Objectius

Aquesta tesi s'ha realitzat en el grup d'estudi de malalties lisosòmiques. L'objectiu general plantejat va ser l'anàlisi de les bases moleculars de dues noves malalties lisosòmiques: la Gangliosidosi GM1 i la malaltia de Morquio B.

Els objectius concrets plantejats van ser:

- Realitzar un estudi mutacional del gen *GLB1* en mostres de pacients procedents principalment d'Espanya i d'Argentina.

- Caracterizar l'efecte específic d'aquelles mutacions que alteren el procés d'*splicing*.

- Realitzar una expressió *in vitro* d'al.lels mutants identificats prèviament, per demostrar la seva patogenicitat.

- Establir correlacions genotip-fenotip.

- Estudiar el possible efecte de les mutacions del gen *GLB1* sobre les interaccions de la β-galactosidasa amb altres proteïnes del complex LMC.

- Analitzar els mecanismes que regulen el procés d'*splicing* alternatiu del gen *GLB1* mitjançant l'ús de minigens, així com l'efecte de l'NMD en aquest procés.

Resultats

Informe sobre la contribució del doctorand en els resultats d'aquesta Tesi Doctoral:

CAPÍTOL 1

Article 1: *Human Mutation* 27: 1060 (2006).

Autors: Raül Santamaria, Amparo Chabás, Maria Josep Coll, Clara Sa Miranda, Lluïsa Vilageliu i Daniel Grinberg.

La feina presentada en aquest article ha estat bàsicament realitzada pel doctorand a partir d'un grup de pacients diagnosticats bioquímicament a l'Institut de Bioquímica Clínica (Barcelona, Espanya) pels grups de les doctores Amparo Chabás i Maria Josep Coll i a l'Instituto de Genética Médica Jacinto de Magalhães (Porto, Portugal) pel grup de la Dra. Clara Sa Miranda. Igualment, les dades clíniques del pacients presentades a l'article van ser facilitades pels metges que s'indiquen als agraïments.

Article 2: *Clinical Genetics* 71: 273–279 (2007).

Autors: Raül Santamaria, Mariana Blanco, Amparo Chabás, Daniel Grinberg i Lluïsa Vilageliu.

La feina presentada en aquest article va ser realitzada pel doctorand a partir d'un grup de pacients diagnosticats bioquímicament al Laboratorio de Neuroquímica Dr N. A. Chamoles (Buenos Aires, Argentina) pel grup de la Dra. Mariana Blanco i un pacient diagnosticat a l'Institut de Bioquímica Clínica (Barcelona, Espanya) pel grup de la Dra. Amparo Chabás.

CAPÍTOL 2

Article 3: Article en preparació.

Autors: Raül Santamaria, Amparo Chabás, John W. Callahan, Daniel Grinberg i Lluïsa Vilageliu.

La major part de la feina presentada en aquest article va ser realitzada pel doctorand. L'anàlisi de les activitats enzimàtiques de l'expressió de la proteïna van ser realitzades a l'Institut de Bioquímica Clínica (Barcelona, Espanya) al grup de la Dra. Amparo Chabás. També les mostres de cèl·lules de pacients amb les que es van realitzar els estudis de Western Blot van ser obtingudes de l'Institut de Bioquímica Clínica. Com s'indica a l'article, l'anticòs emprat va ser cedit pel Dr. J. Callahan.

Article 4: Sotmès a *Clinical Genetics*.

Autors: Laura Gort, Raül Santamaria, Daniel Grinberg, Lluïsa Vilageliu i Amparo Chabás.

En aquest article el doctorand va realitzar l'anàlisi mutacional del pacient amb Gangliosidosi GM1 presentat, així com la part relacionada amb l'expressió *in vitro* de la variant presentada (p.R595W). Això inclou el clonatge del cDNA del gen, la mutagènesi dirigida i la transfecció de les cèl·lules COS.

CAPÍTOL 3

Article 5: Article en preparació.

Autors: Raül Santamaria, Amparo Chabás, Lluïsa Vilageliu i Daniel Grinberg.

La totalitat de la feina presentada en aquest article va ser realitzada pel doctorand. Com s'indica a l'article, els plasmidis amb les proteïnes SR clonades emprats van ser cedits pel Dr. A. Kornblihtt.

CAPÍTOL 1

Estudi mutacional de diferents poblacions de malalts amb Gangliosidosi GM1 i malaltia de Morquio B

1.1 Estudi mutacional de pacients espanyols amb Gangliosidosi GM1 i malaltia de Morquio B

Fins al moment de començar aquest treball s'havien descrit poc més de 60 mutacions diferents causants de la gangliosidosi GM1 i de la malaltia de Morquio B. Per altra banda, les mutacions presents al gen *GLB1* en la població espanyola no havien estat mai caracteritzades.

Per això, davant de la disponibilitat de mostres de 30 pacients amb gangliosidosi GM1 i 5 amb Morquio B, es va realitzar l'anàlisi mutacional de tots ells i es van poder caracteritzar els 70 al.lels mutats. Es van identificar fins a 30 mutacions diferents, de les que 21 no havien estat descrites prèviament. Aquests al.lels mutats incloïen 20 mutacions de sentit erroni, 4 mutacions sense sentit, 2 insercions i 4 mutacions que afectaven el correcte *splicing*. Les diferents mutacions d'*splicing* van ser estudiades a nivell de l'RNA per conéixer de quina manera alteraven el procés d'*splicing*.

Entre les noves mutacions trobades, les mutacions p.L264S, p.T420K i p.Y83C es van poder associar a les formes juvenil (tipus II), adulta (tipus III) i Morquio B, respectivament.

A més a més, als cromosomes portadors de la mutació p.R59H es va poder fer una anàlisi d'haplotips mitjançant 5 marcadors polimòrfics que va posar de manifest l'existència d'un haplotip conservat en el grup ètnic dels gitanos. Això indica un possible efecte fundador d'aquesta mutació en aquest grup ètnic.

Treball publicat:

"Twenty-one novel mutations in the GLB1 gene identified in a large group of GM1-gangliosidosis and Morquio B patients: possible common origin for the prevalent p.R59H mutation among Gypsies." *Human Mutation: 27: 1060 (2006).*

1.2 Estudi mutacional de pacients Sudamericans, principalment Argentins, amb Gangliosidosi GM1

Ja que podíem disposar de mostres de pacients amb Gangliosidosi GM1 de Sudamèrica (majoritàriament Argentina) es va realitzar l'anàlisi mutacional del gen *GLB1* en aquesta població. Es van analitzar 19 pacients i es van identificar les 38 mutacions causants de la malaltia. En aquesta ocasió es van trobar 22 mutacions diferents de les que 14 no havien estat descrites prèviament. De les 14 noves mutacions descrites, 7 eren de sentit erroni (p.G134V, p.L155R, p.L162S, p.S434L, p.D491Y, p.P549L i p.G554E), una era una mutació sense sentit (p.W576X), una afectava un lloc donador d'*splicing* (c.1068+1G>T) i les 5 restants eren delecions.

Trobar 5 delecions va ser sorprenent perquè fins al moment únicament s'havia descrit una deleció al gen *GLB1*. D'aquestes delecions, dues eren delecions d'un sol nucleòtid (c.845_846delC i c.1706_1707delC), una era de 3 nucleòtids (c.435_440delTCT), una de 15 nucleòtids (c.1131_1145del15) i, finalment, una va resultar ser una deleció de 1529 nucleòtids que incloïa l'exó 5 deguda a un entrecreuament desigual entre seqüències *Alu* flanquejants d'aquest exó (c.458-401_552+1033del1529).

En aquest treball també es van poder establir algunes correlacions genotip-fenotip, tals com l'associació de la nova mutació p.L155R a la forma adulta de la malaltia, o les mutacions p.S434L i/o p.G554E que van poder ser associades a la forma juvenil.

Treball publicat:
"Identification of 14 novel *GLB1* mutations, including five deletions, in 19 GM1-gangliosidosis patients from South-America." *Clinical Genetics 71: 273–279 (2007).*

CAPÍTOL 2

Anàlisi funcional de variants proteiques de la β-galactosidasa

2.1 Anàlisi funcional de variants mutades del gen *GLB1* trobades a pacients amb Gangliosidosi GM1 o malaltia de Morquio B

Les noves variants identificades als pacients amb Gangliosidosi GM1 o Morquio B, tot i no haver estat detectades en 100 cromosomes d'individus sans, requerien altres comprovacions per demostrar la seva patogenicitat. Una comprovació consisteix en la seva expressió mitjançant un sistema *in vitro* i la caracterització de les proteïnes mutades resultants a nivell de l'activitat enzimàtica residual, així com l'anàlisi del patró obtingut per *Western blot*.

En aquest treball es van expressar 14 variants diferents amb el sistema de cèl.lules COS-7 transfectades amb Lipofectamina. Es va poder demostrar la patogenicitat de 12 d'elles. A més a més, dues variants (p.R521C i p.S532G) presentaven activitats residuals suficientment elevades com per a considerar-les polimorfismes. El canvi p.R521C havia estat descrit prèviament com un polimorfisme i, en un cas, com una mutació patogènica. A l'article es mostra com l'activitat residual trobada per nosaltres (33,2%) el fa compatible amb ser una mutació amb penetrància molt baixa.

També es van realitzar diferents experiments per analitzar l'efecte que les mutacions del gen *GLB1* identificades en els pacients tenien sobre les altres proteïnes del complex multienzimàtic lisosòmic (LMC). Es van dur a terme CoIP de les variants mutades de la β-gal amb l'anticòs anti-neuraminidasa. També es van realitzar diferents transferències *Western* utilitzant extractes proteics procedents de fibroblasts de pacients i anticossos anti-β-gal, anti-Neuraminidasa o anti-PPCA. Els resultats demostren una alteració del LMC en els extractes procedents dels fibroblasts dels pacients, mentre que no s'observa cap efecte clar en el cas de les CoIP.

Treball en preparació:
"Functional Analyses of 14 *GLB1* Mutant Alleles Found in GM1-gangliosidosis and Morquio B patients: Effect on the Lysosomal Multienzyme Complex ."

Functional Analyses of 14 *GLB1* Mutant Alleles Found in GM1-gangliosidosis and Morquio B patients: Effect on the Lysosomal Multienzyme Complex

Raül Santamaria[1,2,3], Amparo Chabás[2,4], John W Callahan[5], Daniel Grinberg [1,2,3], Lluïsa Vilageliu [1,2,3]

[1]Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona

[2]Centre for Biomedical Research on Rare Diseases (CIBERER), ISCIII, Barcelona

[3]Institut de Biomedicina de la Universitat de Barcelona (IBUB)

[4]Institut de Bioquímica Clínica, Hospital Clínic, Barcelona, Spain

[5]Department of Pediatric Laboratory Medicine and Biochemistry, The Hospital for Sick Children, University of Toronto, Toronto, Canada

ABSTRACT

GM1-gangliosidosis and Morquio B disease are lysosomal storage disorders caused by β-galactosidase deficiency due to mutations in the *GLB1* gene. Upon arriving to the endosomal-lysosomal compartment the β-galactosidase protein associates with the PPCA and neuraminidase proteins to form the Lysosomal Multienzyme Complex (LMC). The correct interaction of these proteins in the complex is essential for their activity.

More than one hundred mutations have been described in GM1-gangliosidosis and Morquio B patients, but few have been further characterized. We expressed 12 mutations suspected to be pathogenic, one known polymorphic change (p.S532G) and a variant controversially described both as a pathogenic or as a polymorphic change (p.R521C). Ten of them had not been expressed before. The expression analysis confirmed the pathogenicity of the 12 mutations while the relatively high activity of p.S532G is consistent with its definition as a polymorphism. The results for p.R521C suggest that this change could be considered as a low-penetrant disease-causing allele.

Furthermore, the effect of these β-galactosidase changes in the LMC was also studied by coimmunoprecipitations and Western blotting. The alteration of neuraminidase and PPCA patterns in several of the Western blotting analyses performed on patients' protein extracts indicated that the LMC is affected in at least some GM1-gangliosidosis and Morquio B disease patients.

KEYWORDS

β-galactosidase; GM1-gangliosidosis; heterologous expression; Lysosomal Multienzyme Complex (LMC)

INTRODUCTION

The deficiency of lysosomal β-galactosidase (β-gal) due to mutations in the *GLB1* gene is the cause of the rare lysosomal storage disorders GM1-gangliosidosis (MIM# 230500) and Morquio B disease (MIM# 253010). Natural substrates for β-galactosidase are ganglioside GM1, keratan sulphate and different glycopeptides, which accumulate in the patients. Three major clinical forms according to age of onset and severity of symptoms have been established in GM1-gangliosidosis patients: type I (infantile form), type II (late infantile/juvenile form) and type III (adult form). The different clinical forms are mainly due to different residual activities of the mutant enzymes and, thus, to the different level of substrate accumulation in tissues. Residual β-galactosidase activity in fibroblasts from patients, measured using the artificial 4-methylumbelliferyl β-galactopyranoside (MUB) substrate, varies from 0.07-1.3% of control values in infantile patients, 0.3-4.8% in the juvenile form and up to 9% in the adults. Morquio B patients have no neurological involvement but display severe skeletal dysostosis multiplex due to high accumulation of keratan sulphate (Okumiya, et al., 2003).

The human *GLB1* gene, localized to chromosome 3 at 3p21.33 (Takano and Yamanouchi, 1993), contains 16 exons that give rise to two alternatively spliced mRNAs: a major transcript of 2.5 kb which encodes the β-galactosidase and a minor transcript of 2 kb encoding the elastin-binding protein (EBP). The β-galactosidase protein is targeted to the lysosome as a 88 kDa polypeptide that is then proteolytically processed to the mature 64 kDa form by the protective protein/cathepsin A (PPCA). EBP is a major component of the non-integrin cell surface elastin receptor complex, in which EBP binds and protects tropoelastin (Privitera, et al., 1998). This complex consists of 3 units: EBP, neuraminidase and PPCA. The β-galactosidase protein forms also a complex, the Lysosomal Multienzyme Complex (LMC): upon arriving to the endosomal-lysosomal compartment the enzyme associates with the pre-associated PPCA and neuraminidase. This complex also includes the *N*-acetylgalactosamine-6-sulphate sulphatase enzyme. The correct interaction of these proteins in the complex is essential for their correct activity (Pshezhetsky and Ashmarina, 2001). Within the LMC, PPCA has not only a protective role for both glycosidases against lysosomal proteolysis but it is also necessary for the correct maturation of β-galactosidase to the 64 kDa form (Bonten and d'Azzo, 2000; Morreau, et al., 1992; Rudenko, et al., 1995).

Our group has recently described mutations in more than 40 GM1-gangliosidosis and Morquio B patients (Santamaria, et al., 2007; Santamaria, et al., 2006) most of them missense. Although their absence in 100 control chromosomes strongly suggests that they are disease-causing mutations, the functional analysis presented here has confirmed their pathogenicity. *In vitro* expression in COS-7 cells followed by the enzymatic activity analyses were performed for 12 mutations, a functional polymorphism (p.S532G) and a change for which the pathogenic status was controversial (p.R521C). Ten of these changes had not been expressed before. While the 12 mutations caused a complete absence or a great reduction of the enzyme activity, p.S532G and p.R521C showed only a partial decrease of the activity. In most cases a good correlation between residual activity and phenotype could be established. In some cases, an effect of the β-galactosidase alterations on other members of the LMC was observed.

MATERIALS AND METHODS

Patients

The mutant alleles, samples and cells used in this work correspond to GM1-gangliosidosis or Morquio B patients already described by our group (Santamaria, et al., 2007; Santamaria, et al., 2006). Genotypes and clinical types of some representative patients are listed in Table 1.

Vector Construction

The whole coding region of the *GLB1* cDNA was PCR-amplified in 2 fragments that were ligated and cloned in a pUC18 vector. Mutagenesis was always performed in the *GLB1*-pUC plasmid. For protein expression, the wild-type and mutated cDNAs were subcloned into the pcDNA3.1 expression vector (p.BGALwt and p.BGALmut, respectively). Coding region of human PPCA and human neuraminidase-1 (or sialidase) were also cloned using the same procedure (p.NEU and p.PPCA, respectively).

Site-directed Mutagenesis

All mutations were generated by site-directed mutagenesis using the *QuickChange™ Site-Directed Mutagenesis XL kit* (Stratagene, La Jolla, CA) following the manufacturer's instructions. In the case of the double mutation p.L162S+p.R521C, the p.R521C mutation was generated in a vector which already bore the p.L162S mutation. All constructs were resequenced to ensure that no spurious mutation had been introduced.

Cell Culture and Transfection

COS-7 cells were cultured in 100-mm diameter tissue culture dishes with DMEM medium (GIBCO, BRL Grand Island, NY, USA), 10% fetal bovine serum (GIBCO, BRL Grand Island, NY, USA) and antibiotics. For transfection with wild-type and mutant β-galactosidase cDNAs, $30x10^5$ cells were plated. Twenty-four hours later (when the cells were at 90% of confluency), 2 μg of corresponding plasmid mixed with 15 μl of *Lipofectamine™ 2000 Reagent* (Invitrogen, Carlsbad, CA) was added. As a negative control, an empty pcDNA3.1 vector and/or a pcDNA vector carrying an antisense β-galactosidase cDNA was transfected. Cells were collected 48 hours after transfection by scraping and centrifugating for 5 min at 300 g. Cellular pellets were washed twice with phosphate-buffered saline (PBS). About a 10% of the volume was set aside for Western blotting. The remaining cellular pellet was stored at -80ºC until the enzymatic analysis was performed.

Patients' fibroblasts were cultured as previously described (Santamaria, et al., 2006) and whole protein extract was obtained by scraping the cells in lysis buffer (1% Triton X-100, 10% glycerol, 20 mM Sodium acetate and 150 mM NaCl) containing the protease inhibitor cocktail *Complete Mini, EDTA free* (Roche, Basel, Switzerland).

Enzymatic Analysis

β-galactosidase activity was measured with the fluorogenic substrate 4–methylumbelliferyl- β-D- galactopyranoside (MUB).

SDS-PAGE and Western Blot Analysis

Proteins were subjected to SDS-PAGE (12.5% polyacrilamide) and electrophoretically transferred onto nitrocellulose membranes. For transfected cells, about 30 µg of protein extract were loaded, whereas for patients' fibroblasts, about 40 µg were loaded. Western blots were performed as previously described (Montfort, et al., 2004). Polyclonal anti-human β-galactosidase was kindly provided by Dr. Callahan. Anti-human PPCA and anti-human neuraminidase-1 antibodies were purchased at Rockland (Gilberstville, PA). Immunoreactive bands were detected by incubating the membrane for 2 min in the following solution: 10 ml 100 mM Tris-HCl pH 9, 50 µl 45 mM p-Coumaric Acid, 50 µl luminol and 10 µl 30% H_20_2.

Coimmunoprecipitation (CoIP)

The interaction between β-galactosidase and neuraminidase-1 was studied by CoIP. COS-7 cells were cotransfected with 2 µg of each plasmid (p.BGALwt/p.BGALmut, p.PPCA and p.NEU) which were mixed with 20 µl of *Lipofectamine™ 2000 Reagent* in the conditions described above. Correct expression of each protein in each sample was checked by Western blotting (data not shown). CoIP was performed with *Rabbit IgG Trueblot™ Set* (eBioscience, Boston, MA), according to manufacturer's recommendations except for the following changes. Cell lysate was quantified and 200 µg of total protein extract was used for each CoIP with 6 µl of commercial anti-neuraminidase antibody. The cell lysis buffer contained 1% Triton X-100, 10% glycerol, 20 mM Sodium acetate and 150 mM NaCl, while the washing buffer contained 0.01% NP-40, 10% glycerol, 20 mM sodium acetate and 150 mM NaCl. CoIP was performed at pH 5.5, as previously described (van der Spoel, et al., 1998). CoIP samples were subjected to Western blotting and detected with the specific primary antibody and using the *Rabbit IgG Trueblot™* as secondary antibody, to avoid detection of the immunoprecipitated antibody, following manufacturer's recommendations.

Statistical Analysis

For each expression experiment, transfections were performed three times for each mutation as well as for the wild-type construct and for an empty vector used as negative control. Activity of mutant alleles was expressed as percentage of the wild type enzymatic activity. In both cases, mutant and the wild-type, the negative control activity was subtracted. For each mutant allele, at least two independent experiments, as those described above, were carried out. The U de Mann-Whitney test was used to analyze significant differences in enzymatic activities between pairs of mutant and/or wild-type enzymes.

RESULTS

Expression and Enzymatic Activity of Wild-Type and Mutant β-galactosidase in COS-7 Cells

Different alleles of the *GLB1* gene were expressed in COS-7 cells in order to establish their residual enzymatic activities. The expressed changes were: c.176G>A (p.R59H), c.248A>G (p.Y83C), c.485T>C (p.L162S), c.518T>C (p.L173P), c.602G>A (p.R201H), c.1258A>C (p.T420P), c.1259C>A (p.T420K), c.1321G>A (p.D441N), c.1331A>G (p.Y444C), c.1480G>A (p.G494S), c.1561C>T (p.R521C), c.1594A>G (p.S532G), c.1768C>T (p.R590C) and the double mutant c.[485T>C;1561C>T] (p.[L162S;p.R521C]). The average expressed activity of wild-type β-galactosidase in COS-7 cells was 2009.6 ± 533.4 nmol/hr/mg, whereas the average endogenous activity (COS-7 cells transfected with an empty pcDNA3) was 408 ± 94.7 nmol/hr/mg. The residual enzymatic activity for each protein, given as percentage of the wild-type activity, is shown in Figure 1. The following changes resulted in lack of activity: p.R59H, p.L162S, p.L173P, p.T420P, p.D441N, p.Y444C, p.G494S, p.R590C, p.[L162S;R521C]; whereas p.Y83C, p.R201H, p.T420K, p.R521C and p.S532G produced enzymes with 6.9%, 15.9%, 10.3%, 33.2% and 60.2% of wild-type activity, respectively. Table 2 shows the statistical analysis between the activity values of all pairs of mutant enzymes with detectable levels of activity and between each of them and the wild-type protein. All changes displayed residual enzymatic activities statistically different from the wild type enzyme. No significant differences were detected for all pair-wise comparisons for mutations p.Y83C, p.R201H and p.T420K. The change p.R521C, previously described both as a polymorphism (Silva, et al., 1999) and as a mutation (Caciotti, et al., 2005), differs statistically from all other mutations but also from the p.S532G polymorphism, which displays the highest residual activity.

Western Blot Analysis of Expressed Changes

To confirm that a correct synthesis of the proteins from the transfected construct took place in COS-7 cells, particularly in those cases were absence of activity was found, immunoblots of the normal and the mutant proteins were performed. A unique band of approximately 84 kDa was detected in all cases using a polyclonal anti-human β-galactosidase antibody. This band corresponds to the precursor protein, indicating a lack of mature protein in COS-7 cells, as previously suggested (Morreau, et al., 1989; Oshima, et al.,

1988). In any case, this maturation seems not to be necessary to detect the *in vitro* activity of the enzyme. The 84-kDa band was observed in all the expressed alleles (Fig. 2).

Coimmunoprecipitation

Association of β-galactosidase with other proteins in COS cells has been previously reported (van der Spoel, et al., 1998). Here we coexpressed the wild-type and mutant β-gal constructs with neuraminidase and PPCA, and CoIP β-gal with an anti-neuraminidase antibody to check if the mutations found in the patients can alter the formation of this complex. The results are shown in Figure 3. Controls confirmed that the observed β-gal band came from the specific co-immunoprecipitation, since no band was observed neither when the anti-neuraminidase antibody was not included (-), nor when the pNEU (-NEU) or the pBGALwt (-βGAL) plasmid was not co-expressed. For all the mutations, a band of 64 kDa could be detected, indicating that in these conditions at least some amount of protein remains associated with the neuraminidase, in the complex. However, differences in the intensity are clearly observed which suggest that some of the mutations could affect the interaction ability of the enzyme.

Western Blot Analysis of Patients' Fibroblasts

To see the effect of mutations on the proteins of the LMC, Western blots were performed using antibodies against human β-galactosidase, neuraminidase and PPCA on whole protein extracts from fibroblasts of different patients (Fig. 4). Western blots were repeated at least twice. Patients' characteristics are listed on Table 1.

For β-galactosidase, the wild-type sample gave rise to a main band of 64 kDa corresponding to the mature form of the protein. A faint band is also observed, which corresponds to the 84-kDa immature protein. All the patient samples showed the 64-kDa band, although for patient GM4, the 64-kDa band seems to be fainter than the rest. Most samples showed a weak 84-kDa immature band (except for GM8) and some samples also showed smaller bands, which may correspond to degradation products.

The Western blot detected with anti-Neuraminidase antibodies reveals the previously described pattern of 2 bands due to different glycosylation states of the enzyme (Bonten, et al., 1996; Milner, et al., 1997). All patients display this pattern except for GM4, for whom the neuraminidase bands were nearly undetectable.

Finally, the Western blot for PPCA confirms the ability of the anti-PPCA antibody to recognize the immature form of the enzyme (54 kDa) as well as the two mature bands (32

and 20 kDa). In most of the samples, a band of a higher MW (above 100 kDa) was observed. It could correspond to a not completely denatured multimeric complex. Interestingly, different band patterns were observed for some of the patients. The 32-kDa band is not observed in GM1 and MB3 samples. Patient GM4 seems to lack the 20-kDa band while patient MB4 shows a reduced amount of the 54-kDa immature protein.

DISCUSSION

More than 80 mutations causing GM1-gangliosidosis or Morquio B disease have been described in the *GLB1* gene but few of them have been expressed *in vitro* in order to establish their residual enzymatic activity. This is necessary because *in vivo* residual activities found in patients' fibroblasts or leukocytes do not always correlate well with the observed phenotypes (as shown in Table 1). Moreover, most of the patients are compound heterozygotes making it difficult to know the effect of each mutation individually.

COS cells have been widely and successfully used to express different lysosomal enzymes and, in particular, expression of lysosomal β-galactosidase in COS cells has been previously reported. For this reason we decided to use this expression system to study the residual enzyme activities of different mutant alleles found in our series of patients.

Table 3 shows the activities of previously expressed mutations found in the same codons of the mutations that we have expressed, which results are also included. Our results are consistent with previous data, in spite of the different expression systems used. It should be noted, however, that few of these studies actually demonstrated the production of expressed protein. The only case for which our results are different from those previously described is the one of p.R201H. Studies by two groups reported very different results and we obtained an intermediate value. This mutation was found in patients with a relatively mild phenotype, which makes the result by Kaye et al. (1997) difficult to understand. These authors reported a very low value for the wild type enzyme, which may explain the null result for the mutant construct. In fact, as they did not demonstrate the production of expressed protein, there is no way to correlate the low level of enzyme activity obtained with the level of protein synthesized. Our results for p.R201H activity (and those of Ishii et al., 1995b) are higher than those described for the p.R201C mutation, which is consistent with the phenotypes associated with these two genetic changes.

The case of p.R521C change is particularly interesting because it was originally described as a polymorphism and found in 4% of a Brazilian control population (Silva, et al., 1999) and in 1% of Spanish control individuals (Santamaria, et al., 2006). Recently, Caciotti et al. (2005) described a patient homozygous for this change and, after expressing it in COS-1 cells, they reported that it was likely to be the disease-causing mutation. In our case, the change p.R521C has resulted in a 33.2% of residual activity. This activity is statistically different from the other polymorphisms but also different from all the other disease-causing

mutations (see Table 2). This result could explain the apparent paradox that 4% of the Brazilian population bears the p.R521C mutation but no patients with such change have been reported to date as the cause of the disease in that population. Individuals bearing p.R521C, an allele showing a relatively high residual activity, could often remain assymptomatic. Only few cases, like the one reported by Caciotti et al., probably due to particular differences in genetic background or environmental factors, will develop the disease. Thus, the p.R521C change could be considered as a low penetrant mutation. A fact that supports this idea is that, the p.T420K mutation, which in our experiments has shown 10.3% of residual activity, was found in homozygosity in patient GM20, a mild adult patient. This patient had the first symptoms of the disease at age of 26 years and a mild course of the disease. Accordingly, a mutant allele, such as p.R521C, with 3-fold higher activity only would be pathogenic in a particular context. In addition, the p.S532G polymorphic change was also expressed. This allele resulted in 60.2% of activity, confirming this is a non-pathogenic change.

In general, in the present study, good correlations could be established between the expressed mutations and the phenotypes of the patients that bore them. The changes p.R59H, p.T420P, p.D441N, p.R590C and p.[L162S;R521C], found in infantile patients, had null residual activity. The mutations found in adult patients, p.R201H and p.T420K, had 15.9%, 10.3% of residual activity, respectively. Mutations found in Morquio B are expected to have some residual activity, because they can barely cleave terminal β-galactosyl residue of keratan sulfate, but keep some enzymatic activity against the ganglioside GM1 (Okumiya, et al., 2003). This is true for p.Y83C which shows 6.9% of activity. The p.L173P mutation, found in 2 Morquio B patients, had a null enzymatic activity, suggesting that the accompanying mutation (p.T500A) should be the Morquio B causing allele, as previously reported (Santamaria, et al., 2006).

By contrast, it was difficult to establish any correlation for mutations p.Y444C and p.G494S found in another Morquio B patient. It was expected that at least one of them would have shown some residual activity. However, p.G494S showed null activity and, although p.Y444C had a residual activity different than 0 (0.65%), this was too low to confirm it as the Morquio B causing change. Further cases should be analyzed to establish this correlation.

The differences of residual activity found for all the expressed amino acid changes cannot be attributed neither to the lack of correct synthesis nor to a complete protein degradation, as shown by Western blot analyses (Fig. 2). However, this system does not

allow determining the cause of this absence of activity of the mutant proteins, which could be either a direct effect on the catalytic site or a misfolding and/or erroneous trafficking of the protein. The misfolding/trafficking hypothesis seems to be more probable since all the mutations expressed in the present work affect residues that are far from those described as essential for the catalytic process: Glu268 (McCarter, et al., 1997) and Asp 332 (Zhang, et al., 2000).

The possible effect of the β-gal mutations and polymorphisms on the interactions with the other proteins within the LMC was studied by coimmunoprecipitating β-gal with anti-neuraminidase antibody (Fig. 3). Apparently all of the changes allow some level of interaction between β-gal and neuraminidase. The differences in the intensity suggest that some of the mutations could affect the interaction ability of the enzyme.

A complementary approach to study the effect of the mutations on the protein interactions within the LMC consisted in the analysis by Western blots of protein extracts from patients' fibroblasts (Fig. 4), using antibodies against either β-gal or neuraminidase or PPCA.

The results for β-gal revealed that mutations on patients GM4 and GM8 surprinsingly allowed the synthesis of a β-gal protein of about 64 kDa. This is surprising because these patients both carry either insertions or mutations affecting the position +1 of a donor splice site, which would be expected to give rise to proteins with different MW. In fact, these patients, as indicated in Table 1, also showed some residual enzymatic activity in their fibroblasts, especially GM4. In the case of GM8, both mutations (c.245+1G>A /c.1572_1577insG) are expected to cause nonsense-mediated mRNA decay (NMD), but both alleles were amplified by RT-PCR (data not shown). In particular, the c.1572_1577insG allele, if not degraded, is expected to give rise to a truncated protein of about 64 kDa (similar in size to the mature form of the wild type protein). The truncated protein corresponding to the other allele (with exon 2 skipping) would be very small if synthesized. The exon skipping caused by the mutation present in homozygosity in patient GM4 does not generate a frameshift. However, the absence of exon 14 implies a reduction of about 4.8 kDa in the protein MW, which does not correspond to any of the bands in the Western blot. An altered maturation and/or a modified post-translational process could be the explanation for the observed bands.

Western blots performed with anti-neuraminidase and anti-PPCA antibodies showed that in some patients the band patterns for these two proteins were also altered. Patient GM4

showed a clear reduction of neuraminidase bands intensity and a lack of the 20 kDa PPCA band, whereas for patients GM1 and MB3 the 32 kDa PPCA band is not observed. Compared to the rest of the bands, patient MB4 also displays a great reduction of the 54 kDa PPCA band intensity. The altered band patterns observed for the PPCA and neuraminidase assays indicate that some mutations in the β-gal protein affect the stability of the other proteins in the LMC. In fact, most GM1-gangliosidosis patients also show a reduced neuraminidase activity (5%-18% of control mean), although higher than in sialidosis or galactosialidosis patients which is another sign of this LMC alteration.

REFERENCES

Bonten E, van der Spoel A, Fornerod M, Grosveld G, d'Azzo A. 1996. Characterization of human lysosomal neuraminidase defines the molecular basis of the metabolic storage disorder sialidosis. Genes Dev 10:3156-69.

Bonten EJ, d'Azzo A. 2000. Lysosomal neuraminidase. Catalytic activation in insect cells is controlled by the protective protein/cathepsin A. J Biol Chem 275:37657-63.

Boustany RM, Qian WH, Suzuki K. 1993. Mutations in acid beta-galactosidase cause GM1-gangliosidosis in American patients. Am J Hum Genet 53:881-8.

Caciotti A, Bardelli T, Cunningham J, D'Azzo A, Zammarchi E, Morrone A. 2003. Modulating action of the new polymorphism L436F detected in the GLB1 gene of a type-II GM1 gangliosidosis patient. Hum Genet 113:44-50.

Caciotti A, Donati MA, Boneh A, d'Azzo A, Federico A, Parini R, Antuzzi D, Bardelli T, Nosi D, Kimonis V and others. 2005. Role of beta-galactosidase and elastin binding protein in lysosomal and nonlysosomal complexes of patients with GM1-gangliosidosis. Hum Mutat 25:285-92.

Ishii N, Oohira T, Oshima A, Sakuraba H, Endo F, Matsuda I, Sukegawa K, Orii T, Suzuki Y. 1995a. Clinical and molecular analysis of a Japanese boy with Morquio B disease. Clin Genet 48:103-8.

Ishii N, Oshima A, Sakuraba H, Osawa M, Suzuki Y. 1995b. Beta-galactosidosis (genetic beta-galactosidase deficiency): Clinical and genetic heterogeneity of the skeletal form. Dev Brain Dysfunct 8:40-50.

Kaye EM, Shalish C, Livermore J, Taylor HA, Stevenson RE, Breakefield XO. 1997. beta-Galactosidase gene mutations in patients with slowly progressive GM1 gangliosidosis. J Child Neurol 12:242-7.

McCarter JD, Burgoyne DL, Miao S, Zhang S, Callahan JW, Withers SG. 1997. Identification of Glu-268 as the catalytic nucleophile of human lysosomal beta-galactosidase precursor by mass spectrometry. J Biol Chem 272(1):396-400.

Milner CM, Smith SV, Carrillo MB, Taylor GL, Hollinshead M, Campbell RD. 1997. Identification of a sialidase encoded in the human major histocompatibility complex. J Biol Chem 272:4549-58.

Montfort M, Chabas A, Vilageliu L, Grinberg D. 2004. Functional analysis of 13 GBA mutant alleles identified in Gaucher disease patients: Pathogenic changes and "modifier" polymorphisms. Hum Mutat 23:567-75.

Morreau H, Galjart NJ, Gillemans N, Willemsen R, van der Horst GT, d'Azzo A. 1989. Alternative splicing of beta-galactosidase mRNA generates the classic lysosomal enzyme and a beta-galactosidase-related protein. J Biol Chem 264:20655-63.

Morreau H, Galjart NJ, Willemsen R, Gillemans N, Zhou XY, d'Azzo A. 1992. Human lysosomal protective protein. Glycosylation, intracellular transport, and association with beta-galactosidase in the endoplasmic reticulum. J Biol Chem 267:17949-56.

Nishimoto J, Nanba E, Inui K, Okada S, Suzuki K. 1991. GM1-gangliosidosis (genetic beta-galactosidase deficiency): identification of four mutations in different clinical phenotypes among Japanese patients. Am J Hum Genet 49:566-74.

Okumiya T, Sakuraba H, Kase R, Sugiura T. 2003. Imbalanced substrate specificity of mutant beta-galactosidase in patients with Morquio B disease. Mol Genet Metab 78:51-8.

Oshima A, Tsuji A, Nagao Y, Sakuraba H, Suzuki Y. 1988. Cloning, sequencing, and expression of cDNA for human beta-galactosidase. Biochem Biophys Res Commun 157:238-44.

Oshima A, Yoshida K, Itoh K, Kase R, Sakuraba H, Suzuki Y. 1994. Intracellular processing and maturation of mutant gene products in hereditary beta-galactosidase deficiency (beta-galactosidosis). Hum Genet 93:109-14.

Privitera S, Prody CA, Callahan JW, Hinek A. 1998. The 67-kDa enzymatically inactive alternatively spliced variant of beta-galactosidase is identical to the elastin/laminin-binding protein. J Biol Chem 273:6319-26.

Pshezhetsky AV, Ashmarina M. 2001. Lysosomal multienzyme complex: biochemistry, genetics, and molecular pathophysiology. Prog Nucleic Acid Res Mol Biol 69:81-114.

Rudenko G, Bonten E, d'Azzo A, Hol WG. 1995. Three-dimensional structure of the human 'protective protein': structure of the precursor form suggests a complex activation mechanism. Structure 3:1249-59.

Santamaria R, Blanco M, Chabas A, Grinberg D, Vilageliu L. 2007. Identification of 14 novel GLB1 mutations, including five deletions, in 19 patients with GM1 gangliosidosis from South America. Clin Genet 71(3):273-9.

Santamaria R, Chabas A, Coll MJ, Miranda CS, Vilageliu L, Grinberg D. 2006b

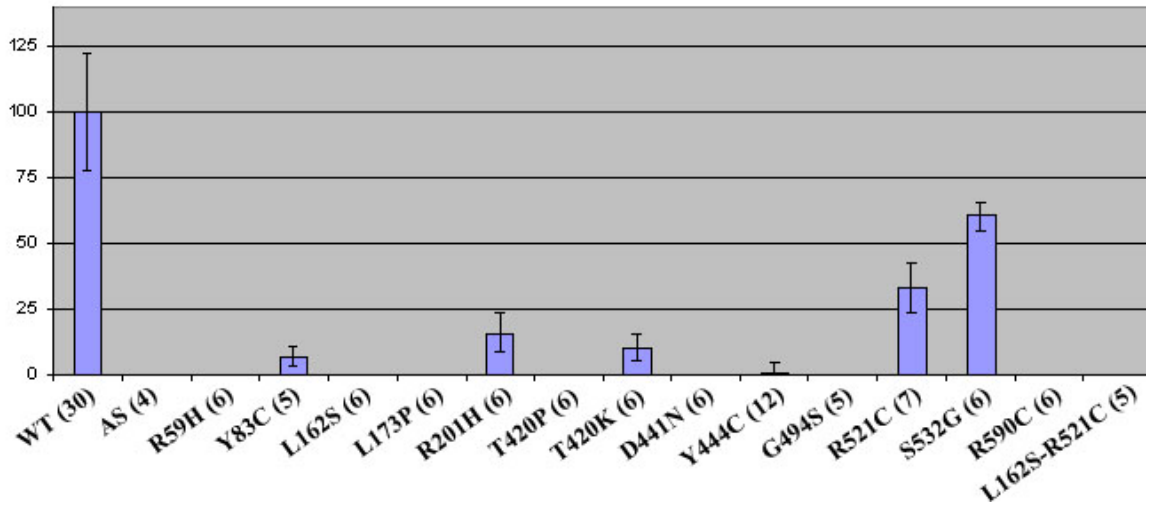


FIGURE 1. Residual β-galactosidase enzymatic activity of the proteins from mutant alleles, expressed as the percentage of the wild-type activity. Number of replicas for activity measurements is indicated in brackets. Bars correspond to standard deviation.

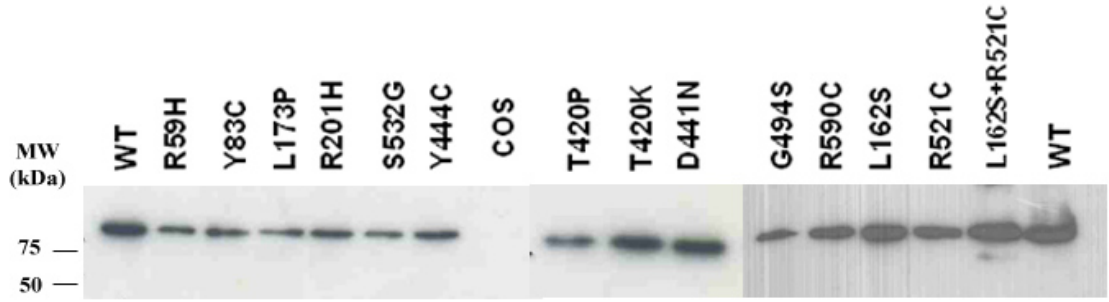


FIGURE 2. Western blotting of expressed mutations in COS cells. COS lane are the cells transfected with an empty pcDNA3.1 plasmid. 30 µg of protein extract were loaded in each lane.

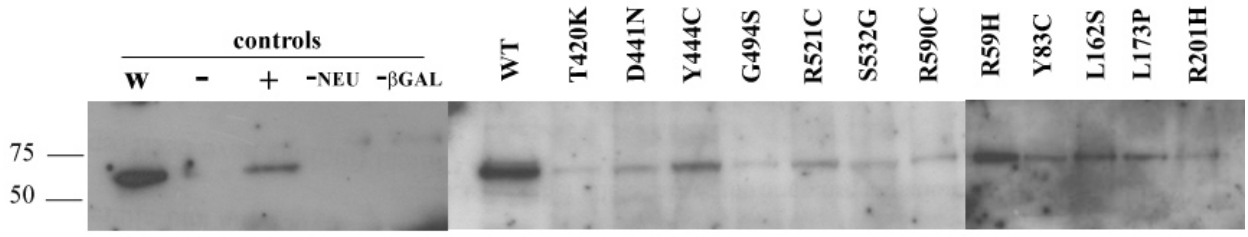


FIGURE 3. Coimmunoprecipitation. Figures shown are Western blots using anti-β-gal antibody. Controls are: **W**:Western control: whole protein extract of coexpressed β-gal, PPCA and neuraminidase; **—**:CoIP without anti-neuraminidase antibody; +: Positive control: CoIP of coexpressed β-gal, PPCA and neuraminidase, using anti-neuraminidase antibody; **–NEU**: CoIP of coexpressed β-gal and PPCA without expressed NEU; **–β-gal**: CoIP of coexpressed NEU and PPCA without expressed β-gal. On the right are the different mutations coexpressed and coimmunoprecipitated. 200 µg of total protein extract was used for each CoIP.

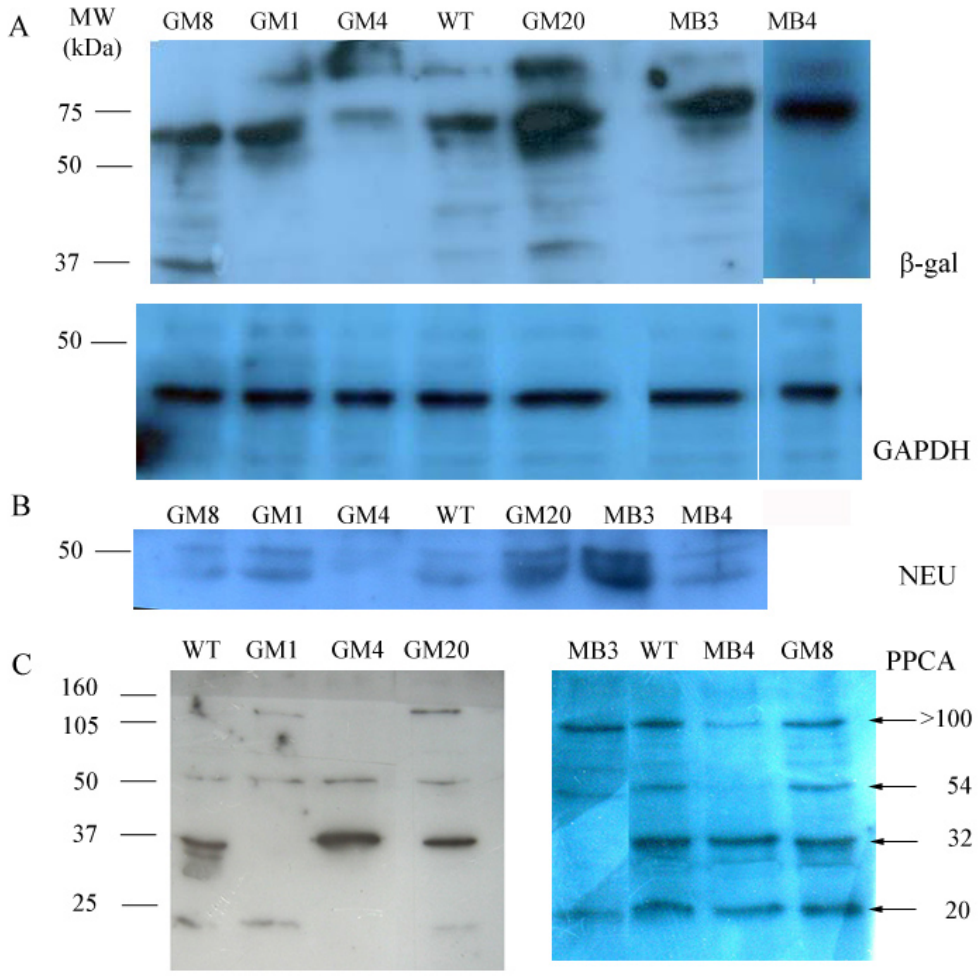


FIGURE 4. Western blottings on patients' fibroblasts using antibodies against β-galactosidase (A), neuraminidase (B) and PPCA (C). On the left molecular weights (kDa) are displayed and on the right, antibodies used. Arrows indicate expected molecular weight for the observed PPCA bands. Above each lane, patient's sample loaded is shown. 40 µg of protein extract were loaded in each lane.

TABLE 1. Genotype, clinical type and residual enzyme activity of patients included in this study.

Patient[a]	Mutations		Clinical Type	Activity[b]
GM5, 6, 10, 14, 16, 17, 22	p.R59H	p.R59H	I	0.05-3.5%
GM13	p.R59H	p.D441N	I	2.5%
GM9	p.D441N	p.D441N	I	0.7%
GM19	p.D198X	p.T420P	I	0.8%
GM7	p.R590C	p.R590C	I	0.9%
A1	p.[L162S;R521C]	p.[L162S;R521C]	I	3.2-1.2%
GM1	p.R201H	p.[G272D;S532G]	III	1.2%
GM20	p.T420K	p.T420K	III	3.2%
GM4	c.1479+1G>T	c.1479+1G>T	I	4%
GM8	c.245+1G>A	c.1572_1577insG	I	1.2%
MB1, MB5	p.L173P	p.T500A	Morquio B	1.8-5%
MB2	p.R201H	p.R201H	Morquio B	2.4%
MB3	p.Y83C	p.D441N	Morquio B	11.9%
MB4	p.Y444C	p.G494S	Morquio B	3.3%

[a]All patients are described in Santamaria et al. (2006) except for A1, described in Santamaria et al. (2007).
[b] % of residual enzyme activity in patient's fibroblasts.

TABLE 2. Significance of the difference in pair-wise comparisons between mutant enzymes with detectable levels of activity and the wild-type enzyme.

Y83C	R201H	T420K	R521C	S532G	Alleles	n
< P=0.001	< P=0.001	< P=0.001	< P=0.001	< P=0.001	WT	30
	NS	NS	P=0.003	P=0.004	Y83C	5
		NS	P=0.005	P=0.002	R201H	6
			P=0.001	P=0.002	T420K	6
				P=0.001	R521C	7
					S532G	6

n: number of replicas; NS: not significant

TABLE 3. Previously reported activities of the changes found in codons expressed in this paper and activities of the changes in those codons expressed in this study.

Mutation	Activity	Expression System[a]	Reference
p.R59H	0.06 %	COS cells-DEAEdextran	Caciotti et al. (2005)
p.R59H	0%	COS cells- Lipofectamine™	This study
p.R59C	0.23 %	COS cells-DEAEdextran	Caciotti et al. (2005)
p.Y83H	2-5 %	$ASVG_{M1}$ cells-Calcium phosphate	Ishii et al. (1995a)
p.Y83C	6.9%	COS cells- Lipofectamine™	This study
p.R201C	8.3%	$ASVG_{M1}$ cells-Calcium phosphate	Yoshida et al. (1991)
p.R201C	3.4 %		Ishii et al. (1995b)
p.R201C	12'9%	COS cells-DEAEdextran	Caciotti et al (2003)
p.R201C	8.5%	$ASVG_{M1}$ cells-Calcium phosphate	Oshima et al. (1994)
p.R201C	0%	COS cells- Lipofectin™	Nishimoto et al. (1991)
p.R201H	46.5%		Ishii et al. (1995b)
p.R201H	0%	COS cells- Lipofectamine™	Kaye et al. (1997)
p.R201H	15.9%	COS cells- Lipofectamine™	This study
p.R521C	24%	COS cells-DEAEdextran	Caciotti et al. (2005)
p.R521C	33.2%	COS cells- Lipofectamine™	This study
p.S532G	81.0%	COS cells-Adenovirus mediated	Zhang et al. (2000)
p.S532G	75-97%	CHO cells-Lipofectin™	Zhang et al. (2000)
p.S532G	60.2%	COS cells- Lipofectamine™	This study
p.R590H	<1%	COS cells- Lipofectamine™	Boustany et al. (1993)
p.R590C	0%	COS cells- Lipofectamine™	This study

[a]Cell line and transfection system are shown

2.2 El cas d'una nova pseudodeficiència: Estudi de la variant p.R595W del gen *GLB1*

En aquest article es descriu una nova variant del gen *GLB1*, el canvi p.R595W. Es tracta d'un canvi identificat en el pare d'un pacient amb Gangliosidosi GM1 en l'al.lel no patogènic. L'activitat enzimàtica residual en les cèl.lules d'aquest individu va resultar ser inferior a l'activitat mitjana de les cèl.lules dels individus portadors d'un canvi patogènic. Per aquesta raó el canvi va ser expressat en un sistema *in vitro* en el que vàrem utilitzar cèl.lules COS i lipofectamina per transfectar un plasmidi pcDNA3.1 que duia el cDNA del gen *GLB1* amb el canvi p.R595W clonat. L'activitat residual *in vitro* per aquest canvi va ser d'un 33-59% del valor control.

Això va permetre classificar-lo com un nou al.lel pseudodeficient, és a dir, una variant no patogènica que es troba a la població general però que té un efecte sobre l'activitat enzimàtica (la redueix) de manera que podria provocar errors en el diagnòstic.

Per comprovar la importància d'aquesta variant a l'hora de fer diagnòstics es va estudiar la seva freqüència a la població general. En concret, com que la variant s'havia trobat en un pacient d'ascendència basca es van analitzar 62 individus control d'origen basc i 118 individus espanyols no bascos. Les freqüències trobades van ser d'un 3,2% a la població basca i d'un 0,8% a la població espanyola.

Treball sotmès a publicació:
"Identification of a novel pseudodeficiency allele in the *GLB1* gene in a carrier of GM1-gangliosidosis." *Clinical Genetics*

Identification of a novel pseudodeficiency allele in the *GLB1* gene in a carrier of GM1-gangliosidosis.

Laura Gort [1,2], Raül Santamaria [2,3,4], Daniel Grinberg [2,3,4], Lluïsa Vilageliu [2,3,4], Amparo Chabás [1,2]

[1]Institut de Bioquímica Clínica, Hospital Clínic, Barcelona, Spain
[2]Centre for Biomedical Research on Rare Diseases (CIBERER), ISCIII, Barcelona
[3]Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona
[4]Institut de Biomedicina de la Universitat de Barcelona (IBUB)

Running title:
A novel GM1 pseudodeficiency allele

Correspondence to:
Dr. Amparo Chabás
Institut de Bioquímica Clínica,
Hospital Clínic, Corporació Sanitària Clínic,
C/ Mejía Lequerica s/n, Edifici Helios III.
E-08028 Barcelona. Spain.
Tel. +34 93 227 5671
Fax. +34 93 227 5668
E-mail: achabas@clinic.ub.es

ABSTRACT

The term “pseudodeficiency” is used in lysosomal storage diseases to denote the situation in which individuals show greatly reduced enzyme activity but remain clinically healthy. Pseudodeficiencies have been reported for several lysosomal hydrolases. GM1-gangliosidosis is a rare autosomal recessive lysosomal storage disorder caused by β-galactosidase hydrolase deficiency due to mutations in the *GLB1* gene. Until now, two variants altering the β-galactosidase activity have been described, p.Arg521Cys and p.Ser532Gly. Here we report the new variant p.Arg595Trp in the *GLB1* gene which markedly reduces β-galactosidase activity when expressed in COS-7 cells. The variant was identified in the healthy father of a girl with GM1-gangliosidosis. He was a heterozygous compound with p.Arg595Trp in *trans* with one of the disease-causing mutations identified in his daughter; in leukocytes and plasma he showed lower β-galactosidase activity than that observed in GM1-gangliosidosis carriers. As this family came from the Basque Country in the north of Spain, we decided to analyze individuals of Basque and non-Basque origin, finding the p.Arg595Trp allele in 3.2% of Basque and in 0.8% of non-Basque control alleles. The detection of the presence of alterations resulting in pseudodeficient activity in leukocytes and plasma is important for the correct diagnosis of GM1-gangliosidosis.

KEYWORDS

β-galactosidase; GM1-gangliosidosis; *GLB1* gene; pseudodeficiency; modifier variant.

INTRODUCTION

The term “pseudodeficient enzyme activity” is used in lysosomal storage disorders to denote situations in which individuals show greatly reduced enzyme activity but without pathological implications, thus complicating the diagnosis of carriers and affected patients. Pseudodeficiencies have been reported for several lysosomal hydrolases such as arylsulfatase A, β-hexosaminidase A and B, galactosylceramidase, α-galactosidase A, β-glucuronidase, α-L-iduronidase, acid α-glucosidase and α-L-fucosidase (Thomas, 1994).

GM1-gangliosidosis (MIM 230500) is a rare autosomal recessive lysosomal storage disorder caused by β-galactosidase hydrolase (GLB1; E.C.3.2.1.23) deficiency due to mutations in the *GLB1* gene. This enzyme is involved in the degradation of GM1 and GA1 gangliosides, lactosylceramide, asialofetuin, oligosaccharides carrying terminal β-linked galactose and keratan sulphate (Suzuki et al. 2001).

So far, two polymorphic changes with reduced β-galactosidase activity have been reported, p.Arg521Cys (Silva et al.1999, Cacciotti et al. 2005) and p.Ser532Gly (Zhang et al. 2000). Here we discuss the identification of a new molecular variant in the *GLB1* gene, presenting with pseudodeficient β-galactosidase activity in plasma and leukocytes of an asymptomatic individual. Expression studies and population analysis are also presented.

MATERIAL AND METHODS

Probands and control samples

The first individual identified as a carrier of the pseudodeficiency variant reported here was the healthy father (case GM33) of a GM1-gangliosidosis patient. This family came from the Basque Country in northern Spain.

Fifty DNA control samples of individuals of Basque origin were kindly provided by Dr. A. López de Munain (Hospital Donostia, San Sebastián, Spain) who confirmed the Basque ancestry of the samples. Twelve more samples of Basque origin corresponding to asymptomatic parents of patients with an inborn error of metabolism not related to GM1-gangliosidosis, sent by different Basque hospitals, were also screened. As controls, 118 DNA samples from individuals of non-Basque ancestry provided by different hospitals were analyzed.

Molecular analysis

Genomic DNA was extracted from leukocytes or cultured skin fibroblasts using the Puregene™ DNA isolation kit of Gentra Systems (Minneapolis). Each exon and flanking intron regions of the *GLB1* gene (GenBank: NM_000404.1; Ensembl gene ID: ENSG00000170266) was amplified by polymerase chain reaction (PCR) using self designed oligonucleotides and then sequenced using the dye-terminator technology, according to the protocol supplied by the manufacturer (Perkin Elmer, USA). After identification of the p.Arg595Trp variant, part of exon 16 and flanking intron 15 were PCR-amplified using oligonucleotides GM16F.-TTCTTTCCTCCATTGCTGCT and GM16R.-GTGAAATGTGGCATGACAGG. PCR product size was 480bp. PCR reaction conditions were as follows: 200 ng of DNA were amplified in a total volume of 25 μl containing 0.2 mM dNTPs, 0.4 mM of each primer, 1.5 mM $MgCl_2$, x1 of Taq DNA polymerase buffer (Roche, Germany) and 1 U of Taq DNA Polymerase (Roche). Reaction parameters were 94°C for 5 min, then 35 cycles at 94°C for 30 sec followed by 40 sec at 60°C, and a final extension step of 72ºC for 3 min. The presence of p.Arg595Trp variant was confirmed by restriction analysis using the elimination of an *AvaI* natural site.

Expression studies

The whole coding region of the *GLB1* cDNA was cloned into the pcDNA3.1 expression vector. Polymorphism p.Arg595Trp was generated by site-directed

mutagenesis using the *QuickChange™ Site-Directed Mutagenesis XL kit* (Stratagene, La Jolla, CA) following the manufacturer's instructions and also cloned into the pcDNA3.1 vector (p.BGAL_R595W). COS-7 cells were cultured in 100-mm diameter tissue culture dishes with DMEM medium (GIBCO, BRL Grand Island, NY, USA), 10% fetal bovine serum (GIBCO, BRL Grand Island, NY, USA) and antibiotics. For transfection with wild-type and mutant β-galactosidase cDNAs, $30x10^5$ cells were plated. Twenty-four hours later (when the cells were at 90% of confluency), 2 µg of corresponding plasmid mixed with 15 µl of *Lipofectamine™ 2000 Reagent* (Invitrogen, Carlsbad, CA) was added. Cells were collected 48 hours after transfection and stored at -80ºC until the enzyme analysis was performed. Each transfection experiment was performed in triplicate.

Enzyme analysis

β-galactosidase activity was assayed with the fluorogenic substrate 4-methylumbelliferyl-β-D-galactopyranoside.

RESULTS

The parents of a patient diagnosed with GM1-gangliosidosis were analyzed to confirm the carrier status. After enzyme analysis in leukocytes and plasma, the patient's mother showed reduced activity in leukocytes (77% of control values) and normal activity in plasma, whereas the father (case GM33) showed a significant reduction of β–galactosidase activity in leukocytes and plasma compared with other GM1-gangliosidosis carriers (Table 1). The molecular lesions identified in the affected daughter were mutations [c.588_591insT] + [p.Glu517X] (patient GM1.26 in Santamaria et al, 2006). The patient's mother carried the insertion c.588_591insT, and GM33 carried the p.Glu517X. Screening of the rest of the *GLB1* gene in GM33 revealed a new variant in exon 16, the p.Arg595Trp (c.1783C>T) change, not present in his daughter. Therefore, GM33 is a compound heterozygote for changes [p.Glu517X] + [p.Arg595Trp].

The c.1783C>T-bearing allele was transiently expressed in COS-7 cells and the β–galactosidase activity was measured in three independent experiments. Enzyme activity values (as mean±SD) were 491.3±39.1; 564.3±18.6 and 879.3±136.3 nmol/h/mg, accounting for reductions of 33% to 59% of activity in cells transfected with wild-type cDNA (mean±SD: 1493±623 nmol/h/mg, for 17 replicas measured).

The family of GM33 came from the Basque Country and their surnames were all of Basque origin. Therefore, 62 Basque controls were screened in order to detect the new change. The p.Arg595Trp variant was detected in 4 out of 124 alleles (frequency of 3.2%). After screening 118 non-Basque Spanish controls, two p.Arg595Trp alleles were identified (frequency 0.8%). Chi-square test demonstrated a p.Arg595Trp association with the Basque alleles ($p<0.5$).

DISCUSSION

Several pseudodeficiencies of hydrolase enzymes have been described in the context of lysosomal storage disorders (Thomas et al, 1994). The enzyme values observed in some carriers of the pseudodeficient allele overlap with those obtained for true disease carriers. On the other hand, a compound heterozygote for a disease-causing mutation and a pseudodeficient allele may be misdiagnosed as an affected case and so, detection of the pseudodeficiency status, has become an important objective to avoid diagnosis errors. Identification of the specific molecular lesion causing the pseudodeficiency in each disease is the main tool to rule out this condition when the enzyme values and urine analysis are not clear. Many molecular changes have been reported to cause pseudodeficiency in lysosomal storage diseases; for instance, the ASApd allele ([p.Asn350Ser; g.2723A>G,polyA-]) in metachromatic leukodystrophy where the change that decreases the enzyme activity is one affecting the polyadenylation signal (Gieselmann et al.1989); the p.Ala300Thr change in mucopolysaccharidosis type I (Aronovich et al. 1996); the p.Arg247Trp and p.Arg249Trp alleles in Tay-Sachs disease (Cao et al, 1997); the p.Asp313Tyr change in Fabry disease (Yasuda et al. 2003) and the p.Asp152Asn variant in mucopolysaccharidosis typeVII (Vervoort et al 1998). All of them reduce the enzyme values to 10-50% of normal activity.

Some alterations may modulate the effect of a disease-causing mutation when in *cis* with a damaging mutation. They can be considered as “modifier variants” and it has been reported for the p.Glu326Lys polymorphism in Gaucher disease (Montfort et al. 2004) and the p.Leu436Phe polymorphism in GM1-gangliosidosis (Caciotti et al 2003). When present as a single change in one allele, these last variants may slightly reduce the enzyme activity or not at all.

At present, two polymorphisms have been reported which partially reduce the β-galactosidase activity, when expressed in COS cells: p.Ser532Gly (Zhang et al. 2000) and p.Arg521Cys (Cacciotti et al. 2005). The identification of the new pseudodeficient allele p.Arg595Trp in the *GLB1* gene shows that this gene is prone to present pseudodeficient and modulating variants. Nevertheless, more studies are required to establish whether the new molecular change has also a modulating effect since, until now, it has not been identified in *cis* with a disease-causing

mutation. These results show that detection of low β-galactosidase activity makes analysis of these pseudodeficient alleles recommendable in order to avoid misdiagnosis in patients and carriers of GM1-gangliosidosis.

The p.Arg595Trp variant has been found more frequently in the Basques (3.2%) than in the rest of Spanish population (0.8%). The higher frequency of the novel pseudodeficient allele found in Basques supports the observations made by several authors regarding the significant differences in the genetic background of this population (Calafell and Bertranpetit 1994).

In conclusion, the identification of a new variant with intracellular activity at around 45% of normal levels and pseudodeficient activity in plasma and cells stresses the importance of analysing the complete gene sequence to explain atypical biochemical results.

ACKNOWLEDGMENTS

The authors thank Dr. López de Munain (Hospital Donostia, San Sebastián, Spain) for referral of DNA Basque control samples and Dr Ruiz Benito (Hospital Donostia) for sending blood samples. The excellent technical assistance of J. Jarque and H. Sellés is also acknowledged. The authors are grateful to R.Rycroft for revising the English. This study was supported in part by FIS (Redes temáticas G03/054 REDEMETH, PI051182) and CICYT (SAF 2003-00386). R. Santamaria is the recipient of a fellowship from the Spanish Ministerio de Educación y Cultura.

REFERENCES

Aronovich EL, Pan D, Whitley CB. Molecular genetic defect underlying alpha-L-iduronidase pseudodeficiency. Am J Hum Genet 1996: 58: 75-85

Caciotti A, Bardelli T, Cunningham J, d'Azzo A, Zammarchi E, Morrone A. Modulating action of the new polymorphism L436F detected in the GLB1 gene of a type-II GM1-gangliosidosis patient. Hum Genet 2003: 113: 44-50

Cacciotti A, Donati MA, Boneh A et al. Role of beta-galactosidase and elastin binding protein in lysosomal and nonlysosomal complexes of patients with GM1-gangliosidosis. Hum Mutat 2005: 25: 285-292

Calafell F, Bertranpetit J. Principal component analysis of gene frequencies and the origin of Basques. Am J Phys Anthropol 1994: 93: 201-215

Cao Z, Petroulakis E, Salo T, Triggs-Raine B. Benign HEXA mutations, C739T (R247W) and C745T (R249W), cause beta-hexosaminidase A pseudodeficiency by reducing the alpha-subunit protein levels. J Biol Chem 1997: 272: 14975-14982.

Gieselmann V, Polten A, Kreysing J, von Figura K. Arylsulfatase A pseudodeficiency: loss of a polyadenylylation signal and N-glycosylation site. Proc Natl Acad Sci USA 1989: 86: 9436-40.

Martiniuk F, Bodkin M, Tzall S, Hirschhorn R. Identification of the base-pair substitution responsible for a human acid alpha glucosidase allele with lower "affinity" for glycogen (GAA 2) and transient gene expression in deficient cells. Am J Hum Genet 1990: 47: 440-445

Montfort M, Chabás A, Vilageliu L, Grinberg D. Functional analysis of 13 GBA mutant alleles identified in Gaucher disease patients: Pathogenic changes and "modifier" polymorphisms. Hum Mutat 2004: 23: 567-575

Santamaria R, Chabás A, Coll MJ, Miranda CS, Vilageliu L, Grinberg D. Twenty-one novel mutations in the GLB1 gene identified in a large group of GM1-gangliosidosis and Morquio B patients: possible common origin for the prevalent p.R59H mutation among gypsies. Hum Mutat 2006: 27: 1060

Silva CM, Severini MH, Sopelsa A, Coelho JC, Zaha A, d'Azzo A, Giugliani R. Six novel beta-galactosidase gene mutations in Brazilian patients with GM1-gangliosidosis. Hum Mutat 1999: 13: 401-409

Suzuki Y, Oshima A, Namba E. β-galactosidase deficiency (β-galactosidosis): GM1-gangliosidosis and Morquio B disease. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. The Metabolic and Molecular Bases of Inherited Disease. McGraw-Hill, New York, 2001: 3775-3809

Thomas GH. “Pseudodeficiencies” of lysosomal hydrolases. Am J Hum Genet 1994: 54: 934-940

Vervoort R, Gitzelmann R, Bosshard N, Maire I, Liebaers I, Lissens W. Low beta-glucuronidase enzyme activity and mutations in the human beta-glucuronidase gene in mild mucopolysaccharidosis type VII, pseudodeficiency and a heterozygote. Hum Genet 1998: 102: 69-78.

Yasuda M, Shabbeer J, Benson SD, Maire I, Burnett RM, Desnick RJ. Fabry disease: Characterization of alpha-galactosidase A double mutations and the D313Y plasma enzyme pseudodeficiency allele. Hum Mutat 2003: 22: 486-492.

Zhang S, Bagshaw R, Hilson W et al. Characterization of beta-galactosidase mutations Asp332-->Asn and Arg148-->Ser, and a polymorphism, Ser532-->Gly, in a case of GM1 gangliosidosis. Biochem J 2000: 348: 621-632

TABLE 1. β-Galactosidase activity of case GM33, his affected daughter (GM1.26) and other GM1-gangliosidosis patients and carriers.

	Leukocytes (nmol/h/mg)	Plasma (nmol/h/ml)	Fibroblasts (nmol/h/mg)
Case GM33	25.2*	7.8*	NA
Patient GM1.26	NA	NA	16.4
Other GM1-Gangl. patients	12.6±9.7 (n=12)	2.30±0.9 (n=11)	
GM1-Gangl. carriers	106.5±44.7 (n=12)	19.5±5.4 (n= 5)	
Controls	175.9±43.8	22.2±15.6	1439±320

Enzyme activity expressed as mean±SD
*mean activity value from 2 aliquots of a single sample; NA: not available

CAPÍTOL 3

Estudi del procés d'*splicing* alternatiu del gen *GLB1*

3.1 Anàlisi dels mecanismes implicats en el procés d' alternatiu del gen *GLB1*

Fins al moment de la realització d'aquest treball s'havien descrit diversos estudis sobre les proteïnes β-galactosidasa i EBP, però en cap d'ells s'havien analitzat els possibles mecanismes reguladors de l'*splicing* alternatiu del gen *GLB1*. Per aquesta raó vam dissenyar una sèrie d'experiments per intentar esbrinar els mecanismes que feien que existís un transcrit que no inclou els exons 3, 4 i 6 simultàniament.

En tractar cèl.lules HeLa i fibroblasts humans amb cicloheximida (CHX) per inhibir el procés de *Nonsense-Mediated Decay* (NMD) es va posar de manifest l'existència de transcrits amb altres combinacions d'exons. Per tant, aquests transcrits se sintetitzen però no s'observen perquè són degradats pel sistema d'NMD. A més a més, en tractar cèl.lules de ratolí amb CHX vam aconseguir que es posés de manifest el transcrit homòleg al de l'EBP en aquestes cèl.lules. Això suggerí que a ratolí, on no hi ha EBP, l'absència del transcrit de l'EBP es deu bàsicament a la presència d'un codó d'aturada prematur (PTC) a la pauta de lectura alternativa de l'exó 5 en aquesta espècie.

Per altra banda, vam construir un minigèn que incloïa els exons implicats en l'*splicing* alternatiu del gen *GLB1* humà. Amb aquest minigèn es va constatar que la feblesa dels llocs 3' d'*splicing* dels exons 3 i 4 no era causa suficient per explicar l'*skipping* d'aquests exons en l'*splicing* alternatiu, així com la capacitat que tenen algunes proteïnes SR de regular la proporció dels diferents transcrits que se sintetitzen.

Treball en preparació:
"Analysis of the alternative splicing of the *GLB1* gene: effect of the NMD and the SR proteins."

Analysis of the alternative splicing of the *GLB1* gene: effect of the NMD and the SR proteins.

Raül Santamaria[1,2,3], Amparo Chabás[2,4], Lluïsa Vilageliu[1,2,3], Daniel Grinberg[1,2,3]

[1]Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Barcelona

[2]Centre for Biomedical Research on Rare Diseases (CIBERER), ISCIII, Barcelona

[3]Institut de Biomedicina de la Universitat de Barcelona (IBUB)

[4]Institut de Bioquímica Clínica, Hospital Clínic, Barcelona, Spain

ABSTRACT

The human *GLB1* gene is known to give rise to two alternatively spliced mRNAs, which encode two different proteins: lysosomal β-galactosidase (β-gal) and elastin-binding protein (EBP). The β-gal transcript is 2.5 kb long and includes the 16 exons of the *GLB1* gene. In the EBP transcript, exons 3, 4 and 6 are skipped, while exon 5 has a different reading frame. Different experiments were carried out to study how this alternative splicing is generated. Treatment of cultured cells (HeLa and human fibroblasts) with cycloheximide showed up the presence of transcripts that are usually degraded by nonsense-mediated decay (NMD). These transcripts had combinations of exons 3, 4 and 6 different to those present in the transcripts of EBP and β-Gal, and bore premature termination codons (PTCs).

A minigene carrying the exons involved in the alternative splicing of the *GLB1* gene was constructed. Improving the 3' splice site scores of exons 3 or 4 increased the relative inclusion of these exons, but did not prevent that they were skipped in some transcripts. Thus, exon skipping could not be explained only by the weak 3' splice sites. Overexpression of different SR proteins altered the relative proportion of the different transcripts produced by the minigene, indicating a possible way of regulating the *GLB1* alternative splicing. Finally, a comparison of this gene between different species was performed. The existence of a transcript with the EBP reading frame was possible in all species analyzed except for rodents, where a PTC would be generated in exon 5. Treatment of RAW264.7 mouse cells with cycloheximide allowed the detection by RT-PCR of the EBP-like transcript, indicating that the absence of this transcript was only due to the presence of the PTC. This is the first time that the mechanisms underlying the alternative splicing of the *GLB1* gene are addressed.

KEYWORDS

GLB1 gene; Nonsense-mediated decay (NMD); SR protein; alternative splicing

INTRODUCTION

Splicing of eucaryotic RNAs has been known for more than 25 years (Berget, et al., 1977; Chow, et al., 1977). Alternative splicing, which could affect 30 to 50% of the genes in higher eukaryotes (Modrek and Lee, 2002; Roberts and Smith, 2002), allows the generation of multiple transcripts by different mechanisms such as exon skipping/inclusion, alternative 5' or 3' splice sites, intron retention or mutually exclusive exons. Thus, analysis of alternative splicing of genes has become an important landmark to unravel the knowledge given by the human genome sequence. The mechanisms underlying alternative splicing are not known in detail but a general outline is understood. Alternative exons often have suboptimal splice sites and/or suboptimal length when compared with constitutive exons. Besides the proteins that are present in constitutive splicing, a group of proteins that bind to pre-mRNA sequences are also necessary in alternative splicing (Cartegni, et al., 2002). The sequences where these proteins bind are called exonic or intronic enhancers (ESEs or ISEs) and silencers (ESSs or ISSs). In the case of ESEs, most of them are recognized by SR proteins. These proteins contitute a family of proteins that has an RS domain (rich in arginines and serines) and 1 or 2 RRM domains (which are RNA-binding domains) and are essential, multifunctional splicing factors required both for spliceosome assembly and for alternative splicing (Graveley, 2000). On the other hand, the best characterized ESSs and ISSs are recognized by members of the heterogeneus nuclear ribonucleoprotein (hnRNP) family, which are highly abundant RNA-binding proteins that lack an RS domain (Zhu, et al., 2001). It is known that these proteins have a dose-dependent effect and that the binding of SR and hnnRNP proteins on the pre-mRNA allows the splicing machinery to act in a tissue or in a developmental stage -specific manner (Caceres and Kornblihtt, 2002).

The nonsense-mediated mRNA decay (NMD) is a well known mechanism that degrades mRNAs harbouring premature termination codons (PTCs), not only of *de novo* nonsense mutations but also of those found in physiologic transcripts. This means that NMD has a function in muting the genomic noise, thus regulating the expression of thousands of physiologic transcripts (Mendell, et al., 2004). The coupling between NMD and alternative splicing has become therefore more evident (Lewis, et al., 2003).

The *GLB1* gene, located on chromosome 3 at 3p21.33 and organized in 16 exons, was found to be the gene coding for the β-galactosidase protein (E.C.3.2.1.23) (Oshima, et al., 1990). This enzyme cleaves terminal β-galactoses from different substrates in the

lysosome. A second transcript generated by alternative splicing was described (Morreau, et al., 1989), but it was not until 10 years later that the protein coded by this second transcript, the elastin-binding protein (EBP), was identified. While the β-galactosidase cDNA is 2.5 kb long, the second transcript is only 2 kb long. In the 2-kb transcript, exons 3, 4 and 6 are missing, and exon 5 has a different reading frame that gives rise to a unique stretch of 32 amino acids that allows EBP to bind tropoelastin and to contribute to correct elastogenesis in cells (Privitera, et al., 1998). Therefore, the only differences between the 2 proteins are the absence of exons 3, 4 and 6 and the different reading frame of exon 5. This was the first example of such a configuration in a mammalian gene (Morreau, et al., 1989).

Mutations in the *GLB1* gene that result in absence or reduced activity of the lysosomal enzyme β-galactosidase produce two different diseases: GM1-gangliosidosis (MIM# 230500) and Morquio B (MIM# 253010). A role of EBP in these diseases has also been suggested but is still under discussion (Caciotti, et al., 2005; Morrone, et al., 2000).

Efforts have been made to establish the function and interactions of both proteins. In contrast, no studies have been reported on the way alternative splicing of the *GLB1* gene is regulated. Here we report a preliminar approach to establish the mechanisms underlying this alternative splicing.

MATERIALS AND METHODS

Vector construction

The *in vitro* studies of the alternative splicing of the *GLB1* gene were performed using a minigene construct carrying exons 2, 3, 4, 5, 6 and 7 (construct BX; see Figure 1A). Since the introns involved are very long (more than 15 kb), only the exons and their flanking sequences (about 200 bp on each side) were cloned in a pcDNA3.1 plasmid. Fragments were PCR-amplified with primers including terminal restriction sites useful for cloning (primer sequences are available on demmand). The pCGT7 plasmids bearing cDNAs enconding SR proteins (ASF/SF2, SRp20, SRp40, SRp55, 9G8) and hnRNPA1 were a kind gift of Dr. Kornblihtt and have been described elsewhere (Caceres, et al., 1997).

Cell Culture and Transfection

HeLa cells, human fibroblasts and RAW264.7 mouse cells were cultured in the presence of DMEM medium (GIBCO, BRL Grand Island, NY, USA) with 10% fetal bovine serum (GIBCO, BRL Grand Island, NY, USA) and antibiotics, at 37ºC and 5% CO_2. The cells were also treated with 500 µg/ml of cycloheximide (CHX). When cells were at 90% of confluency, CHX was added to the medium for 4 hours and then RNA was isolated.

For transfection, cells were plated at 50% of confluency in 6-well culture plates and 24 hours later, at 90% of confluency, 500 ng of the corresponding plasmid were mixed with 4µl of *Lipofectamine™ 2000 Reagent* (Invitrogen, Carlsbad, CA), according to the manufacturer's recommendations. In cotransfections, 500 ng of each plasmid were used. As a control, the minigene plasmid (BX) was mixed with 500 ng of an empty pUC vector. Cells were collected 48 hours after transfection. RNA isolation, cDNA synthesis, RT-PCR fragment purification and sequencing were performed as previously described (Santamaria, et al., 2006). Correct transfection and expression of SR proteins was always checked by RT-PCR of the specific transcript.

PCR Amplification

Endogenous and minigene transcripts were analyzed by normal and semiquantitative RT-PCR, respectively. For the minigene transcripts amplification, a transcript-specific primer and a plasmid-specific primer (T7 or SP6) were used to avoid amplification of endogenous transcripts. PCR was performed using 1 U of Taq DNA Polymerase (Promega, Madison, WI, USA), 10 pmols of each primer (see Table 1), 2.5 mM $MgCl_2$, 200 µM

dNTPs, 0.1 µl ^{32}α-P dATP (10µCi/µl) (in semiquantitative PCRs) and 5-10 ng of cDNA in the recommended buffer. PCR conditions were as follows: 4 min of denaturation at 95ºC, 25 (or 35 in the non-quantitative PCRs) cycles of denaturation at 94ºC for 30s, annealing for 30s in the temperature indicated in Table 1 and extension for 30s at 72ºC.

Transcript specificity was checked for the EBP primers. EBP_F primer was complementary at its 5' end to exon 2 and at its 3' end to exon 5. EBP_R primer was complementary at its 5' end to exon 7 and at its 3' end to exon 5. As a control of transcript specificity, "scrambled" primers were used. The scrambled-EBP_F primer was synthesized with the same exon 5 sequence as EBP_F primer, but with the nucleotides corresponding to exon 2 scrambled. The same was performed for the scrambled EBP_R primer, where nucleotides corresponding to exon 7 were scrambled. The absence of any amplification in the PCRs performed with scrambled primers, confirmed the specificity of the EBP primers.

To quantify the resulting semiquantitative PCR products, 6 µl of each sample were loaded on a non-denaturing 5% polyacrilamide gel and run for 1h 30min at 14 mA. Gels were then dried and visualized in a *Molecular Imager FX* (Biorad). Quantification was performed using the *Quantity One*® (Biorad) software.

Site-directed Mutagenesis

Changes at the 3' splice sites (3'ss) of exons 3 and 4 were carried out by site-directed mutagenesis on the minigene construct, using the *QuickChange*^{TM} *Site-Directed Mutagenesis XL kit* (Stratagene, La Jolla, CA) following the manufacturer's instructions. In particular, two nucleotide changes were introduced on each 3'ss. Changes introduced in the 3'ss of exon 3 were: c.246-10G>C and c.246-12G>C (BX3 construct); and in the 3'ss of exon 4 were: c.397-3T>C and c.397-7G>C (BX4 construct) (Fig. 1B).

Informatic support and Statistical Analysis

The 3' splice sites scores were analyzed using the software developed by Zhang and Yada at http://rulai.cshl.edu/new_alt_exon_db2/HTML/score.html. Substitution rates for synonymous positions were calculated using the SNAP program at http://hiv-web.lanl.gov/content/hiv-db/SNAP/WEBSNAP/SNAP.html.

In semiquantitative PCRs, each transfection experiment was performed at least twice. At least 3 PCRs of each transfection were performed to quantify the relative proportion of each transcript. The U de Mann-Whitney test was used to analyze significant differences in band intensities between those corresponding to control and treated cells.

RESULTS

Cycloheximide treatment

As a result of the alternative splicing of the *GLB1* gene, only two transcripts are naturally found: the β-Gal transcript, including the 16 exons, and the EBP transcript, where exons 3, 4 and 6 are skipped. Other "intermediate transcripts" (i.e. combinations of exons 3, 4 and 6 different to those in transcripts of EBP and β-Gal) are not found in cells. In case these transcripts were generated, they would bear PTCs (premature termination codons), because they would produce new reading frames. Thus, we treated HeLa cells with CHX to inhibit a possible NMD (nonsense-mediated mRNA decay) of these transcripts. RT-PCR using EBP-specific primers revealed the presence of these intermediate transcripts in the CHX-treated cells, as shown in Figure 2. The experiment was also performed using human control fibroblasts to rule out a possible effect, specific of the transformed HeLa cells and the same results were obtained (Fig.2).

A primer that should not anneal to any of the two natural transcripts was designed (2_4F, see Table 1). RT-PCR with 2_4F/C3R primers on cDNA from CHX-treated and not treated HeLa cells displayed a band only in the CHX-treated cells (data not shown).

Minigene studies

In order to study the alternative splicing of the *GLB1* gene, a minigene including the alternatively spliced region of the gene (exons 2, 3, 4, 5, 6 and 7) was cloned (Fig.1A). This construct was transfected into HeLa cells and RT-PCR was carried out using plasmid-specific primers (T7/SP6) to know if some of the deleted intronic sequences were essential for the correct splicing of any of the exons cloned. A single band was obtained (data not shown). Sequencing of this PCR product showed that it corresponded to a transcript which included the 6 cloned exons, indicating that intron regions in the BX construct bored the necessary *cis* sequences for the correct splicing. On the other hand, if the EBP splicing was taking place it should be at a very low level (as in the endogenous splicing), as no band corresponding to this transcript was detected. In order to check if this splicing ocurred at all, RT-PCR was performed with EBP specific primers. Both T7/EBPR and EBPF/SP6 pairs of primers produced a PCR band corresponding to the EBP transcript. However, extra bands corresponding to other intermediate transcripts, as shown in the BX lane in Figure 3A, were also observed.

Site-directed mutagenesis

In silico analysis of the 3' splice sites (3'ss) of the *GLB1* gene revealed a low score of 2.5 for the site flanking exon 4, a relatively low score of 6 for the site adjacent to exon 3 and a high score of 10 for the 3'ss flanking exon 6, compared to the average score of constitutive exons (7.9). To check if the skipping of exons 3 and 4 in the alternative splicing was due to the weak 3'ss, 2 positions on each 3'ss were mutagenized. These changes increased the 3'ss scores on exons 3 and 4, from 6 to 9.5 and from 2.5 to 7, respectively (Fig. 1B). Analysis of the transcripts generated by these plasmids (BX3 and BX4) showed that although the transcripts lacking both exons still appeared, an increase of the transcript including only exon 3 (in BX3 lane) and of the transcript including only exon 4 (in BX4 lane), were clearly observed (Fig.3A). Upon quantification of the transcripts bearing exon 3 (or exon 4) versus transcripts lacking exon 3 (or exon 4) some small but significant increases were detected. As expected, plasmid BX3 produced proportionally more transcripts bearing exon 3 than BX (116%)(Fig.3B), and plasmid BX4 produced more bearing exon 4 (120%)(Fig.3C).

SR transfection

Another factor involved in the regulation of alternative splicing is the action of the SR and hnRNP A/B proteins. These proteins can bind the pre-mRNA and alter the way it is spliced. That is why we decided to cotransfect HeLa cells with the BX plasmid together with a plasmid expressing one of the following proteins: ASF/SF2, SRp20, SRp40, SRp55, 9G8 or hnRNPA1. We focused on the inclusion/exclusion of exons 3 and 4. The different transcripts detected by RT-PCR using primers T7/EBP_R were quantified. As shown in Figure 4, a specific effect on the relative proportion of the different transcripts due to the overexpressed proteins could be observed. As a control, transfection of BX plasmid plus an empty plasmid was used (BX lane). The apparent absence of effect of NMD on the minigene transcripts showed that different exonic combinations were also generated by splicing.

RT-PCR with 2_4F/SP6 primers on cDNA from HeLa cells cotransfected with BX and different SR plasmids, although not in quantitative conditions, showed again a specific, repetitive pattern (Fig.5) and had a general effect of favouring the inclusion of exons 5 and 6.

Species comparison

It had already been noted that the alternative splicing described for the human *GLB1* gene could not take place in mouse (*Mus musculus*), because the correct translation of exon 5 in mouse is only possible in the reading frame that codes for the lysosomal β-galactosidase

(Morreau, et al., 1991). In fact, this also happens in rat (*Rattus norvegicus*) but not in other mammals such as dog (*Canis familiaris*), cat (*Felis catus*), crab-eating macaque (*Macaca fascicularis*) or rabbit (*Oryctolagus cuniculus*) and also in birds like chicken (*Gallus gallus*), where both reading frames are possible. Therefore, in all the species analyzed, except for rat and mouse, a reading frame where exons 3, 4 and 6 are skipped is possible.

Mouse RAW264.7 cells were cultured in the presence and in the absence of CHX. β-Gal (C1Fm/C1Rm primers) and EBP-like (EBPFm/EBPRm primers) -specific RT-PCR were carried out. As shown in Figure 6, the EBP-like transcript is detected only in the CHX-treated cells.

Synonymous substitution rates (Ks) were calculated for exon 5 between human and the species analyzed. Results are presented in Table 2, where it can be observed that mouse and rat show higher Ks (even greater than chicken) than the rest of studied animals. These *in silico* results seem to point out that all the studied animals have the potential to produce both proteins from the same gene and only the rodents (*Murinae*) have lost this ability.

DISCUSSION

One of the conclusions that raised the analysis of the human genome was that a limited number of genes generated the large proteomic complexity. This fact led to a renewed interest in the post-transcriptional mechanisms that generate this protein diversity, especially alternative splicing.

The *GLB1* gene was known to produce two different coding transcripts since it was described (Morreau, et al., 1989), but it was not until 1998 that the protein coded by the shorter transcript, the EBP, was assigned (Privitera, et al., 1998). Although many protein studies have been developed, no studies on the regulation of the alternative splicing in this gene have been reported.

The fact that the skipping of exons 3, 4 and 6 was the only alternative combination of these exons that did not generate a PTC, led us to analyze the effect of NMD on some other possible splicing products. Indeed, we found these intermediate transcripts after treating cells with CHX, indicating that these transcripts are in fact generated by splicing but that are afterwards degraded by NMD (Fig.2).

To further study this alternative splicing we constructed a minigene bearing exons 2, 3, 4, 5, 6 and 7 (Fig.1A) and confirmed that the exons and their 200-250 bp flanking sequences were enough for the correct splicing. In fact, as NMD does not act on the transcripts produced by the minigene (no NMD was expected to occur in the different minigene transcripts because several possible ORFs prevented the different transcripts to suffer NMD), many other combinations, including the EBP one, are generated. This was observed for exons 3 and 4 (with primers T7/EBP_R) and for exons 5 and 6 (with primers 2_4F/SP6). On the other hand, it should be reminded that these combinations are only observed when using primers EBP_F, EBP_R or 2_4F, indicating that they are present only in small amounts. Thus, all the exons seem to be skipped to some extent in some of the minigene transcripts.

We focused on the analysis of the splicing of exons 3 and 4. In order to check if a weak 3' splice site was favouring the skipping of exons 3 or 4 in the alternative splicing, we mutagenized the minigene construct to increase the scores of the 3'ss flanking these exons. This resulted in a slight but significant increase of inclusion of exons 3 and 4 respectively. Plasmid BX4 also presented more transcripts bearing exon 3 than BX (124%), but this was not significant. Nevertheless, the fact that some transcripts with skipped exons 3 and/or 4

were still produced (Fig.3A), points out that these weak splice sites are not sufficient to explain why these exons are skipped in some transcripts.

The SR proteins are known to be essential in constitutive splicing but also necessary to achieve a high level of specificity in alternative splicing. These proteins are *trans*-acting factors that can bind to some pre-mRNA sequences present in the exons or in the introns and favour exon inclusion in the mature mRNA. It has also been described that overexpression of these splicing factors regulates the splicing pattern of alternatively spliced exons (Caceres, et al., 1994). To test a possible effect of these proteins on the *GLB1* alternative splicing we cotransfected HeLa cells with the BX minigene together with a plasmid coding for a SR protein or hnRNPA1. None of these proteins could completely restore the inclusion of one of the exons (3 or 4) but important variations on the relative proportion between the different transcripts could be detected (Fig.4A-D). In general, all SR proteins proportionally reduced the amount of the EBP transcript (exon 3 and 4 skipped, in Fig.4D) and also other intermediate transcripts to some extent (Fig.4B and C). This indicates a function for SRs proteins favouring the inclusion of these exons. On the other hand, hnRNPA1, which has been described to inhibit splicing, here seems to favour the inclusion of both exons too. Previous studies have found a close similarity between the RRMs of proteins 9G8 and SRp20, which can recognize some identical sequences (Cavaloc, et al., 1999), and moreover, it has also been described that 9G8 can participate in the splicing regulation as a functional SRp20 homologue in certain conditions (Lynch and Maniatis, 1996). In our study, both proteins produced the same change of pattern, consistent with these previous results.

SR proteins have a broad and moderate specificity of ESE recognition, which can allow either competitive binding of SR proteins to the same element or cooperative binding to clusters of elements within an ESE. The overexpression of the SR proteins in the minigene experiments, which is far from being a physiological situation, could alter the interactions between the different SR proteins and, thus, produce the observed pattern alteration. This clear change of pattern observed for each protein could be a sign of their implication in the constitutive and alternative splicing of these exons. However, the characteristics of the minigene, such as the CMV promoter or the lack of some of the introns sequence (elements that are known to influence the splicing), should also be taken into account.

Finally, it was known that in mouse there was a PTC in the EBP-like reading frame of exon 5 (Morreau, et al., 1991). We were interested in determining whether the lack of the

EBP-like transcript in mouse was due to the absence of the EBP-like alternative splicing or was due to the presence of the PTC in the transcript, which caused NMD. Thus, we searched for this transcript after treatment with CHX in RAW264.7 mouse cells. Although only a slight band could be observed (Fig.6), the existence of this EBP-like transcript suggests that just the presence of the PTC (a change that is synonymous in the β-gal frame) is enough for the lack of this transcript in rodents (*Murinae*). On the other hand, the Ks analysis (Table 2) supported the hypothesis that in the other species analyzed the EBP transcript should be present, because the highest Ks values were obtained comparing human with rodents (higher even than the comparison between human and chicken). This could be explained by the presence on the other species of a selective force over the second reading frame of exon 5 that is not present in the rodents.

All these results suggest possible explanations for the alternative splicing of the *GLB1* gene. The relatively weak splice sites in the exons involved in the alternative splicing could allow the skipping of the exons in some transcripts. Therefore different combinations of transcripts with skipped exons are generated but only those with a correct reading frame are not degraded. This would also explain the lower presence of the EBP transcript in the cells, as only a few of the transcripts that skip some exons have the correct combination of exons (that is, absence of exons 3, 4 and 6). The fact that the different SR proteins could also have an effect in the process of skipping of these exons indicates a possible way of regulation of this alternative splicing.

REFERENCES

Berget SM, Moore C, Sharp PA. 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci U S A 74(8):3171-5.

Caceres JF, Kornblihtt AR. 2002. Alternative splicing: multiple control mechanisms and involvement in human disease. Trends Genet 18(4):186-93.

Caceres JF, Misteli T, Screaton GR, Spector DL, Krainer AR. 1997. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. J Cell Biol 138(2):225-38.

Caceres JF, Stamm S, Helfman DM, Krainer AR. 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science 265(5179):1706-9.

Caciotti A, Donati MA, Boneh A, d'Azzo A, Federico A, Parini R, Antuzzi D, Bardelli T, Nosi D, Kimonis V and others. 2005. Role of beta-galactosidase and elastin binding protein in lysosomal and nonlysosomal complexes of patients with GM1-gangliosidosis. Hum Mutat 25(3):285-92.

Cartegni L, Chew SL, Krainer AR. 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3(4):285-98.

Cavaloc Y, Bourgeois CF, Kister L, Stevenin J. 1999. The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. Rna 5(3):468-83.

Chow LT, Gelinas RE, Broker TR, Roberts RJ. 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell 12(1):1-8.

Graveley BR. 2000. Sorting out the complexity of SR protein functions. Rna 6(9):1197-211.

Lewis BP, Green RE, Brenner SE. 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci U S A 100(1):189-92.

Lynch KW, Maniatis T. 1996. Assembly of specific SR protein complexes on distinct regulatory elements of the Drosophila doublesex splicing enhancer. Genes Dev 10(16):2089-101.

Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC. 2004. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet 36(10):1073-8.

Modrek B, Lee C. 2002. A genomic view of alternative splicing. Nat Genet 30(1):13-9.

Morreau H, Bonten E, Zhou XY, D'Azzo A. 1991. Organization of the gene encoding human lysosomal beta-galactosidase. DNA Cell Biol 10(7):495-504.

Morreau H, Galjart NJ, Gillemans N, Willemsen R, van der Horst GT, d'Azzo A. 1989. Alternative splicing of beta-galactosidase mRNA generates the classic lysosomal enzyme and a beta-galactosidase-related protein. J Biol Chem 264(34):20655-63.

Morrone A, Bardelli T, Donati MA, Giorgi M, Di Rocco M, Gatti R, Parini R, Ricci R, Taddeucci G, D'Azzo A and others. 2000. Beta-galactosidase gene mutations affecting the lysosomal enzyme and the elastin-binding protein in GM1-gangliosidosis patients with cardiac involvement. Hum Mutat 15(4):354-66.

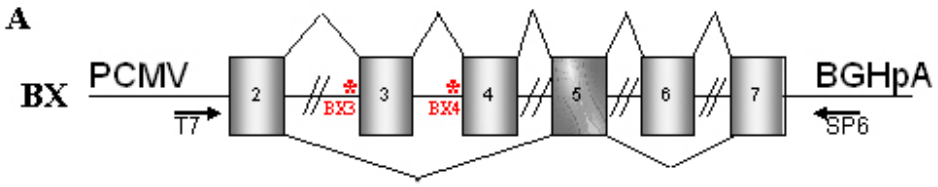
Oshima A, Itoh K, Nagao Y, Sakuraba H, Suzuki Y. 1990. Beta-galactosidase-deficient human fibroblasts: uptake and processing of the exogenous precursor enzyme expressed by stable transformant COS cells. Hum Genet 85(5):505-8.

Privitera S, Prody CA, Callahan JW, Hinek A. 1998. The 67-kDa enzymatically inactive alternatively spliced variant of beta-galactosidase is identical to the elastin/laminin-binding protein. J Biol Chem 273(11):6319-26.

Roberts GC, Smith CW. 2002. Alternative splicing: combinatorial output from the genome. Curr Opin Chem Biol 6(3):375-83.

Santamaria R, Chabas A, Coll MJ, Miranda CS, Vilageliu L, Grinberg D. 2006. Twenty-one novel mutations in the GLB1 gene identified in a large group of GM1-gangliosidosis and Morquio B patients: possible common origin for the prevalent p.R59H mutation among gypsies. Hum Mutat 27(10):1060.

Zhu J, Mayeda A, Krainer AR. 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol Cell 8(6):1351-61.



B

cons	3'ss sequence		scores
---	tt ttccc tccag|G	consensus	14.2 (max)
BX	gt gtctt ggcag|G	exon 3	6
BX3	ct ctctt ggcag|G	exon 3-mut	9.5
BX	tc cctga actag|G	exon 4	2.5
BX4	tc cctca accag|G	exon 4-mut	7

FIGURE 1.A. Scheme of the BX minigene construct. Exons: grey boxes. Introns: horizontal lines. Two lines cutting introns indicate that the whole intron was not cloned. Lines over and under the exons show the naturally ocurring splicing events. * points the localization of the changes introduced by site-directed mutagenesis in BX3 and BX4 constructs. Arrows indicate the localization of T7 and SP6 primers. PCMV: CMV promotor; BGHpA: BGH polyadenilation signal. **B.** 3'ss wild-type and mutated sequences and calculated splicing scores. Changes introduced by mutagenesis are shown in red. cons: plasmid construct.

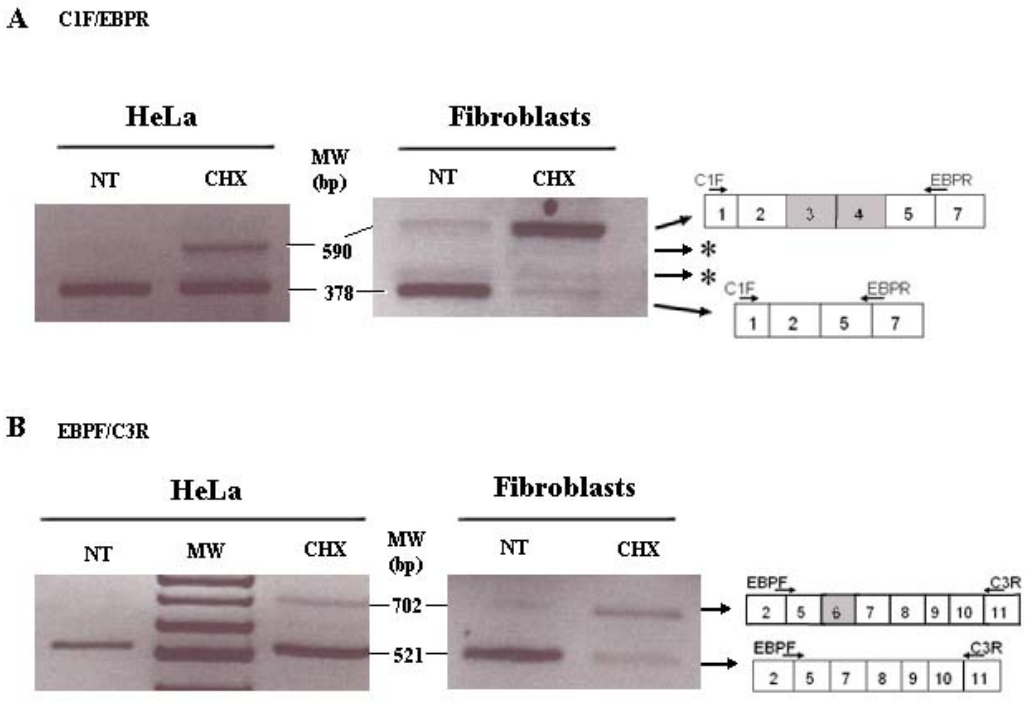


FIGURE 2. RT-PCR on cDNA from not treated (NT) and cycloheximide-treated (CHX) cells in HeLa cells and human fibroblasts. On the right, the transcripts corresponding to each band are outlined. **A**. PCR using C1F/EBPR as primers. **B.** PCR using EBPF/C3R as primers. * points bands for possible transcripts including only exon 3 or only exon 4.

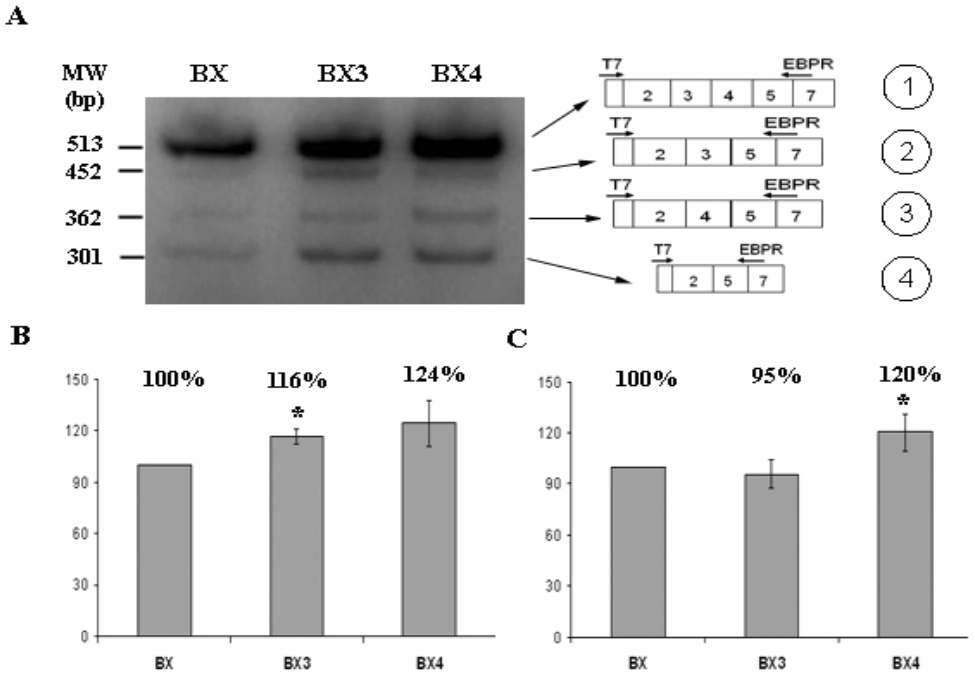


FIGURE 3.A. Semiquantitative RT-PCR on cDNA from transfected HeLa cells with plasmids BX, BX3 and BX4 using T7/EBPR primers. On the right, a scheme of the transcript corresponding to each band is shown. **B**. Percentage of transcripts including exon 3 versus transcripts lacking it, i.e. (1+2)/(3+4). **C**. Percentage of transcripts including exon 4 versus transcripts lacking it, i.e. (1+3)/(2+4). Each value is the mean±S.E. of two independent experiments, each of which was replicated three times. * when $p<0.05$.

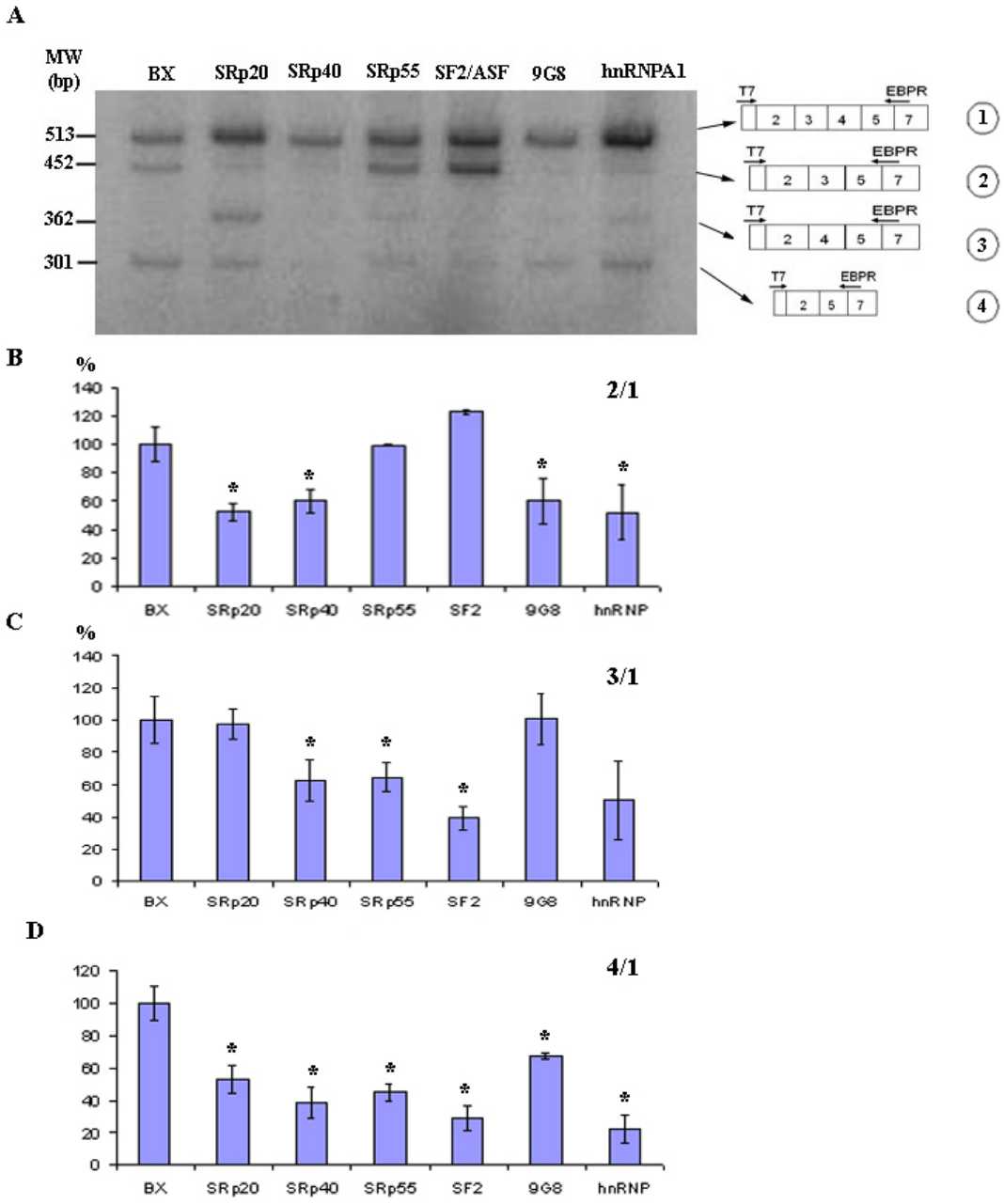


FIGURE 4.A. Semiquantitative RT-PCR using primers T7/EBP_R on cDNA from HeLa cells cotransfected with BX plasmid and one plasmid codifying one of the following proteins: ASF/SF2, SRp20, SRp40, SRp55, 9G8 and hnRNPA1. BX lane is the control cotransfected with BX plasmid plus an empty pUC vector. On the right, the transcripts corresponding to each band are drawn. **B, C and D**. Quantification of radio-labelled RT-PCR for the different transcripts (**B** for 2, **C** for 3 and **D** for 4) standardized by expression of transcript 1 and considering BX transfection as 100% in each case. Each value is the mean±S.E. of two independent experiments, each of which was replicated three times. * when $p<0.05$.

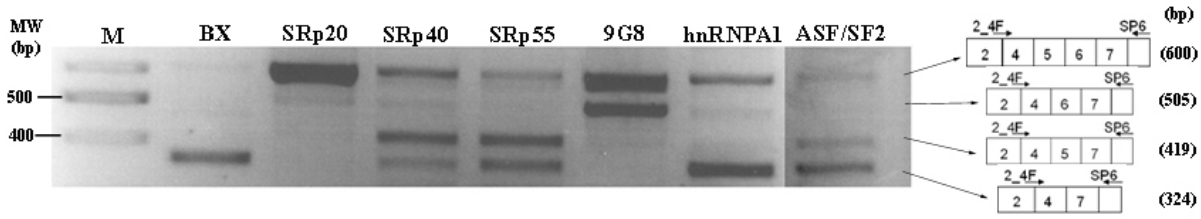


FIGURE 5. RT-PCR using primers 2_4F/SP6 on cDNA from HeLa cells cotransfected with BX plasmid and one plasmid codifying one of the following proteins: ASF/SF2, SRp20, SRp40, SRp55, 9G8 and hnRNPA1. On the right, the transcripts corresponding to each band are drawn. Not quantitative conditions.

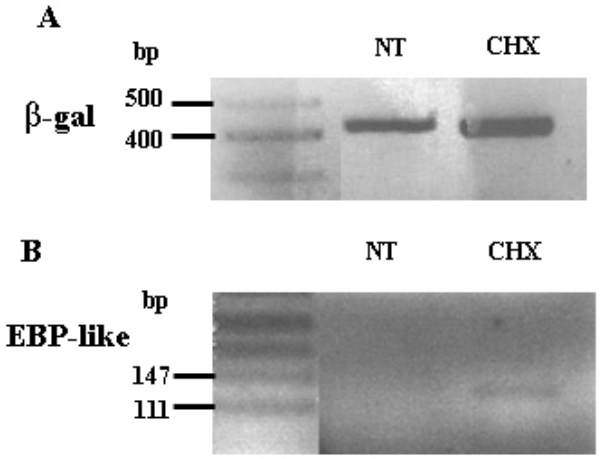


Fig.6 RT-PCR on cDNA from mouse RAW264.7 cells treated (CHX) or not treated (NT) with cycloheximide. **A**. PCR with primers C1Fm and C1Rm which only amplify the β-Gal transcript. **B.** PCR with primers EBPFm and EBPRm which only amplify the EBP-like transcript.

TABLE 1. Primers and RT-PCR conditions.

Primers employed in the RT-PCRs				RT-PCRs conditions	
Primer	Sequence (5'->3')	*GLB1* Exon[a]	Transcript specificity	Pairs of primers employed	AT (ºC)
C1F	ACTGCAGAGCCGGGAGGCTGGT	1	EBP/β-Gal	C1F/EBP_R	58
C3R	CCAAAGTGACCTTTCCATATG	11	EBP/β-Gal	EBP_F/C3R	56.5
EBP_F	GGCTGAACGCCATCCAGACATT	2-5	EBP	T7/SP6	55
EBP_R	TGATGTTGCTGCCTGCACTG	7-5	EBP	T7/EBP_R	56.5
2_4F	ACGCCATCCAGACGGAGGAT	2-4	NN	EBP_F/SP6	58
T7	TAATACGACTCACTATAGGG	--	pcDNA3	2_4F/C3R	59
SP6	GATTTAGGTGACACTATAG	--	pcDNA3	2_4F/SP6	58.3
C1Fm	GAGACCCCATCGTGGCGCGA	1	m EBP/β-Gal	C1Fm/C1Rm	59
C1Rm	CTCTAGTAGCCAAGCAGGTAAGC	4	m β-Gal	EBPFm/EBPRm	58
EBPFm	GGGCTGAATGCTATCCAGATAC	2-5	m EBP		
EBPRm	TGTGATATTGTTGCCTGCACGGT	7-5	m EBP		

AT: annealing temperature; NN: not naturally ocurring; m:mouse

[a]*GLB1* exon or exons containing sequences complementary to the indicated primers

TABLE 2. Ks of exon 5 in the β-gal reading frame.

Species compared	Ks
Human-macaque	0.00
Human-dog	0.33
Human-cat	0.20
Human-mouse	1.00
Human-rat	0.63
Human-rabbit	0.38
Human-chicken	0.56
Rabbit-mouse	0.85