicated in splicing regulation [1–4]. In contrast, GRSF1 has also been found outside the nucleus and has been identified as an RNA-binding protein (RBP) that recognizes specific G-rich stretches of RNA [1]. In the cytoplasm GRSF1 functions as translational activator for both, eukaryotic and viral RNA transcripts and has been implicated in recruiting RNA molecules to active ribosomes [5–9]. Interestingly, the influenza virus takes advantage of the GRSF1 activity to stimulate the biosynthesis of viral proteins by the host cells [5,6]. Moreover, GRSF1 plays a role in mitochondrial RNA metabolism. In fact, GRSF1 associates with nascent RNA in mitochondrial RNA granules, where it may play a role in the maturation of transcripts originating from the mitochondrial genome [10]. More recently, GRSF1 has been implicated in the expression regulation of regulatory RNAs [11–13]. During ontogenesis the *Grsf1* gene is already expressed in developing embryos and siRNA-induced expression silencing induces abnormal brain development by elevating oxidative stress [7,14].

Human GRSF1 is encoded for by a single copy gene, which is located on the short arm of chromosome 4 (4q13.3). The corresponding mouse ortholog has been mapped to a syntenic region of chromosome 5 (88,659,448–88,676,171 bp). Unpublished database (NCBI BLAST) searches suggested that the *GRSF1* gene is highly conserved in lower and higher mammals but has not been detected in the genomes of lower model organisms such as *E. coli, D. melanogaster* and *C. elegans*. However, *GRSF1* is present in *D. rerio* [15], which is frequently used as model organism of vertebrates. The gene locus of human GRSF1 has been correlated with familial mesial temporal lobe epilepsy (FMTLE) but the involvement of GRSF1 in the pathogenesis of this disorder has not been investigated in detail [16]. The human *GRSF1* gene has been suggested to give rise to several isoproteins, which differ from each other with respect to their N-terminal sequences [15,1]. At least one of these isoforms (NM_001098477.1) lacks the N-terminal alanine-rich

* Corresponding author at: Institute of Biochemistry, University Medicine Berlin-Charité, Charitepl. 1, CCO-Building, Virchowweg 6, D-10117 Berlin, Germany.
E-mail address: christoph.ufer@charite.de (C. Ufer).

https://doi.org/10.1016/j.bbagen.2017.12.008
Received 4 July 2017; Received in revised form 12 December 2017; Accepted 22 December 2017

domain but all isoforms carry the three strictly conserved RNA-binding domains (RBDs) [10]. Although the detailed molecular function of the Ala-rich domain remains unclear, it was recently suggested to carry a mitochondrial targeting signal, which is responsible for mitochondrial localization of GRSF1 isoforms involving this domain [10]. In contrast, GRSF1 isoforms lacking the alanine-rich domain are mainly localized to the cytoplasm [10]. In addition, an acidic domain, which is rich in glutamate residues, can be found between the two C-terminal RBDs [15,1]. Unfortunately, the function of this auxiliary domain is not known. The longest GRSF1 isoform (NM_002092.3) consists of 480 amino acids and contains N-terminal mitochondrial targeting signal, the two auxiliary domains (Ala-rich domain, acidic domain) as well as all three RBDs. This GRSF1 isoform has mainly been detected as constituent of mitochondrial RNA granules and has been implicated in mitochondrial RNA processing [10]. Similar to other members of the hnRNP F/H family the RBDs of GRSF1 resemble quasi-RNA recognition motifs (qRRMs). They are denoted quasi-RRMs because conserved aromatic and positively charged residues found in the corresponding RNA-binding regions of classical RRMs are poorly conserved [17].

GRSF1 specifically interacts with G-rich RNA sequences [7,1,18]. Such G-rich stretches have a strong tendency for self-association and potentially fold into four-stranded secondary structures called G-quadruplexes [19–21]. In fact, RNA substrates of GRSF1 have been shown to fold into G-quadruplex structures [18] but owing to the low number of mRNA species for which GRSF1 interactions has been reported, it is currently impossible to decide whether G-quadruplex formation is a necessary precondition for GRSF1 binding. G-quadruplexes do not only occur in mRNAs but they have also been identified as part of the chromosomal telomeres [22]. However, so far only one viral (influenza NP) and two cellular (GPx4, Use1) RNA substrates have been characterized for GRSF1 [5,7,18] and several coding and non-coding mitochondrial RNA transcripts have been suggested to be associated with GRSF1 [10,13,23].

The genetic variability of the human *GRSF1* gene has not been explored in detail and it remains unclear whether naturally occurring amino acid exchanges affect (impairment or enhancement) the interaction between the GRSF1 and its target RNAs. To shed light on this topic we first searched several online sequence databases for naturally occurring mutations in the human *GRSF1* gene. Selected GRSF1 mutants with potentially defective phenotypes were expressed in *E. coli*, purified by affinity chromatography and their RNA-binding capacities were tested. Our results demonstrate that some of the naturally occurring mutants exhibit significantly impaired RNA-binding affinity although the general structures of the corresponding proteins were not affected.

2. Methods

2.1. Chemicals

The chemicals used were obtained from the following sources: Chloramphenicol and ampicillin from Roth (Karlsruhe, Germany), Kanamycin from Sigma-Aldrich (Steinheim, Germany), isopropyl-*β*-*D*-thiogalacto-pyranoside (IPTG) from Roth (Karlsruhe, Germany). The restriction enzymes *Eco*RI/*Hin*dIII and *Xho*I/*Bam*HI were purchased from Thermoscientific (Darmstadt, Germany). Oligonucleotides were synthesized by Biotez (Berlin, Germany) and DNA sequencing was performed at Eurofins MWG Operon (Ebersberg, Germany). The *E.coli* strain BL21 (DE3) and the fluorescent dye Sypro Orange® were purchased from Life Technologies (Carlsbad, USA). The alkaline-phosphatase (AP) labeled antibody was purchased from Roche Diagnostics (Manheim, Germany) and horseradish-peroxidase labeled antibody from Sigma (Steinheim, Germany). The maleic acid and blocking buffer for RNA gel shift assays were purchased from Roche Diagnostics (Manheim, Germany).

2.2. Recombinant expression and purification of full-length GRSF1 and its domains

The full-length cDNA encoding for human GRSF1 was amplified from an RNA extract of human embryonic kidney 293 (HEK293). For the RNA-binding studies we expressed a truncated version of human GRSF1, which lacks the N-terminal Ala-rich-domain but involved the three RNA-binding motifs qRRM1, qRRM2 and qRRM3. This construct was called ΔE1-hGRSF1 and was employed as template for all mutagenesis studies. For recombinant expression of the different GRSF1 constructs the ΔE1-hGRSF1 cDNA was cloned into the bacterial expression plasmid pET-42a, which contains an N-terminal glutathione-*S*-transferase tag (GST-tag) and expression was carried out using the EnPresso System according to the manufacturer's instructions (Biosilta, Berlin, Germany). In short, competent *E. coli* cells [BL21 (DE3), Life Technologies, Carlsbad, USA] were transformed with the recombinant plasmids and selected bacterial clones were grown at 30 °C in a 50 mL bacterial liquid culture containing kanamycin (50 μg/mL) as selection marker until an $OD_{600} > 5$ was reached. Expression of the recombinant proteins was then induced by the addition of 1 mM (final concentration) of isopropyl-*β*-*D*-thiogalacto-pyranoside (IPTG) and bacteria were further cultured for 15–18 h at room temperature under vigorous agitation to avoid shortage of oxygen. The cells were harvested by centrifuging (4.000 r.p.m., 4 °C, 15 min), the pellet was reconstituted in phosphate buffered saline (PBS) and sonicated twice (3 times; 10 s; 20% maximal intensity) using a Branson tip-sonifier (Heinemann, Schwabisch Gmund, Germany). For affinity chromatographic purification of the recombinant proteins the cell lysate was centrifuged (20.000 × *g*, 4 °C, 20 min) and the supernatant was incubated with glutathione-coupled agarose (Machery-Nagel, Düren, Germany) at 4 °C for 1 h under gentle agitation. Unbound proteins were washed away four times with washing buffer (140 mM NaCl; 2.7 mM KCl; 10 mM Na_2HPO_4; 1.8 mM KH_2PO_4; pH 7.3). Bound GST-tagged proteins were competitively eluted with elution buffer (50 mM Tris-HCl, pH 9.5 containing 10 mM reduced L-glutathione). The elution fractions were combined and concentrated to the final concentration of 0.5–1 mg/mL using a protein concentrator (Corning, England, UK) with a cutoff limit of 5 kDa. The purified protein was either snap frozen in liquid nitrogen or mixed with 10% *v*/v glycerol, frozen and then stored at − 80 °C.

2.3. Site-directed mutagenesis

For the production of human GRSF1 point mutants, mutagenesis experiments were carried out using the QuickChange® Site-Directed Mutagenesis Stratagene (La Jolla, USA) following manufacturer's instructions.

2.4. In vitro transcription

RNA probes used in the RNA gel shift assays were synthesized by *in vitro* transcription using the T7 Megashortscript Kit (Ambion, Huntingdon, UK) in the presence of digoxigenin-UTP (molar ratio of UTP:digoxigenin-UTP was 15:1) (Roche Diagnostics, Mannheim, Germany). Thus, statistically each RNA molecule contains a maximum of one labeled uracil nucleotide. RNA probes were purified using spin column chromatography (Bio-Rad, California, USA) and analyzed under denaturing conditions by 10% urea gel electrophoresis.

2.5. RNA electrophoretic mobility shift assays (REMSA)

The RNA probe (5′-GCC GAC GCG CGU CCA UUG GUC GGC UGG ACG AGG GGA GGA GCC GCU GGC UCC CAG-3′) was generated by *in vitro* transcription as described above. This RNA probe corresponds to the 5′-UTR of *GPx4* mRNA [7] and was employed as a GRSF1-binding substrate for REMSA. RNA probes were labeled with Digoxigenin-11-