Faculdade de Ciências
Universidade do Porto

Molecular, Biochemical and Functional Studies in Genes Determining Missorting of Lysosomal Proteins

Maria Francisca Coutinho

2013

Faculdade de Ciências da Universidade do Porto

Maria Francisca
de Lima Magriço Coutinho

Molecular, Biochemical and Functional Studies in Genes Determining Missorting of Lysosomal Proteins

Dissertação apresentada à Faculdade de Ciências da Universidade do Porto para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Biologia, realizada sob orientação da Doutora Sandra Alves, Investigadora Auxiliar do Departamento de Genética Humana, INSA, I.P e da Professora Doutora Maria João Prata, Professora Associada da Faculdade de Ciências da Universidade do Porto.

“What is a scientist after all?
It is a curious man looking through a keyhole,
the keyhole of nature, trying to know what's going on.”

(Jacques Yves Cousteau, 1910-1997)

„Alice looking behind the curtain”, original illustration by John Tenniel for Lewis Carroll´s “Alice in Wonderland”, 1866

Agradecimentos

A realização deste trabalho dependeu do interesse, boa vontade e disponibilidade de muitos. Não poderia, portanto, deixar de dizer a *todos* o meu mais sincero *"muito obrigada"*!

Em primeiro lugar, um agradecimento às instituições e respectivos chefes de departamento que me permitiram a realização deste trabalho: Instituto Nacional de Saúde Dr. Ricardo Jorge (INSA), na pessoa do Professor Pereira Miguel, como Director da Instituição. Agradeço ainda à Doutora Glória Isidro, Coordenadora do Departamento de Genética e ao Doutor João Lavinha, Coordenador da Unidade de Investigação e Desenvolvimento.

À Fundação para a Ciência e Tecnologia (FCT) pela atribuição de bolsa individual de doutoramento e financiamento destes estudos.

Gostaria de agradecer, de uma forma especial:

À Doutora Sandra Alves (CGMJM, INSA), minha companhia diária e meu apoio ao longo de todos estes anos... Muito tenho a agradecer, desde a confiança que em mim deposita à sua constante e inteira disponibilidade... E, um agradecimento muito particular, pelo carinho e amizade que sempre me demonstra e tanto me estimulam e fortalecem.

À Professora Doutora Maria João Prata (IPATIMUP; FCUP) pela atenção, disponibilidade e generosidade que sempre tem para comigo. Um "obrigada" muito especial pelo cuidado e minúcia com que leu cada artigo e, em particular esta tese, na intenção de sempre a valorizar. Nunca esquecerei, também, todo carinho e apoio que, em momentos de grande desalento, generosamente me prestou.

To Doctor Mirella Filocamo, for the opportunity to perform training at the "*Diagnosi Pre-Postnatale Malattie Metaboliche*" Laboratory, *Istituto Gianina Gaslini*, Genova. This was my first experience abroad ever and all the team was extremely nice with me. Special thanks to Barbara Tappino and, particularly, Stefano Regis for all the patience and support with the lab work. I will remember you all, forever. *Grazie mille!*

To Professor Thomas Braukle for the opportunity to work and learn in a top research lab and to all his team at the "Biochemistry Section, Children´s Hospital, University Medical Center Hamburg", who always made me feel welcome and supported. *Vielen Dank für alles!* Special thanks to Sandra Pohl, for super-coordination, Joahnnes Brand for super-technical assistance, and Raffaella De Pace, for continuing the studies after I came back to Portugal. To her, *piccolla Raffaellina*, and her husband, Francesco Massei, I owe another, greater acknowledgement, for receiving me at their place and making me fell home. *Grazie mille!*

À Doutora Sandra Ribeiro (IBMC) pela preciosa ajuda nos estudos *in silico*. Agradeço a pronta resposta ao sem-número de *e-mails* com que fui enchendo a sua caixa de correio sempre que me surgia uma dificuldade e a paciência com que ouviu as minhas tímidas "teorias" sobre o efeito de cada mutação.

Às meninas do "*muquifo*" e a todos os que por ele já passaram e com quem convivi diariamente ao longo de vários anos. De todos guardo boas recordações. Agradeço, principalmente, os momentos de descontração, sem os quais a minha cabeça já teria, certamente, estourado. Um obrigada muito especial à Carla Andreia, pela paciência e minúcia com que leu esta tese, e à Liliana Matos, mais eficiente do que o próprio EndNote a conferir referências.

À equipa de Investigação e Desenvolvimento e aos membros da antiga Unidade de Enzimologia do Centro de Genética Médica Doutor Jacinto de Magalhães (CGMJM) chefiada pela Doutora Lúcia Lacerda, por todo o apoio e colaboração.

Por último, e acima de tudo, à minha família. Aos meus pais, pela confiança que sempre depositaram em mim e no resultado de todo o meu esforço; e, muito especialmente à minha avó e tia, sem as quais não teria chegado até aqui.

Summary

Lysosomes are specialized cellular organelles for the degradation of endocytosed and intracellular material and essential regulators of cell homeostasis. Inside them, a multitude of soluble hydrolases ensures proper digestion of macromolecules. The importance of such organelles is well illustrated by the increasing number of human diseases related to defects in lysosomal functioning, the so-called lysosomal storage disorders (LSDs). To fulfill their degradative function lysosomes must firstly receive specific proteins, which after being synthesized in the endoplasmic reticulum (ER) need to be directed to the *trans*-Golgi network (TGN) for further processing and lysosomal targeting. The major pathway for lysosomal enzyme delivery, the M6P-dependent pathway, through which most enzymes reaches the lysosome, requires the recognition of a specific mannose 6-phosphate signal (M6P marker) selectively added to lysosomal enzymes in the TGN. Nevertheless, increasing evidence indicates the existence of additional or alternative pathways, M6P-independent, from the TGN to lysosomes. Presently, two alternative receptors which mediate that transport have been described: the lysosomal integral membrane protein LIMP-2 and sortilin.

In this work, we focused our attention on the genes that contribute for a proper sorting of lysosomal enzymes, by coding for proteins which are directly involved in the above-referred transport pathways. The work is divided in two parts, each having specific aims. The first part addresses the molecular genetics and diagnosis of disorders related with malfunction of M6P pathway and the second reports on the studies addressing the role of the M6P-independent trafficking pathways in LSDs.

Impairments in the M6P pathway are long known to cause two well-defined LSDs: Mucolipidosis types II (MLII) and III (MLIII), which are caused by total or partial loss of GlcNAc-phosphotransferase activity, respectively. This enzyme is responsible for the first catalytic step leading to the formation of the M6P recognition marker. Far from being a simple protein, GlcNAc-phosphotransferase is a hexameric transmembrane enzyme, composed by three different subunits, $\alpha_2\beta_2\gamma_2$, encoded by two independent genes: *GNPTAB*, which codes for the aminoacidic precursor of the α- and β-subunits and *GNPTG*, which codes for the γ-subunit. Depending on the gene which is affected in each patient, mucolipidoses may be classified as alpha/beta (mutations in the *GNPTAB* gene) or gamma (mutations in the *GNPTG* gene). Even though ML II and ML III share similar clinical features, including skeletal abnormalities, ML II is the more severe in terms of phenotype with early onset of symptoms and death occurring, usually, between 5 and 8 years of age. Mutations in *GNPTAB* may underlie either ML II alpha/beta or ML III alpha/beta while defects in the *GNPTG* gene are only causative of ML III gamma.

For this study, a set of 23 unrelated ML II and III cases from several origins was screened, leading to the identification of 18 different mutations: 16 in the *GNPTAB* gene and 2 in the *GNPTG* gene. Of those, 13 were novel: 11 in the *GNPTAB* gene and 2 in the *GNPTG* gene. All mutations identified were further analyzed, paying special attention to the novel alterations. Whenever RNA samples were available real time PCR studies were performed, whose results suggested the possible existence of feedback regulation mechanisms between the α/β- and the γ-subunits.

From the analyzed series of patients, some had particularly unusual molecular lesions and/or phenotypes, which triggered additional analyses.
One of those patients was shown to harbor a homozygous genomic lesion in *GNPTAB* leading to the entire loss of exon 19 and some of its surrounding intronic regions (c.3435-386_3602+343del897). Analysis of the deletion breakpoints indicated that the mutation was generated by an unequal homologous recombination process between highly similar *Alu* elements located in two regions within the gene where such kind of sequences were demonstrated to be present. Subsequent cDNA analysis revealed the presence of three abnormal transcripts created by this gross deletion at mRNA level: one without exon 19 (p.Lys1146_Trp1201del); another with an additional loss of exon 20 (p.Arg1145Ser*fs*X2), and a third in which exon 19 was substituted by a pseudoexon inclusion consisting of a 62 bp fragment from intron 18 (p.Arg1145Ser*fs*X16). No other large deletions have been reported in the *GNPTAB* gene so far.
Another patient had a peculiar phenotype, with severe prenatal skeletal abnormalities that fitted prior descriptions of a rare phenotype known as Pacman dysplasia. Retrospective analysis of the parents, inferred him to be heterozygous for two pathogenic mutations in the *GNPTAB* gene: the well-known c.3503_3504delTC (p.L1168Q*fs*X5) and a novel frameshift variant, c.1701delC (p.F566L*fs*X5), thus supporting previous assumptions that at least some cases of Pacman dysplasia phenotype are, in fact, severe prenatal forms of ML II alpha/beta.

Whenever possible, genotype-phenotype correlations were established. For the majority of patients analysed, assessment of severity was rather straightforward, with a strong correlation between the observed clinical phenotype and the type of molecular lesion assessed through molecular analysis. Nevertheless, the identification of three patients, carrying homozygous missense mutations but presenting with early onset severe symptoms (two patients with p.W81L and one with p.R986C), defied previous assumptions that such mutations would underlie a mild phenotype, while only nonsense and frameshift mutations would give rise to the more severe ML II alpha/beta phenotype. To explore the causes underlying these unusual genotype/phenotype relationships further analyses were performed. A panel of *GNPTAB* mutations, including the atypically severe missense ones, was expressed in model cell lines and analyzed through Western blot and Immunofluorescence. In general, the combined analysis of the mutants' expression levels and subcellular location appeared to explain the associated phenotypes: mutants harboring the severe deletions revealed to be

retained in the ER, presenting only the non-cleaved α/β-precursor form; mild missense mutations were correctly located in the Golgi apparatus but presented reduced levels of the mature α- and β-subunits, when compared to the wild-type and, finally, severe missense mutations were either retained in the ER on the non-cleaved and inactive precursor form (p.W81L) or not expressed at all (p.R986C). The particularly severe consequences demonstrated for the aminoacid substitutions p.W81L and p.R986C, additionally suggested that residues W^{81} and especially R^{986} may assume crucial roles on the function and/or stability of GlcNAc-phosphotransferase.

The observation that the *GNPTAB* deletion c.3503_3504delTC, already reported in several populations worldwide, was the most common disease allele among the Portuguese patients, has instigated us to perform a haplotypic analysis to investigate its origin. A total of 44 patients and 16 carriers of this deletion, from different geographic regions, was analyzed for 3 intragenic polymorphisms and 2 microsatellite markers flanking the *GNPTAB* gene. A common haplotype was identified in all chromosomes bearing the deletion, providing evidence that c.3503_3504delTC was due to a unique founder molecular lesion. Additionally, the level of diversity observed at the most distant microssatelite indicated it to be a relatively ancient mutation, whilst its geographical distribution further suggested it to have arisen in a peri-Mediterranean region.

Recently, impairments of the M6P-independent travelling route were also associated with disease. From the two alternative receptors that have been described to be involved in lysosomal transport (LIMP-2 and sortilin), only LIMP-2 has been associated with disease. LIMP-2 is responsible for the delivery of β-glucocerebrosidase (GCase, the defective enzyme in patients with Gaucher disease) to the lysosomes. Mutations in *SCARB2*, the gene that codes for this receptor, have been reported to underlie a serious autosomal genetic disorder presently known as action myoclonus renal failure (AMRF). Regarding sortilin (encoded by *SORT1*), which has been suggested to mediate transport of the sphingolipid activator proteins (SAPs) prosaposin (PSAP) and GM2 activator protein (GM2AP); acid sphingomyelinase (ASM) and cathepsins D and H, no storage disorder caused by its loss of function has been reported yet.

In the second part of this work, the contribution of defects in the genes that code for proteins involved in the M6P-independent trafficking pathways was addressed. A sample of patients with symptoms suggestive of LSD but without definitive diagnose after routine tests was analysed for genetic impairments in *SCARB2* and *SORT1* genes. Attention was focused in uncharacterised patients whose clinical manifestations overlapped those that could predictably be due to loss of function of GCase, saposins, GM2AP and/or ASM which, theoretically, could be resultant from failures in M6P-independent pathways through which those proteins reach the lysosome.

In a set of 120 individuals with clinical suspicion of LSD but without definitive biochemical and/or molecular diagnosis, no novel mutations were detected either on the *SCARB2* or on

the *SORT1* genes. So far, no evidence came that *SORT1* deficiencies may be associated to LSD phenotypes. Still, enlarged sample sizes are needed in order to draw more reliable conclusions on the topic. Most probably, final conclusions may only be drawn when individuals carrying gene variants which disrupt protein function are described and their phenotype evaluated.

Finally, a complementary point was also addressed: the role of variations in *SCARB2* on the broad phenotype spectrum observed for Gaucher disease (GD) patients. GD has a huge phenotypic range and, even though some specific genotype-phenotype correlations have been established for a few of its causing mutations, the pathological cascade leading to the great variety of phenotypes among GD patients remains elusive, being still unclear why patients carrying identical *GBA* genotypes present with disparate phenotypes. Being the most common LSD and the first to have available enzyme replacement therapy (ERT), GD has been extensively studied over the years but, still, little is known on the factors that modify its associated phenotypes. Recently, mutations in the LIMP-2 coding gene were also reported to affect the severity of GD phenotype. To evaluate the role of *SCARB2* mutations on the GD phenotype, the whole cohort of Portuguese GD patients (which totalizes 91 individuals) was screened and 1 novel mutation in the *SCARB2* gene identified, reinforcing previous evidence that mutations in the gene that codes for the GCase transporter can act as GD modifiers. In the particular case of the Portuguese patient in whom this variant was identified, even though the clinical phenotype could, theoretically, be explained solely by the *GBA* genotype, the child had biochemical evidences suggesting poor response to treatment, which may be justified by the additional GCase carrier deficiency. From the results obtained in this study, however, it became clear that mutations in *SCARB2* are not frequent modifiers of the GD phenotype; other genetic, epigenetic and/or environmental factors must exist that influence GD phenotypic manifestations.

In general, this study reinforces the biological importance of a proper targeting of lysosomal proteins to their final destination, by highlighting the severe consequences of impairments in proteins involved in that process.

Key Words

Lysosomal targeting; Mucolipidosis type II (ML II); Mucolipidosis type III (ML III); Action Myclonus Renal Failure (AMRF); Gaucher Disease (GD).

Resumo*

Os lisossomas são organelos celulares especializados na degradação de material intracelular ou endocitado, desempenhando também um papel essencial na regulação da homeostasia celular. No seu interior, encontra-se um grande número de hidrolases solúveis que assegura uma eficiente digestão de todos os tipos de macromoléculas. A importância destes organelos é demonstrada pelo número crescente de doenças genéticas relacionadas com o deficiente funcionamento lisossomal, as denominadas doenças lisossomais de sobrecarga (DLS). De forma a cumprir as suas funções degradativas, os lisossomas têm de receber enzimas específicas que, após serem sintetizadas no retículo endoplasmático (RE), precisam de ser dirigidas até à face *trans* do complexo de Golgi para posterior processamento e transporte para o lisossoma. A principal via de transporte das enzimas lisossomais, a via dependente da manose 6-fosfato (M6P), por meio da qual a maioria das enzimas atinge o lisossoma, requer o reconhecimento de um sinal específico de manose-6-fosfato (marcador de M6P), que é selectivamente adicionado às enzimas lisossomais na face trans do Golgi. Recentemente, surgiram estudos revelando a existência de vias adicionais ou alternativas, independentes da M6P, de transporte para os lisossomas. Até ao momento, conhecem-se já dois receptores alternativos capazes de mediar esse transporte: a proteína integral da membrana lisossomal LIMP-2 e a sortilina.

Este trabalho teve como principal alvo de estudo os genes que contribuem para o correcto transporte das enzimas lisossomais, isto é, aqueles que codificam proteínas diretamente envolvidas nas vias de transporte acima referidas. O trabalho foi estruturado em duas partes, cada uma apresentando objectivos específicos. A primeira aborda a genética molecular e o diagnóstico das doenças relacionadas com disfunção da via da M6P enquanto a segunda envolve estudos sobre o possível papel das vias independentes de M6P no desenvolvimento de DLS.

Deficiências ao nível da via da M6P são há muito conhecidas, estando associadas a duas DLS bem definidas: as mucolipidoses tipo II (MLII) e III (MLIII) causadas, respectivamente, pela perda total ou parcial de actividade da enzima GlcNAc-fosfotransferase. Esta enzima é responsável pelo primeiro de dois passos catalíticos que resultam na formação do marcador de reconhecimento M6P. A GlcNAc-fosfotransferase é uma enzima transmembranar hexamérica, composta por três subunidades diferentes, $\alpha_2\beta_2\gamma_2$, codificadas por dois genes independentes: *GNPTAB*, codificador do precursor aminoacídico das subunidades α e β e *GNPTG*, que codifica a subunidade γ. Dependendo do gene que é afectado, as Mucolipidoses podem ser classificados como alfa/beta (mutações no gene *GNPTAB*) ou gama (mutações no gene *GNPTG*). Embora ML II e ML III partilhem características clínicas semelhantes, incluindo anomalias esqueléticas, a ML II é mais grave em termos de fenótipo, sendo caracterizada pela

precocidade dos sintomas e morte prematura dos doentes, geralmente entre os 5 e os 8 anos de idade. Mutações no gene *GNPTAB* podem estar na base de ML II alfa/beta e ML III alfa/beta, enquanto que mutações no gene *GNPTG* apenas originam ML III gama.

Entre os principais objectivos deste trabalho incluem-se a identificação dos defeitos moleculares causadores de ML II e III nos doentes referenciados da população Portuguesa, a avaliação dos efeitos de cada mutação ao nível da proteína/mRNA, e ainda o estudo da relação entre presença de uma dada mutação e a sintomatologia clínica.

No decurso do trabalho, para além de todos os doentes portugueses conhecidos, foi ainda possível estudar vários doentes de outras origens, por terem sido referenciados ao nosso grupo para estudo molecular. No total, foi caracterizada uma amostra de 23 casos não aparentados de ML II e III, de várias origens, que conduziu à identificação de 18 mutações diferentes: 16 no gene *GNPTAB* e 2 no gene *GNPTG*. Dessas, 13 foram descritas pela primeira vez: 11 no gene *GNPTAB* e 2 no gene *GNPTG*. Todas as mutações identificadas, particularmente as novas, foram estudadas utilizando diversas abordagens. De referir, entre outras, a realização de estudos PCR em tempo real em amostras de RNA, disponíveis para a maior parte dos casos, cujos resultados sugeriram a possível existência de mecanismos de feedback na regulação da expressão das subunidades α/β e γ.

Alguns dos doentes analisados apresentavam lesões moleculares e/ou fenótipos particularmente invulgares, o que desencadeou análises dirigidas.

Num desses doentes, de origem palestiniana e com forma grave de MLII alfa/beta, a caracterização molecular levou à identificação em homozigotia de uma grande deleção, a única até agora conhecida no gene *GNPTAB,* que envolvia a perda de todo o exão 19 e algumas das suas regiões intrónicas adjacentes (c.3435-386_3602 + 343del897). A análise dos pontos de quebra da deleção permitiu deduzir o mecanismo explicativo subjacente ao seu aparecimento, podendo estar relacionado com um processo de recombinação homóloga desigual entre elementos *Alu* com elevado grau de homologia, localizados em duas das três regiões do gene onde se verificou estarem presentes este tipo de sequências. Posteriormente, uma análise do RNA do mesmo doente permitiu encontrar três transcritos anormais, resultantes, assim, da presença desta grande deleção: um, sem o exão 19 (p.Lys1146_Trp1201del); outro, com uma perda adicional do exão 20 (p.Arg1145Ser*fs*X2); e, um terceiro, no qual o exão 19 foi substituído pela inclusão de um pseudoexão que consistia num fragmento de 62 pb do intrão 18 (p.Arg1145Ser*fs*X16).

Outro doente, de origem indiana, também apresentava um fenótipo peculiar, muito grave (morte ao quarto dia de vida), envolvendo anomalias esqueléticas pré-natais severas compatíveis com descrições anteriores de um fenótipo raro conhecido como displasia de Pacman. A análise dos seus pais permitiu inferir retrospectivamente que o doente seria heterozigótico composto para duas mutações patogénicas no gene *GNPTAB*: a deleção c.3503_3504delTC (p.L1168Q*fs*X5), já bem conhecida, e uma nova deleção, c.1701delC (p.F566L*fs*X5). Este resultado apoia hipóteses anteriores de que, pelo menos alguns casos de displasia de Pacman seriam, de facto, formas graves e pré-natais de ML II alfa/beta.

Sempre que possível, foram investigadas correlações genótipo-fenótipo. Para a maioria dos doentes, a avaliação do potencial patogénico de cada uma das mutações foi bastante coerente, tendo-se observado uma forte correlação entre o fenótipo clínico observado e o tipo de lesão molecular previamente identificada. No entanto, em três doentes, portadores de mutações *missense* em homozigotia, cujos sintomas, além de graves, surgiram muito precocemente (dois doentes p.W81L/p.W81L e um p.R986C/p.R986C), os seus fenótipos contrariavam pressupostos anteriores de que mutações do tipo das que apresentavam, *missense*, estariam associadas a fenótipos leves, enquanto apenas mutações *nonsense* e *frameshift* dariam origem aos fenótipos mais graves de ML II α/β. Para explorar as causas subjacentes a estas relações genótipo/fenótipo, expressou-se, em linhas celulares modelo (HEK e HeLa), um painel de mutações *GNPTAB*, incluindo as duas *missense* atipicamente graves, que posteriormente foram sujeitas a análise de Western Blot e Imunofluorescência. Como controlo, expressaram-se também duas deleções graves e três mutações missense associadas a fenótipos leves. No global, a análise combinada dos níveis de expressão e localização subcelular das proteínas mutantes, pareceu dar, de forma muito satisfatória, sentido aos fenótipos que lhes estavam associados: as deleções graves ficavam retidas no retículo endoplasmático (RE), apresentando apenas o precursor α/β não-clivado; as mutações *missense* leves ficavam correctamente localizadas no complexo de Golgi mas apresentavam níveis reduzidos das subunidades α e β maduras, quando comparados com o *wildtype*. Finalmente, as mutações *missense* graves, ou ficavam retidas no RE na forma precursora, não-clivada e inactiva (p.W81L), ou não eram expressas de todo (p.R986C).
Atendendo às consequências particularmente graves associadas às substituições aminoacídicas p.W81L e p.R986C, é de admitir que o resíduo W^{81} e, especialmente o resíduo R^{986}, possam assumir papeis chave na função e/ou estabilidade da GlcNAc-fosfotransferase.

A observação de que a deleção c.3503_3504delTC no gene *GNPTAB*, já descrita em várias populações a nível mundial, era o alelo mais comum entre os doentes portugueses, levou à realização de um estudo haplotípico com o objectivo de investigar a sua origem. Foi possível recrutar um total de 44 doentes e 16 portadores da deleção, provenientes de diferentes regiões geográficas, tendo sido analisados para três SNPs intragénicos e 2 microssatélites que flanqueiam o gene *GNPTAB*. A detecção de um haplótipo comum em todos os cromossomas portadores da deleção, constituiu forte indicação de que a mutação c.3503_3504delTC foi originada por uma lesão molecular fundadora única. Por outro lado, se o nível de diversidade associado a um dos microssatélites evidenciou tratar-se de uma mutação relativamente antiga, a sua distribuição geográfica apontou para a região peri-mediterrânica como sendo aquela onde provavelmente a mutação surgiu.

Recentemente, deficiências ao nível da via de transporte independente da M6P foram também relacionadas com consequências patogénicas. No entanto, dos dois receptores alternativos já descritos (LIMP-2 e sortilina), apenas a perda de função da LIMP-2 foi inequivocamente associada a doença. A LIMP-2 é responsável pelo transporte da enzima β-

glucocerebrosidase (cuja deficiência está associada a doença de Gaucher), para os lisossomas. Mutações em *SCARB2*, gene que codifica esta proteína, têm sido descritas como causais de uma doença genética autossómica grave, actualmente conhecida como epilepsia mioclónica com síndrome nefrótico (*action myoclonus renal failure*, AMRF). Relativamente à sortilina (codificada pelo gene *SORT1*), que se pensa mediar o transporte de proteínas activadoras de esfingolípidos (*sphingolipid activator proteins*, SAPs) prosaposina (PSAP) e proteína activadora de GM2 (*GM2 activator protein*, GM2AP); esfingomielinase ácida (*acid sphingomyelinase*, ASM) e das catepsinas D e H, nenhuma doença foi até agora conotada com a sua perda de função.
Na segunda parte deste trabalho, abordámos a possível contribuição de alterações nos genes que codificam estes dois receptores alternativos para a génese das doenças lisossomais de sobrecarga. Foram pesquisadas mutações nos genes *SCARB2* e *SORT1* numa amostra constituída por doentes com sintomas sugestivos de DLS mas carecendo de diagnóstico definitivo após realização dos exames de rotina. Entre os doentes nesta situação, foram seleccionados aqueles cujas manifestações clínicas eram sobreponíveis às espectáveis por perda de função de β-glucocerebrosidase (GCase), saposinas, GM2AP e/ou ASM, partindo da hipótese de que os respectivos fenótipos poderiam, teoricamente, ser resultantes de falhas no transporte destas proteínas para o lisossoma.
Porém, como num total de 120 doentes com suspeita clínica de DLS, mas sem diagnóstico bioquímico e/ou molecular definitivo, não foram detectadas mutações novas quer no gene *SCARB2* quer no gene *SORT1,* até ao momento não se encontraram sinais de que alterações no gene *SORT1* possam estar associadas a fenótipos de DLS. Sendo ainda necessário ampliar o tamanho da amostra, conclusões mais fiáveis provavelmente só poderão ser tiradas quando indivíduos portadores de variantes genéticas que perturbem a função da proteína forem encontrados e o seu fenótipo cuidadosamente avaliado.

Finalmente, foi ainda abordado um ponto complementar: o papel de variantes no gene *SCARB2* no espectro fenotípico da doença de Gaucher (*Gaucher disease*, GD). A GD apresenta uma enorme variabilidade fenotípica. Embora já tenham sido estabelecidas algumas correlações genótipo-fenótipo para certas mutações causais, a cascata patológica que leva à grande variabilidade de fenótipos entre doentes com GD não está completamente definida, estando por esclarecer a razão pela qual doentes com o mesmo genótipo *GBA* apresentam fenótipos distintos. Sendo a DLS mais comum e a primeira a ter terapia de substituição enzimática disponível (*enzyme replacement therapy*, ERT), a GD tem sido extensivamente estudada ao longo dos anos mas, ainda assim, pouco se sabe sobre os fatores que contribuem para a variabilidade observada ao nível dos fenótipos dos doentes. Recentemente, surgiram indícios de que mutações no gene que codifica LIMP-2 podem também afetar a gravidade do fenótipo de GD. Para avaliar o papel das mutações *SCARB2* no fenótipo de GD, analisou-se a população de doentes portugueses com GD (91 indivíduos), tendo sido identificada uma nova mutação no gene *SCARB2*, reforçando assim evidências anteriores de que mutações no gene que codifica o transportador da GCase podem funcionar como modificadores de GD. No caso particular do doente português em que a alteração em *SCARB2* foi encontrada, embora o

fenótipo clínico pudesse, aparentemente, ser explicado pelo respectivo genótipo *GBA*, a criança apresentava indícios bioquímicos sugestivos de fraca resposta ao tratamento, o que poderá ser justificado pela deficiência adicional ao nível do transportador. Este estudo deixa, no entanto, claro que mutações no gene *SCARB2* não são modificadores frequentes do fenótipo de GD, devendo existir outros fatores genéticos, epigenéticos e/ou ambientais que influenciem mais as manifestações fenotípicas de GD.

Em termos gerais, este trabalho reforça a importância biológica do correcto transporte das proteínas lisossomais até ao destino apropriado, ao destacar as graves consequências que advêm de deficiências em proteínas envolvidas nesse processo.

Palavras Chave

Transporte lisossomal; Mucolipidose tipo II (ML II); Mucolipidose tipo III (ML III); Epilepsia Mioclónica com Síndrome Nefrótico (AMRF); Doença de Gaucher (GD).

* O autor escreve de acordo com a antiga ortografia.

Abbreviation List

α-gal	α-galactosidase
β-gal	β-galactosidase
β-gluc	β-glucocerebrosidase
α-man	α-mannosidase
ALS	amyotrophic lateral sclerosis
AMRF	action myoclonus renal failure
apo	apolipoprotein
ASM	acid sphingomyelinase
AV	autophagic vacuole
BGT	biochemical genetic testing
BMT	bone marrow transplantation
CAD	coronary artery disease
CD-MPR	cation-dependent mannose-6-phosphate receptor
CESD	cholesteryl ester storage disease
CI-MPR	cation-independent mannose-6-phosphate receptor
CNS	central nervous system
CO-IP	co-immunoprecipitation
EE	early endosome
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum-associated degradation
ERT	enzyme replacement therapy
ES	embryonic stem cells
ESE	exon-splicing enhancer
ESS	exon-splicing supressor
FTLD	frontotemporal lobar degeneration
GCase	β-glucocerebrosidase
GD	Gaucher disease
GM2AP	GM2 activator protein
GWAS	genome wide association studies
HDL	high density lipoprotein
Hex A	β-hexosaminidase A
HexT	hexosaminidase T
hMSC	human mesenchymal stem cell
IGF	insulin-like growth factor
KO	knock-out

LDL	low density lipoprotein
LDLR	low density lipoprotein receptor
LE	late endosome
LIMP-2	lysosomal integral membrane protein 2
LINE	long interspersed elements
LPL	lipoprotein lipase
LSD	lysosomal storage diseases
M6P	mannose-6-phosphate
MGT	molecular genetic testing
MI	myocardial infarction
ML	Mucolipidosis
MPR	mannose-6-phosphate receptor
MPS	Mucopolysaccharidosis
MRH	mannose-6-phosphate receptor homology domain
NCL	Neuronal Ceroid Lipofuscinosis
NGF	nerve growth factor
NMD	nonsense mediated mRNA decay
NPC	Niemann-Pick C
ORF	open reading frame
OTCD	ornithine transcarbamylase deficiency
PD	Pacman dysplasia
PGRN	progranulin
PME	progressive myoclonus epilepsy
PNGaseF	peptide N-glycosidase F
PSAP	prosaposin
PTC	premature termination codon
qRT-PCR	quantitative real time polymerase chain reaction
RAP	receptor-associated protein
S1P	site-1-protease
SAP	sphingolipid activator protein
SINE	short interspersed element
SNP	single nucleotide polymorphism
TGN	*trans*-Golgi network
UCE	uncovering enzyme
VHL	Von Hippel-Lindau
VLDL	very low density lipoprotein
WT	wild type

Chapter 1

Introduction

1. LYSOSOME AND LYSOSOMAL STORAGE DISORDERS

1.1. THE LYSOSOME

Discovered in 1955 by Christian René de Duve as membrane-closed acidic compartments with acid phosphatase activity [De Duve *et al.*, 1955], lysosomes (Figure 1.1) are key organelles in the regulation of cellular homeostasis, being responsible for the degradation of a multitude of proteins. Lysosomes contain more than 60 soluble hydrolases (e.g. proteases, phosphatases, glycosidases, lipases, nucleases, sulfatases) that are capable of degrading macromolecules and even membranes into their monomeric components [reviewed in Schröder *et al.*, 2010].

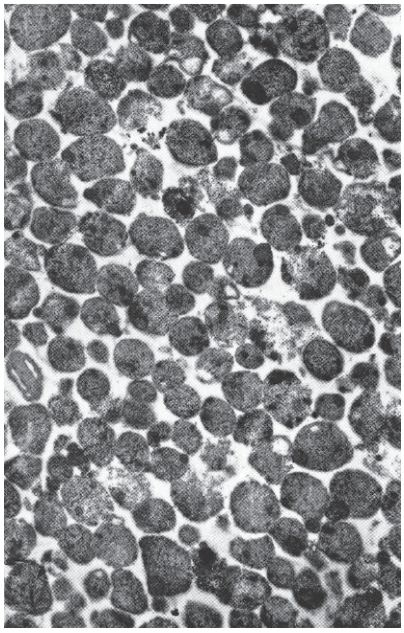


Figure 1.1: Purified lysosomes from rat liver, ×18,400.

[de Duve, 2005, reproduced from *The Journal of Cell Biology* (**26,** 219–243; 1965)]

Lysosomes receive their substrates through different internalization pathways: endocytosis, phagocytosis and autophagy (Figure 1.2) and are involved in a wide range of physiological functions from cell death and signaling to cholesterol homeostasis and plasma membrane repair [Saftig and Klumperman, 2009]. Lysosomal hydrolases are confined in the lumen of the lysosome and require an optimum pH (i.e., pH 4.5) to work. This acidic pH compared with the slightly alkaline pH of the cytosol (i.e., ~pH 7.2) is maintained by the activity of lysosomal integral membrane proteins (LIMPs, that represent the second class of lysosomal proteins), including the V-type proton (H+)-ATPase [Williamson and Hiesinger, 2010] and the chloride ion channel CLC74 that pumps protons from the cytosol across the lysosomal membrane [Weiss, 2012; reviewed in Schröder *et al.*, 2010].

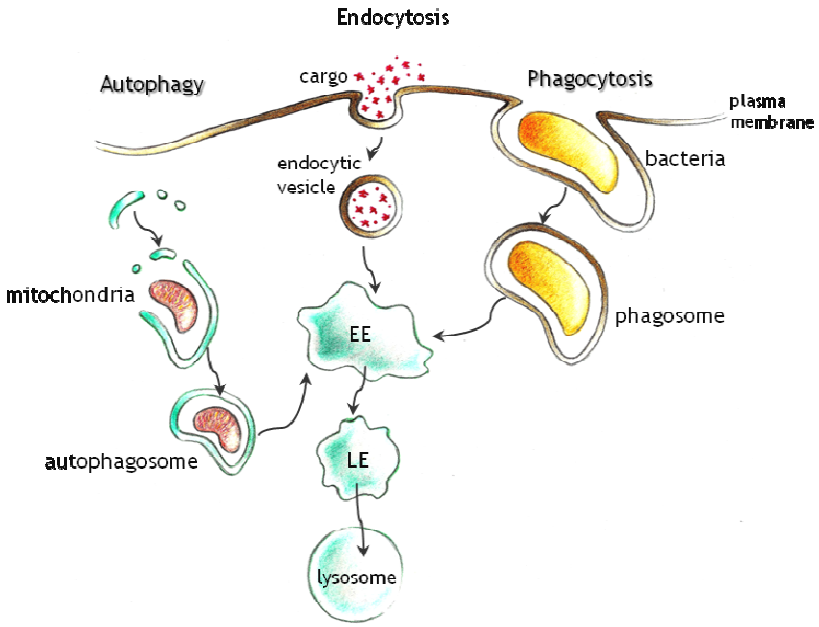


Figure 1.2: Intracellular pathways to the lysosome.

[*adapted from* Pryor and Luzio, 2009]

Endocytic cargo is internalised from the plasma membrane and delivered to early endosomes. Maturation of the early endosome gives rise to a late endosome/multi-vesicular body where the cargo destined for degradation has been sorted into intralumenal vesicles. The limiting membranes of late endosomes and lysosomes can fuse to form a hybrid organelle where degradation of endocytosed macromolecules commences. Lysosomes are reformed from the hybrid organelle by membrane retrieval and condensation reactions. Autophagic and phagocytic pathways both feed into the endocytic pathway with the lumenal contents of the autophagosome and phagosome eventually being delivered to the lysosome for degradation. (EE, early endosome; LE, late endosome)

The biological relevance of lysosomes and of the repertoire of degradative enzymes they harbor is illustrated by the existence of over 50 different genetic disorders caused by a deficiency of a catabolic enzyme [Lübke *et al.*, 2009] that as a whole are estimated to affect 1:4000-9000 live births [Michelakakis *et al.*, 1995; Poorthius *et al.*, 1999; Meikle *et al.*, 1999; Poupětová *et al.*, 2010; Pinto *et al.*, 2004]: the lysosomal storage diseases (LSDs).

1.2. LYSOSOMAL STORAGE DISORDERS

1.2.1. General description

Most LSDs result from a deficiency of specific lysosomal enzymes. In a few cases, though, non-enzymatic lysosomal proteins or non-lysosomal proteins involved in lysosomal biogenesis are deficient [reviewed in Futerman and van Meer, 2004; Filocamo and Morrone, 2011]. Typically, LSDs are recessively inherited, monogenic and progressive diseases [reviewed in Futerman and van Meer, 2004; Filocamo and Morrone, 2011]. While the majority of LSDs is autosomal, a few also exist, such as MPS type II (Hunter syndrome) and Fabry disease,

which are X-linked recessive genetic disorders.

The common biochemical hallmark of LSDs is the accumulation of un- or partially digested metabolites in the lysosome. This can arise through several mechanisms as a result of defects in any process underlying the lysosomal biology that hampers the catabolism of molecules in the lysosome, or the egress of naturally occurring molecules from the organelle: 1) defects in either glycosaminoglycan, lipid, or protein degradation; 2) transport across the lysosomal membrane; or 3) endosome-lysosome trafficking. Actually, any disruption of lysosomal function can lead to the accumulation of undegraded substrate(s) in endosomes and lysosomes, eventually compromising cellular function [Parkinson-Lawrence *et al.*, 2010]. Lysosomal accumulation activates a variety of pathogenetic cascades that result in complex clinical pictures characterized by multi-systemic involvement [Wraith, 2002; Futerman and van Meer, 2004; Vellodi, 2005; Ballabio and Gieselmann, 2009; Vitner *et al.*, 2010]. Presently, and even though some models have been suggested (Figure 1.3), we still lack a clear picture of the relevant events linking disease causing mutations to the symptoms of the disease which are determined by mechanisms operating not only at the cellular level, but also in tissues and organs. So, one of the greatest challenges is to understand how substrate storage impacts on function of cells, tissues and organs, causing disease pathogenesis.

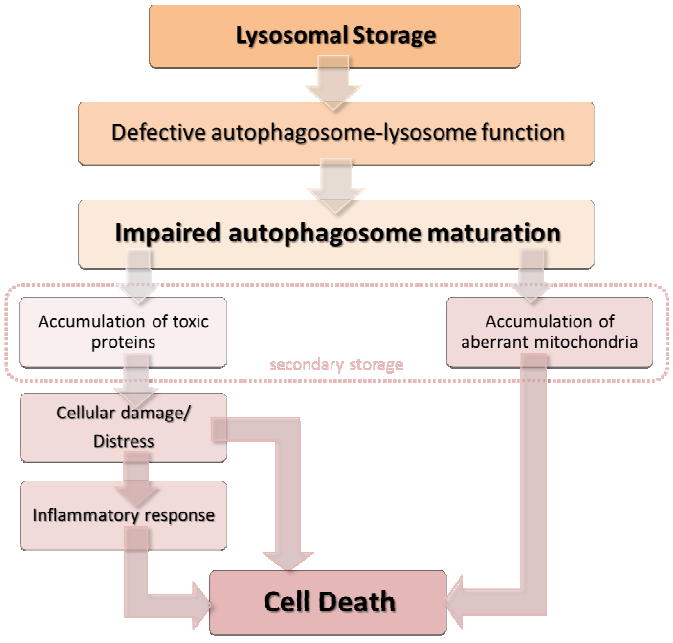


Figure 1.3: A proposed model for the pathogenesis of LSDs.

[*adapted from* Ballabio and Gieselmann, 2009]

Lysosomal storage leads to a reduction on the ability of lysosomes to fuse with autophagosomes. Consequently, there is a block (either total or partial) of autophagy maturation and defective degradation. As a result, it occurs an accumulation of authophagy substrates such as polyubiquitinated protein aggregates and dysfunctional mitochondria, promoting cell death, which is further triggered by the inflammatory response to cell damage.

Phenotypic expression is extremely variable both in terms of the clinical manifestations and of their severity, as it depends on the specific macromolecule which is accumulated, the site of production and degradation of the specific metabolites, the residual enzymatic activity and the general genetic background of the patient (*See section 1.2.3*).
In general, over the last 25 years research on LSDs has been focused on the genetics of this group of disorders. During this period the majority of lysosomal enzymes-coding genes was cloned and a huge number of disease-causing mutations identified. Also our vision of the lysosome has changed tremendously. The lysosome is no longer viewed as just an end-point degradative compartment but rather as part of a very complex and interactive set of intracellular organelles that have a wide array of specialist functions. Lysosomes are integrally involved in phagocytosis, autophagy, exocytosis, receptor recycling and regulation, intracellular signaling, immunity, pigmentation, bone biology, and neurotransmission (Figure 1.4).
This new view on the lysosome role brought the notion that defects in any one of the processes in which it is involved can be associated with lysosomal disease, leading then to the recognition that LSDs may be better viewed as conditions of molecular deprivation that result in specific pathogenic cascades [Walkley, 2009; Walkley and Vanier, 2009; Parkinson-Lawrence *et al.*, 2010].

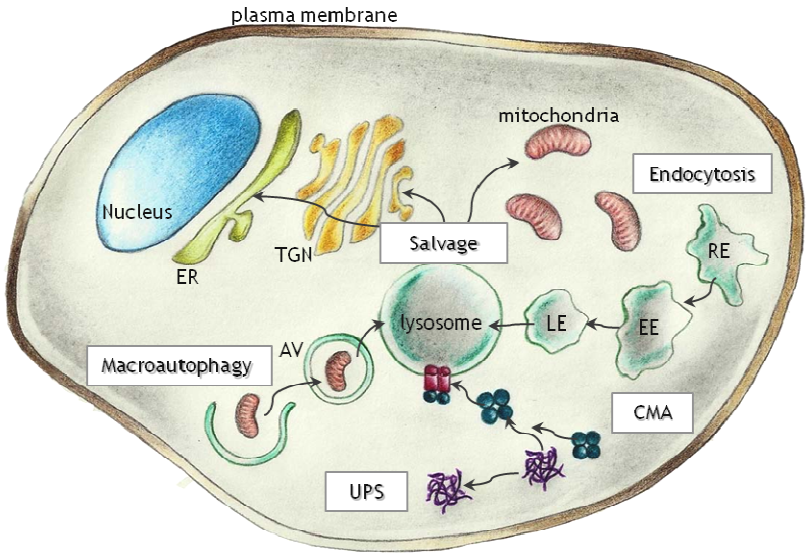


Figure 1.4: The lysosome as a central element of the cell.

[*adapted from* Walkley and Vanier, 2009]

Lysosomes link to the endocytic pathways, to the macroautophagy stream and its close allies, the ubiquitin-proteosomal system (UPS) and the chaperone-mediated autophagy (CMA) components. What flows into this system must also leave in some form, depicted here as the salvage pathway with delivery to the *trans*-Golgi network (TGN), mitochondria, and other sites in the cell. The complexity of the disease cascades in lysosomal disorders is conjectured to emanate in part from disruption of these interrelated components of the greater lysosomal system. (EE, early endosome; RE, recycling endosome; LE, late endosome; AV, autophagic vacuole; ER, endoplasmic reticulum).

Ultimately, a better comprehension of pathogenic pathways will likely afford clues to identify therapeutic targets which, in the future, can be used for the benefit of the patients [Ballabio and Gieselmann, 2009].

To date, no worldwide epidemiological data are available on LSDs. There are, however, several studies in distinct populations focusing either LSDs as a group or some specific LSDs which are known to present a high prevalence in specific populations such as the Ashkenazi Jewish population, which is at high risk for Gaucher disease [Beuttler and Gradowsky, 2001], Tay–Sachs disease and Niemann–Pick disease [Vallance and Ford, 2003], or the Finnish population which is characterized by high incidence of aspartylglucosaminuria [Arvio *et al.*, 1993] and infantile/juvenile NCLs [Santavuori, 1988]. Reports on LSDs as a whole are only available for a few countries: Greece (Michelakakis *et al.*, 1995), The Netherlands (1:7,000; Poorthius *et al.*, 1999), Australia (1:7700; Meikle *et al.*, 1999), Czech Republic (1:8000; Poupětová *et al.*, 2010) and Portugal (1:4000; Pinto *et al.*, 2004) [reviewed in Filocamo and Morrone, 2011].

1.2.2. Classification: from the nature of the primary stored material to the type of molecular defect

Classically, LSDs are classified on the basis of the nature of the accumulated substrate(s). From a clinical perspective, this classification is very useful and well accepted [Ballabio and Gieselmann, 2009]. Thus, disorders in which glycosaminoglycans (also called mucopolysaccharides) accumulation prevails are classified as Mucopolysaccharidoses*; those dominated by lipid storage are Lipidoses; the ones in which the accumulation of sphingolipids prevails are Sphingolipidoses; those mostly characterized by the storage of oligosaccharides are Oligosaccharidoses and so on [Schultz *et al.*, 2011].

This classification, even though highly intuitive and well-accepted both by clinicians and researchers, difficultly incorporates all the knowledge coming to light as the biochemical basis for LSDs started being unveiled. So, a novel classification is emerging, focused on the recent understanding of LSDs´ molecular basis, that also includes groups of disorders due to: (i) non-enzymatic lysosomal protein defects; (ii) transmembrane protein defects (transporters and structural proteins); (iii) lysosomal enzyme protection defects; (iv) post-translational processing defects of lysosomal enzymes; (v) trafficking defects in lysosomal enzymes; and (vi) polypeptide degradation defects (Table 1.1). Finally, another group includes the NCLs, which are considered to be lysosomal disorders, despite with distinct characteristics [reviewed in Filocamo and Morrone, 2011].

* For an extensive review on MPSs see Appendix 1, review paper 1: Coutinho MF, Lacerda L, Alves S. **Glycosaminoglycan storage disorders: a review.** *Biochem Res Int*. 2012 [*in press*]

Table 1.1.: Lysosomal storage disorders.

[*adapted from* Futerman and van Meer, 2004; Filocamo and Morrone, 2011 and Haltia and Gobel, 2012]

Sphingolipidoses		
Fabry	α-Galactosidase A	Globotriasylceramide and blood-group-B substances
Farber lipogranulomatosis	Acid Ceramidase	Ceramide
Gaucher [a]	-Glucosidase Saposin-C activator	Glucosylceramide Glucosylceramide
Niemann–Pick A and B	Sphingomyelinase	Sphingomyelin
Sphingolipid-activator deficiency	Sphingolipid activator	Glycolipids
GM1 gangliosidosis	β-Galactosidase	GM1 ganglioside
GM2 gangliosidosis (Tay–Sachs)	β-Hexosaminidase A	GM2 ganglioside and related glycolipids
GM2 gangliosidosis (Sandhoff)	β-Hexosaminidase A and B	GM2 ganglioside and related glycolipids
GM2 gangliosidosis (GM2-activator deficiency)[b]	GM2-activator protein	GM2 ganglioside and related glycolipids
Krabbe's disease [c]	Galactocerebrosidase β-galactosidase	Galactosylceramide
Metachromatic leucodystrophy[d]	Arylsulfatase A/Saposin B activator	Sulphated glycolipds/ Sulphated glycolipds and GM1 ganglioside
Mucopolysaccharidoses (MPS)		
MPS I (Hurler, Scheie, Hurler/Scheie)	α-Iduronidase	Dermatan sulphate and heparan sulphate
MPS II (Hunter)	Iduronate-2-sulphatase	Dermatan sulphate and heparan sulphate
MPS IIIA (Sanfilippo A)	Heparan *N*-sulphatase (sulphamidase)	Heparan sulphate
MPS IIIB (Sanfilippo B)	*N*-Acetyl-α-glucosaminidase	Heparan sulphate
MPS IIIC (Sanfilippo C)	Acetyl-CoA:α-glucosamide *N*-acetyltransferase	Heparan sulphate
MPS IIID (Sanfilippo D)	*N*-Acetylglucosamine-6-sulphatase	Heparan sulphate
Morquio-A disease	*N*-Acetylgalactosamine -6-sulphate-sulphatase	Keratan sulphate, chondroitin-6-sulphate
Morquio-B disease	β-Galactosidase	Keratan sulphate
MPS VI (Maroteaux–Lamy)	*N*-Acetylgalactosamine-4-sulphatase (arylsulphatase B)	Dermatan sulphate
MPS VII (Sly)	β-Glucuronidase	Heparan sulphate, dermatan sulphate, chondroitin-4- and -6-sulphates
Oligosaccharidoses (glycoproteinoses)		
Pompe (glycogen-storage-disease II)	α-Glucosidase	Glycogen
Aspartylglucosaminuria	Aspartylglucosaminidase	Aspartylglucosamine
Fucosidosis	α-Fucosidase	Fucosides and glycolipids
α-Mannosidosis	α-Mannosidase	Mannose-containing oligosaccharides
ß-Mannosidosis	β-Mannosidase	Man (β→4) GlcNAc
Mucolipidosis I (Sialidosis)	Sialidase (Neuraminidase)	Sialyloligosaccharides and sialylglycopeptides
Schindler disease	α-*N*-Acetylgalactosaminidase	Glyco-conjugates containing α-*N*-acetylgalactosaminyl

Glycogenoses		
Glycogenosis II/ Pompe	α-1,4- Glucosidase	Glycogen
Lipidoses		
Wolman/CESD	Acid lipase	Cholesterol esters
Non-enzymatic lysosomal protein defects		
Gaucher [a]	Saposin C	Glucosylceramide
GM2 gangliosidosis (GM2-activator deficiency) [b]	GM2-activator protein	GM2 ganglioside and related glycolipids
Krabbe [c]	Saposin A	Galactosylceramide
Metachromatic leucodystrophy [d]	Saposin B	Sulphatides
Integral membrane protein defects		
Transporters		
Cystinosis	Cystinosin	Cystine
Infantile sialic-acid-storage disease and Salla disease	Sialin	Sialic acid
Niemann–Pick C (NPC)	NPC1 and NPC2	Cholesterol and sphingolipids
Structural proteins		
Danon disease	LAMP2	Cytoplasmic debris and glycogen
Mucolipidosis (ML) IV	Mucolipin-1	Lipids and acid mucopolysaccharides
Lysosomal enzyme protection defects		
Galactosialidosis	Cathepsin A	Sialyloligosaccharides
Post-transcriptional processing defects		
Multiple sulphatase deficiency	Cα-formylglycine-generating enzyme	Sulphatides
Trafficking defects in lysosomal enzymes		
I Cell (ML II)	UDP-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosaminyl-1-phosphotransferase – α/β subunits	Oligosaccharides, mucopolysaccharides and lipids
Pseudo-Hurler polydystrophy (ML III)	UDP-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosaminyl-1-phosphotransferase – α/β subunits or γ subunit	Oligosaccharides, mucopolysaccharides and lipids
Polypeptide degradation defects		
Pycnodysostosis	Cathepsin K	Bone proteins
Neuronal ceroid lipofuscinoses		
NCL 1	CLN1 (protein palmitoylthioesterase-1)	Lipidated thioesters (saposins A and D)
NCL 2	CLN2 (tripeptidyl amino peptidase-1)	Subunit c of the mitochondrial ATP synthase
NCL 3	Arginine transporter	Subunit c of the mitochondrial ATP synthase
NCL 5	CLN 5, soluble lysosomal protein	Subunit c of the mitochondrial ATP synthase
NCL 6	CLN 6, transmembrane protein of ER	Subunit c of the mitochondrial ATP synthase
NCL 7	CLN 7, lysosomal chloride channel	Subunit c of the mitochondrial ATP synthase
NCL 8	CLN 8, transmembrane protein of ER	Subunit c of the mitochondrial ATP synthase

NCL 10	Cathepsin D	Saposins A and D
NCL11	Progranulin	[e]
NCL12	P type ATPase	[e]
NCL13	Cathepsin F	[e]
NCL14	Potassium channel tetramerisation domain -containing protein 7	[e]

[a,b,c,d] GD caused by deficiencies in saposin C, GM2 gangliosidosis, Krabbe's disease caused by deficiency of saposin A and Metachromatic Leucodystrophy were traditionally classified as sphingolipidoses but recent classifications based on the nature of the enzyme/protein defect, tend to skip them from the larger sphingolipidoses group and gather them in a more restrict group of LSDs due to non-enzymatic lysosomal protein defects. [e] The human NCL category was recently expanded to incorporate four novel pathologies (NCL11-14) but, even though the underlying genes have already been identified, these variants have not yet been recognized as independent clinical entities through OMIM annotation and data is still scare on their associated storage materials.

1.2.3. Symptoms: from severe neuropathology to mild somatic disease

Virtually all known LSDs are characterized by considerable phenotypic heterogeneity (Figure 1.5). Nevertheless, though, some common symptoms may be listed (Table 1.2).

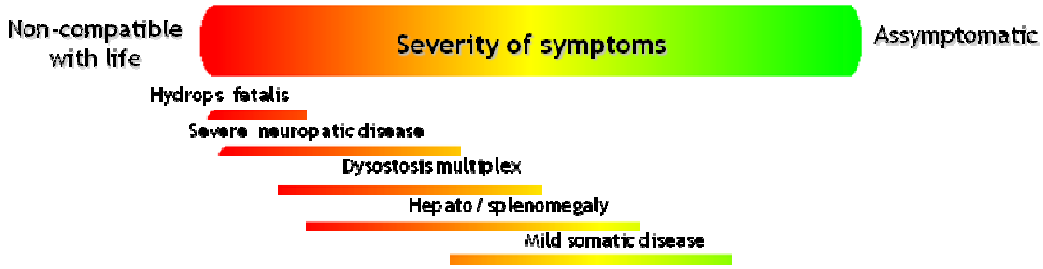


Figure 1.5: Schematic representation of the phenotypic continuum observed for LSDs.

In general, a common characteristic of LSDs is the relentless development and progression of signs and symptoms. There is high morbidity which shortens, sometimes dramatically, the life span [Beck, 2001; Wraith, 2002]. The disease onset, however, may vary tremendously, not only between different diseases but even within the same disease. So, LSDs´phenotypes may range from extremely severe and early presentations with multisystemic and neuronal involvement, leading to death in the first years of life or even in the neo-natal period, to later onset milder disease with patients reported to survive until the 8th decade of life.

Probably one of the most severe clinical presentations of LSDs is also the one with the earliest onset: hydrops fetalis, a condition in the fetus that is characterized by an accumulation of fluid in at least two fetal compartments. Recurrent non-immune hydrops fetalis is an uncommon but important presentation of LSDs. It has been already reported for several LSDs including Farber's disease [Kattner *et al.*, 1997], GM1-Gangliosidosis [Denis *et al.*, 1996], Galactosialidosis [Haverkamp *et al.*, 1996], Niemann-Pick disease type C [Meizner *et al.*, 1990], infantile free sialic acid storage disease, Mucolipidosis II [Appelman *et al.*, 1988], type 2 Gaucher disease [Reissner *et al.*, 1998], MPSs IVA [Applegarth *et al.*, 1987] and VII [Van

Dorpe *et al.*, 1996] and Sialidosis [Ovali *et al.*, 1998]. Since a series of conditions exist which may lead to non-immune hydrops fetalis while post-mortem analyses are frequently inconclusive, it is actually possible that cases due to lysosomal dysfunctions are underestimated. So, investigation of recurrent non-immune hydrops should include screening for LSDs [Piraud *et al.*, 1996; reviewed in Vellodi, 2005]. In general, the disorders that can present with hydrops fetalis can also present in the newborn period with dysmorfism. Infants in the newborn period with a "Hurler-like" phenotype usually have Mucolipidosis II, infantile sialic acid storage disease or GM1-Gangliosidosis [reviewed in Wraith, 2002].

The radiological abnormalities associated with LSDs are known as dysostosis multiplex. These are defects which are usually absent in the neonatal period but become progressively more evident during the first or second years of life. Different radiological abnormalities may be seen in different disorders, one of the most significant examples being the characteristic features seen in ML II patients who, in addition to a severe dysmorphism, have usually pronounced gingival hyperplasia [reviewed in Wraith, 2002].

Table 1.2: General symptomatology of LSDs.

[*adapted from* Wilcox, 2004]

Obvious physical abnormalities	Nonimmune hydrops fetalis; hepato/splenomegaly; macroglossia; coarse facies; skeletal disease, deformations, dysplasias with or without short stature, skeletal disproportion; hirsutism; macrocephaly; inguinal and umbilical hernias; skin lesions (angiokeratomas)
Physical symptoms	Joint stiffness; episodic crises of pain radiating from extremities, especially in children; acroparesthesias; heat and cold sensitivity; exercise intolerance; hoarse voice; recurrent otitis media, sinusitis, pneumonia
Abnormal laboratory or imaging findings	Unexplained proteinuria or isosthenuria; anemia; thrombocytopenia; vacuolated lymphocytes; white matter lesions; hydrocephalus on brain magnetic resonance imaging
Neurologic	Neurologic degeneration; peripheral neuropathy; progressive mental retardation; progressive dementia; behavioral abnormalities and psychosis; hyperactivity; excessive startle; seizures; hypotonia; weakness; spasticity; myoclonic jerks; ataxia; dystonia
Ophthalmologic	Corneal clouding; whorled corneal opacity; cataract; cherry red spot; macular degeneration; ophthalmoplegia, especially upward gaze palsy; optic atrophy; strabismus
Cardiac	Unexplained cardiomyopathy; valvular heart disease; arrhythmia
Gastrointestinal	Attacks of abdominal pain; diarrhea; constipation
Renal	Proteinuria; isosthenuria; tubular dysfunction
Cutaneous	Angiokeratomas (raised reddish-purple cutaneous vascular lesions); subcutaneous nodules; ichthyosis; hypohidrosis; coarse skin and hair

In synthesis: *1*) within each disorder, there is usually a clinical spectrum that reflects the impact of different mutations on the function of the lysosomal compartment; *2*) different disorders have remarkable similarities in clinical phenotype, especially when the stored substrates are similar (as is the case with some of the MPSs and the Sphingolipidoses), or when similar cells and tissues are involved (e.g., organomegally, bone pathology, and neuropathology); *3*) there are some characteristic clinical symptoms that can be used to distinguish disorders from each other (e.g. gingival hyperplasia in ML II disease). Thus, despite the similarities in the clinical phenotype between different lysosomal storage disorders, no two disorders have identical pathophysiology [reviewed in Parkinson-Lawrence *et al.*, 2010].

1.2.4. Diagnosis of LSDs

LDSs are sometimes hard to recognize due to the wide spectrum of phenotypes that the patients may present (Figure 1.5). Indeed, the recognition of LSD clinical features requires clinical expertise, as most of them are not specific and can be caused by defects in other metabolic pathways (mitochondrial and peroxisomal), or by environmental factors [reviewed in Filocamo and Morrone, 2011]. Therefore, it is common to elapse a long period between the onset of the first symptoms and the definitive diagnosis. And even in the presence of typical clinical signs and symptoms, samples and diagnostic tests are different for each group of lysosomal disorders, being often specific of a given disease. This is one of the major reasons why definitive diagnosis of LSDs requires a close collaboration between laboratory specialists and clinicians, in order to reach a correct diagnose in the shortest time.

Biochemical genetic testing (BGT), which includes the determination of the enzymatic activity of lysosomal hydrolases is feasible for most LSDs and essential for the diagnosis of primary lysosomal enzyme deficiency. Actually, the majority of patients is initially screened by enzyme assays and only after that the molecular studies are performed to determine the disease-causing mutation(s). In families with an already identified causative molecular defect, mutation analysis may be performed directly. Mutation analysis is done through molecular genetic testing (MGT), which can be performed either on DNA or RNA and comprises a range of different approaches for investigating the entire gene-coding regions and exon-intron boundaries, as well as the 5´and 3´untranslated regions (UTRs). It is not only useful but frequently necessary to confirm the enzymatic diagnosis of a LSD. For LSDs resulting from non-enzymatic lysosomal proteins, MGT is absolutely essencial for a proper diagnosis, the same happening for the analysis of post-mortem samples [reviewed in Filocamo and Morrone, 2011].

Recently, another possibility is being discussed for the diagnosis of LSDs that foresees the

inclusion of some of them (e.g. Krabbe disease; Pompe disease; Fabry disease and several MPSs such as type I, IIIA, IIIB and VI) in neonatal screening programs [Giugliani, 2012; Simonaro *et al.*, 2005] which would now be possible thanks to the significant advances that were made in last decade since dried blood spot technology was introduced for enzymatic assays and lysosomal protein profile was developed. The ongoing development of enzyme replacement therapy (ERT) and other treatments for several LSDs, combined with the growing evidence that early commencement of therapy improves outcomes, has increased the pressure for the introduction of newborn screening programs and a number of pilot studies are ongoing for LSDs [Simonaro *et al.*, 2005; Spada *et al.*, 2006; Lin *et al.*, 2009; Chien *et al.*, 2008; Hwu *et al.*2009; Duffner *et al.* 2009a; Duffner *et al.* 2009b; Fuller *et al.*, 2011].

1.2.5. Treatment: from bone marrow transplantation to gene therapy

Currently, there is no cure available for any LSD, even though treatment strategies do exist for some of them. In general, treatments for LSDs can be broadly divided into those that reduce the levels of biosynthesis of the accumulating substrate (substrate reduction therapy, SRT) and those aiming at directly addressing either the functional loss of the enzyme or its clinical effects (Figure 1.6).

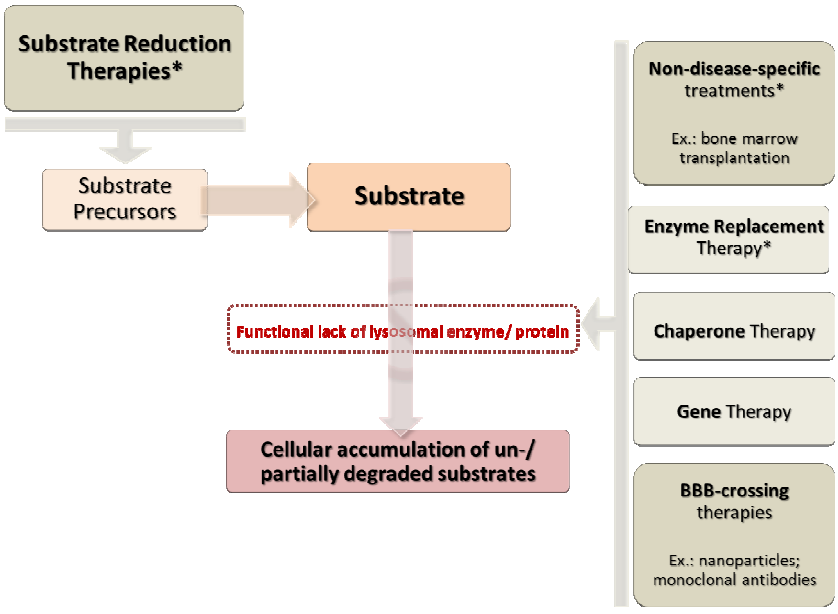


Figure 1.6: Therapies for lysosomal storage disorders.

LSD treatments can be divided into those that reduce the levels of biosynthesis of the accumulating substrate (left side of the image) and those aiming at directly addressing either the functional loss of the enzyme or its clinical effects (right side of the image). The later include (a) general treatments that deal mainly with the symptoms (ex.: bone marrow transplantation); (b) ERT and (c) potential new therapies such as the use of chemical chaperones, gene therapy or BBB-crossing therapies. Treatments marked with an * are in clinical use at present.

Of the treatments for symptoms, splenectomy used to be relatively common, at least for Gaucher-disease (GD) patients. Splenectomy can reduce the symptoms, such as thrombocytopenia (low platelet count) and anemia that accompany some LSDs, but does not cure the disease. Bone-marrow transplantation (BMT) is another option which sometimes is carried out. Umbilical cord blood transplantation is also being performed at specialized centers for a number of these diseases. The reason for such approaches lies on having found out that a small, but significant, fraction of lysosomal enzymes are secreted by the donor normal cells. So, transplanted bone marrow will eventually lead to the production of normal cells that secrete normal enzymes which may subsequently be internalized by the affected cells. Other approaches are also common to ameliorate symptoms of some particular LSDs. For example, chronic hemodialysis and/or renal transplantation are used in Fabry disease because renal insufficiency is a frequent complication. It should be stressed, however, that none of these treatment approaches are either specific or effective. Actually, no treatments are yet available targeting most of the rarer LSDs and overall the medical options are mostly directed towards disease management rather than therapy.

Nevertheless, some rather specific and effective treatments are available for a restrict number of LSDs. Those are not the above referred general treatments for the symptoms, but instead, therapeutic approaches that deal directly with the cause of the disease, either by replacing the defective gene or enzyme or by directly targeting the systems that are affected by the accumulation of undegraded metabolites. Such approaches were only possible when remarkable progresses in genetic engineering allowed the production of recombinant proteins that could replace the missing enzyme providing specific treatment for a growing number of LSDs. This process started with intravenous ERT for GD in the 1990s, and was later extended to Fabry disease, Pompe disease and MPSs I, II and VI [Giugliani, 2012]. This therapy is likely to be extended to other LSDs in the near future and some clinical trials are already in course (e.g. MPS IVA). In general, ERT is more effective on later onset or milder variants of LSDs, as it has no impact on the neuropathology of the diseases, since the recombinant enzyme is unable to cross the blood brain barrier and so to reach the brain.

The ineffectiveness of ERT and, in some cases BMT, to overcome neurological symptoms is one of the main reasons why additional therapeutic approaches are being investigated addressing the needs of the more severe types. SRT, for example, is currently being evaluated for some of these diseases (e.g. MPS I, II, IIIA and IIIB). Also chaperone therapy, which holds the potential to stabilize the defective enzymes produced by patients, or gene enhancement and gene therapy, techniques in which specific viral vectors may be used to directly transduce central nervous system (CNS) cells that would thus become an endogenous source of high levels of functional enzyme [Gritti, 2011], are also under

evaluation and are expected to provide a cure for many of these genetic diseases in the future [Beck, 2010].

Despite the prospects of innovative therapeutic developments, without a detailed understanding of the precise mechanisms operating in each LSD, and without easy and early means of diagnosis, new and successful treatments will not become available [reviewed in Futerman and van Meer, 2004]. So, the major efforts at the moment should be focused both at unveiling the underlying pathological mechanisms that lead to LDSs phenotypes (*See section 1.2.3*) and at developing sensitive and early methods of diagnosis for such disorders (*see Section 1.2.4*).

2. TRAVELLING ROUTES TO THE LYSOSOME

2.1. THE MANNOSE-6-PHOSPHATE PATHWAY*

Like several other proteins, lysosomal hydrolases are synthesized in the endoplasmic reticulum (ER) and cotranslationally glycosylated on selected asparagine residues. Soluble lysosomal proteins are synthesized as pre-pro-proteins with an N-terminal signal sequence (20-25 aa), which translocates the precursor protein into the lumen of the ER [Hasilik *et al.*, 2009]. The signal peptides are removed and initial protein glycosylation takes place. The signal motif for *N*-glycosylation is Asn-X-Ser/Thr ("X" being any amino acid except proline), which is a motif evolutionary highly conserved [Braulke and Bonifacino, 2009]. The oligosaccharide branch $Glc_3Man_9(GlcNAc)_2$ is attached to the amino group of the asparagine side chain amide, being then processed and modified in the ER and Golgi apparatus. Mature glycoproteins can contain three different types of *N*-glycans: high mannose, complex and hybrid sugars. Another post-translational modification taking place in the Golgi complex is *O*-glycosylation, where sugar residues can be transferred to the OH group of an amino acid side chain, mostly of serine or threonine. This type of glycosylation is a single monosaccharide transfer process initiated by specific GalNAc transferases (over 24 described). In contrast to *N*-glycosylation, *O*-glycosylation is not known to occur at a single consensus sequence. However, the various GalNAc transferases are highly specific and control the glycosylation process. Addition and modification of sugar residues also play an important role in protein folding, quality control and transport [Helenius and Aebi, 2001; Ruddock and Molinari, 2006]. When moving to the secretory pathway, these proteins are also selectively recognized by a phosphotransferase, UDP-*N*-acetylglucosamine-1-phosphotransferase (GlcNAc-1-phosphotransferase), which initiates the two-step reaction that results in the generation of the M6P marker (Figure 2.1) on specific *N*-linked oligosaccharides.

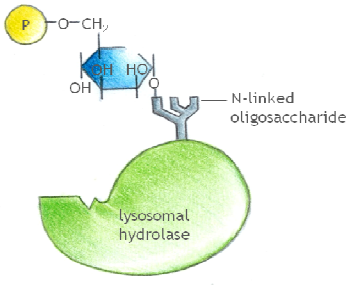


Figure 2.1: The Mannose-6-Phosphate recognition signal.

* For an extensive review on the M6P travelling route see Appendix 1, review paper 2: Coutinho MF, Prata MJ, Alves S. **The mannose-6-phosphate pathway: Function and dysfunction**. *Mol Genet Metab*. 2012; 105:542-50.

GlcNAc-1-phosphotransferase catalyses the transfer of a GlcNAc-1-phosphate residue from UDP-GlcNAc to C6 positions of selected mannoses in high-mannose type oligosaccharides of the hydrolases. The second step in the formation of the M6P marker involves the removal of the terminal GlcNAc by an *N*-acetylglucosamine-1-phosphodiester α-*N*-acetyl-glucosaminidase, also known as "uncovering enzyme" (UCE), exposing the M6P recognition signal. Next, the modified proteins are recognized by two independent receptors that bind the M6P residue of the newly synthesized lysosomal hydrolases in the TGN and finally, the ligand-receptor complex is packaged into clathrin-coated transport vesicles for delivery to endosomes and lysosomes (Figure 2.2) [Braulke and Bonifacino, 2009].

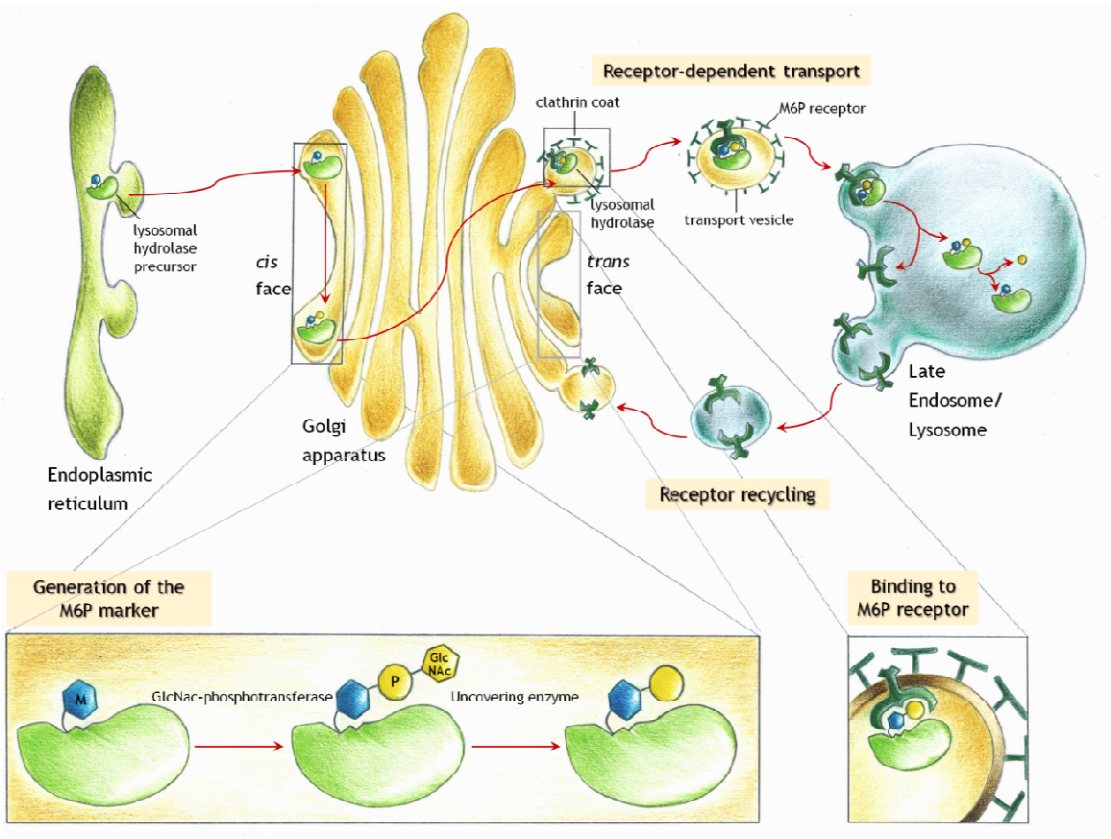


Figure 2.2: Transport of newly synthesized lysosomal hydrolases to lysosomes.

Lysosomal hydrolases are synthesized in the endoplasmic reticulum and move to the *cis* Golgi network, where they are covalently modified by the addition of mannose 6-phosphate (M6P) groups. The formation of this marker depends on the sequential effect of two lysosomal enzymes: UDP-*N*-acetylglucosamine 1-phosphotransferase (GlcNAc-phosphotransferase) and *N*-acetylglucosamine-1-phosphodiester α-*N*-acetyl-glucosaminidase (UCE). GlcNAc-phosphotransferase catalyses the transfer of a GlcNAc-1-phosphate residue from UDP-GlcNAc to C6 positions of selected mannoses in high-mannose type oligosaccharides of the hydrolases. Then, the uncovering enzyme removes the terminal GlcNAc, exposing the M6P recognition signal. At the *trans*-Golgi network, the M6P signal allows the segregation of lysosomal hydrolases from all other types of proteins through selective binding to the M6P receptors. The clathrin-coated vesicles produced bud off from the *trans*-Golgi network and fuse with late endosomes. At the low pH of the late endosome, the hydrolases dissociate from the M6P receptors and the empty receptors are recycled to the Golgi apparatus for further rounds of transport.

2.1.1. FUNCTIONAL COMPONENTS OF THE MANNOSE-6-PHOSPHATE PATHWAY

2.1.1.1. GlcNAc-1-phosphotransferase*

GlcNAc-phosphotransferase is a Golgi-resident 540-kDa hexameric transmembrane enzyme composed by three subunits, $\alpha_2\beta_2\gamma_2$ (Figure 2.3).

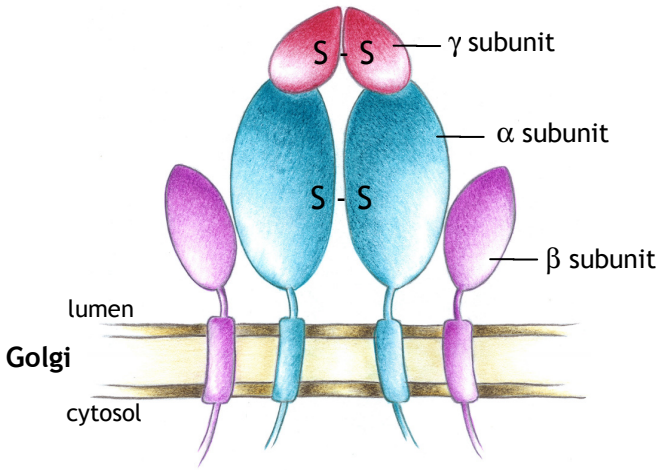


Figure 2.3: Model of the GlcNAc-1-phosphotransferase complex.

[*adapted from* Braulke *et al.*, 2008 ; Encarnação *et al.*, 2011]

GlcNAc-1-phosphotransferase is a hexameric complex of three subunits ($\alpha_2\beta_2\gamma_2$). The α- and γ-subunits form disulfide-linked dimers.

This enzyme is coded by two different genes: *GNPTAB* and *GNPTG*, encoding for the α/β-subunits and the γ-subunit, respectively. The *GNPTAB* gene contains 21 exons and spans 85 kb on chromosome 12q23.3. It encodes a protein of 1256 amino acids with a predicted molecular mass of 144 kDa (α/β-precursor). Hydrophobicity analysis showed two transmembrane domains and 19 potential *N*-glycosylation sites (17 in the α-subunit and 3 in the β-subunit) [Tiede *et al.*, 2005].

Proteolytic processing of the α/β-precursor between Lys^{928} and Asp^{929} generates the individual α- and β-subunits and is a prerequisite for the enzymatic activity of the GlcNAc-1-phosphotransferase complex [Tiede *et al.*, 2005]. Recently Marschner *et al.* [2011] provided evidence that the α/β-subunit precursor of GlcNAc-1-phosphotransferase is cleaved by the site-1 protease (S1P; also known as subtilisin kexin isoenzyme–1, SKI-1). This enzyme is encoded by the *MBTPS1* gene and is a membrane bound serine protease. The prototypical membrane-bound S1P substrates are the sterol regulatory element-binding proteins SREBP1 and 2, which play a major role in lipid metabolism and cholesterol homeostasis [referred in Marschner *et al.*, 2011]. These authors verified that S1P-deficient

* For an extensive review on GlcNAc-1-phosphotransferase see Appendix 2, book chapter: Coutinho MF. ***N*-acetylglucosamine-1-phosphate transferase (*GNPTAB*, *GNPTG*) *in*** Taniguchi N., Honke K., Fukuda M., Narimatsu H., Yamaguchi Y., Angata T. (Ed.) "Handbook of Glycosyltransferases and Related Genes". SpringerReference [*in press*]

cells failed to activate the α/β-subunit precursor and exhibited a mucolipidosis II-like phenotype. This implies a new focus for molecular diagnosis, since mutations in the gene that codes for this protease may be present in individuals with genetically undefined mucolipidosis II-like phenotypes.

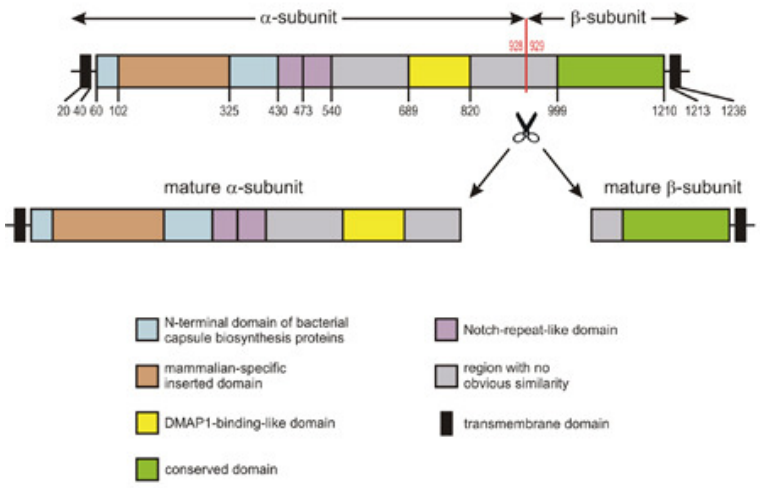


Figure 2.4: Modular structure of the human GlcNAc-1-phosphotransferase α/ß-precursor.

[adapted from Marschner *et al.*, 2011]

Schematic presentation of the modular structure of the α/β-precursor of the human GlcNAc-1-phosphotransferase and its proteolytic processing to mature α- and β-subunits.

The *GNPTG* gene, spanning 11 exons on chromosome 16p13.3, encodes the γ-subunit, which is a soluble glycoprotein of 305 amino acids, with a molecular weight of 97kDa capable to form disulfide-linked dimers [Raas-Rothschild *et al.*, 2000]. The initial 305 aminoacid polypeptide is processed into a mature form by proteolytic cleavage of a signal peptide with 24 amino acids; then the mature protein can form disulfide bonds [Tiede *et al.*, 2004]. The γ-subunit contains two potential *N*-glycosylation sites at N^{88} and N^{115} [Encarnação *et al.*, 2011].

Functionally, the α/β-subunits constitute the catalytic portion of the enzyme and carry the substrate binding sites for UDP-GlcNAc and high mannose-type lysosomal enzymes [Kudo *et al.*, 2005; Lee *et al.*, 2007]. Whilst being in the α-subunit that are located the UDP-GlcNAc binding site and homology region to the bacterial capsular polymerases [Tiede *et al.*, 2004; Kudo *et al.*, 2005; Bao *et al.*, 1996], the β-subunit also appears to be necessary to the enzyme function, as demonstrated by the failure of transfection of the α-subunit alone to generate GlcNAc-1-phosphotransferase activity [Kudo *et al.*, 2006a] plus the genetic evidence that truncation of the β-subunit results in GlcNAc-1-phosphotransferase deficiency [Tiede *et al.*, 2004; Kudo *et al.*, 2006a; Paik *et al.*, 2005]. Moreover, the transmembrane domains on the α- and β-subunits are required for natural processing of the GlcNAc-1-phosphotransferase precursor [Kudo *et al.*, 2006a].

Analysis of a library of cathepsin D/pepsinogen chimeric proteins along with mutant cathepsin L has shown that GlcNAc-1-phosphotransferase recognizes common conformation-dependent protein structures of lysosomal enzymes in which lysine residues are the major determinants [Baranski *et al.*, 1990; Cantor *et al.*, 1992; Cuozzo *et al.*, 1995; Qian *et al.*, 2010], as was also proven through antibody inhibition experiments with arylsulfatase A [Sommerlade *et al.*, 1994]. In addition, the interaction of both subunits with the protein determinant of acid hydrolases stimulates the catalytic function of the transferase [Qian *et al.*, 2010].

In 2007, Gelfman and colleagues generated a mouse model for *GNPTAB* deficiencies by microinjection of embryonic stem (ES) cell clones *GNPTAB*$^{-/-}$ into host blastocysts. Homozygous mice lacking α/β-subunits of GlcNAc-1-phosphotransferase presented with growth retardation, retinal degeneration and secretory cell lesions that do not exactly mimic the human disease. Nevertheless, the less severe phenotype of the *GNPTAB*-knockout mouse enables studies designed to understand alternative molecular mechanisms involved in the trafficking of lysosomal enzymes to lysosomes and provides a valuable tool for assessing the role of proper lysosome function in the maintenance of the retina and the secretory cells of exocrine glands [Gelfman *et al.*, 2007].

The role of the γ-subunits of the GlcNAc-1-phosphotransferase remained elusive [Tiede *et al.*, 2005; Kudo *et al.*, 2006a] up to recently, when studies begun to appear demonstrating that the γ-chains are important to facilitate the proper folding of the full set of subunits in the GlcNAc-1-phosphotransferase and to maintain them in the right conformation, turning the enzyme competent for substrate recognition and binding. The γ-chains are also necessary to regulate the activity and expression of the α/β-subunits [Lee *et al.*, 2007; Qian *et al.*, 2010; Pohl *et al.*, 2009; Pohl *et al.*, 2010]. Pohl and colleagues [2009], after having identified, in a boy with MLIII gamma, a novel homozygous mutation in the *GNPTG* gene resulting in a truncated but stable γ-subunit protein, performed *in vitro* and *in vivo* experiments to characterize the mutation, obtained compelling support that the γ-subunit was involved in the regulation of the GlcNAc-1-phosphotransferase activity. Latter, the same group [Pohl *et al.*, 2010] showed that in human macrophages the proteolytic cleavage of the γ-subunit of the GlcNAc-1-phosphotransferase leads to the formation of unique oligomers that fail to associate with the other GlcNAc-1-phosphotransferase subunits in higher molecular mass complexes, which might explain why almost all proteins in human macrophages lack M6P residues. The limited proteolysis of the γ-subunits seems to be the regulator mechanism that controls the GlcNAc-1-phosphotransferase activity and subsequent sorting efficiency of lysosomal enzymes. In mice lacking the γ-subunit, the extent of acid hydrolases phosphorylation relative to wild-type animals was investigated revealing that the γ-subunit increases the activity of

the α/β-subunits toward protein acceptors to variable extents [Qian *et al.*, 2010]. The same study afforded evidence that the γ-subunit serves at least two roles. One is to interact with the high mannose oligosaccharides of the acceptor hydrolase, facilitating the addiction of the second GlcNAc-P to the molecule. This GlcNAc-P is transferred to a specific mannose residue on the 3' arm of the high mannose unit, whereas the first GlcNAc-P almost always is added to a mannose on the 6' arm of the glycan. The authors also postulate that a specific domain of the γ-subunit, the mannose 6-phosphate receptor homology (MRH) domain binds the high mannose oligosaccharide in such a way that the transfer of the second GlcNAc-P is enhanced, explaining the role of the γ-subunit in increasing the overall phosphorylation in a subset of the acid hydrolases [Qian *et al.*, 2010]. This second role is in line with previous data from Lee *et al.* [Lee *et al.*, 2007], who through the analysis of brain tissue of knock-out mice for the γ gene, observed that, for example, α-mannosidase and β-glucuronidase were similarly phosphorylated as in the wild-type brain, whereas β-hexosaminidase and β-galactosidase were not or were weakly phosphorylated [Lee *et al.*, 2007].

2.1.1.2. The uncovering enzyme: *N*-acetylglucosamine-1-phosphodiester α-*N*-acetyl-glucosaminidase

N-acetylglucosamine-1-phosphodiester α-*N*-acetylglucosaminidase, also known as the uncovering enzyme (UCE) excises the covering GlcNAc from the GlcNAc-P-Man diesters, to form the M6P monoester recognition signal on lysosomal acid hydrolases (Figure 2.5).

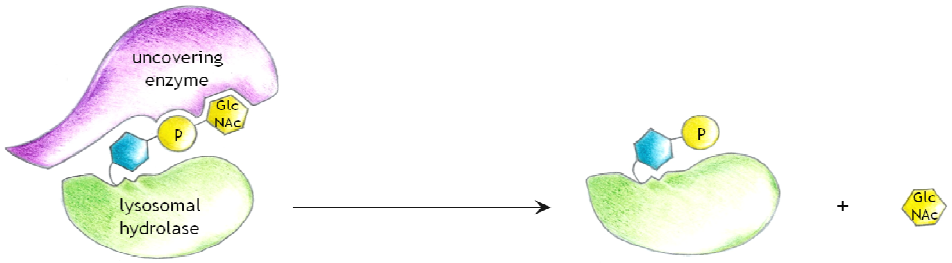


Figure 2.5: Schematic presentation of UCE action on the M6P recognition signal formation.

UCE is a type I transmembrane glycoprotein of 515 amino acids which exists as a tetramer that cycles between the TGN and the plasma membrane [Varki and Kornfeld, 1981; Kornfeld *et al.*, 1999; Lee *et al.*, 2002]. This enzyme is coded by the *NAGPA* gene that contains 10 exons, being located on chromosome 16 [Kornfeld *et al.* 1999]. UCE presents six potential *N*-glycosylation sites [Kornfeld *et al.* 1999], being synthesized as an inactive precursor that is activated upon proteolytic cleavage of a 24-aminoacid propiece by the endoprotease furin. As furin is localized in the TGN, newly synthesized UCE is inactive until reaching the terminal Golgi compartment [Do *et al.*, 2002].

To date no pathological conditions have been associated with the loss of UCE activity. Speculation on the existence or inexistence of such pathologies was maintained over years, with some authors arguing that an alternative mechanism of conversion of the M6P diesters to monoesters should exist, able to correct malformation of the recognition signal induced by UCE deficiency, while others argued that inactivating mutations in the UCE gene should be so rare events that could explain the non-observation of UCE related phenotypes. However, a few years ago, Chavez and collaborators [2007], shed some light on why UCE loss of function might not be associated with overt pathological condition, when demonstrating that the fifth repeat domain, of a total of 15 that constitute the cation-independent mannose-6-phosphate receptor (CI-MPR), is able to bind GlcNAc-P-Man diesters. This binding occurred with much lower affinity than the binding to M6P monoesters, which take place in domains 1-3 and 9 of the same receptor [Chavez *et al.*, 2007]. The first evidence came from *in vitro* studies but next they were corroborated by *in vivo* experiments with mice lacking UCE, created through insertional mutagenesis in the gene that codes for UCE [Boonen *et al.*, 2009]. Despite $UCE^{-/-}$ mice were viable, grew normally and lacked detectable histological abnormalities, the plasma levels of six monitored lysosomal hydrolases were 1.6 to 5.4-fold elevated over wild type levels. The secreted hydrolases contained GlcNAc-P-Man diesters, exhibited a decreased affinity for the CI-MPR receptor and no affinity at all to the cation-dependent mannose-6-phosphate receptor (CD-MPR). On the grounds of these results, Boonen and colleagues [2009] proposed that, in the absence of UCE, the weak binding of the lysosomal hydrolases to the CI-MPR allows enough sorting to lysosomes to prevent the tissue abnormalities seen in GlcNAc-1-phosphotransferase deficiency. The ability of the CI-MPR to recognize the product of GlcNAc-1-phosphotransferase implies that deficiencies in the UCE probably will not be lethal [Boonen *et al.*, 2009]. Furthermore, it allows delivery to the lysosome of those acid hydrolases that may be poor substrates for the uncovering enzyme in quantities sufficient to prevent the severe phenotypes usually associated with LSDs [Olson *et al.*, 2010].

2.1.1.3. Mannose-6-Phosphate receptors

Once anchored in the *N*-linked oligosaccharides of newly synthesized hydrolases, the M6P recognition marker has to be recognized by specific receptors, so that the labeled hydrolases are correctly transported to the lysosome (Figure 2.2) [reviewed in Kornfeld and Mellman, 1989]. This role is carried out by the mannose-6-phosphate receptors (MPRs), which are type I transmembrane glycoproteins that bind their specific cargo proteins at pH 6,5-6,7, in the TGN and release them at pH 6, the typical pH inside late endosomes. Inside these compartments, lysosomal hydrolases dissociate from the MPRs, and when the pH drops further, during endosomal maturation, and reaches 5 - the typical acidic pH of

lysosomes which is maintained by a membrane ATP-driven H^+ pump -, hydrolases begin to digest the endocytosed material delivered from early endosomes [reviewed in Alberts *et al.*, 2002].

There are two different M6P receptors, both of them extensively characterized over the last decades [Kornfeld *et al.*, 1992; Ludwig *et al.*, 1995; reviewed in Ghosh *et al.*, 2003]. The first is the cation-independent M6P receptor (CI-MPR or MPR300), a transmembrane glycoprotein of approximately 300-kDa that also binds with high affinity other ligands such as the insulin-like growth factor II (IGF II) that is involved in early growth and development. This is why the CI-MPR is also called the IGF II receptor. The extracytoplasmic domain of this receptor contains fifteen homologous repeating units of around 150 amino acids [Lobel *et al.*, 1988; Oshima *et al.*, 1988]. The third and the ninth of them provide two M6P binding sites, whereas the eleventh element provides a binding site to IGF II [Westlund *et al.*, 1991; Dahms *et al.*, 1993; Dahms *et al.*, 1994]. Additionally, CI-MPR also contains one phosphodiester (Man6P-GlcNAc) binding site [Marron-Terada *et al.*, 2000; Hancock *et al.*, 2002]. The other receptor is the cation-dependent M6P receptor (CD-MPR or MPR46). The extracytoplasmic domain of this 46-kDa transmembrane glycoprotein is homologous to each of the 15 repeats found in the extracytoplasmic domain of CI-MPR [Dahms *et al.*, 1987] and contains one M6P binding site [Tong and Kornfeld, 1989]. The detergent-solubilized CD-MPR is a homodimer which may form tetramers in Golgi membranes [Dahms and Kornfeld, 1989; Waheed and von Figura, 1990].

Although related to one another, and even having similar domains, both MPRs are essential and non-synonymous components of the M6P-dependent targeting system. This was proven by gene disruption experiments in mice. Animals with disrupted CI-MPRs or CD-MPRs genes were partially impaired in intracellular lysosomal enzyme sorting [Köster *et al.*, 1993; Ludwig *et al.*, 1993]. Mutant mice lacking the CD-MPR are viable, while mice lacking the CI-MPR accumulate high levels of IGF-II and usually die perinatally, whereas mice that lack both IGF-II and CI-MPR are viable [Sohar *et al.*, 1998]. A partial missorting of phosphorylated lysosomal enzymes was also observed in embryonic fibroblasts derived from mouse embryos lacking either the CI-MPR or the CD-MPR [Munier-Lehman *et al.*, 1996]. Further experiments enlightened the fact that each mutant cell line secretes, in a large part, different phosphorylated ligands and that high levels of one MPR do not fully compensate for the absence of the other, demonstrating that the two MPRs have complementary targeting functions. One possible explanation is that each MPR recognizes different features of the lysosomal enzymes [Munier-Lehman *et al.*, 1996]. Furthermore, primary fibroblasts from embryos lacking both MPRs are almost totally impaired in lysosomal enzyme sorting. As a consequence, cells accumulate undigested material in their endosomes/lysosomes [Ludwig *et al.*, 1994] presenting a phenotype quite similar to the one characteristic of ML II fibroblasts, which have a lack of synthesis of the M6P

signal due to defects in GlcNAc-1-phosphotransferase.

The idea that the MPRs are complementary in function was reinforced by studies of receptor deficient fibroblasts showing that three different subgroups of M6P containing proteins could be distinguished: (i) those which are preferentially secreted in the absence of the CD-MPR and, therefore, do not interact well with the remaining CI-MPR; (ii) those that are preferentially secreted in the absence of the CI-MPR and, therefore, do not interact well with CD-MPR and, finally, (iii) those which are only secreted in the absence of both MPRs, thus interacting equally well with either MPR [Munier-Lehmann *et al.*, 1996].

In terms of localization, the MPRs are distributed over several intracellular compartments including the TGN, the plasma membrane and the endosomes. At steady level, the majority of MPRs reside in the endosomes [Kornfeld and Mellman, 1989]. When located in the TGN, the MPRs divert the newly synthesized lysosomal hydrolases from the secretory pathway to the endocytic pathway, after which they unload their bound ligands. However, not all the hydrolases that are tagged for delivery to the endosome/lysosome compartment arrive to their proper destination. Actually, some escape binding to MPR in the TGN being directed, by default, to the cell surface, where they are secreted to the extracellular fluid. When this happens some MPRs can take a detour back to the plasma membrane, where they recapture the escaped lysosomal enzymes returning them to lysosomes by receptor-mediated endocytosis via early and late endosomes [*in* Alberts *et al.*, 2002]. The CI-MPR localized at the plasma membrane may also contribute to the partial rescue of those hydrolases [Marron-Terada *et al.*, 2000; Dahms *et al.*, 1987].

2.1.2. IMPAIRMENTS IN THE MANNOSE-6-PHOSPHATE PATHWAY AND DISEASE

2.1.2.1. Defective GlcNAc-1-phosphotransferase causing Mucolipidosis II and III

Defective GlcNAc-1-phosphotransferase causes two distinct human lysosomal storage diseases, Mucolipidosis II (ML II) and Mucolipidosis III (ML III), which are among the few lysosomal storage disorders related to defects in non-lysosomal proteins. ML II, also known as I-cell disease, is characterized by a total loss of GlcNAc-1-phosphotransferase activity whether ML III, often referred to as pseudo-Hurler polydystrophy, manifests when enzymatic activity is reduced [Leroy and Spranger, 1970; Spranger and Wiedemann, 1970]. In both ML II and ML III patients, newly synthesized lysosomal enzymes fail to be correctly sorted to the endosome/lysosome compartment due to the absence or weak equipment in M6P residues. As a consequence, lysosomal dysfunction develops leading to accumulation of non-degraded material, the hallmark of this group of diseases. Unlike the majority of lysosomal storage disorders, which involve single enzymes acting in a catabolic pathway,

MLII and III results from impaired sorting of multiple enzymes to lysosomes that instead are over-secreted from cells. The excessive accumulation of non-degraded substrates results in the subsequent formation of large inclusion bodies.

While ML II and ML III share similar clinical features, including skeletal abnormalities, ML II is the more severe in terms of phenotype [Leroy and Spranger, 1970]. In this pathology the skeletal system is severely affected, with abnormalities in both cartilage and bone. Linear growth decelerates during the first year of life almost stopping during the second year and death usually occurs between 5 and 8 years of age [Kornfeld and Sly *in* Scriver, 2001]. ML III is a much milder disorder, being characterized by latter onset of clinical symptoms and slower progressive course, which may allow the survival into the eighth decade. Usual clinical findings in ML III patients include restricted joint mobility, short stature and mild Hurler-like dysmorphism, among other less severe features [Spranger and Wiedemann, 1970; Umehara *et al.*, 1997]. Only 50% of the patients present with mental retardation [Umehara *et al.*, 1997].

Once GlcNAc-1-phosphotransferase is an hexameric complex whose protein subunits are encoded by two genes, depending on which of them harbors the causal mutation(s) and simultaneously on the severity/clinical course of the disease, the associated pathologies are classified as ML II alpha/beta (OMIM: 252500) and ML III alpha/beta (OMIM: 252600) if mutations are present in the *GNPTAB* gene, or ML III gamma (OMIM: 252605) if mutations occur in the *GNPTG* gene [Tiede *et al.*, 2005; Raas-Rothschild *et al.*, 2000].

To date, more than 100 different *GNPTAB* mutations have been reported, causing either ML II alpha/beta or ML III alpha/beta, including 30 missense, 20 nonsense, 32 small deletions, 25 small insertions, 2 small indels and 14 splice site mutations [HGMD and references therein; reviewed in Kollman *et al.*, 2010]. Large genomic rearrangements appear to be rare (1,6%) although two gross insertions [Otomo *et al.*, 2009; Tapino *et al.*, 2008] have already been detected. Most of these mutations are private or rare. Uncommonly, however, the microdeletion c.3503_3504delTC, presents a remarkably wide geographical distribution, having been detected among Israeli and Palestinian Arab-Muslim, Turkish, Irish traveler [Bargal *et al.*, 2006], Italian [Tapino *et al.*, 2009] and U.S. patients [Kudo *et al.*, 2006b]. That deletion was found to be a frequent mutation in a French Canadian founder population [Plante *et al.*, 2008].

Concerning the *GNPTG* gene, the mutations until now reported associated to Mucolipidosis type III gamma include 4 missense, 2 nonsense, 5 small deletions, 4 small insertions, 4 splice site mutations and 2 gross deletions [HGMD and references therein, reviewed in Kollman *et al.*, 2010].

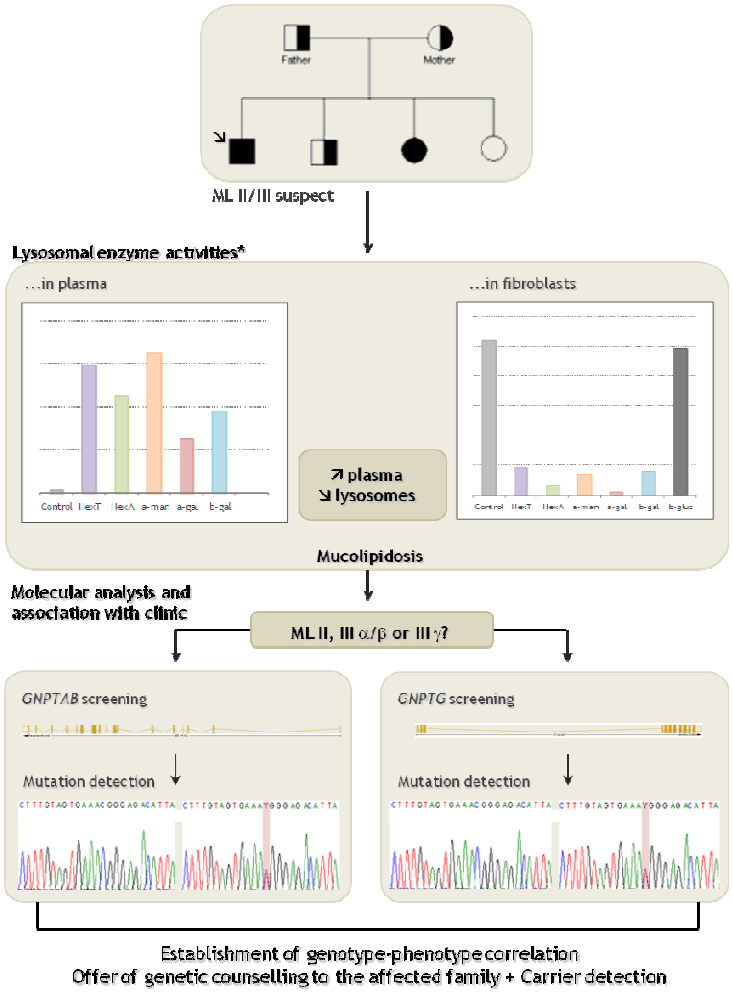


Figure 2.6: Diagnostic strategy for MLII alpha/beta, ML III alpha/beta and ML III gamma.

Even though direct assessment of GlcNAc-1-phosphotransferase is also possible, usually ML II and III biochemical confirmation is attained indirectly through detection of activities of standard lysosomal enzymes in plasma or in cultured fibroblasts. In general, a 10- to 20-fold increase in serum of lysosomal enzymes, together with decreased intralysosomal levels assessed in fibroblasts (or amniocytes in cases of prenatal diagnosis) is indicative of these disorders. Importantly, ML II cannot be distinguished from ML III on the basis of residual enzyme activity. Only subsequent molecular analysis and association with information on the clinical symptoms and age of onset allow definitive establishment of diagnosis. HexT - hexosaminidase T; Hex A - hexosaminidase A; α-man - α-mannosidase; α-gal - α-galactosidase; β-gal - β-galactosidase; β-gluc - β-glucocerebrosidase

2.1.2.2. Mutations in the lysosomal enzyme-targeting pathway causing stuttering

Recently, it was disclosed an unexpected relation between mutations in the lysosomal enzyme-targeting pathway and persistent stuttering.

Stuttering, is a disorder in which speech fluency can be severely compromised, but its primary causes, the high rates of spontaneous recovery and the etiological differences between persistent and resolved forms have resisted explanation. In the last years, genomewide scans provided signals of linkage at multiple chromosomal sites with

stuttering [Shugart *et al.*, 2004; Riaz *et al.* 2005; Suresh *et al.*, 2006; Wittke-Thompson *et al.*, 2007], but the strongest evidence was with a locus on chromosome 12q, which was found in a study involving 46 consanguineous families from Pakistan [Riaz *et al.*, 2005]. Afterwards, when Kang and colleagues [2010], focusing on the largest of those families, carried out a search for mutations across the chromosomal interval showing the strongest linkage, they ended up identifying a single nucleotide change in the *GNPTAB* gene (c.3598G>A). Then, the authors went on to sequence *GNPTG*, as well as *NAGPA*, in which mutations causing changes at highly conserved amino acid residues were also identified, as occurred in *GNPTAB*. A strong association was established between persistent stuttering and four mutations in the *GNPTAB* gene; three mutations in the *GNPTG* gene and three mutations in the *NAGPA* gene. This study unveiled an unsuspected culprit to explain stuttering in some cases, due to the nature of the implicated biologic pathway whose impairment had been essentially connected to lysosomal storage diseases. Yet, the affected subjects studied by Kang *et al.* [2010] stuttered but were otherwise normal, without presenting any of the typical symptoms of lysosomal malfunction, even being homozygous for some of the identified mutations on the *GNPTAB* and *GNPTG* genes. Eventually, this can be explained admitting that the efficiency of lysosomal targeting was only partially reduced in the presence of the detected mutations. Recently, Lee *et al.* [2011] performed a biochemical evaluation of the effect of the three *NAGPA* mutations which were associated with persistent stuttering and found that each mutation leads to lower cellular UCE activity [Lee *et al.*, 2011].

Additional studies in this area may contribute to a better understanding of the neural structures and functions within the brain that coordinate human speech, which are still poorly elucidated and to a better comprehension on the role of the M6P-dependent pathway for lysosomal enzyme routing in this area.

2.2. THE MANNOSE-6-PHOSPHATE INDEPENDENT PATHWAY*

Although M6P receptors play a major role in the intracellular transport of newly synthesized lysosomal enzymes in mammalian cells, there are, as previously referred several lines of evidence suggesting the existence of an alternative mechanism of lysosomal targeting [Ni *et al.*, 2006], the first of which have come from analysis of cells of patients with I-cell disease, who despite lacking M6P-containing lysosomal enzymes, demonstrate normal lysosomal enzyme levels in several organs.

* For an extensive review on the M6P travelling route see Appendix 1, review paper 3: Coutinho MF, Prata MJ, Alves S. **A Shortcut to the Lysosome: The Mannose-6-Phosphate Independent Pathway**. *Mol Genet Metab*. 2012; 107:257-66.

Other experimental approaches have also proven that there are soluble enzymes as well as non-enzymatic proteins that are transported to lysosomes in a M6P-independent manner, specifically by means of two alternative receptors: the lysosomal integral membrane protein (LIMP-2) and sortilin. LIMP-2 was implicated in the delivery of β-glucocerebrosidase (GCase, the defective enzyme in patients with Gaucher disease) to the lysosomes [Reczek *et al.*, 2007]. Sortilin, on the other hand, is a multifunctional receptor capable of binding different ligands, which has been suggested to mediate Golgi-to-lysosome transport of the sphingolipid activator proteins (SAPs) prosaposin (PSAP) and GM2 activator protein (GM2AP); acid sphingomyelinase (ASM); and cathepsins D and H [Canuel *et al.*, 2008; Nielsen *et al.*, 1999; Lefrancois *et al.*, 2003; Canuel *et al.*, 2008].

2.2.1. LIMP-2

2.2.1.1. A specific lysosomal receptor

LIMP-2 was first described in 1995 [Calvo *et al.*, 1995], but its role as a sorting receptor was only unveiled several years later when Reczek and colleagues [2007] unexpectedly realized it was responsible for the transport of GCase from the TGN to the endosomal/lysosomal compartment. LIMP-2 binds GCase in a pH-dependent fashion, enabling its association in the ER and transport to the lysosome, where the complex likely dissociates due to the acidic pH [Reczek *et al.*, 2007]. In fact, during the passage through the various organelles of the secretory pathway, the receptor-ligand complex experiences a gradient of decreasing pH: the nearly neutral milieu of the ER allows GCase to associate with LIMP-2, whereas the low pH of the endosomes/lysosomes leads to the dissociation of GCase from LIMP-2 [Zachos *et al.*, 2012]. Additional studies from Blanz *et al.* [Blanz *et al.*, 2010] demonstrated that, through disruption of either the helical arrangement or the amphiphatic nature of the LIMP-2 coiled-coil domain (residues 152 to 167), GCase binding was abolished. Furthermore, a synthetic peptide comprising the coiled-coil domain of LIMP-2 was shown to display pH-selective multimerization properties. All these data point out the importance of an intact coiled-coil structure for the interaction of LIMP-2 and GCase.

Further analysis of the pH dependency of LIMP-2 and GCase binding by perturbing the intra-lysosomal pH in cultured cells revealed that lumenal acidification mediated by the vacuolar (H^+)-ATPase triggers the dissociation of LIMP-2 and GCase in late endosomal/lysosomal compartments [Zachos *et al.*, 2012]. Moreover, the same authors identified a single histidine residue in LIMP-2 (H^{171}) that likely serves as the critical pH sensor. H^{171} is in close proximity to the previously described coiled-coil motif in the luminal domain of LIMP-2, which is necessary for GCase binding. Interestingly, that residue is not conserved in chicken or *Xenopus*, species that are also devoid of the

proposed coiled-coil domain, another important GCase binding factor. Thus, such species may be of great value to decipher additional roles of LIMP-2 other than transporting GCase to lysosomes [Zachos *et al.*, 2012].

2.2.1.2. LIMP-2 and disease: action myoclonus-renal failure syndrome

Mutations in the gene coding for this receptor prevent GCase from reaching the lysosome and from assuming its function inside the lysosome. Not surprisingly, *SCARB2* mutations underlie a serious autosomal-recessive disorder presently known as action myoclonus-renal failure syndrome (AMRF; OMIM#254900) [Berkovic *et al.*, 2008; Balreira *et al.*, 2008]. This disease combines progressive myoclonus epilepsy (PME), associated with storage material in the brain, and focal glomerulosclerosis, frequently with glomerular collapse.

At the biochemical level, this condition was shown to be characterized by pathological levels of GCase activity in fibroblasts, normal or slightly reduced levels in leukocytes, increased levels in plasma and the absence of other markers of Gaucher disease (GD), such as elevated chitotriosidase activity [Dardis *et al.*, 2009].

Recently, Hopfner and colleagues [Hopfner *et al.*, 2011], by screening a German AMRF family with these features associated with renal failure, provided data indicating that demyelinating polineuropathy and dilated cardiomyopathy are part of the AMRF syndrome. The findings were confirmed almost simultaneously by Dibbens *et al.* [2011], who described a patient with PME and demyelinating peripheral neuropathy, also carrying a *SCARB2* mutation that could explain the phenotype. This particular case had previously been published in the *Archives of Neurology* and presented with PME, preserved intellect and a nonprogressive generalized demyelination neuropathy [Costello *et al.*, 2009]. By that time, failure to establish the cause of the phenotype, despite the extensive evaluation, led the authors to describe a novel PME syndrome, but due to the absence of renal failure, they withdrew the initial AMRF diagnosis.

Currently, the presentation of PME without renal impairment [Dibbens *et al.*, 2009] prompts additional screenings of the *SCARB2* gene for PME causative mutations, as exemplified by the study of Rubboli *et al.* [2011], which reports on *SCARB2* mutations in five non-related Italian patients who presented with PME of unknown origin of adolescent or adult onset.

So far, no *SCARB2* genotype-phenotype correlations have been established. From the inspection of the *SCARB2* mutations found to date, no clear distinction emerges between the class (e.g., missense or nonsense) or position of the mutation and the patient phenotype. The phenotypic heterogeneity encompasses a wide range of tissues, but it is still unclear why some patients with *SCARB2* mutations develop PME with or without renal failure, or why others develop mild hearing impairment or peripheral neuropathy [Rubboli *et al.*, 2011].

Other symptoms may also occur, as suggested by the clinical manifestations observed in LIMP-2-deficient mice [Gamp *et al.*, 2003], which presented increased postnatal mortality and uni- or bilateral hydronephrosis caused by an obstruction of the ureteropelvic junction. In addition, LIMP-2-deficient mice suffered from a peripheral demyelinating neuropathy [Gamp *et al.*, 2003]. A subsequent study on that model revealed that the potassium channel KCNQ1 and its β-subunit KCNE1 were almost completely lost in the luminal part of marginal cells in the stria vascularis, affecting first higher and then lower frequency processing cochlear turns. The observation that LIMP-2 deficient mice displayed progressive hearing loss was particularly important because it suggested an important and previously unsuspected role for LIMP-2 in controlling the localization and the level of apically expressed membrane proteins such as KCNQ1, KCNE2 and megalin in the stria vascularis [Knipper *et al.*, 2006]. Extrapolation of mouse studies to humans is still premature, but it is tempting to hypothesize that LIMP-2 may assume additional roles in the endocytic pathway. Nevertheless, because for the moment only a few clinical and neurophysiologic descriptions of AMRF syndrome have been reported [Berkovic *et al.*, 2008; Balreira *et al.*, 2008; Horoupian *et al.*, 1977; Andermann *et al.*, 1986; Rothdach *et al.*, 2001; Badhwar *et al.*, 2004; Vadlamudi *et al.*, 2006], there is still much to learn about the clinical presentation and molecular basis of this disease. In the near future, more research exploring the function of *SCARB2* gene and its role in cerebral and renal function is warranted [Rubboli *et al.*, 2011] as well are new insights into its function in the endocytic pathway and its unidentified interacting partners [Knipper *et al.*, 2006] are also needed.

2.2.1.3. LIMP-2 and disease: *SCARB2* mutations as modifiers of Gaucher disease

Recently, *SCARB2* mutations have also been demonstrated to act as modifiers in GD. Considering that LIMP-2 is the only known receptor for the enzyme that is deficient in Gaucher disease, Velayati and co-workers [2011] hypothesized that *SCARB2* mutations could impact the Gaucher phenotype. In order to test that hypothesis, they re-evaluated a previously reported case of two siblings with GD with very disparate phenotypes [Eyal *et al.*, 1991; Ron *et al.*, 2008]. One sibling had been followed closely for over a decade with PME and dementia, while the second was subsequently diagnosed through family screenings, having only a few disease manifestations throughout his life, with no neurologic involvement. Both shared three *GBA* alterations: the maternally inherited c.535G>C (p.D140H) and c.10936G>A (p.E326K) and the paternally inherited c.586A>C (p.K157Q) [Eyal *et al.*, 1991]. Extensive analyses, including studies of endoplasmic reticulum-associated degradation (ERAD), had previously been performed on fibroblasts from both siblings, demonstrating that the one with myoclonic epilepsy had increased ERAD of GCase and elevated intracellular cholesterol [Ron *et al.*, 2008]. Velayati *et al.* [2011] screened this sib-pair for mutations in *SCARB2* and found one novel, heterozygous,

maternally-inherited mutation, c.1412A>G (p.E471G), in the brother with GD and myoclonic epilepsy, which was absent in his sibling and controls. This finding demonstrated that LIMP-2 could serve as a modifier in GD. Although myoclonic epilepsy had previously been reported in heterozygotes with S*CARB2* mutations [Reczek *et al.*, 2007], only in this case was it possible to demonstrate that the *SCARB2* mutation alone was not responsible for the development of such a phenotype, as it was also present in the sib's mother, who had a normal clinical phenotype [Velayati *et al.*, 2011]. In the severely affected brother, however, the deficient GCase activity, accompanied by a mutation in the transporter, resulted in mistrafficking of the enzyme and contributed to the observed phenotype.

2.2.2. SORTILIN

2.2.2.1. Structural features: insights into sortilin function and evolution

Sortilin is one of five members of a Vps10p domain receptor family, often found in the TGN and early endosomes: sortilin, SorLA, SorCS1, SorCS2 and SorCS3 [Willnow *et al.*, 2008]. The Vps10 domain is a 10-bladed β-propeller consisting of three structural domains: the N-terminal domain (residues 45-576), followed by two smaller domains named 10CC-a (residues 577-633) and 10CC-b (residues 634-716) [Willnow *et al.*, 2008]. Being the protein with the largest number of blades (10) at the Vps10 domain, sortilin has an increased potential to accommodate larger ligands when compared to similar proteins. With several highly conserved regions in the inner rim of the tunnel that correspond to different binding sites, sortilin has a unique structural organization that allows it to bind a series of different ligands. Furthermore, it is translated as an inactive propeptide, which is only activated after furine cleavage [Munck Petersen *et al.*, 1999]. Its simple structural organization seems to be a key element of protein regulation because the sortilin propeptide blocks premature binding of any ligand, preventing simultaneous binding of ligands with conflicting functions [reviewed in Willnow *et al.*, 2008]. Both sortilin and SorLA shuttle between the TGN and late endosomes through interactions between their cytoplasmic domains and various adaptors, including adaptor-protein-1 (AP1), Golgi-localized, γ ear-containing, ARF-binding proteins (GGA1, GGA2 and GGA3) and the retromer complex [Nielsen *et al.*, 2001; Nielsen *et al.*, 2007]. Furthermore, the cytosolic tail of sortilin closely resembles the one of the MPRs, containing motifs known to be involved in trafficking from the Golgi to the endosome and *vice versa*. The retrograde recycling of sortilin involves an interaction with the retromer complex through a specific site in the cytosolic tail [reviewed in Canuel *et al.*, 2009]. Finally, four particular amino acids in the cytoplasmic domain, which are shared by both sortilin and SorLA (the MVIA motif), seem to be crucial for binding to the cytosolic adaptors GGA1, 2 and 3 [Klinger *et al.*, 2011].

Table 2.1: Pleiotropic activities of sortilin in various biological pathways.

[*adapted from* Dubé *et al.*, 2011]

Pathway	Sortilin mediated role
Lysosomal transport	Mediates Golgi-to-lysosome transport of sphingolipid activator proteins (prosaposin and GM2 activator protein) [Lefrancois *et al.*, 2003], acid sphingomyelinase [Ni *et al.*, 2006] and of cathepsins D and H [Canuel *et al.*, 2008].
Neurotrophic signalling	Binds proNGF forming an apoptotic signalling complex with p75NTR facilitating neuronal cell death [Nykjaer *et al.*, 2004]; Transports TrkA, TrkB and TrkC to nerve synapse terminals enhancing their interaction with neurotrophins; Facilitates neuronal differentiation, survival and plasticity [Vaegter *et al.*, 2011].
Fronto-temporal lobar degeneration	Modulates extracellular PGRN levels via binding and endocytosis; pathogenesis via decreased extracellular PGRN levels through haploinsufficiency [Hu *et al.*, 2010].
Carbohydrate metabolism	Binds and translocates Glut4 to the plasma membrane from Glut4 storage vesicles in response to insulin signalling [Shi *et al.*, 2007].
Osteoblast differentiation	Limits LPL-mediated suppression of mineralization activity in differentiating hMSC [Maeda *et al.*, 2002]; Identified as a marker of late osteogenesis [Kulterer *et al.*, 2007].
Lipoprotein metabolism	Binds LDL RAP [Tauris *et al.*, 1998]; Surface-expression mediates binding to LPL and subsequent degradation by endocytosis [Nielsen *et al.*, 1999]; *In vitro* binding and endocytosis of apo A-V; facilitates lysosomal degradation of apo A-V [Nilsson *et al.*, 2008]; Robust association between *SORT1* variation (chromosome 1p13) and LDL-C concentrations [Samani *et al.*, 2007]; Expression assays show regulatory effects on plasma LDL-C concentrations [Linsel-Nitschke *et al.*, 2010; Musunuru *et al.*, 2010; Kjolby *et al.*, 2010]; Implicated in apo B-100 metabolism and VLDL secretion [Kjolby *et al.*, 2010].

There are some clues suggesting that the sortilin sorting pathway may represent an evolutionary mechanism for lysosomal transport older than the M6P pathway. Actually, even though the MPRs are usually described as analogous to the yeast Vps10p receptor, sortilin shares more similarities with Vps10p than the MPRs since it contains the Vps10 domain, which is not present in the MPRs [Tauris *et al.*, 1998]. Additionally, sortilin and

Vps10p interact with their ligands through proteinaceous interactions, while the MPRs recognize their ligands through an oligosaccharide side-chain, the M6P tag [Marcusson *et al.*, 1994; Kornfeld and Kornfeld, 1985; Baldwin *et al.*, 1993; reviewed in Canuel *et al.*, 2009]. It appears, therefore, that the M6P pathway emerged as a very specific mechanism for sorting soluble lysosomal hydrolases. Nevertheless, it did not completely substitute the mechanism mediated by sortilin, which retained the ability to translocate different cargo proteins (*see Section 3.4.3.*).

Albeit sortilin seems to be an ancient sorting receptor, during the evolutionary process it has developed additional functions [Table 2.1] that are gradually being discovered. In various pathways sortilin is revealing to play key roles, which can range from important functions in normal biological processes, such as glucose metabolism [Maeda *et al.*, 2002; Shi and Kandror, 2005] and neuronal cell proliferation and death [Willnow *et al.*, 2008; Nykjaer *et al.*, 2004], to influencing pathological processes like those underlying Alzheimer's disease [Mufson *et al.*, 2010]. Sortilin is also acknowledged to play a part in lipoprotein metabolism by interacting with receptor associated protein (RAP) [Petersen *et al.*, 1997] and binding lipoprotein lipase (LPL) [Nielsen *et al.*, 1999] and apolipoprotein (apo) A-V [Nilsson *et al.*, 2008].

2.2.2.2. A multiligand lysosomal receptor?

In 2003, the accumulated evidence of a possible role of sortilin as an alternative sorting receptor in the M6P-independent transport of newly synthesized hydrolases lead Lefrancois and colleagues [Lefrancois *et al.*, 2003] to investigate whether sortilin could be involved in the sorting of SAPs. Earlier reports firmly demonstrated that SAPs (prosaposin and GM2AP) could reach the lysosomes in an M6P-independent manner. A few years earlier, a carbohydrate-independent pathway of native GM2AP had been characterized [Rigat *et al.*, 1997]. Formerly, immunocytochemical analysis of tunicamycin-treated cells had shown that non-glycosylated prosaposin could be targeted more efficiently to lysosomes, whereas biochemical analysis revealed that prosaposin was associated with Golgi membrane fractions in an association not disrupted by free mannose 6-phosphate [Igdoura *et al.*, 1996]. When Lefrancois *et al.* [2003] performed their first study on sortilin's role as a lysosomal sorting receptor, they applied three approaches: a dominant-negative competition experiment, siRNA and co-immunoprecipitation (Co-IP) assays, which as a whole seemed to suggest that the SAPs (prosaposin and GM2AP) constituted a class of proteins whose lysosomal sorting and trafficking were mediated by sortilin. The dominant-negative construct was a truncated sortilin lacking the acidic cluster dileucine signal (Δcytosolic), which is implicated in the binding of GGAs [Nielsen *et al.*, 2001]. That construct was retained in the Golgi apparatus and localized to the same compartment as

the Golgi marker. Full-length sortilin, in contrast, was not restricted to the perinuclear region but extended into punctuate structures. A C-terminal myc-tag was added to the construct, to discriminate between truncate and endogenous sortilin. The overexpression of such a construct abolished the punctuate staining seen in non-transfected COS-7 cells, suggesting that the truncated sortilin competed for the binding of prosaposin and abolished its transport to the lysosomal compartment. Through siRNA assays, the authors saw that ablation of sortilin increases the secretion of prosaposin into the culture media. Finally, Co-IP assays showed that sortilin interacts with both SAPs [Lefrancois *et al.*, 2003].

In 2006, with the same dominant-negative sortilin construct and using confocal microscopy, Co-IP and pulse-chase experiments, Ni and colleagues [Ni *et al.*, 2006] suggested that sortilin was also involved in the lysosomal targeting of acid sphingomyelinase (ASM).

Two years later, the same group also implicated sortilin in the lysosomal targeting of cathepsins D and H [Canuel *et al.*, 2008]. Interestingly, when analyzing the cathepsins' interaction with the CI-MPR, the authors observed that, unlike cathepsin H, which bound exclusively to sortilin, cathepsin D was also able to bind to the CI-MPR. To better understand the role of sortilin and the M6P in the trafficking of cathepsins D and H, cells with both the sortilin and the M6P pathways blocked were examined, demonstrating that the transport of cathepsin D to the lysosomes occurs through both sorting pathways. In conclusion, sortilin appears to be the sole receptor necessary for lysosomal translocation of cathepsin H. Alternatively, the efficient transport of cathepsin D requires both the MPR and sortilin.

Other soluble proteases that are known to be transported in a M6P-independent fashion exist, but this transport is not mediated by sortilin. This is the case of cathepsins K and L, where evidence of M6P-independent transport are well established. Nevertheless, when both cathepsins were studied using an approach similar to that applied to cathepsins D and H, it was observed that neither cathepsins K nor L interacted and immunoprecipitated with the MPR, nor did they interact with sortilin. Furthermore, their subcellular localization remained unaltered in cells transfected with truncated sortilin when compared to untransfected cells [Canuel *et al.*, 2009].

Yet, many aspects concerning sortilin's function as an alternative receptor for lysosomal transport remain unclear and further studies are still needed to clarify sortilin's subcellular location and interaction with its putative ligands.

2.2.2.3. Sortilin and Disease: the cardiovascular risk association*

Recently, genome-wide association studies (GWAS) brought into connection sortilin and cardiovascular risk.

GWAS have been reasonably successful in the identification of single nucleotide polymorphisms (SNPs) within or near genes associated with variations in plasma lipid and lipoprotein levels. More than 100 loci were reported to account for a significant part of the genetic variation in triglyceride, low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol levels. In coronary artery disease or myocardial infarction (MI), GWAS have identified a smaller number of genetic loci, some of which are also associated with changes in traditional lipoprotein risk factors. Notably, within a widely replicated chromosome 1p13 locus associated with both myocardial infaction (MI) and LDL cholesterol levels [Kathiresan *et al.*, 2008; Kathiresan *et al.*, 2009; Musunuru *et al.*, 2010; reviewed in Tall and Ai, 2011], the relevant SNPs were found to be located in a gene cluster containing four genes, one of them being *SORT1*.

Subsequently, three independent teams [Linsel-Nitschke *et al.*, 2010; Musunuru *et al.*, 2010; Kjolby *et al.*, 2010] tried to understand the biological mechanism relating *SORT1* expression to LDL-cholesterol (LDL-C) levels and, ultimately, to risk of coronary artery disease (CAD). To this end, they used different approaches and, interestingly, they also came to different conclusions (Figure 2.7). The first study was developed by Linsel-Nitschke and colleagues [2010] who based on their findings, proposed that the overexpression of sortilin increases the internalization of LDL, with a consequent decrease of its plasma levels [Figure 2.7 (A) and (B)]. Soon after, through studies in human cohorts, hepatocytes and mice, Musunuru *et al.* [2010] reported an inverse relationship between sortilin expression and circulating LDL-C levels and further advanced with an explanatory mechanism through transcriptional regulation (liver-specific) of the *SORT1* gene by the transcription factor C/EBPα [Figure 2.7 (C) and (D)]. Conversely, Kjolby and his team [2010] observed a direct relationship between the expression of *Sort1* and the concentration of circulating LDL, suggesting it could result from increased VLDL secretion [Figure 2.7 (E) and (F)].

Several explanations have been given to justify the discrepancy between these results but the answer seems to rely in sortilin itself, which appears to be a multifaceted protein that may assume different functions depending on the circumstances. Overall however, the three studies presented strong evidence demonstrating that *SORT1* is a regulatory element of plasma LDL-C levels, adding a significant and previously unsuspected role to the sortilin-coding gene.

* For an extensive review on sortilin and the cardiovascular risk association see Appendix 1, review paper 4: Coutinho MF, Prata MJ, Alves S. **Sortilina e risco de doença cardiovascular** (Portuguese/English version) [*in press*].

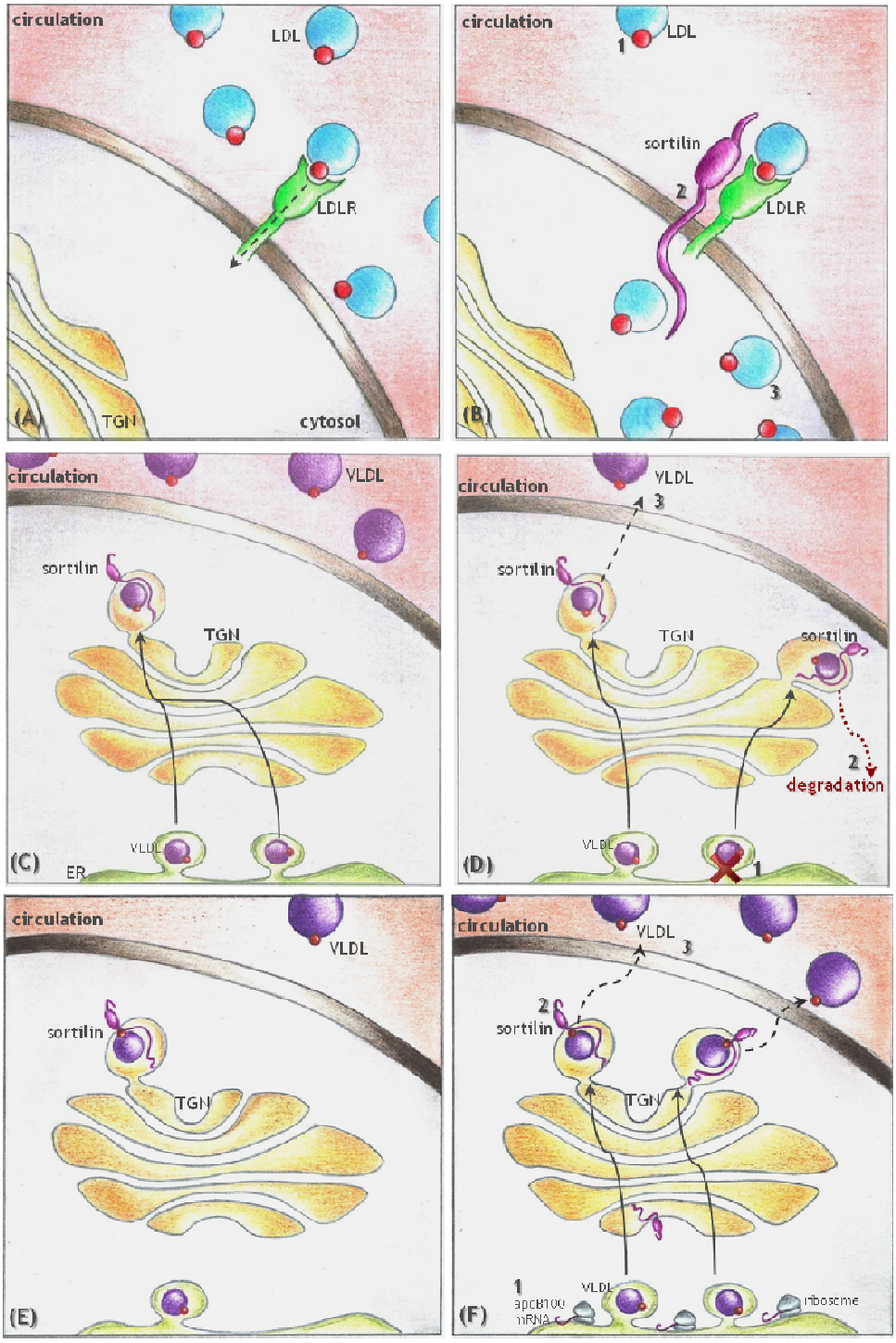


Figure 2.7: Schematic representation of the three proposed models of sortilin function

[*adapted from* Dubé *et al.*, 2011]

1. Linsel-Nitschke's model: sortilin increases LDL metabolism. From their observations Linsel-Nitschke *et al.* [2010] proposed that the overexpression of sortilin increases the internalization of LDL, with a consequent decrease of its plasma levels. **(A)** Basal circulating LDL levels in human embryonal kidney cell line. **(B)** Transient expression of *SORT1* in the same cell line was associated to increased internalization of extracellular labeled LDL particles. The authors proposed that increased interactions between sortilin and the LDLR with consequent increase of receptor-mediated endocytosis, were behind such observations. The upstream mechanistic basis for the upregulation of the LDLR was unclear, but could involve the LDLR directly, or related proteins like proprotein convertase subtilisin/kexin type 9.

The lowering of circulating LDL levels would be consistent with the expected clinical biochemical consequence of increased sortilin expression. **2. Musunuru's model: sortilin decreases the production of LDL precursors.** Musunuru *et al.* [2010] reported an inverse relationship between sortilin expression and circulating LDL-C levels and proposed an explanatory mechanism through transcriptional regulation (liver-specific) of the *SORT1* gene by the transcription factor C/EBPα. **(C)** Basal circulating LDL levels in a humanized mouse model. **(D)** Increased *SORT1* expression in the same cell line lowered LDL levels through a decrease in the production of very-low-density lipoprotein (VLDL), the precursor of LDL. While the upstream mechanistic basis of this decrease remain unclear, possibilities include disruption of lipidation of nascent lipoproteins (1) or shunting of nascent lipoproteins into intracellular degradation pathways (2). The lowering of circulating LDL precursors (3) would be consistent with the expected clinical and biochemical consequences of increased sortilin expression. **3. Kjolby's model: sortilin facilitates VLDL secretion, increasing circulating LDL levels.** (1) Apo B-100 is partially lipidated and transported to the TGN. (2) Binding of apo B-100 to sortilin facilitates maturation of pre-VLDL and secretion of VLDL from hepatocytes. *Sort1* expression increases potential for a sortilin apo B-100 complex, increases lipidation of nascent lipoproteins and VLDL formation. (3) Sortilin facilitates VLDL secretion from hepatocytes, with resulting increased circulating VLDL and ultimately LDL levels.

2.2.2.4. Sortilin and Disease: molecular basis for familial amyotrophic lateral sclerosis?

Recently, Belzil and co-workers [2012] hypothesized that, by reducing progranulin levels and promoting neurodegeneration (*See Table 2.1*), *SORT1* mutations or *SORT1* aberrant splicing could cause amyotrophic lateral sclerosis (ALS).

That hypothesis was based on additional independent assumptions. First, causative mutations had just been reported in the *TARDBP* and *FUS* genes, encoding respectively for TAR DNA binding protein TDP-43 and RNA binding protein FUS (both involved in the regulation of transcription and RNA splicing) [Kabashi *et al.*, 2008; Sreedharan *et al.*, 2008; Kwiatkowski *et al.*, 2009; Vance *et al.*, 2009], in both ALS and frontotemporal lobar degeneration (FTLD) [Broustal *et al.*, 2010; Huey *et al.*, 2012]. Second, it had also been reported that depletion of TDP-43 in the mouse adult brain modifies the expression levels of *FUS* and *GRN* (gene encoding progranulin) and alters the splicing of *SORT1* [Polymenidou *et al.*, 2011]. Finally, sortilin had also recently been identified as a receptor that mediates progranulin uptake [Hu *et al.*, 2010]. Thus, the authors speculated that *TARDBP*, *GRN*, and *SORT1* act in a common pathological pathway and that mutations in any of those genes could result in a similar phenotype.

To test this reasoning, they screened the *SORT1* coding sequence for mutations in a cohort of 112 unrelated individuals with familial ALS and tested for aberrant splicing by RT-PCR using RNA samples from cell lines expressing six different ALS-associated *TARDBP* mutations [Belzil *et al.*, 2012]. They identified one unique missense and two silent mutations; none of them were predicted to have functional effects. No aberrant *SORT1* splicing events were observed. From this study the authors concluded that *SORT1* mutations are not a common cause of familial ALS.

Chapter 2

Aims

A proper knowledge of the different trafficking mechanisms responsible for the sorting of lysosomal proteins is essential to a better understanding of lysosomal function and may also unveil the molecular defects underlying LSD phenotypes.

The major mechanism of transport of newly synthesized lysosomal enzymes to the endosomal/lysosomal compartment is mediated by the M6P recognition signal and its dysfunction results in the LSDs ML II and III. Alternative mechanisms of transport, M6P independent, were recently described but, for most cases it is still unclear how and if they might affect the function of the lysosome or the course of LSDs.

The present work is focused in the study of genes involved in both M6P and M6P independent trafficking routes to the lysosome, relating them with the occurrence of LSDs.

This thesis is divided in two major parts, each having specific aims. Part 1 addresses the molecular genetics, pathophysiology and diagnosis of disorders related with malfunction of M6P pathway and part 2 reports the study of the role of the M6P independent trafficking pathways in LSDs.

Part 1

Molecular genetics, pathophysiology and diagnosis of disorders related with malfunction of the Man-6P pathway

1. Regarding the study of M6P pathway, the main objectives were:

- To perform the molecular characterization of the Portuguese patients with Mucolipidoses II and III for *GNPTAB* and *GNPTG* genes and to identify the molecular defects underlying these LSDs in Portugal, by so contributing to a better knowledge of their mutational spectrum and geographical distribution;
- To enlarge our sample of ML II and ML III patients by contacting foreign institutions;
- To develop effective molecular genetic testing techniques for ML II and ML III confirming and refining the enzymatic diagnosis by elucidating the sub-type of the disease (alpha/beta or gamma), according to the affected gene;
- To carry out retrospective clinical analyses to assess the role of genetic variations on the clinical symptoms and outcome of ML II and III patients;
- To establish, whenever possible, genotype-phenotype correlations that may be used for future genetic counseling by clinicians;
- To study the impact of *GNPTAB* and *GNPTG* DNA mutations at mRNA and protein levels and obtain insights on their disease-causing mechanism.

2. The observation that the *GNPTAB* deletion c.3503_3504delTC was the most common disease allele among the Portuguese patients, while also showing a worldwide broad distribution, led to pursuing two additional goals:

- To investigate whether the worldwide spread of this mutation was due to a unique founder molecular lesion or if the deletion arose more than once through a recurrent mutational event;
- To estimate the age of this deletion and its place of origin.

3. Finally, the identification of two missense mutations (p.W81L and p.R986C) which were able to trigger a severe, early onset phenotype, contrarily to the usual phenotypes associated with missense mutations in the *GNPTAB* gene, prompted further analyses:

- To study the impact of the two mutations at protein expression levels and at the subcellular distribution of GNPTAB mutant proteins in order to get insights on the atypical genotype-phenotype relations observed.

Part 2

The role of Man-6P independent trafficking pathways in LSDs

1. Regarding the M6P independent lysosomal trafficking the main objectives were:

- To analyze molecularly *SORT1* and *SCARB2*, the genes that encode sortilin and LIMP-2, respectively, in patients with clinical phenotypes suggestive of having a lysosomal storage disease but still without specific diagnosis;
- To develop effective DNA-based methods for detection of mutations in the *SCARB2* and *SORT1* genes through both cDNA and gDNA analysis.

2. Taking into account that LSDs are often the cause of hydropsis fetalis, a severe fetal condition, another investigation was also planned having as aim:

- To analyze the *SCARB2* gene in the available samples of non-explained cases of hydrops fetalis;

3. The very recent evidence that a mutation in the gene that codes for LIMP-2 can act as a modifier in GD, led to address the issue through a study whose goal was:

- To perform the molecular screening of the *SCARB2* gene in GD patients in order to understand in what extent may *SCARB2* variants influence the severity of GD phenotype.

Chapter 3

Results and Discussion

Part 1

The M6P-dependent pathway

Here we present a summary of the major results obtained in the studies focused on the molecular, biochemical and genotype-phenotype correlation analysis of patients suffering from impairments on the mannose 6-phosphate travelling pathway: ML II and ML III. For a more detailed description see papers 1 to 6.

1: Molecular Characterization of ML II and ML III patients

We started by performing the molecular analysis of the *GNPTAB* and *GNPTG* genes in 13 Mucolipidosis II and III patients (10 Portuguese, 1 Finnish, 1 Spanish of Arab origin and 1 Indian), having identified 11 different mutations, 8 of which were novel: 6 in the *GNPTAB* gene (c.121delG, c.440delC, c.2249_50insA, p.W81L, p.I403T and p.E667X) and 2 in the *GNPTG* gene (c.610-1G>T and c.639delT). *GNPTAB* and *GNPTG* mRNA expression levels assessed through qRT-PCR analysis. The results suggested the possible existence of feedback regulation mechanisms between α/β- and the γ-subunits.
In general, there was a strong correlation between the observed clinical phenotype and the genotype assessed through molecular analysis. Severely affected ML II alpha/beta patients, with progressive postnatal forms and death occurring before 10 years of age, presented with nonsense or frameshift mutations in homozigousity in the *GNPTAB* gene whereas the presence of at least one mild mutation in the *GNPTAB* gene or of mutations in the *GNPTG* gene was associated with the mildest pathologies ML III alpha/beta or ML III gamma, respectively. The only exception was one patient harboring the *GNPTAB* missense mutation c.242G>T (p.W81L) in homozygosity, who showed a more severe phenotype. (See paper 1: Encarnação *et al.*, 2009).

As a result from our initial study, molecular genetic testing for ML II and ML III became available in our Institute, allowing carrier detection and prenatal molecular diagnosis in several families and contributing to an effective improvement of genetic counseling. Being one of the first laboratories to provide molecular genetic testing for ML II and III, we received several samples from other countries, whose clinicians and/or centers requested molecular characterization of affected individuals. In total, we received 2 patients from Denmark (1 of Palestinian origin), 3 Turkish patients and 2 of Indian origin. For all of them the molecular characterization was successful (Table 3.1), but the most interesting molecular defects were detected in the Palestinian and in the Indian patients, all suffering from ML II alpha/beta.

The Palestinian patient, referred by a specialized Danish center, carried a homozygous deletion of exon 19 (c.3435-386_3602+343del897). This was the first case of a large homozygous intragenic *GNPTAB* gene deletion in a ML II patient. Long-range PCR and sequencing methodologies were used to refine the characterization of this rearrangement, leading to the identification of a 21 bp repetitive motif in introns 18 and 19. Further analysis revealed that both the 5'and 3'breakpoints were located within highly homologous *Alu* elements suggesting that this deletion probably resulted from *Alu-Alu* unequal homologous recombination (See paper 2: Coutinho *et al.*, 2011a).

The first Indian ML II alpha/beta patient harbored a novel homozygous mutation, p.R986C, in homozygousity. This was the second case in our lab in which a missense mutation was not associated, as usually, to a mild phenotype (ML III alpha/beta), but instead presented with severe phenotypes (ML II alpha/beta) (See paper 3: Coutinho *et al.*, 2012a).

The second Indian patient, who died on the fourth day of life after respiratory distress, had a peculiar phenotype with severe prenatal skeletal abnormalities that fitted prior descriptions of a rare phenotype known as Pacman dysplasia. Trough retrospective analysis of the parents, he was inferred to be heterozygous for two pathogenic mutations in the *GNPTAB* gene: the well-known c.3503_3504delTC (p.L1168Q*fs*X5) and a novel frameshift variant, c.1701delC (p.F566L*fs*X5), thus supporting previous assumptions that at least some cases of Pacman dysplasia phenotype are, in fact, severe prenatal forms of ML II α/β (See paper 4: Coutinho *et al.*, *under preparation*).

Concerning the entire set of patients under study (Table 3.1), in general full characterization was possible applying the most conventional molecular approach, with the exception of two Portuguese subjects that were initially screened through cDNA sequencing but in whom only one of the mutant alleles was identified (See paper 1: patients P1 and P3). Latter, however, extensive gDNA and additional cDNA analyses were performed, eventually leading to a full characterization of the two individuals. Both were found to harbor NMD-triggering mutations: one nonsense mutation (p.E667X) and one small deletion (c.3503_3504delTC). These cases represent paradigmatic examples of tricky molecular diagnoses that could easily lead to a misclassification of the studied individuals, with tremendous consequences especially for prenatal diagnosis. Such cases highlight the importance of performing both biochemical and molecular genetic testing at both gDNA and cDNA levels whenever possible.

2: Biochemical Studies of ML II-causing mutations

The identification of three patients carrying homozygous missense mutations but presenting with early onset severe symptoms, defies previous assumptions that such mutations would usually underlie a mild phenotype, while only frameshift mutations would give rise to the more severe ML II alpha/beta phenotype. It became, then, mandatory to explore the causes underlying this unusual genotype/phenotype correlation in order to establish reliable genotype-phenotype predictions.

The pathogenic mutations c.1581delC, c.3503_3504delTC, p.W81L, p.S399F, p.I403T, p.R986C and p.K1236M were inserted into a *GNPTAB* miniconstruct ($\alpha\beta$-mini 1-430+849-

1256 in pcDNA3.1). Plasmids containing wild-type and mutant GNPTAB proteins were transfected into HEK and HeLa cells and analysed through Western Blot and Immunofluorescence, respectively.

In general, the combined analysis of the mutants' expression levels and subcellular location perfectly explains the phenotype: mutants harboring the severe deletions (c.1581delC and c.3503_4delTC) are retained in the ER, presenting only the non-cleaved α/β-precursor form; mild missense mutations (p.S399F, p.I403T and p.K1236M) are correctly located in the Golgi apparatus but present reduced levels of the mature α and β subunits, when compared to the wild-type and, finally, severe missense mutations were either retained in the ER on the non-cleaved and inactive precursor form (p.W81L) or not present at all (p.R986C) (See paper 5: Coutinho *et al., under preparation*).

3: Haplotypic Analysis of the frequent ML II-causing mutation c.3503_3504delTC

Another interesting result provided by the molecular study of ML II alpha/beta patients was the detection of a common mutation among our series of patients - the deletion c.3503_3504delTC. This mutation had already been detected in a series of populations and reports showed it to have a wide geographical distribution. To investigate whether its worldwide spread was due to a unique founder molecular lesion or if the deletion arose more than once through recurrent mutational event(s), we analysed 44 patients and 16 carriers sharing this deletion but being from diverse geographic regions, for 3 intragenic polymorphisms and 2 microsatellite markers flanking the *GNPTAB* gene. We identified a common haplotype in all chromosomes bearing the c.3503_3504delTC implying a common origin of the mutation. In addition, while the level of diversity observed at the most distant microsatellite locus indicated that the mutation is relatively ancient (~2063 years old), its geographical distribution further suggested that it probably arose in a peri-Mediterranean region (See paper 6: Coutinho *et al.*, 2011b).

ꕥ

As a complementary point to our study of patients suffering from diseases related to dysfunctional lysosomal traffic of M6P-tagged proteins we published:

- A minireview on the M6P travelling route to the lysosome with special focus on its role in lysosomal function and dysfunction (Appendix 1, review paper 2: Coutinho *et al.*, 2012).

	Mucolipidosis classification	Origin	Gene	Mutation (cDNA)	Mutation (protein)	Exon/ Intron affected	Original reference	
P1	ML III α/β	Portuguese	*GNPTAB*	c.1196C>T/c.3503_4delTC[a]	p.S399F/p.L1168Q*fs*X5[a]	e7/e19	Bargal *et al.*, 2006	Paper 1
P2	ML III α/β	Portuguese		c.1196C>T/c.1999G>T [a]	p.I403Tb/p.E667X [a]	e10/ e10	This study	Paper 1
P3	ML III α/β	Portuguese		c.1196C>T/c.1196C>T	p.S399F/p.S399F	e7/e7	Bargal *et al.*, 2006	Paper 1
P4	ML III α/β	Portuguese		c.242G>T/c.242G>T	p.W81L/p.W81L	e3/e3	Kudo *et al.*, 2006	Paper 1
P5	ML II α/β	Portuguese		c.3503_4delTC/c.1999G>T	p.L1168Q*fs*X5/p.E667X	e19/e13	Kudo *et al.*, 2006/ This study	Paper 1
P6	ML II α/β	Portuguese		c.3503_4delTC/c.3503_4delTC	p.L1168Q*fs*X5/p.L1168Q*fs*X5	e19/e19	Kudo *et al.*, 2006	Paper 1
P7	ML II α/β	Portuguese		c.3503_4delTC/c.3503_4delTC	p.L1168Q*fs*X5/p.L1168Q*fs*X5	e19/e19	Kudo *et al.*, 2006	Paper 1
P8	ML II α/β	Portuguese		c.3503_4delTC/c.3503_4delTC	p.L1168Q*fs*X5/p.L1168Q*fs*X5	e19/e19	Kudo *et al.*, 2006	Paper 1
P9	ML II α/β	Indian		c.440delC/c.440delC	p.A147A*fs*X5/p.A147A*fs*X5	e5/e5	This study	Paper 1
P10	ML II α/β	Portuguese		c.3503_4delTC/c.3503_4delTC	p.L1168Q*fs*X5/p.L1168Q*fs*X5	e19/e19	Kudo *et al.*, 2006	Paper 1
P11	ML II α/β	Arab		c.121delG/c.121delG	p.V41FfsX42/p.V41F*fs*X42	e2/e2	This study	Paper 1
P12	ML II α/β	Finish		c.1581delC/c.2249_50insA	p.S527S*fs*X20/p.N750K*fs*X8	e12/e13	Kudo *et al.*, 2006/ This study	Paper 1
P13	ML II α/β	Palestinian		c.3435-386_3602+343del897/ c.3435-386_3602+343del897	*3 abnormal transcripts*	i18/i18	This study	Paper 2
P14	ML II α/β	Danish		c.3503_4delTC/c.3503_4delTC	p.L1168Q*fs*X5/p.L1168Q*fs*X5	e19/e19	Kudo *et al.*, 2006	Not published
P15	ML II α/β	Turkish		c.3503_4delTC/c.542C>G	p.L1168Q*fs*X5/p.S181X	e19/e5	Kudo *et al.*, 2006/ This study	Not published
P16	[b]	Turkish		c.3503_4delTC/c.2822T>C	p.L1168Q*fs*X5/p. I941T	e19/e14	Kudo *et al.*, 2006/ This study	Not published
P17	ML II α/β	Turkish		c.3503_4delTC/c.3503_4delTC	p.L1168Q*fs*X5/p.L1168Q*fs*X5	e19/e19	Kudo *et al.*, 2006	Not published

P18	[b]	Turkish		c.1124delG/c.1124delG	p.R357Q*fs*X2/ p.R357Q*fs*X2	e10/e10	This study	Not published
P19	ML II α/β	Indian		c.2956C>T/c.2956C>T	p.R986C/p.R986C	e15/e15	This study	Paper 3
P20	ML II α/β[c]	Indian		c.3503_4delTC/c.1701delC	p.L1168Q*fs*X5/p.F566L*fs*X5	e19/e13	Kudo *et al.*, 2006/ This study	Paper 4
P21	ML II α/β	Indian		c.637-1G>A/c.637-1G>A	*3 abnormal transcripts*	i6/i6	Cathey *et al.*, 2010	Not published
P22	ML II α/β	Portuguese		c.242G>T/c.242G>T	p.W81L/p.W81L	e3/e3	Kudo *et al.*, 2006	Paper 5
P23	ML III λ	Portuguese	*GNPTG*	c.639delT/c.610-1G>T (IVS8-1G.T)	p.F213L*fs*X7/3 transcripts	e8/i8	This study	Paper 1

[a] When first published (Paper 1: Encarnação *et al.*, 2009) only one of the mutations had been identified for both patients through cDNA analysis. Subsequent gDNA sequencing and NMD-inactivation studies led to the identification of the second mutant allele on both patients.

[b] No clinical data was available for this patient.

[c] Pacman dysplasia phenotype.

Note: Patients are divided into two groups according to the mutated gene and samples are numbered according to reference in the text. These numbers are not in accordance with the papers.

1. Molecular Characterizations

Paper 1: Molecular analysis of the *GNPTAB* and *GNPTG* genes in 13 patients with Mucolipidosis type II or type III – identification of eight novel mutations

Paper 2: *Alu-Alu* recombination underlying the first large genomic deletion in GlcNAc-phosphotransferase α/β (*GNPTAB*) gene in a MLII alpha/beta patient

Paper 3: Mucolipidosis Type II alpha/beta with a homozygous missense mutation in the *GNPTAB* Gene

Paper 4: Neonatal Mucolipidosis type II overlapping Pacman dysplasia phenotype in an Indian family

Clin Genet 2009: 76: 76–84
Printed in Singapore. All rights reserved

CLINICAL GENETICS
doi: 10.1111/j.1399-0004.2009.01185.x

Short Report

Molecular analysis of the *GNPTAB* and *GNPTG* genes in 13 patients with mucolipidosis type II or type III – identification of eight novel mutations

Encarnação M, Lacerda L, Costa R, Prata MJ, Coutinho MF, Ribeiro H, Lopes L, Pineda M, Ignatius J, Galvez H, Mustonen A, Vieira P, Lima MR, Alves S. Molecular analysis of the *GNPTAB* and *GNPTG* genes in 13 patients with mucolipidosis type II or type III – identification of eight novel mutations.
Clin Genet 2009: 76: 76–84. © Blackwell Munksgaard, 2009

Mucolipidosis II (ML II) and mucolipidosis III (ML III) are diseases in which the activity of the uridine diphosphate (UDP)-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase (GlcNAc-phosphotransferase) is absent or reduced, respectively. In the absence of mannose phosphorylation, trafficking of lysosomal hydrolases to the lysosome is impaired. In these diseases, mistargeted lysosomal hydrolases are secreted into the blood, resulting in lysosomal deficiency of many hydrolases and a storage-disease phenotype. GlcNAc-phosphotransferase is a multimeric transmembrane enzyme composed of three subunits (α, β and γ) encoded by two genes – *GNPTAB* and *GNPTG*. Defects in *GNPTAB* result in ML II and III whereas mutations in *GNPTG* were only found in ML III patients. We have performed a molecular analysis of the *GNPTAB* and *GNPTG* genes in 13 mucolipidosis II and III patients (10 Portuguese, one Finnish, one Spanish of Arab origin and one Indian). Mutations were identified by the study of both cDNA and gDNA. The *GNPTAB* and *GNPTG* mRNA expressions were determined by quantitative reverse transcriptase polymerase chain reaction (qRT-PCR). The study led to the identification of 11 different mutations. Eight of these mutations are novel, six in the *GNPTAB* gene [c.121delG (V41FfsX42), c.440delC (A147AfsX5), c.2249_50insA (N750KfsX8), c.242G>T (W81L), c.1208T>C (I403T) and c.1999G>T (p.E667X)] and two in the *GNPTG* gene [c.610-1G>T and c.639delT (F213LfsX7)]. With regard to the mRNA expression studies, the values obtained by qRT-PCR indicate the possible existence of feedback regulation mechanisms between α/β and the γ subunits.

M Encarnação[a], L Lacerda[a], R Costa[a], MJ Prata[b], MF Coutinho[a], H Ribeiro[a], L Lopes[a], M Pineda[c], J Ignatius[d], H Galvez[c], A Mustonen[d], P Vieira[e], MR Lima[a] and S Alves[a]

[a]National Health Institute Dr Ricardo Jorge, Center of Medical Genetics Dr Jacinto Magalhães and [b]Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal, [c]Department of Neurology, Hospital Sant Joan de Deu and Center for Biomedical Research on Rare Diseases (CIBERER), Barcelona, Spain, and [d]Departments of Clinical Genetics and [e]Child Neurology, Oulu University Hospital, Oulu, Finland

Key words: feedback regulation mechanism – GlcNAc-phosphotransferase – *GNPTAB* – *GNPTG* – mucolipidosis II and III – mutation analysis

Corresponding author: Sandra Alves, Centro de Genética Médica Jacinto Magalhães, Praça Pedro Nunes, 88, 4099-028 Porto, Portugal.
Tel.: +351 226070300;
fax: +351 226070399;
e-mail: sandra.alves@igm.min-saude.pt; alvessandra@hotmail.com

Received 13 October 2008, revised and accepted for publication 26 January 2009

Mucolipidosis type II (ML II, I-cell disease; MIM 252500) and mucolipidosis type III (ML III, pseudo-Hurler polydystrophy; MIM 252600) are two autosomal recessive diseases of lysosomal hydrolase trafficking caused by total or near total deficiency of UDP-N-acetylglucosamine (UDP-GlcNAc): lysosomal enzyme N-acetylglucosamine-1-phosphotransferase (GlcNAc-phosphotransferase, EC 2.7.8.17).

In higher eukaryotes the trafficking of most lysosomal hydrolases is mediated by a mannose-6-phosphate (M6P)-dependent pathway, in which

an M6P residue is attached to asparagine-linked oligosaccharides on newly synthesized lysosomal hydrolases. GlcNAc-1-phosphotransferase catalyzes the first step of mannose residue phosphorylation of lysosomal enzymes, transferring GlcNAc-1-phosphate from UDP-GlcNAc to certain terminal or penultimate mannoses on high-mannose-type glycans.

In patients suffering from mucolipidosis II and III, the correct targeting of lysosomal hydrolases is impaired, leading to a massive secretion and an intracellular deficiency of acid hydrolases. ML II patients are characterized by dwarfism, skeletal abnormalities, developmental delay and cardiomegaly leading to death between 5 and 8 years of age (1, 2). ML III is a much milder disorder with later onset of clinical symptoms and more slowly progressing course allowing the survival into the eighth decade (3).

ML II and ML III are both rare diseases, in Portugal their birth prevalence is 0.16 and 0.08 cases per 100,000 live births, respectively (4).

The GlcNAc-phosphotransferase has been reported to be composed of three subunits, α, β and γ (5), encoded by two genes (1). It is believed that the α and/or β subunits of the enzyme contain the catalytic portion whereas the γ subunit is predictably involved in the recognition of lysosomal enzymes (6). However, the failure of lysosomal proteins to bind to a γ-subunit affinity matrix (7) supported the assumption that γ-subunits are not involved in interactions with specific substrates.

The gene coding for the α and β subunits (*GNPTAB*) contains 21 exons and spans 85 kb on chromosome 12q23.3. It encodes a protein of 1256 amino acids with a predicted molecular mass of 144 kDa (α/β precursor). Hydrophobicity analysis showed two transmembrane domains and 19 potential N-glycosylation sites. Proteolytic processing of the α/β precursor generates the individual α and β subunits (7, 8).

The gene that codes for the γ subunit (*GNPTG*) is located on chromosome 16, contains 11 exons and encodes a soluble protein containing 305 amino acids with a predicted molecular mass of 34 kDa (6, 9).

To date approximately 40 mutations have been described in the *GNPTAB* gene (7, 8, 10–17) and eight mutations in the *GNPTG* gene (6, 9, 18). Most of these mutations are private or rare. According to a new classification proposed by Cathey et al. (19) mutations in the *GNPTAB* gene are associated with ML II α/β and ML III α/β whereas mutation in the *GNPTG* gene cause ML III γ.

Here, we have undertaken a molecular study of the *GNPTAB* and *GNPTG* genes in 13 mucolipidosis II and III patients (ten Portuguese, one Finnish, one Spanish of Arab origin and one Indian) that lead to the detection of six novel mutations in *GNPTAB*, two novel mutations in *GNPTG* and three previously described mutations in the *GNPTAB* gene. The impact of each mutation on mRNA and protein levels as well as on the genotype/phenotype correlation was also addressed.

Material and methods

Clinical diagnosis/patients

Our study enrolled 13 unrelated patients with ML II and III, whose clinical diagnosis (Table 1) had been biochemically confirmed by demonstrating that the activities of several lysosomal enzymes were increased in the serum and decreased in cultured fibroblasts. Glucocerebrosidase activity was used as a reference because this enzyme is transported to lysosomes by an M6P-independent pathway.

Cell culture

Cell lines from patients were established from skin biopsies. Human fibroblasts were grown in minimal essential medium (MEM; Gibco Invitrogen, Carlsbad, CA) with 20% fetal calf serum and antibiotics (10,000 U/ml streptomycin and penicillin, 1 mg/ml fungizone), at 37°C in a humidified atmosphere with 5% CO_2.

Molecular diagnosis

Total cellular RNA was extracted from cultured fibroblasts using the 'High Pure RNA Isolation Kit' (Roche, Basel, Switzerland) and reverse transcribed using the 'First-strand cDNA synthesis kit' (Amersham Biosciences, Munich, Germany).

The *GNPTAB* and *GNPTG* cDNAs were amplified with specific primers in six and three overlapping fragments, respectively (primer sequences and PCR conditions available on request).

Genomic DNA was isolated from cultured skin fibroblasts and peripheral blood total leukocytes according to standard procedures.

In order to confirm the presence of each mutation found in *GNPTAB* and *GNPTG* genes, PCR amplification of the corresponding exons as well as of the adjacent intronic regions was performed with specific primers (primer sequences and PCR conditions available on request).

Samples were submitted to automated DNA sequencing, performed on an ABI PRISM

Table 1. Clinical findings of the patients studied

Patient	Clinical classification	Age of onset	Clinical symptoms	Other information
P1	ML III	3 years	-	At least the patient reached the age of 20 years
P3	ML III	7 years	Polydystrophic dwarfism	-
P5	ML III	9 years	-	-
P11	ML II	4 months	Psychomotor retardation, multiple congenital malformations, high forehead, trigonocephaly, gingival hyperplasia, micrognathia, flat nasal bridge, pectus excavatum, bilateral inguinal hernia, diastasis recti, camptodactyly, arthrogryposis, generalized hypotonia, severe growth retardation, hip subluxation, generalized cortical atrophy.	Death at 23 months
P2	ML II	<1 year	Dysmorphic facial features, gingival hyperplasia, generalized hypotonia.	-
P7	ML II	<1 year	Dysmorphic facial features, gingival hyperplasia, radiographic skeletal changes	-
P8	ML II	<1 year	Coarse facial features, altered body trunk, psychomotor retardation	-
P9	ML II	-	-	-
P10	ML II	<1 year	Delayed milestones since birth, distorted skull since birth, cannot stand/walk, non-development of dentition, craniosynostosis, hypopigmented skin, large tongue, restricted joint movement of all four limbs	-
P4	ML II	Prenatal diagnosis	-	Two affected siblings. Both presented severe clinical and radiological manifestations characteristic of ML II. One of them survived two and half months and the other only two days
P12	ML II	22 months	Dysplasic hips, coarse features, short body trunk, thin and fragile hair. Facial dysmorphism with orbital hypertelorism, very populated eyebrows, big mouth with gross lips and severe gingival hypertrophy. Lumbar scoliosis of the spine. No cardiac affectation, nor organomegaly. Articulated movements of the hip, elbows and knees were decreased. Marked axial hypotonia with normal strength, osteotendinous reflexes were decreased. Absence of cephalic control nor sitting and no visual follow up.	-

Table 1. Continued

Patient	Clinical classification	Age of onset	Clinical symptoms	Other information
P13	ML II	<1 year	Evident psychomotor retardation, enlarged liver (at 2 years) and large kidneys.Coarse facial features, iliac spines were broad and the medial phalanxes of the hand were thick. In the skull the bone development was advanced and there was early closure of the anterior fontanelle. The base of the skull appeared sclerotic. MRI showed thin corpus callosum.	-
P6	ML III	10 years	Without dysmorphic features, without psychomotor retardation, difficulty in closing hands	At least the patient reached the age of 18 years

3130xl. Results were analyzed with the sequence analysis software FINCHTV, version 1.3.1.

The presence of each mutation was always confirmed in two independent experimental assays and when available also in patient's relatives. In addition, all novel mutations were screened in 50 unrelated Portuguese controls.

Real-time reverse transcriptase polymerase chain reaction

The *GNPTAB* and *GNPTG* mRNA levels were determined by real-time quantitative reverse transcriptase polymerase chain reaction techniques. PCR reactions were prepared in a final volume of 50 µl using 1× Taqman Universal PCR Master Mix (Applied Biosystems), 1× TaqMan gene expression assays-on-demand kit (which included the primers and the probes) Hs00225647_m1 for the *GNPTAB* gene and Hs01126100_gH for the *GNPTG* gene and 1 µg of RNA converted into cDNA. Thermal cycling conditions comprised an initial UNG incubation at 50°C for 2 min, 95°C for 10 min, 40 cycles of denaturation at 95°C for 15 s, annealing and extension at 60°C for 1 min. Each measurement was performed in triplicate and the threshold cycle (Ct) – the fractional cycle number at which the amount of amplified target reached a fixed threshold – was determined. Relative quantification of gene expression was performed using the standard curve method comprising four serial dilution points (ranging from 0.05 to 50 ng). Glyceraldehyde 3-phosphate dehydrogenase (GAPDH, Applied Biosystems, #431098) and human phosphoglycerate kinase 1 (PGK1, Applied Biosystems, #4326318E) gene expression was used for data normalization. The relative quantification of the RNA was determined through the ratio of the normalized expressions of the target sample and the normalized control sample. qRT-PCR analyses were performed on an Applied Biosystems ABI PRISM 7000 Sequence Detection System.

Other analysis

Values obtained for the relative quantification of *GNPTAB* and *GNPTG* mRNA in patients and control samples were compared using Student's test. Differences were considered significant when $p < 0.05$.

To evaluate the potential effect of *missense* mutations, two bioinformatics tools were used: the PolyPhen (Polymorphism Phenotyping;

http://coot.embl.de/PolyPhen/) and the Sorting Intolerant From Tolerant (SIFT; http://blocks.fhcrc.org/sift/SIFT.html).

The prediction of the scores for the splice-site junctions was obtained with the MAXENT program (http://genes.mit.edu/burgelab/maxent/Xmaxentscan scoreseq acc.html). Ideal MAXENT splice-site scores are: 11.81 (5_ss) and 13.59 (3_ss).

Results/Discussion

Identification and characterization of mutations

In this study we have analyzed the entire coding regions of *GNPTG* and *GNPTAB* genes in 13 mucolipidosis II and III patients (10 Portuguese, one Finnish, one Spanish of Arab origin and one Indian).

In 11 of them, molecular characterization was successfully completed and in the remaining two patients (P1 and P3) only one mutation could be detected (Table 2).

For some of the patients (P1, P2, P3, P10, P11 and P12) it was also possible to study their relatives and in all cases mutated alleles were present in both parents.

All the detected variant alleles were excluded to be present in a minimum of 50 Portuguese healthy individuals (100 control alleles), confirming that none of them was a polymorphic variation at least in the Portuguese population.

GNPTAB gene

As summarized in Table 2, we have identified nine different mutations in the *GNPTAB* gene, therefore revealing a high molecular heterogeneity, which is in fair agreement with the mutational spectrum previously reported (7, 8, 10–17). The majority of the detected mutations (seven out of nine) were found to be located between exons 2 and 13.

Six of the eight mutations are novel: c.121delG (V41FfsX42), c.440delC (A147AfsX5), c.2249_50insA (N750KfsX8), c.242G>T (W81L), c.1208T>C (I403T) and c.1999G>T (p.E667X) while the others were previously described: c.1196C>T (S399F) (13), c.1581delC (S527SfsX20) (8) and c.3503_4delTC (L1168QfsX5) (8).

As the impact on protein function has been already assessed for the two previously described mutations, herein we focused only on the six newly discovered mutations, two of the six mutations [c.242G>T (W81L) and c.1208T>C (I403T)], leading to amino acid replacements. The mutation c.242G>T (W81L) corresponds to the replacement of a tryptophan by a leucine and involves an amino acid residue evolutionary highly conserved among the compared species. The mutation c.1208T>C (I403T) occurs the substitution of the non-polar isoleucine residue by the polar amino acid threonine, which alters a conserved motif in a domain homologous to the N-terminal bacterial capsule biosynthesis proteins (7). These two missense mutations were submitted to the predictive methods implemented in Polyphen and SIFT bioinformatic tools and both were expected to be potentially damaging to the protein function with a high probability of functional impairment.

The nonsense mutation c.1999G>T (p.E667X) creates a premature stop codon in the *GNPTAB* mRNA open reading frame (ORF) leading to a truncated GlcNAc-phosphotransferase precursor lacking 590 amino acids, thus implying that none of the subunits (i.e. neither α nor β) will be formed.

The other novel mutations, the deletions c.121delG (V41FfsX42), c.440delC (A147AfsX5) and the insertion c.2249_50insA (N750KfsX8) were also found to be responsible for the alteration of *GNPTAB* mRNA ORF, causing the appearance of premature stop codons. These mutations are thus presumed to severely affect the formation of the GlcNAc-phosphotransferase precursor, as all of them give rise to very truncated proteins with 81, 151 and 756 amino acids, respectively.

The most frequent *GNPTAB* mutation in ML II Portuguese patients was found to be the c.3503_4delTC (L1168QfsX5), accounting for 45% of the mutant alleles detected in the Portuguese series. Interestingly, this mutation was also reported to be the most frequent in ML II and ML III α/β patients of Arab–Muslim origin (13) and the unique mutation present in a French-Canadian founder population (17).

GNPTG gene

One patient was found to be a compound heterozygous for two novel mutations in the *GNPTG* gene, namely c.610-1G>T and c.639delT (F213LfsX7).

When the respective *GNPTG* cDNA fragments were amplified, the presence of four bands was observed. Three bands of this unusual four-banded pattern correspond to alternative transcripts that are originated by the presence of the splicing mutation c.610-1G>T in intron 8. The mutation alters a conserved 3′ splice site in intron 8 (see Fig. 1). In the normal sequence, the score value predicted by MAXENT software for this

Table 2. Mutations identified in 13 mucolipidosis patients

Patient	Mucolipidosis classification	Origin	Gene	Mutation (cDNA)	Mutation (protein)	Exon/intron affected	Prevision of the effect of the mutation on protein function		Phenotype	Reference
							Poliphen	SIFT		
P1	ML III α/β	Portuguese	*GNPTAB*	1196C>T/–[a]	S399F/–[a]	e7	Probably deleterious	P< 0.05 not tolerated	Mild	(13)
P3	ML III α/β	Portuguese		1208T>C/–[a]	I403T[b]/–[a]	e10	Probably deleterious	P< 0.05 not tolerated	Mild	This study
P5	ML III α/β	Portuguese		1196C>T/1196C>T	S399F/S399F	e7	Probably deleterious	P< 0.05 not tolerated	Mild	(13)
P11	ML III α/β	Portuguese		242G>T/242G>T	W81L[b]/W81L[b]	e3	Probably deleterious	P< 0.05 not tolerated	Severe	This study
P2	ML II α/β	Portuguese		3503_4delTC/1999G>T	L1168QfsX5/E667X	e19/e13	NA	NA	Severe	(8)/This study
P7	ML II α/β	Portuguese		3503_4delTC/3503_4delTC	L1168QfsX5/L1168QfsX5	e19	NA	NA	Severe	(8)
P8	ML II α/β	Portuguese		3503_4delTC/3503_4delTC	L1168QfsX5/L1168QfsX5	e19	NA	NA	Severe	(8)
P9	ML II α/β	Portuguese		3503_4delTC/3503_4delTC	L1168QfsX5/L1168QfsX5	e19	NA	NA	Severe	(8)
P10	ML II α/β	Indian		440delC/440delC	A147AfsX5/A147AfsX5	e5	NA	NA	Severe	This study
P4	ML II α/β	Portuguese		3503_4delTC/3503_4delTC	L1168QfsX5/L1168QfsX5	e19	NA	NA	Severe	(8)
P12	ML II α/β	Arab		121delG/121delG	V41FfsX42/ V41FfsX42	e2	NA	NA	Severe	This study
P13	ML II α/β	Finish		1581delC[c]2249_50insA	S527SfsX20/ N750KfsX8	e12/e13	NA	NA	Severe	(8)/This study
P6	ML III γ	Portuguese	*GNPTG*	639delT/610-1G>T (IVS8-1G>T)	F213LfsX7/3 transcripts	e8/i8	NA	NA	Mild	This study

NA, not applicable.
[a]For the patients P1 and P3 one mutation remains unknown.
[b]The evidence for novel missense mutations to actually cause deficiency is only hypothetical. Expression studies are necessary to prove them as causal mutations.
[c]This mutation is designed in reference 8 as c.1744delC because numbering of cDNA in this reference is different from others.

Encarnação et al.

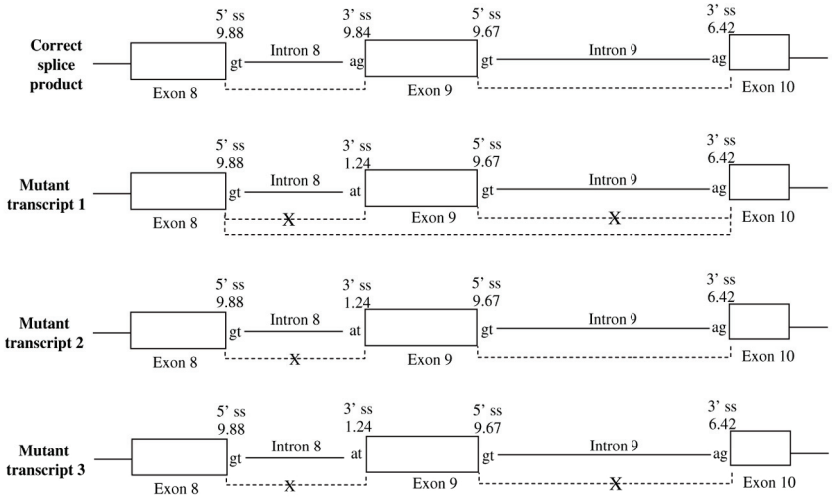


Fig. 1. Schematic representation of the transcripts obtained in the presence of c.610-1G>T substitution. X, splicing mechanism represented is not used.

splice site is 9.84. The replacement of guanine by thymidine in this position dramatically reduces the score value to 1.24 and as a consequence three different aberrant transcripts may be produced. In one of the mutant transcripts exon 9 is completely skipped, in another intron 8 becomes incorporated and turns to be the predominant transcript, while in the third transcript both introns 8 and 9 are incorporated. It is not clear why this splicing mutation occurring in intron 8 also affects the splicing of intron 9. However, finding that it really takes place indicates that the disruption of the splice site in intron 8 strongly impairs the correct recognition mechanism of normal splice sites in the nearby intron located upstream.

The presence of the c.639delT (F213LfsX7) in the other allele is responsible for the alteration of the *GNPTG* mRNA ORF leading to the appearance of a premature stop codon, which leads to a protein with only 218 amino acids.

Impact of identified mutations at *GNPTAB* and *GNPTG* mRNA levels

Results on mRNA relative quantification are shown in Table 3. The values obtained indicate that all the identified mutations in the *GNPTAB* gene (deletions, non-sense or mis-sense) are associated with a significant decrease in mRNA *GNPTAB* levels compared with the median controls values. Both deletions and nonsense mutations generate premature termination codons. So the observed reduction in mRNA levels might reasonably be explained by the involvement of the transcripts in the nonsense-mediated mRNA decay (NMD) pathway. However, this response is not triggered by the missense substitutions harbored in homozygosity by patients P5 and 11, meaning that other factors inducing enhanced mRNA instability and/or degradation may be responsible for the significant reduction (although minor) in their *GNPTAB* mRNA levels.

Also patients P1 and P3, heterozygous for missense mutations (despite in each patient one causative mutation remains unidentified) showed significant reduction in the *GNPTAB* mRNA levels.

In the patient with alterations in the *GNPTG* gene [c.610-1G>T/c.639delT (F213LfsX7)] both *GNPTAB* and *GNPTG* mRNA levels were significantly decreased (2.4- and 10-fold, respectively) compared with controls.

The reduction of the *GNPTG* mRNA levels can be easily explained by the nature of the two mutations, a deletion and a splicing mutation as both result in transcripts that are predictably targets of the NMD mechanism. On the other hand, the observed low-level expression of *GNPTAB* might be related to the existence of feedback mechanism between the *GNPTG* and *GNPTAB* encoded subunits. Evidence that a reciprocal feedback mechanism might be acting between α/β and γ subunits was recently reported by Ho et al. (20) when studying the expression patterns of the *GNPTAB* and *GNPTG* subunits in human mucolipidosis cell lines.

The existence of such a mechanism also enables to explain the reduction of the mRNA *GNPTG* levels now observed in three patients who presented mutations in the *GNPTAB* gene, being patient P11 [homozygous c.242G>T (W81L)], patient P1 [c.1196C>T (S399F)/unknown] and patient P3 [c.1208T>C (I403T)/unknown]. In the remaining patients, no significant alteration of the mRNA *GNPTG* levels was observed, which may mean that the relationship between the expression patterns of *GNPT* genes depends on the presence of specific mutations.

Table 3. Relative quantification of mRNA *GNPTAB* and mRNA *GNPTG* by real time quantitative reverse transcriptase polymerase chain reaction (qRT-PCR)[a]

Gene	Patient	Genotype	GNPTAB Ratio observed in relation to control sample	mRNA relative quantification[b]	p-value (*t*-test)	GNPTG Ratio observed in relation to control sample	mRNA relative quantification[b]	p-value (*t*-test)
GNPTAB	P1	c.1196C>T/–	−4.0	0.27 ± 0.14	0	−1.9	0.58 ± 0.21	0.015
	P2	c.3503_4delTC/c.1999G>T	−6.3	0.17 ± 0.15	0.007	−1.2	0.93 ± 0.64	0.700
	P3	c.1208T>C /–	−4.1	0.27 ± 0.16	0	−1.9	0.57 ± 0.45	0.062
	P5	c.1196C>T/c.1196C>T	−2.0	0.54 ± 0.22	0.014	+1.1	1.10 ± 0.43	0.939
	P4, P7,P8	3503_4delTC /3503_4delTC	−4.7	0.23 ± 0.13	0	1.0	1.04 ± 0.33	0.716
	P10	c.440delC/c.440delC	—	—	—	—	—	—
	P11	c.242G>T/c.242G>T	−3.0	0.36 ± 0.14	0	−2.0	0.53 ± 0.12	0
	P12	c.121delG/c.121delG	—	—	—	—	—	—
	P13	c.1581delC/c.2249_50insA	−4.1	0.27± 0.19	0	+1.1	1.21 ± 0.40	0.455
GNPTG	P6	c.del639T/ c.610-1G>T	−2.4	0.45 ± 0.24	0.001	−10.8	0.10 ± 0.05	0
		Wt/Wt		1.09 ± 0.06			1.08 ± 0.06	

—, the real-time qRT-PCR was not performed in these patients due to the fact that RNA concentrations of both samples were very low.
[a]For patients 1, 3, 11, 8 and 6 two samples were analyzed. For patients 5, 2, 7, 4, 12 and 13 it was only possible to analyze one sample.
[b]Values are expressed in mean ± SD.

However, contrary to the findings of Ho et al. who showed that low levels of both α and β subunits were associated with increased expression of the γ subunit in patients with alterations in *GNPTAB* (ML II and ML III α/β), none of the patients studied here with mutations in that gene showed increased *GNPTG* mRNA levels. Therefore further studies are needed in order to obtain a better understanding of the factors that control the balance between α/β and γ subunits.

Phenotype–genotype relationship

In general, we observed a strong correlation between clinical phenotype and genotype assessed through molecular analysis (see Table 2).

The presence of nonsense and frameshift mutations in *GNPTAB* (P2, P4, P7, P8, P9, P10, P12 and P13) was systematically associated with severe clinical manifestations and patients harboring those mutations were classified as ML II, and presented earlier onset of signs and symptoms. These results are in accordance with previous studies (13, 14).

The majority of patients with missense mutations in at least one allele (P1, P3 and P5) presented a mild phenotype and were clinically classified as ML III (ML III α/β). Only patient P11, homozygous for the missense mutation c.242G>T (W81L), showed a more severe phenotype. He was diagnosed at the age of 4 months presenting clinical features of mucolipidosis type II such as gingival hypertrophy, hypertonia and delayed growth. Death occurred at 23 months of age. As already mentioned, this mutation affects an amino acid in a highly conserved residue of the α chain, which is located in a region of the protein crucial for its normal function. Therefore its presence in homozygosity explains the severity of the clinical manifestations.

Finally, the presence in compound heterozygosity of two *GNPTG* mutations in patient P6 was accompanied by a mild phenotype characterized by later onset of clinical symptoms and survival into adulthood. The patient was classified as ML III (ML III γ), reinforcing previous data on the strong association between ML III and mutations at *GNPTG*, and his milder clinical course can probably be explained by the fact that the γ subunit does not contribute to the catalytic function of the enzyme.

In conclusion, we have observed that patients with progressive postnatal forms and death occurring before 10 years of age – ML II α/β- are almost all associated with the presence of nonsense or frameshift mutations in homozygosity in the *GNPTAB* gene whereas the presence of at least

one mild mutation in the *GNPTAB* gene or mutations in the *GNPTG* gene are associated with ML III – ML III α/β and ML III γ, respectively.

Acknowledgements

This work was partially supported by FCT (Project PIC/IC/83252/2007).

References

1. Kornfeld S, Sly WS. I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localisation. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. Childs B, Kinzler KW, Vogelstein B, assoc. eds. The metabolic and molecular bases of inherited disease, 8th edn. New York: McGrawHill, 2001: 3421–3442.
2. Spranger JW, Brill PW, Poznanski AK. Bone dysplasias: an atlas of genetic disorders of the skeletal development. New York: Oxford University Press, 2002: 57–79.
3. Umehara F, Matsumoto W, Kuriyama M, Sukegawa K, Gasa S, Osame M. Mucolipidosis III (pseudo-Hurler polydystrophy); clinical studies in aged patients in one family. J Neurol Sci 1997: 146: 167–172.
4. Pinto R, Caseiro C, Lemos M et al. Prevalence of lysosomal storage diseases in Portugal. Eur J Hum Genet 2004: 12 (2): 87–92.
5. Bao M, Booth JL, Elmendorf BJ, Canfield WM. Bovine UDP-N-acetyl-glucosamine: lysosomal-enzyme N-acetyglucosamine-1-phosphotransferase, Purification and subunit structure. J Biol Chem 1996: 271: 31437–31445.
6. Raas-Rothschild A, Cormier-Daire V, Bao M et al. Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC). J Clin Invest 2000: 105: 673–681.
7. Tiede S, Storch S, Lübke T, Henrissat B, Bargal R, Raas-Rothschild A, Braulke T. Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase. Nat Med 2005: 11: 1109–1112.
8. Kudo M, Brem MS, Canfield WM. Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha/beta-subunits precursor gene. Am J Hum Genet 2006: 78 (3):451–463.
9. Tiede S, Cantz M, Raas-Rothschild A et al. A novel mutation in UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) in two siblings with mucolipidosis type III alters a used glycosylation site. Hum Mutat 2004: 24: 535–536.
10. Tiede S, Muschol N, Reutter G, Cantz M, Ullrich K, Braulke T. Missense mutations in N-acetylglucosamine-1-phosphotransferase alpha/beta subunit gene in a patient with mucolipidosis III and a mild clinical phenotype. Am J Med Genet A 2005: 137A (3): 235–240.
11. Steet RA, Hullin R, Kudo M, Martinelli M, Bosshard NU, Schaffner T, Kornfeld S, Steinmann B. A splicing mutation in the alpha/beta GlcNAc-1-phosphotransferase gene results in an adult onset form of mucolipidosis III associated with sensory neuropathy and cardiomyopathy. Am J Med Genet A 2005: 132 (4): 369–375.
12. Paik KH, Song SM, Ki CS et al. Identification of mutations in the GNPTA (MGC4170) gene coding for Glc-NAc-phosphotransferase alpha/beta subunits in Korean patients with mucolipidosis type II or type IIIA. Hum Mutat 2005: 26 (4): 308–314.
13. Bargal R, Zeigler M, Abu-Libdeh B et al. When Mucolipidosis III meets Mucolipidosis II: GNPTA gene mutations in 24 patients. Mol Genet Metab 2006: 88 (4): 359–363.
14. Tiede S, Cantz M, Spranger J, Braulke T. Missense mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTA) in a patient with mucolipidosis II induces changes in the size and cellular distribution of GNPTG. Hum Mutat 2006: 27 (8): 830–831.
15. Lam CW, Yan MS, Li CK, Lau KC, Tong SF, Tang HY. DNA-based diagnosis of mucolipidosis type IIIA and mucopolysaccharoidosis type VI in a Chinese family: a chance of 1 in 7.6 trillion. Clin Chim Acta 2007: 376 (1–2): 250–252.
16. Tappino B, Regis S, Corsolini F, Filocamo M. An Alu insertion in compound heterozygosity with a microduplication in GNPTAB gene underlies Mucolipidosis II. Mol Genet Metab 2008: 93 (2): 129–133.
17. Plante M, Claveau S, Lepage P et al. Mucolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTAB) in a French Canadian founder population. Clin Genet 2008: 73 (3): 236–244.
18. Raas-Rothschild A, Bargal R, Goldman O et al.Genomic organisation of the UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) and its mutations in mucolipidosis III. J Med Genet 2004: 41 (4): e52.
19. Cathey SS, Kudo M, Tiede S et al. Molecular order in mucolipidosis II and III nomenclature. Am J Med Genet A 2008: 146A: 512–513.
20. Ho CY, Tang NL, Yeung AK, Lau AK, Hui J, Lo AW. Abnormal expressions of the subunits of the UDP-*N*-acetylglucosamine: lysosomal enzyme, N-acetylglucosamine-1-phosphotransferase, result in the formation of cytoplasmic vacuoles resembling those of the I-cells. J Mol Med 2007: 85 (4): 351–360.

JIMD Reports
DOI 10.1007/8904_2011_83

RESEARCH REPORT

Alu–Alu Recombination Underlying the First Large Genomic Deletion in GlcNAc-Phosphotransferase Alpha/Beta (*GNPTAB*) Gene in a MLII Alpha/Beta Patient

Maria Francisca Coutinho • Liliana da Silva Santos • Lúcia Lacerda • Sofia Quental • Flemming Wibrand • Allan M. Lund • Klaus B. Johansen • Maria João Prata • Sandra Alves

Received: 5 May 2011 / Revised: 26 July 2011 / Accepted: 27 July 2011 / Published online: 20 October 2011

Abstract Mucolipidosis type II α/β is a severe, autosomal recessive lysosomal storage disorder, caused by a defect in the *GNPTAB* gene that codes for the α/β subunits of the GlcNAc-phosphotransferase. To date, over 100 different mutations have been identified in MLII α/β patients, but no large deletions have been reported. Here we present the first case of a large homozygous intragenic *GNPTAB* gene deletion (c.3435-386_3602 + 343del897) encompassing exon 19, identified in a ML II α/β patient. Long-range PCR and sequencing methodologies were used to refine the characterization of this rearrangement, leading to the identification of a 21 bp repetitive motif in introns 18 and 19. Further analysis revealed that both the 5′ and 3′ breakpoints were located within highly homologous *Alu* elements (*Alu-Sz* in intron 18 and *Alu-Sq2*, in intron 19), suggesting that this deletion has probably resulted from *Alu–Alu* unequal homologous recombination. RT-PCR methods were used to further evaluate the consequences of the alteration for the processing of the mutant pre mRNA *GNPTAB,* revealing the production of three abnormal transcripts: one without exon 19 (p.Lys1146_Trp1201del); another with an additional loss of exon 20 (p.Arg1145-Serfs*2), and a third in which exon 19 was substituted by a pseudoexon inclusion consisting of a 62 bp fragment from intron 18 (p.Arg1145Serfs*16). Interestingly, this 62 bp fragment corresponds to the *Alu-Sz* element integrated in intron 18.

This represents the first description of a large deletion identified in the *GNPTAB* gene and contributes to enrich the knowledge on the molecular mechanisms underlying causative mutations in ML II.

Communicated by: Verena Peters.

Competing interests: None declared.

L. da Silva Santos · S. Alves (✉)
Research and Development Unit, Department of Genetics, CGMJM, INSA, Porto Portugal
e-mail: sandra.alves@insa.min-saude.pt; alvessandra@hotmail.com

M.F. Coutinho
Research and Development Unit, Department of Genetics, CGMJM, INSA/IPATIMUP, Porto Portugal

L. Lacerda
Biochemical Genetics Unit, Department of Genetics, CGMJM, INSA, Porto Portugal

S. Quental
Quental, IPATIMUP, Porto Portugal

F. Wibrand · A.M. Lund
Department of Clinical Genetics, Rigshospitalet, Copenhagen Denmark

K.B. Johansen
Department of Pediatrics, Kolding Hospital, Kolding Denmark

M.J. Prata
IPATIMUP/ Department of Biology, Faculty of Sciences, Porto Portugal

Introduction

Mucolipidosis II α/β (ML II; MIM# 252500), ML III α/β (MIM# 252600) and ML III γ (MIM# 252605) are rare lysosomal storage diseases which share similar clinical features, including skeletal abnormalities. Among them, mucolipidosis type II α/β is clinically the most severe. The skeletal system is profoundly affected in all patients, who usually present with abnormalities in both cartilage and

bone. Linear growth decelerates during the first year of life and almost ceases during the second year. Death usually occurs between 5 and 8 years of age (Kornfeld and Sly 2001), although early death in utero as well as in the first years of life is not uncommon.

The diseases are caused by deficiencies in activity of the uridine diphosphate (UDP)-N-acetylglucosamine:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase (GlcNAc-phosphotransferase), which may be reduced or absent. GlcNAc-phosphotransferase is a multimeric enzyme encoded by two genes: *GNPTAB* (α/β subunits) and *GNPTG* (γ subunit). Mutations in the *GNPTAB* gene are associated with ML II α/β and ML III α/β, whereas those in the *GNPTG* gene cause ML III γ (Raas-Rothschild et al. 2000; Tiede et al. 2004).

The *GNPTAB* gene (GenBank accession number: NG_021243.1), which codes for the α and β subunits, contains 21 exons and spans 85 kb on chromosome 12q23.3. It encodes a protein of 1,256 amino acids with a predicted molecular mass of 144 kDa (α/β precursor). Proteolytic processing of the α/β precursor generates the individual α and β subunits (Tiede et al. 2004).

To date, more than 100 different ML II α/β and ML III α/β disease-causing mutations have been identified, including 30 missense, 20 nonsense, 32 small deletions, 25 small insertions, 2 small indels, and 14 splice site mutations (Human Gene Mutation Database website [HGMD, http://www.hgmd.org] and references therein). Large genomic rearrangements were rarely reported (1.6%) including an *Alu* retrotransposition in *GNPTAB* exon 5 (Tappino et al. 2008) and the exon 2 duplication (Otomo et al. 2009), resulting from recombination between homologous regions of introns 1 and 2.

In this work, we report on a large deletion found in the *GNPTAB* gene. After performing a fine molecular screening in an ML II α/β patient, we found a homozygous genomic lesion in *GNPTAB* leading to the entire loss of exon 19 and some of its surrounding intronic regions (c.3435-386_3602 + 343del897). Analysis of the deletion breakpoints indicated that the mutation may have been generated by an unequal homologous recombination process between highly similar *Alu* elements located in two out of three regions within the gene where such kind of sequences were demonstrated to be present. Subsequent cDNA analysis revealed the presence of three abnormal transcripts created by this gross deletion at mRNA level: one without exon 19 (p.Lys1146_Trp1201del); another with an additional loss of exon 20 (p.Arg1145Serfs*2), and a third in which exon 19 was substituted by a pseudoexon inclusion consisting of a 62 bp fragment from intron 18 (p.Arg1145Serfs*16).

To the best of our knowledge, this represents the first description of a large deletion identified in the *GNPTAB* gene. Furthermore, the work adds on the knowledge of the molecular mechanisms underlying causative mutations in ML II, contributing to improved genetic counselling and prenatal diagnosis in this devastating disease.

Material and Methods

Case Description

This boy was born at 38 weeks of gestation after a caesarean section. His parents were healthy, consanguineous Palestinians who also had a healthy 3-year-old son.

Birth weight was 2,085 g, length 42 cm and head circumference 32 cm. Apgar scores were 9/1 and 10/5, but shortly after birth he developed signs of respiratory insufficiency with oxygen saturation at 70% and was treated with continuous positive airway pressure (CPAP) and oxygen.

At ultrasonography during pregnancy, the foetus was noted to have short extremities, and at birth the newborn presented with short, slightly bowed arms, and legs and a large, protruding abdomen as well as mandibular hypoplasia and gingival hyperplasia. Bilateral inguinal hernia developed at one month of age.

Initial investigations showed low phosphate at 0.8 mmol/L (reference values: 1, 4–2, 3 mmol/L), increased alkaline phosphatase at 1,518 U/L (reference values: 55–515 U/L) increasing to 1,967 U/L at age 18 days. Parathyroid hormone (PTH) was very high at 91, 2 pmol/L (reference values: 1, 1–6, 9 pmol/L). Calcium was initially slightly low, but normalised quickly without treatment. 25-hydroxy-vitamin D2 + D3 was below detection and investigations in his mother also disclosed low 25-hydroxy-vitamin D2 + D3 < 10. Thus, the initial suspicion was that the child had a severe congenital rachitis or hyperparathyroidism and X-rays were thought compatible with that (Fig. 1).

Measurement of lysosomal enzymes revealed grossly elevated hexosaminidase A and B activities in plasma, consistent with mucolipidosis type II.

The child continued to have respiratory problems and died in respiratory insufficiency at the age of 2 months.

Mutation Screening of *GNPTAB* from a gDNA Sample

Genomic DNA was isolated from fibroblasts according to standard procedures. PCR amplifications of the 21 exons and adjacent intronic regions of *GNPTAB* gene were performed using specific primers (available on request). PCR amplification was carried out using approximately 40 ng of genomic DNA, 1 × PCR reaction mix ImmoMix Red (Bioline, London, UK) and 0.25 μM of each primer.

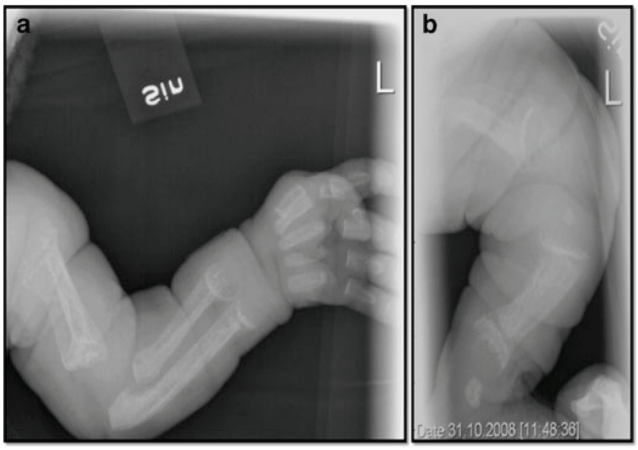


Fig. 1 *X-rays of left arm (**a**) and left leg (**b**) at birth.* In both X-rays, shortening of bones, osteopenia, and a bone-in-bone appearance are seen. The metaphyses are expanded with cupping and fraying

Thereafter, samples were heated to 95°C for 7 min, followed by 35 cycles of denaturation (30 s at 94°C), annealing (30 s at 58°C to exons 2, 3, 4, 6, 11, 12, 14, 15, 16, 17, 18, 20, and 21; 30 s at 59°C to exon 13; 30 s at 60°C to exons 1, 5, 7, 8–10, and 19) and extension (45 s at 72°C). The final extension was completed by 7 min at 72°C.

Fragments were purified with ExoSap-IT (GE Healthcare, United Kingdom) according to the manufacturer's instructions and sequenced using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on an ABI PRISM 3130xl Genetic Analyser (Applied Biosystems, Foster City, CA, USA). Results were analysed with the sequence analysis software FINCH TV (Geospiza, Seattle, WA, USA) version 1.3.1.

Long-Range PCR and Characterization of Deletion Breakpoints

Genomic *GNPTAB* was submitted to long-range PCR using the following primers: primer forward (5′–TGGATGTTGAGTCCACTACGG–3′), designed to anneal across intron 17 and primer reverse (5′–TCATTTTCTAAAACATTCAGATGC–3′), that anneals at intron 20 (see Table 1). PCR mixture contained 500 ng of genomic DNA, 1X PCR 3.75 U Expand Long Taq-System (Roche), reaction buffer 3, 350 μM of each dNTP and 3 μM of each primer. Thereafter, the samples were heated to 94°C for 2 min, followed by 30 cycles of denaturation (10 s at 94°C), annealing (30 s at 58°C) and extension (6 min at 68°C). In the last 20 cycles, the extension time had an increment of 20 s in each one. The final extension was completed by 7 min at 68°C.

The long-range PCR fragment was purified with ExoSap-IT (GE Healthcare, United Kingdom) and sequenced. In order to detect the deletion breakpoints, several internal primers were designed to sequence introns 18 and 19 (see Table 1).

cDNA Analysis

GNPTAB cDNA analysis was performed with specific primers (see Table 1) according to the previously reported conditions (Encarnação et al. 2009).

Bioinformatic Analysis

Considering that several genomic rearrangements are frequently caused by recombination events promoted by repetitive elements present in the human genome, the DNA sequences flanking the 5′ and 3′ deletion breakpoints were screened for the presence of such elements with the RepeatMasker software (http://repeatmasker.org).

The scores for the splice site junctions were obtained with the Maxent programme (http://genes.mit.edu/burgelab/maxent/Xmaxentscanscoreseqacc.html) (Yeo and Burge 2004). Ideal Maxent splice site scores are: 11.81 (5′ss) and 13.59 (3′ss).

Mutation Nomenclature

As reference for the *GNPTAB* gene, we used the sequence available at http://www.ncbi.nlm.nih.gov/nuccore/NG_021243.1 (GenBank accession number: NG_021243.1).

As reference for the *GNPTAB* mRNA, we used the sequence available at http://www.ncbi.nlm.nih.gov/nuccore/NM_024312.4 (GenBank accession number: NM_024312.4).

Table 1 Complete list of primers used to perform this work

gDNA amplification of exon 19 and its intronic boundaries					
	Localization	Sequence	Annealing temperature	PCR product size	Reference
Forward	Intron 18	5′-CCCATAGCTAAAAGGCCATCTACC-3′	60°C	436 bp	Encarnação et al. (2009)
Reverse	Intron 19	5′-GTATACACTCACCCACACACATGC-3′			Encarnação et al. (2009)
Forward 2	Intron 18	5′-TTTGGAATCCACATCCTTGTT-3′	60°C	391 bp	This study
Reverse 2	Intron 19	5′-TGGGCAACAAGAACAAAACTC-3′			This study
gDNA long range PCR					
	Localization	Sequence	Annealing temperature	PCR product size	Reference
Forward	Intron 17	5′–TGGATGTTGAGTCCACTACGG–3′	58°C	8,392 bp	This study
Reverse	Intron 20	5′–TCATTTTCTAAAACATTCAGATGC–3′			This study
Optimised sequencing of the long range amplicon					
	Localization	Sequence	Annealing temperature	PCR product size	Reference
Forward 1	Intron 18	5′–TGGGCTCAAGCAATCCTCC–3′	Not applicable		This study
Forward 2	Intron 18	5′–CCAGGTTTAGAGAAAGATGAA–3′			This study
Forward 3	Intron 18	5′–TTTTGGCGAATCTACTTCAAAAG–3′	Not applicable		This study
Forward 4	Intron 18	5′–CCACATCTGGCTAATTTTCATA–3′			This study
Reverse 1	Intron 19	5′–AGGCGAGAGGGTATGAACTG–3′	Not applicable		This study
Reverse 2	Intron 19	5′–TAGCAGCAATATTCATCCTAAT–3′			This study
Reverse 3	Intron 19	5′–CAATATTGCTAGTGATTATTCCCACA–3′	Not applicable		This study
Reverse 4	Intron 19	5′–CAGTTGAATAAATGAAGTTCC–3′			This study
cDNA amplification of a fragment ranging from exons 17–21					
	Localization	Sequence	Annealing temperature	PCR product size	Reference
Forward	Exon 17	5′-CCAGTAACTGACAAAATCCA-3′	55°C	734 bp	Encarnação et al. (2009)
Reverse	Exon 21	5′-ACAGGTCCATGAGCAAATTC-3′			Encarnação et al. (2009)

Results and Discussion

The molecular examination of a ML II patient was initially assessed through routine procedures, which involved amplification and direct sequencing of *GNPTAB* exons and flanking intronic regions (Encarnação et al. 2009). However, the failure to amplify the fragment correspondent to exon 19 and its intronic boundaries from the patient's DNA in face of normal amplification from control DNA, led us to suspect that a large deletion could underlie the molecular defect. To exclude that the nonamplification of the fragment containing exon 19 was due to any polymorphic variation in the primer's annealing region, a second pair of primers was used that again resulted in unsuccessful amplification of exon 19 (see primer sequences in Table 1). No mutations were found in the remaining regions sequenced.

To investigate whether a large deletion was indeed involved, a long-range PCR was designed to amplify *GNPTAB* genomic DNA encompassing exons 18–20 and respective intronic flanking sequences. This approach resulted in amplification of different-sized fragments in the patient and controls. The electrophoretic pattern in 2% agarose gel of the long-PCR products clearly indicated that the amplified fragment in the patient was shorter (approximately 7,000 bp) than the expected size (8,392 bp) of the control (data not shown). To detect the deletion junction region and disclose the exact size of the gene loss, four internal primers annealing in intron 18 and another four annealing in intron 19 were used for optimal sequencing of the amplicon obtained through long-range PCR (Table 1).

The sequencing results allowed the identification of the deletion breakpoints and the determination of the deletion

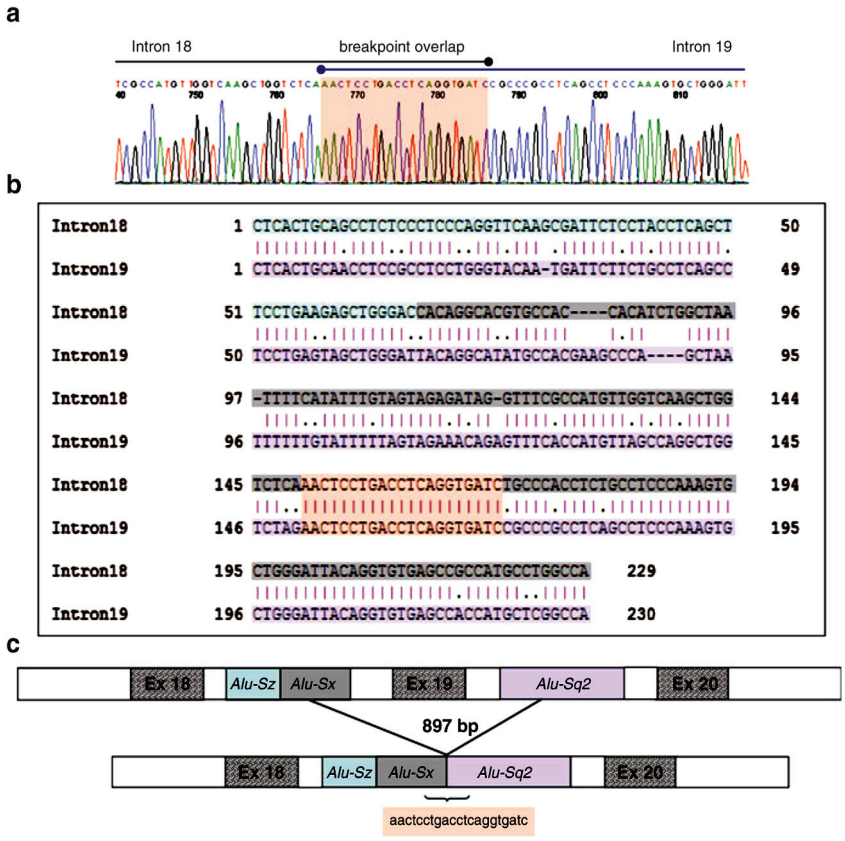


Fig. 2 *Characterization of the deletion breakpoints.* (**a**) Electropherogram evidencing the junction fragment resulting from the large deletion in the affected individual. Partial representation of introns 18 and 19 with the 21 bp breakpoint overlap highlighted in *pink*. (**b**) Nucleotide sequence alignment of the genomic sequence from introns 18 and 19 of the *GNPTAB* gene. Sequence gaps are indicated by *dashed lines*, *short vertical lines* indicate matched bases between both introns. Sequence highlighted in *aquamarine* corresponds to the *Alu-Sz* element of intron 18; sequence highlighted in *grey* corresponds to the *Alu-Sx* element of the same intron and, finally, the one highlighted in *light violet*, indicates the *Alu-Sq2* element of intron 19. (**c**) Schematic representation of the deletion breakpoints and their flanking *Alu* elements. C1 represents a normal gDNA fragment and C2 schematizes the mutated gDNA with a deletion of 897 bp including the last 386 nucleotides of intron 18, exon 19 (168 bp), and the first 343 bp of intron 19

extension which was 897 bp and included the last 386 nucleotides of intron 18, exon 19 (168 bp), and the first 343 bp of intron 19 (c.3435-386_3602 + 343del897). Furthermore, we verified that a 21 bp perfect sequence overlap ("aactcctgacctcaggtgatc"; Fig. 2a, b) occurred at the two deletion breakpoints: the same sequence stretch was present in intron 18 and intron 19. Additional sequence analysis revealed that, in both introns, the "aactcctgacctcaggtgatc" motif was flanked by highly homologous sequences (Fig. 2b). The sequence alignment of the genomic regions (introns 18 and 19), surrounding the 21 bp overlapping segment, showed high homology: in particular, the 149 nucleotides upstream and the 59 nucleotides downstream of the deletion breakpoints shared 76.5% and 89.8% homology, respectively. Since this high level of sequence homology suggested the presence of *Alu* elements, both full-length intron sequences were submitted to analysis by the Repeat Masker Web Server to search for interspersed repeat elements. The "in silico" analysis revealed that two next *Alu* sequences (*Alu-Sz* followed by *Alu-Sx*) were present in intron 18 and another (*Alu Sq2*) in intron 19. It was therefore evident that both the 5′ and 3′ breakpoints of the deletion were located within highly homologous Alu elements (Fig. 2c). Additionally, when the "in silico" analysis was extended to the entire gene, we observed that the 3′UTR region not only was highly homology to introns 18 and 19 but also showed an Alu element (*Alu-Sx* family) bearing the motif "aactcctgacctcaggtgatc" (data not shown).

Alu elements are short interspersed elements (SINEs), normally located within introns, and 3′ untranslated regions of genes, which are considered mutational "hotspots" for large gene rearrangements.

Thus, it is highly probable that the mutation here detected has arisen after an *Alu–Alu* unequal recombination event. During DNA replication, the primer and template strands might have transiently dissociated and then reassociated in a misaligned configuration mediated by the

"aactcctgacctcaggtgatc" sequences within introns 18 and 19, generating the large deletion (Fig. 2c).

The mechanism involving unequal homologous recombination between highly similar *Alu* elements has been described as contributing to several genetic diseases (Deininger and Batzer 1999), such as ornithine transcarbamylase deficiency (OTCD) (Quental et al. 2009), Von Hippel-Lindau disease (VHL) (Franke et al. 2009), Kindler syndrome (Has et al. 2006) and maple syrup urine disease (Quental et al. 2008).

This distribution of *Alu* sequences within *GNPTAB,* indicates that homologous recombination events mediated by similar *Alu* elements might involve introns 18 and 19 and the 3′UTR, increasing the risk of generating gross defective alleles at the gene, some of which may remain unidentified. We can hypothesise that at least some ML II patients from whom no RNA samples were available and, in whom only one mutant allele has previously been found (Encarnação et al. 2009; Tappino et al. 2009; Otomo et al. 2009; Zarghooni and Dittakavi 2009), might be compound heterozygous for large deletions caused by *Alu*-mediated mechanisms. The confirmation of that status would be of obvious implications in genetic counselling and prenatal diagnosis, but in which regards to the partially characterised patients from our series (Encarnação et al. 2009), RNA samples were available for testing and no gross deletions were found in either of them. Another important point that deserves attention is the fact that patients who have already been classified as homozygous for other *GNPTAB* mutations may be misclassified if that mutation happens to occur in a region where one of the alleles carries a gross deletion. Such may be of great importance for a correct establishment of genotype–phenotype correlations.

Finally, to address the impact of the novel deletion upon RNA processing, RT_PCR of the *GNPTAB* cDNA region encompassing exons 17–21 (734 bp) revealed the presence of three abnormal transcripts (Fig. 3a): one without exons 19 and 20 (Fig. 3b); transcript 1; another with the loss of exon 19 alone (Fig. 3b; transcript 2), and a third in which exon 19 was substituted by a pseudoexon inclusion consisting of a 62 bp fragment from intron 18 (Fig. 3b; transcript 3). At a protein level, all three transcripts give rise to aberrant truncated proteins: transcript 1 (Fig. 3) leads to the substitution of an arginine by a serine at position 1145, with a consequent frameshift that results in the introduction of a novel amino acid at position 1146 and of a PTC at position 1147 (p.Arg1145Serfs*2); transcript 2 leads to the formation of a mutant protein with less 56 amino acids between positions 1146 and 1201 (p.Lys1146_Trp1201del); transcript 3 (Fig. 3) also leads to

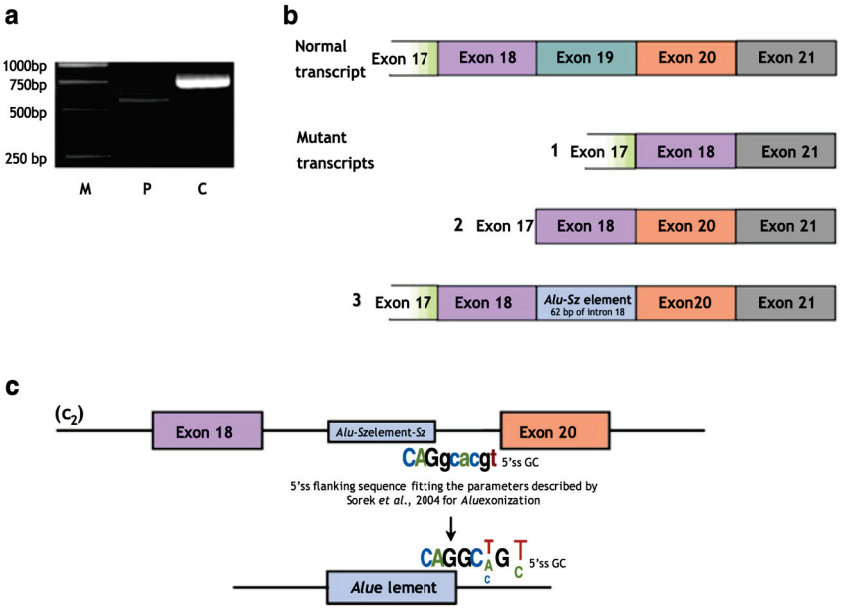


Fig. 3 *Agarose gel showing the GNPTAB transcripts observed in the affected individual together with schematic views of each transcript's constitution.* (**a**) Agarose gel showing the *GNPTAB* transcripts observed in the affected individual. *P* patient, *C* control, *M* marker (GeneRuller™ DNA Ladder 100–10,000 bp, Fermentas). (**b**) Schematic representation of the three abnormal transcripts. The first transcript presented loss of exons 19 and 20 (**b1**). The second transcript presented the loss of exon 19 alone (**b2**) and the third one instead of exon 19 had an inclusion of 62 bp of the *Alu-Sz* element of intron 18 (**b3**). (**c**) Schematic representation of the potential 5′ss of the pseudoexon showing its correspondence to one of the sequences referred to by Sorek et al. (2004) as necessary for the creation of alternatively spliced *Alu* exons

a frameshift after the substitution Arg1145Ser with introduction of 15 novel amino acids followed by a PTC (p.Arg1145Serfs*16).

The finding that this 62 bp fragment corresponded to the majority of the *Alu-Sz* element integrated in intron 18 (Fig. 3c), is quite curious since there are several recent studies in which pseudoexons were characterised as *Alu* or LINE elements (Knebelmann et al. 1995; Meili et al. 2009; Mitchell et al. 1991; Pérez et al. 2009; Vervoort et al. 1998), as well as with recent estimates indicating that up to 5% of human alternative exons could be derived from *Alu* sequences (Sorek et al. 2004; Tazi et al. 2009; Wood et al. 2007). With this in mind, we looked for *Alu* sequences within *GNPTAB* exons, but no repetitive sequences were found in any of them, indicating that exonization was not an evolutionary mechanism at the *GNPTAB* gene. According to the current state of knowledge, exonization of *Alu* sequences depends almost solely on the sequence composition of the potential 3′ and 5′ splice sites. *Alus* containing the 3′ss sequences described by Lev-Maor et al. (2003) or the 5′ss sequences reported by Sorek et al. (2004), are highly scored to become alternatively spliced exons. On these grounds, after analysing within intron 18 the flanking sequences of both potential splice sites of the *Alu-Sz* element, we verified that although the 3′ss sequence was different from that described by Lev-Maor et al. (2003), the residues of intron 18 that are flanking the potential 5′ss corresponded to one of the sequences referred to by Sorek et al. (2004) as necessary for the creation of alternatively spliced *Alu* exons (Fig. 3c). Consistently, the MaxEnt prediction of scores for the splice site junctions pointed towards an activation of a downstream cryptic splice site in intron 18 (5′ss = 9.70) to produce the transcript with the pseudoexon inclusion (Fig. 3b; transcript 3). Taking this data into account, it seems that the novel deletion c.3435-386_3602 + 343del897, here reported activates a cryptical splice site on the *Alu-Sz* element in such a way that it is recognised as an exon.

Altogether, these results clearly demonstrate that this deletion causes a serious aberration in the splicing pattern of the mutant pre mRNA *GNPTAB*, with three mutant transcripts being formed and at the protein level, which would have gone unnoticed if only simple gDNA analyses had been carried out.

Conclusion

In this study, we present the first report of a large deletion in the *GNPTAB* gene (c.3435-386_3602 + 343del897). The mutation was found in a severely affected ML II patient with prenatal onset, severe skeletal changes and hyperparathyroidism and causes homozygous loss of exon 19 and some of its surrounding intronic regions. Given the presence of two highly homologous *Alu* elements in exons 18 and 19, the deletion was probably caused by a process of unequal homologous recombination. Processing of the mutant pre mRNA *GNPTAB* causes production of three abnormal transcripts: one without exon 19 (p.Lys1146_Trp1201del); another with an additional loss of exon 20 (p.Arg1145-Serfs*2) and a third in which exon 19 was substituted by a pseudoexon inclusion consisting of a 62 bp fragment from intron 18 (p.Arg1145Serfs*16). It is important to take into account the possible existence of such large deletions when performing molecular screenings of the *GNPTAB* gene, being aware that when present in heterozygosity, they might easily escape detection through DNA genomic analyses. Without direct cDNA examination, this kind of mutation can remain unidentified.

From the results obtained, we recommend that ML II α/β and ML III α/β cases with only one mutant allele detected at direct sequencing should be further examined for the presence of large heterozygous deletions to strengthen the accuracy of genetic counselling and prenatal diagnosis in the families at risk.

Acknowledgments This work was supported by FCT - project PIC/IC/83252/2007 (http://alfa.fct.mctes.pt/). Coutinho MF and Quental S received grants from the FCT (SFRH/BD/48103/2008; SFRH/BPD/64025/2009).

One Sentence Take-Home Message

First large genomic deletion in the *GNPTAB* gene.

References

Deininger PL, Batzer MA (1999) Alu repeats and human disease. Mol Genet Metab 67:183–93

Encarnação M, Lacerda L, Costa R et al (2009) Molecular analysis of the *GNPTAB* and *GNPTG* genes in 13 patients with mucolipidosis type II or type III - identification of eight novel mutations. Clin Genet 76:76–84

Franke G, Bausch B, Hoffmann MM et al (2009) Alu-Alu recombination underlies the vast majority of large VHL germline deletions: molecular characterization and genotype-phenotype correlations in VHL patients. Hum Mutat 30:776–86

Has C, Wessagowit V, Pascucci M et al (2006) Molecular basis of Kindler syndrome in Italy: novel and recurrent Alu/Alu recombination, splice site, nonsense, and frameshift mutations in the KIND1 gene. J Invest Dermatol 126:1776–83

Knebelmann B, Forestier L, Drouot L et al (1995) Splice-mediated insertion of an Alu sequence in the COL4A3 mRNA causing autosomal recessive Alport syndrome. Hum Mol Genet 4:675–9

Kornfeld S, Sly WS (2001) I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localisation. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease, 8th edn. McGraw-Hill, New York, pp 3421–42

Lev-Maor G, Sorek R, Shomron N et al (2003) The birth of an alternatively spliced exon: 3′ splice-site selection in Alu exons. Science 300:1288–91

Meili D, Kralovicova J, Zagalak J et al (2009) Disease-causing mutations improving the branch site and polypyrimidine tract: pseudoexon activation of LINE-2 and antisense Alu lacking the poly(T)-tail. Hum Mutat 30:823–31

Mitchell GA, Labuda D, Fontaine G et al (1991) Splice-mediated insertion of an Alu sequence inactivates ornithine delta-amino-transferase: a role for Alu elements in human mutation. Proc Natl Acad Sci 88:815–9

Otomo T, Muramatsu T, Yorifuji T et al (2009) Mucolipidosis II and III alpha/beta: mutation analysis of 40 Japanese patients showed genotype-phenotype correlation. J Hum Genet 54:145–51

Pérez B, Rincón A, Jorge-Finnigan A et al (2009) Pseudoexon exclusion by antisense therapy in methylmalonic aciduria (MMAuria). Hum Mutat 30:1676–82

Quental R, Azevedo L, Rubio V et al (2009) Molecular mechanisms underlying large genomic deletions in ornithine transcarbamylase (OTC) gene. Clin Genet 75:457–64

Quental S, Martins E, Vilarinho L et al (2008) Maple syrup urine disease due to a new large deletion at BCKDHA caused by non-homologous recombination. J Inherit Metab Dis 20(4):463–472

Raas-Rothschild A, Cormier-Daire V, Bao M et al (2000) Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC). J Clin Invest 105:673–681

Sorek R, Lev-Maor G, Reznik M et al (2004) Minimal conditions for exonization of intronic sequences: 5′ splice site formation in alu exons. Mol Cell 14:221–31

Tappino B, Chuzhanova NA, Regis S et al (2009) Molecular characterization of 22 novel UDP-N-acetylglucosamine-1-phosphate transferase alpha- and beta-subunit (GNPTAB) gene mutations causing mucolipidosis types IIalpha/beta and IIIalpha/beta in 46 patients. Hum Mutat 30:E956–E973

Tappino B, Regis S, Corsolini F et al (2008) An *Alu* insertion in compound heterozygosity with a microduplication in GNPTAB gene underlies Mucolipidosis II. Mol Genet Metab 93:129–33

Tazi J, Bakkour N, Stamm S (2009) Alternative splicing and disease. Biochim Biophys Acta 1792:14–26

Tiede S, Cantz M, Raas-Rothschild A et al (2004) A novel mutation in UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) in two siblings with mucolipidosis type III alters a used glycosylation site. Hum Mutat 24:535–536

Vervoort R, Gitzelmann R, Lissens W et al (1998) A mutation (IVS8 + 0.6kbdelTC) creating a new donor splice site activates a cryptic exon in an Alu-element in intron 8 of the human beta-glucuronidase gene. Hum Genet 103:686–93

Wood M, Yin H, McClorey G (2007) Modulating the expression of disease genes with RNA-based therapy. PLoS Genet 3:e109

Yeo G, Burge CB (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 11:377–394

Zarghooni M, Dittakavi SS (2009) Molecular analysis of cell lines from patients with mucolipidosis II and mucolipidosis III. Am J Med Genet 149A:2753–61

RESEARCH LETTER

AMERICAN JOURNAL OF medical genetics PART A

Mucolipidosis Type II α/β With a Homozygous Missense Mutation in the *GNPTAB* Gene

Maria Francisca Coutinho,[1,2,3] Liliana da Silva Santos,[1] Katta Mohan Girisha,[4] Kapaettu Satyamoorthy,[5] Lúcia Lacerda,[6] Maria João Prata,[2,3] and Sandra Alves[1]*

[1]Research and Development Unit, Department of Genetics, CGMJM, INSA, Portugal

[2]IPATIMUP, Porto, Portugal

[3]Department of Biology, Faculty of Sciences, Porto, Portugal

[4]Genetics Clinic, Department of Pediatrics, Kasturba Medical College, Manipal University, Manipal, India

[5]Manipal Life Sciences Center, Manipal University, Manipal, India

[6]Biochemical Genetics Unit, Department of Genetics, CGMJM, INSA, Portugal

Received 2 November 2011; Accepted 29 December 2011

References to electronic databases: OMIM disorder/gene accession numbers: OMIM# 252500, MIM# 252600 and MIM# 252605; EC numbers: EC 2.7.8.17.

TO THE EDITOR:

The uridine diphosphate (UDP)-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase (termed GlcNAc-1-phosphotransferase; EC 2.7.8.17) is a Golgi-resident 540-kDa hexameric transmembrane enzyme composed by three subunits, $\alpha_2\beta_2\gamma_2$, which catalyzes the first step in the formation of the M6P recognition marker on lysosomal enzymes. This marker is essential for the proper recognition of newly synthesized lysosomal hydrolases by specific M6P receptors, required for lysosomal targeting. The receptor-enzyme complexes are selectively transported and, upon arrival in the endosomal/lysosomal compartment, dissociate due to the typical low pH. Lysosomal proteins are then delivered to the lysosomes, while the M6P receptors return to the Golgi apparatus for subsequent rounds of [Braulke and Bonifacino, 2009] transport. Impairments in the formation of the M6P recognition marker due to defective GlcNAc-1-phosphotransferase are the basis of two rare lysosomal storage disorders: Mucolipidosis II (ML II) and Mucolipidosis III (ML III). Depending on the affected gene, mucolipidosis may be classified as ML II α/β or ML III α/β if mutations are present in the *GNPTAB* gene, which codes for the α/β subunits, or ML III γ if mutations are present in the *GNPTG* gene which is responsible for coding the γ subunit [Raas-Rothschild et al., 2000; Tiede et al., 2004, 2005]. Mucolipidosis II α/β (ML II; OMIM# 252500], ML III α/β (OMIM# 252600), and ML III γ (OMIM# 252605) share similar manifestations, including skeletal abnormalities. Among them, mucolipidosis type II α/β is clinically the most severe. The skeletal system is severely affected in all patients, who usually present with abnormalities of cartilage and bone. Psychomotor retardation is obvious by 6 months. Linear growth decelerates during the first year of life and almost ceases during the second year. Death usually occurs between 5 and 8 years [Kornfeld and Sly, 2001], though early death in utero or in the first years of life is not uncommon.

How to Cite this Article:
Coutinho MF, da Silva Santos L, Girisha KM, Satyamoorthy K, Lacerda L, Prata MJ, Alves S. 2012. Mucolipidosis type II α/β with a homozygous missense mutation in the *GNPTAB* gene.
Am J Med Genet Part A 158A:1225–1228.

The *GNPTAB* gene, which codes for the α and β subunits, contains 21 exons and spans 85 kb on chromosome 12q23.3. It encodes a protein precursor of 1,256 amino acids with a predicted molecular mass of 144 kDa that, after proteolytic processing, generates the individual α and β subunits [Tiede et al., 2005].

Here we report on a novel genetic defect underlying mucolipidosis type II α/β identified in homozygozity in a patient of Indian origin. The boy was born to consanguineous parents (first cousins; third degree consanguinity) of Indian origin (Fig. 1A–G). The couple had had two earlier first trimester spontaneous abortions. Birth weight was 1.6 kg (<3rd centile). He was noted to have "coarsening" of face and gum hypertrophy by the first month.

Grant sponsor: FCT; Grant number: PIC/IC/83252/2007.
*Correspondence to:
Dr. Sandra Alves, Ph.D., Medical Genetics Center Jacinto Magalhães, INSA
Praça Pedro Nunes, 88, 4099-028 Porto, Portugal.
E-mail: sandra.alves@insa.min-saude.pt
Published online 11 April 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/ajmg.a.35295

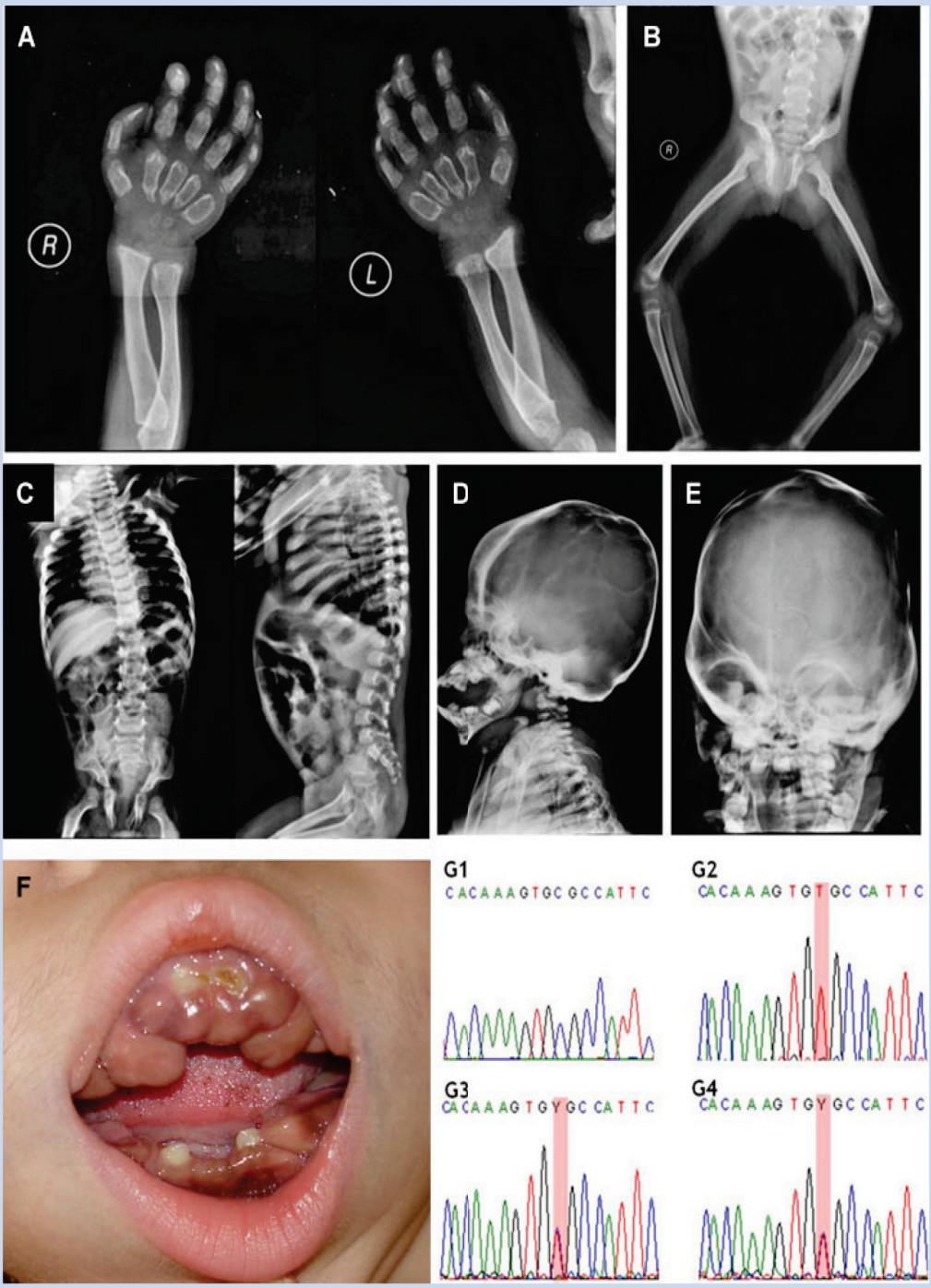


FIG. 1. Phenotypic and genotypic aspects of the patient. A–E: Radiographs taken at age 3 showing dysostosis multiplex: Thickened skull bones, oar shaped ribs, dysplasia of pelvis, anteroposterior shortening of vertebral bodies, diaphyseal widening of radius and ulna, and remodeling deformity of metacarpals and phalanges. There is delay in ossification of carpal bones. F: The patient showing severe gum hypertrophy. G: Electropherograms highlighting the affected residue of control individual (G1), patient (G2), and his parents (G3–mother; G4–father).

He had severe developmental delay. When first seen by us at age 5 years, he had partial head control and could roll over. He was not able to sit or stand. He did not have control over bowel and bladder movements. He was able to understand simple commands. His head circumference was 43 cm, length 71 cm, and weight 7 kg (all <3rd centile for age). He had "coarse" facies, depressed nasal bridge, ecpicanthic folds, severe gum hypertrophy (Fig. 1F), deformed chest, open mouth, broad wrists, contractures of fingers,

and three café au lait spots with the largest measuring 5 × 3 cm over his back. He had clear corneae and there was no organomegaly. Skeletal survey showed dysostosis multiplex (Fig. 1A–E). His enzyme levels in plasma were elevated (in nmol/hr/ml with reference values in parenthesis): Alpha fucosidase 5701 (90-5504), beta hexosaminidase A 994 (30-1504), total hexosaminidase 38890 (620-49904), beta glucuronidase 8671 (30–150), beta galactosidase 170 (4 nmol/hr/ml in healthy control). The same enzymes were within the normal ranges in leukocytes.

After genomic DNA isolation from blood according to standard procedures, we performed a molecular analysis of the *GNPTAB* gene through amplification of all of its 21 exons and the corresponding intronic flanking regions using specific primers (sequences available on request). PCR amplification was carried out using approximately 40 ng of genomic DNA, 1× PCR reaction mix ImmoMix Red (Bioline, London, UK) and 0.5 μM of each primer. Thereafter, samples were heated to 95°C for 7 min, followed by 35 cycles of denaturation (30 sec at 94°C), annealing (30 sec at 58°C to exons 2, 3, 4, 6, 11, 12, 14, 15, 16, 17, 18, 20, and 21; 30 sec at 59°C to exon 13; 30 sec at 60°C to exons 1, 5, 7, 8–10, and 19) and extension (45 sec at 72°C). The final extension was completed by 7 min at 72°C. Fragments were purified with ExoSap-IT (GE Healthcare, UK) according to the manufacturer's instructions and sequenced using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) on an ABI PRISM 3130×l Genetic Analyser (Applied Biosystems, Foster City, CA). Results were analyzed with the sequence analysis software FINCH TV (Geospiza, Seattle, WA) version 1.3.1.

We identified a homozygous mutation: c.2956T>C (Fig. 1F). This substitution leads to the replacement of an arginine by a cysteine at position 986 of the protein (R986C). In order to exclude the possibility of it being a polymorphic variation, we analyzed 50 individuals (100 chromosomes) none with the variation. Interspecific protein alignments were performed with the software ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Using several sequences published in Ensembl (http://www.ensembl.org/), we found that residue R^{986} was highly conserved throughout evolution. To further evaluate the potential effect of these missense mutations, two bioinformatic tools were used: PolyPhen program (Polymorphism Phenotyping; http://coot.embl.de/PolyPhen/) and SIFT (Sorting Intolerant From Tolerant; http://blocks.fhcrc.org/sift/SIFT.html). Both predicted that p.R986C is expected to be damaging to protein function. This homozygous missense mutation as the genetic basis of a phenotype as severe as the one observed is this patient, with such early onset of symptoms and marked skeletal abnormalities (Fig. 1A–E) is not common since, usually, frameshift and nonsense mutations are associated with such severe phenotypes while missense ones are usually related to milder phenotypes. The first reports on the genetic characterization of ML II patients even suggested that only frameshift and nonsense mutations would cause the severe form ML II α/β; missense mutations, on the other hand, would only cause the milder phenotype: ML III α/β [Otomo et al., 2009].

Nevertheless, this is not the first case of a missense mutation associated with a severe phenotype. Others have been reported, some of which were classified as "ML intermediate" [Cathey et al., 2010] and others as typical ML II cases [Tiede et al., 2006; Zarghooni and Dittakavi, 2009]. The first of those cases was reported by Tiede et al. in 2006 with typical ML II findings: "coarse" face; short-trunk dwarfism; flexed elbows, knees and fingers; thick skin and "pithecoid" stance. Latter, in a mutational screening of four ML patients, Zarghooni and Dittakavi [2009] reported two typical ML II cases presenting only missense mutations. Curiously, both cases had one allele with two compound mutations (c.614A>C + c.545T>A/c.2196G>T, p.Gln205Pro + p.Val182Asp/p.Lys732Asn; c.2783A>G + c.2864C>T/c.3160T>G, p.Lys928Arg + p.Ala955Val/p.Leu1054Val) and, even though the interaction between these two compound mutations and the subsequent effects on the protein was not clearly understood, they suggested that those mutations act additively to disrupt different functions, resulting in a more severe ML II phenotype. More recently, Cathey et al. [2010], published a study of 61 ML II and III patients; one patient was homozygous for a missense mutation (c.3053A>G; p. D1018G) with an atypical phenotype. This patient, as others classified as ML intermediate presented with manifestations of both forms of the disease (ML II and ML III]. In general, these authors classified as "ML intermediate" patients who tended to have physical findings similar to, but milder than ML II and a clinical course more reminiscent of ML III [Cathey et al., 2010].

Several attempts have been made to explain the peculiar phenotypes presented by these patients; however, further expression analysis is required to come to a better understanding of the effects of this mutation.

In order to understand the association of this particular mutation to the phenotype of this Indian patient, we checked if it was affecting any particularly important region of the molecule. We checked the glycosylation sites, calcium binding sites, metal binding sites and disulfide bonds but none of them involved R^{986}. Nevertheless, the affected residue is quite near the cleavage site so it is possible that this mutation alters the conformation of that region and blocks the cleavage of the α/β precursor between Lys^{928} and Asp^{929}. Recently, Marschner et al. [2011] demonstrated that 20 amino acids proximal to the cleavage site are required for the proper cleavage of the α/β precursor. R^{986} does not belong to that series of amino acids but it is possible that its change to a cysteine, which, contrary to arginine, is a hydrophobic amino acid capable of forming disulfide bonds, may alter the protein's conformation in such way that the Site-1 protease (S1P) that is responsible for the cleavage is no longer capable of binding to the protein.

Further studies are needed to get a better insight on this mutation's pathological mechanism. Unfortunately, no fibroblast cell line is available on this patient. It would be very interesting to access the expression levels of the α/β subunits, in order to determine if this mutation alters the normal *GNPTAB* mRNA levels. Another interesting aspect would be the analysis of expression of the γ subunit of the phosphotransferase in this patient, since it has been previously shown that mutations in the *GNPTAB* are associated with post-transcriptional modifications in the γ subunit. Expression analysis of ML II patients carrying a homozygous missense mutation located in the α/β subunit (c.3707A>T, p.K1236M) suggested posttranslational proteolytic cleavage of GNPTG, with Western Blot analysis of GNPTG in patients' fibroblasts revealing an altered electrophoretic mobility. In addition, no

mutations or changes in mRNA expression levels were detected in *GNPTG* [Tiede et al., 2006].

Nevertheless, it is of crucial importance to report the existence of patients carrying only missense mutations and presenting severe phenotypes since these types of mutations do not usually correlate with such phenotypes, splicing and frameshift mutations generally associated with the most severe forms of lysosomal storage disorders. The knowledge of the particular missense variations which are associated with such phenotypes is highly relevant for genetic counseling as the identification of novel mutant alleles alone allows no prediction of the phenotype. It is important, though, to report any atypical missense mutations underlying the severe phenotype of this disease that may be devastating.

ACKNOWLEDGMENTS

This work was supported by FCT—project PIC/IC/83252/2007 (http://alfa.fct.mctes.pt/). MF Coutinho is a grantee from the FCT (SFRH/BD/48103/2008).

REFERENCES

Braulke T, Bonifacino JS. 2009. Sorting of lysosomal proteins. Biochim Biophys Acta. 1793:605–614.

Cathey SS, Leroy JG, Wood T, Eaves K, Simensen RJ, Kudo M, Stevenson RE, Friez MJ. 2010. Phenotype and genotype in mucolipidoses II and III alpha/beta: A study of 61 probands. J Med Genet 47:38–48.

Kornfeld S, Sly WS. 2001. I-cell disease and pseudo-Hurler polydystrophy: Disorders of lysosomal enzyme phosphorylation and localisation. In: Scriver CR, Beaudet AL, Sly WS, Valle D, Childs B, Kinzler KW, Vogelstein B, assoc. eds. The metabolic and molecular bases of inherited disease, 8th edition. New York: McGrawHill. 3421–3442.

Marschner K, Kollmann K, Schweizer M, Braulke T, Pohl S. 2011. A key enzyme in the biogenesis of lysosomes is a protease that regulates cholesterol metabolism. Science 333:87–90.

Otomo T, Muramatsu T, Yorifuji T, Okuyama T, Nakabayashi H, Fukao T, Ohura T, Yoshino M, Tanaka A, Okamoto N, Inui K, Ozono K, Sakai N. 2009. Mucolipidosis II and III alpha/beta: Mutation analysis of 40 Japanese patients showed genotype-phenotype correlation. J Hum Genet 54:145–151.

Raas-Rothschild A, Cormier-Daire V, Bao M, Genin E, Salomon R, Brewer K, Zeigler M, Mandel H, Toth S, Roe B, Munnich A, Canfield WM. 2000. Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC). J Clin Invest 105:673–681.

Tiede S, Cantz M, Raas-Rothschild A, Muschol N, Bürger F, Ullrich K, Braulke T. 2004. A novel mutation in UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) in two siblings with mucolipidosis type III alters a used glycosylation site. Hum Mutat 24: 535.

Tiede S, Storch S, Lübke T, Henrissat B, Bargal R, Raas-Rothschild A, Braulke T. 2005. Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase. Nat Med 11:1109–1112.

Tiede S, Cantz M, Spranger J, Braulke T. 2006. Missense mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTA) in a patient with mucolipidosis II induces changes in the size and cellular distribution of GNPTG. Hum Mutat 27:830–831.

Zarghooni M, Dittakavi SS. 2009. Molecular analysis of cell lines from patients with mucolipidosis II and mucolipidosis III. Am J Med Genet Part A 149A:2753–2761.

Paper 4

Neonatal Mucolipidosis type II overlapping Pacman dysplasia phenotype in an Indian family

Coutinho MF[1,2,3*], Aggarwal S[4,5*], Dalal AB[5], Jamal Mohamed Nurul Jain S[5], Prata MJ[2,3], Alves S[1]

[1] *Research and Development Unit, Department of Human Genetics, INSA, Porto, Portugal*

[2] *IPATIMUP, Porto, Portugal*

[3] *Department of Biology, Faculty of Sciences, Porto, Portugal*

[4] *Department of Medical Genetics, Nizam's Institute of Medical Sciences, Hyderabad, India*

[5] *Diagnostics Division, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, India*

[+] These authors equally contributed to the work

(under preparation)

Abstract

Pacman dysplasia is a rare perinatal lethal skeletal dysplasia which has radiological findings overlapping with Mucolipidosis type II alpha/beta (MLII alpha/beta). There have been reports in literature proposing it as a severe prenatal phenotype of MLII alpha/beta. However, no conclusive evidence is available and some geneticists believe it to be a distinct entity. Here we report the identification of mutations in *GNPTAB* gene, which causes MLII alpha/beta, in an Indian family with Pacman dysplasia phenotype.

Key Words: Pacman dysplasia, Mucolipidosis type II alpha/beta (MLII alpha/beta), *GNPTAB*

Introduction

Pacman dysplasia (PD) is a perinatal lethal skeletal dysplasia characterised clinically by presence of short, bowed limbs. Only a few cases of this condition have been reported in literature [Shohat *et al.*, 1993; Wilcox *et al.*, 1998; Miller *et al.*, 2003; Saul *et al.*, 2005]. The disorder is characterised radiologically by presence of periosteal cloaking, epiphyseal stippling and sagittal clefting of vertebra; and histological criteria, once chondro-osseus morphology reveals marrow fibrosis and large osteoclasts containing howship lacunae on their endosteal surface. Mucolipidosis II alpha/beta (ML II alpha/beta) is a lysosomal storage disorder caused by mutations in *GNPTAB* gene which leads to abnormal targeting of lysosomal enzymes to lysosomes [Tiede *et al.*, 2005]. This can be detected biochemically on basis of marked elevation of plasma lysosomal enzymes [Kornfeld and Sly, 2001]. Another milder phenotype caused by mutations in *GNPTAB* gene is Mucolipidosis III alpha/beta (ML III alpha/beta), also known as Pseudo-Hurler dystrophy, which presents later in childhood like mucopolysacharidoses with predominant skeletal findings [Sanklecha *et al.*, 1994; Paik *et al.*, 2005; Kudo *et al.*, 2006].

Clinically, neonates with ML II alpha/beta have coarse facies, joint contractures, deformed bones, developmental delay and inconsistently hepatomegaly. The radiological findings show dysostosis multiplex along with periosteal cloaking and epiphyseal stippling [Kornfeld and Sly, 2001; David-Vizcarra *et al.*, 2010]. Recently, the question was raised of whether PD could not be actually a fetal/neonatal form of ML II alpha/beta presenting with secondary severe manifestations [Unger *et al.*, 2005]. The hypothesis gained further sustainability when Saul *et al.* [2005] confirmed the diagnosis of MLII alpha/beta on basis of enzyme assay in the sib of a fetus previously reported to have Pacman dysplasia. However, Wilcox *et al.* [2005] disputed this by pointing out that these two conditions differed in their chondro-osseus morphology. They also documented normal lysosomal enzyme levels in cultured fibroblast from a patient with Pacman dysplasia [Wilcox *et al.*, 1998; Wilcox *et al.*, 2005]. Therefore, the molecular basis of Pacman dysplasia still remains unconfirmed.

Here we report on a neonate with PD phenotype deceased soon after birth, whose parents attended prenatal counselling for next pregnancy. Screening of the *GNPTAB* gene in the non-consanguineous couple allowed to detect that both members were carriers of pathogenic mutations, a finding that retrospectively gives grounds to a diagnosis of MLII alpha/beta in the proband.

Case Report

A non-consanguineous couple from North-eastern part of India attended for prenatal counselling in view of neonatal death in their previous male offspring. Review of the

antenatal records showed that the child was detected to have short bones in-utero since 31 weeks of gestation. The ultrasound report at 37 weeks gestation showed that all long bones corresponded to 23-24 weeks. There was no evidence of hydrops. The child was born at term with a weight of 2.1kgs. He developed respiratory distress requiring assisted respiration and subsequently died on the fourth day. There were no physician records of the neonatal period. Photographs were brought by the parents, which showed presence of somewhat coarse facies, short limbs, large and dysplastic ears, flat nasal bridge and proptotic eyes. The left leg appeared to be deformed. The diagnostic clue was ascertained on the radiographs which showed short and broad long bones with wide diaphyses and massive periosteal cloaking, dense metaphyses and epiphyseal stippling (Figure 1). These findings led us to suspect a diagnosis of Pacman dysplasia/Mucolipidosis II alpha/beta. The plasma enzyme assays of the couple were non-contributary. No biological tissue from proband was available. DNA was extracted from peripheral blood of parents and the coding region of *GNPTAB* gene sequenced, together with its intronic flanking regions.

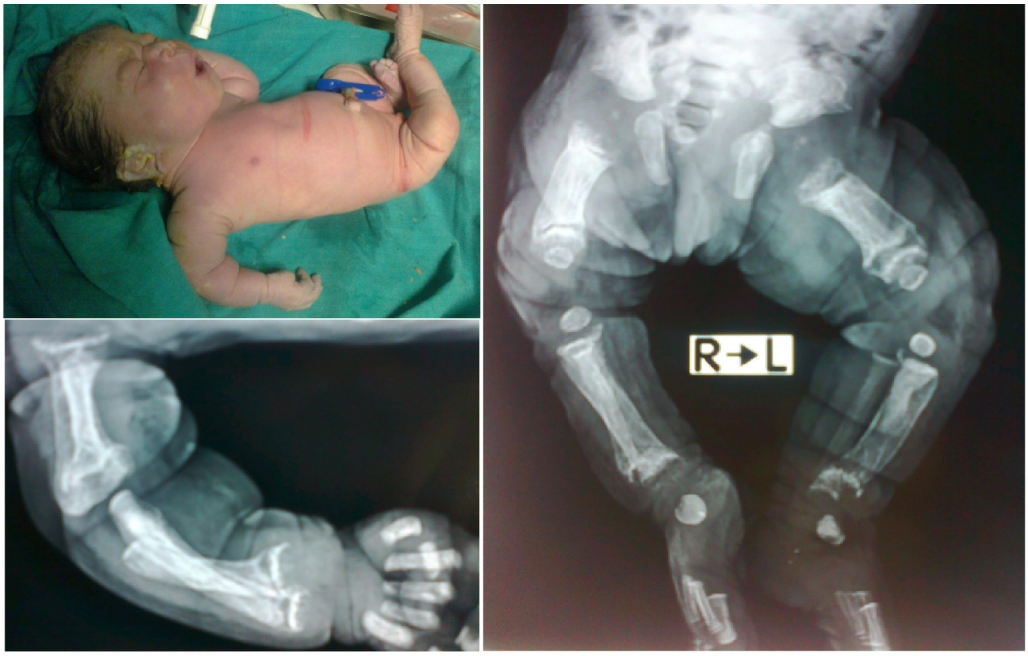


Figure 1: Phenotypic presentation of the patient.

Photographs showing a somewhat coarse facies, short limbs, large and dysplastic ears, flat nasal bridge and proptotic eyes and deformed left leg. X-rays showing short and broad long bones with wide diaphyses and massive periosteal cloaking, dense metaphyses and epiphyseal stippling.

Results

After *GNPTAB* sequencing, the consultant was found to be heterozygous for c.3503_3504delTC (p.L1168Q*fs*X5) mutation in exon 19, which is a previously known

pathogenic mutation. The husband was heterozygous for a novel frameshift mutation c.1701delC (p.F566L*fs*X5) in exon 13.

Discussion

Pacman dysplasia was first described by Shohat *et al*. [1993] in a preterm female fetus of 28 weeks gestation. They recognised it as a novel skeletal dysplasia with characteristic radiographic findings of epihyseal stippling and periosteal cloaking. The chondro-osseus morphology consisted of marrow fibrosis and expansion with numerous osteoclasts containing Howship lacunae on their endosteal surface, findings that were considered indicative of increased bone resorption. Subsequently, Wilcox *et al*. [1998] reported as PD both a male and female sibs terminated at 20 and 16 weeks pregnancy, respectively, on basis of ultrasound findings of short limb skeletal dysplasia. On post-mortem radiological and histopathological examination they also fitted the criteria for PD. Latter, Miller *et al*. [2003] reported a 20 week male fetus with Pacman dysplasia. This fetus had a subsequent female sibling who was clinically, radiographically and biochemically diagnosed to have ML II alpha/beta [Saul *et al*., 2005]. Consequently, Saul *et al*. recovered then the proposal, short before formulated by Unger *et al*. [2005], that Pacman dysplasia could be a prenatally manifesting severe phenotype of ML II alpha/beta. However, Wilcox *et al*. [2005] refuted this hypothesis by noticing that, even though the fetus originally reported by Miller *et al*. [2003] and re-assessed by Saul *et al*. [2005], could indeed be a case of ML II alpha/beta whose chondro-osseus preparation was inadequate to detect lysosomal inclusions (no material was available for electron microscopy or cultured fibroblasts for enzyme assays), the same was not possible for other cases. In fact, detailed revision of the first case reported by Shohat *et al*. [1993], did not demonstrate the presence of abnormal inclusions on the available formalin fixed material whereas atypical inclusions were quite evident in multiple cell types from a *bone fide* ML II alpha/beta case identically prepared. Furthermore, the enzyme assays from cultured fibroblasts of the sibs reported by Wilcox *et al*. were normal, ruling out the diagnosis of ML II alpha/beta [Wilcox *et al*., 2005]. Hence, they opined that Pacman dysplasia and ML II alpha/beta are distinct entities which can be distinguished by chondro-osseus morphology and presence of lysosomal storage, despite similarity in clinical and radiological features. ML II alpha/beta is a lysosomal storage disorder arising from mistrafficking of lysosomal enzymes from their synthesis site (ER) to the lysosomes, which is due to the absence of mannose-6-phosphate (M6P) recognition marker on the enzymes. This defect results from the deficiency of the enzyme UDP-N-acetylglucosamine-1-phosphotransferase (GlcNAc-phosphotransferase), a heterohexameric complex whose α and β subunits are encoded by

the *GNPTAB* gene [reviewed in Braulke and Bonifacino, 2009; Coutinho *et al.*, 2012]. Mutations in this gene lead to a phenotypic spectrum ranging from a neonatal/infantile presentation of ML II alpha/beta to a mild condition with childhood onset called ML III alpha/beta [Tiede *et al.*, 2005; Paik *et al.*, 2005; Kudo *et al.*, 2006]. MLII alpha/beta is a progressive phenotype characterised by onset at birth, coarse facies, gum hypertrophy, developmental delay, short trunk short stature, joint contractures and radiologic findings of dysostosis multiplex [Kornfeld and Sly, 2001]. Patients with a very severe perinatal presentation can have bone deformities in form of bowed long bones, kyphosis, clubfeet etc., which overlap with PD. The x-ray findings in these cases are of periosteal cloaking and epiphyseal stippling, which are also shared by patients with PD. Due to these clinical and radiological similarities, it is difficult to distinguish between PD and the prenatal phenotype of severe ML II alpha/beta, admitting that they actually are distinct entities, which for the moment, still lacks molecular support given the unknowledge of any genetic alterations underlying PD.

In the case here reported, the fetus had clinical and radiographic findings suggestive of PD. Clinically, presentation with short bones prenatally at 31 weeks gestation has never been reported for a confirmed ML II alpha/beta case. Radiologically, the degree of limb shortening at term was severe, whereas the described cases of neonatal MLII alpha/beta have low-normal limb lengths. In spite of both clinical and x-ray findings fitted better into a phenotype associated to PD, such findings are known to be quite similar in patients reported as PD and ML II alpha/beta presenting in neonatal period. Regarding the radiographic features, Miller *et al.* [2003] thought possible to differentiate between the two conditions through the presence of sagittal clefting and shorter metacarpals in cases with Pacman dysplasia, although very probably being gestation dependent radiological findings. Hence, clinically and radiologically our case represents the PD phenotype. Yet, the identification that the parents were both carriers of pathogenic mutations in the *GNPTAB* gene, strongly evinces that the proband was affected with ML II alpha/beta presenting with a PD phenotype.

Genotype-phenotype correlations involving the *GNPTAB* gene suggest that frameshift or nonsense mutations lead to ML II alpha/beta presentation, while missense mutations usually cause the mildest ML III alpha/beta phenotype [Tiede *et al.*, 2006]. In this work, the mutations detected in the non-consanguineous parents were both frameshift. The mother carried the mutation c.3503_3504delTC, which is the most common disease-causing variant in *GNPTAB*, already reported in several populations. When present in homozygousity or in combination with other frameshift or nonsense mutation, c.3503_3504delTC leads to a severe MLII alpha/beta phenotype [Bargal *et al.*, 2006; Kudo *et al.*, 2006; Plante *et al.*, 2008; Tappino *et al.*, 2009; Encarnação *et al.*, 2009]. The paternal mutation, c.1701delC, was a small deletion, detected here for the first time that also

creates a frameshift with consequent premature chain termination potentially leading to a truncated GlcNAc-phosphotransferase α/β-precursor lacking 686 amino acids. It is known that, in the Golgi apparatus, the α/β-precursor is cleaved between Lys^{928} and Asp^{929}, giving rise to the mature α- and β-subunits which are both essential to the phosphotransferase's enzymatic activity. The two identified mutant alleles are predicted to alter the normal open reading frame of the *GNPTAB* mRNA by introducing premature termination codons, which may either trigger the nonsense-mediated mRNA decay mechanism or give rise to the formation of highly truncated non-functional proteins. Hence, if simultaneously present in an individual, this should imply complete absence of GlcNAc-phosphotransferase activity, which most likely was the case of the proband under analysis, therefore explaining his severe phenotype.

The message we sought to share from this study is that at least some cases of Pacman dysplasia phenotype, assessed in the absence of histological examination, are in fact severe prenatal forms of ML II alpha/beta. Nevertheless, we find it easy to admit that there may be casual heterogeneity in PD patients like in the siblings reported by Wilcox *et al*. [1998]. Possible sources may include, for instance, mutations in the gene coding for the site-1-protease (S1P) -the protease that cleaves the α/β-precursor of the GlcNAc-phosphotransferase, as recently suggested by Marschner *et al*. [2011]. Apart from *GNPTAB* and S1P deficiencies, other genetic defects must also be envisioned, since failures on the activity of the phosphotransferase caused by mutations in any of these genes necessarily give rise to mistrafficking of lysosomal enzymes, which has already been shown not to occur in some patients with PD who have normal levels of lysosomal enzymes in plasma.

Deficiencies of osteoclast inhibitor factor(s) have also been suggested to underlie PD phenotypes [Roodman, 1996; Wilcox *et al*., 1998], but until now they have never been documented.

In conclusion, so far, only mutations in the *GNPTAB* gene allowed to explain two of the few reported cases presenting Pacman-like phenotypes. The question may be raised of whether those two patients were accurately diagnosed. It is possible that they escaped correct recognition as ML II alpha/beta due to the lack or unsatisfactory histological examination of their pathological material, which according Saul *et al*. might be critical to make a definitive diagnosis of PD. Still, when histological evaluation is not available, which is certainly a common situation, and the diagnostic must merely rely in clinical and/or radiological findings, ML II alpha/beta must be always considered as a possible cause of Pacman disease presentation. Eventually, only the discovery of the molecular basis of PD will enable the establishment of additional and more reliable features to distinguish between PD and ML II alpha/beta, further illuminating the reason of being, or not, two pathological entities.

Acknowledgments

This work was supported by FCT - project PIC/IC/83252/2007 (http://alfa.fct.mctes.pt/). Coutinho MF is a grantee from the FCT (SFRH/BD/48103/2008).

References

Bargal R, Zeigler M, Abu-Libdeh B, *et al*. When Mucolipidosis III meets Mucolipidosis II: GNPTA gene mutations in 24 patients (2006) *Mol Genet Metab*. 88:359-363.

Braulke T, Bonifacino JS. Sorting of lysosomal proteins (2009) *Biochim Biophys Acta*. 1793:605-14.

Coutinho MF, Prata MJ, Alves S. Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction (2012) *Mol Genet Metab*. 105:542-50.

David-Vizcarra G, Briody J, Ault J, *et al*. The natural history and osteodystrophy of mucolipidosis types II and III (2010) *J Paediatr Child Health*. 46:316-22.

Encarnação M, Lacerda L, Costa R, *et al*. Molecular analysis of the GNPTAB and GNPTG genes in 13 patients with mucolipidosis type II or type III - identification of eight novel mutations (2009) *Clin Genet*. 76:76-84.

Kornfeld S, Sly WS. 2001. I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localisation. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. Childs B, KinzlerKW, Vogelstein B, assoc. eds. The metabolic and molecular bases of inherited disease, 8th edition. New York: McGrawHill. 3421–3442.

Kudo M, Brem MS, Canfield WM. Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha / beta -subunits precursor gene (2006) *Am J Hum Genet*. 78:451-463.

Marschner K, Kollmann K, Schweizer M, *et al*. A key enzyme in the biogenesis of lysosomes is a protease that regulates cholesterol metabolism (2011) *Science*. 333:87-90.

Miller SF, Proud VK, Werner AL, Field FM, Wilcox WF, Lachman RS, Rimoin DL. Pacman dysplasia: a lethal skeletal dysplasia with variable radiographic features (2003) *Pediatr Radiol*. 33:256-60.

Paik KH, Song SM, Ki CS, *et al*. Identification of mutations in the GNPTA (MGC4170) gene coding for GlcNAc-phosphotransferase alpha/beta subunits in Korean patients with mucolipidosis type II or type IIIA (2005) *Hum Mutat*. 26, 308-14.

Plante M, Claveau S, Lepage P, *et al*. Mucolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTAB) in a French Canadian founder population (2008) *Clin Genet*. 73:236-244.

Roodman GC. Advances in bone biology: The osteoclast (1996) *Endocr Rev*. 17:308-332.

Sanklecha MU, Kher A, Naik GG, *et al*. Storage disorders presenting like mucopolysaccharidosis (1994) *Indian J Pediatr*. 61:367-72.

Saul RA, Proud V, Taylor HA, *et al*. Prenatal mucolipidosis type II (I-cell disease) can present as Pacman dysplasia (2005) *Am J Med Genet A*. 135:328-32.

Shohat M, Rimoin DL, Gruber HE, *et al*. New epiphyseal stippling syndrome with osteoclastic hyperplasia (1993) Am J Med Genet. 45:558-61.

Tappino B, Chuzhanova NA, Regis S, *et al*. Molecular characterization of 22 novel UDP-N-acetylglucosamine-1-phosphate transferase alpha- and beta-subunit (GNPTAB) gene mutations causing mucolipidosis types IIalpha/beta and IIIalpha/beta in 46 patients (2009) *Hum Mutat*. 30:E956-973.

Tiede S, Storch S, Lübke T, *et al*. Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase (2005) *Nat Med*. 11:1109-12.

Tiede S, Cantz M, Spranger J, *et al*. Missense mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTA) in a patient with mucolipidosis II induces changes in the size and cellular distribution of GNPTG (2006) *Hum Mutat*. 27:830-1.

Unger S, Paul DA, Nino MC, *et al*. Mucolipidosis II presenting as severe neonatal hyperparathyroidism (2005) *Eur J Pediatr*. 164:236-43.

Wilcox WR, Lucas BC, Loebel B, *et al*. Pacman dysplasia: report of two affected sibs (1998) *Am J Med Genet*. 77:272-6.

Wilcox WR, Wenger DA, Lachman RS, *et al*. Distinguishing Pacman dysplasia from mucolipidosis II: comment on Saul et al. [2005] (2005) *Am J Med Genet A*. 135:333.

2. Biochemical studies

Paper 5: Functional analysis of *GNPTAB* mutations: clues for the establishment of genotype-phenotype correlations

Paper 5

Functional analysis of *GNPTAB* mutations: clues for the establishment of genotype-phenotype correlations

(under preparation)

Coutinho MF[1,2,3], De Pace R[4], Koch-Nolte F[5], Haag F[5], Prata MJ[2,3], Braulke T[4], Alves S[1] and Pohl S[4]

1 Research and Development Unit, Department of Genetics, CGMJM, INSA, Portugal
2 IPATIMUP, Porto, Portugal
3 Department of Biology, Faculty of Sciences, Porto, Portugal
4 Section Biochemistry, Children´s Hospital, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany
5 Institute of Immunology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

Abstract

Mucolipidoses type II and III (ML II and III) are genetic disorders respectively caused by completely absent or reduced activity of the enzyme GlcNAc-1-phosphotransferase, which participates in the synthesis of mannose 6-phosphate recognition marker. This heterohexameric enzyme ($\alpha_2\beta_2\gamma_2$) is a product of two distinct genes: *GNPTAB* and *GNPTG*. Mutations in *GNPTAB* result in ML II alpha/beta and ML III alpha/beta, while mutations in *GNPTG* are only associated with ML III gamma. It has been proposed that splicing and frameshift mutations are associated with more severe ML II alpha/beta phenotype while missense mutations cause the mildest ML III alpha/beta disease. Recent reports of ML II alpha/beta patients carrying homozygous missense mutations are defying this assumption. Here we have explored the expression levels and subcellular location of a panel of GlcNAc-1-phosphotransferase mutants and correlated the observed results with the severity of the phenotype presented by patients harboring them. 6 *GNPTAB* mutations were expressed in a wildtype α/β-subunit miniconstruct: the missense mutations reported in severely affected patients (p.W81L and p.R986C), 2 severe deletions (c.1581delC and c.3503_3504delTC) and 2 missense mutations reported in mildly affected individuals (p.S399F and p.K1236M). The overexpressed mutants were analysed by Western blot and immunofluorescence microscopy. In general, the combined analysis of the mutants' expression levels and subcellular location not only explains but also allows for a prediction of the severity of the phenotype: severe deletions are retained in the ER, in the non-cleaved inactive α/β-precursor form; mild missense mutations are correctly located in the Golgi apparatus but present reduced levels of the mature α- and β-subunits and, finally, severe missense mutations were either retained in the ER on the non-cleaved and inactive form (p.W81L) or not expressed at all (p.R986C).

Key Words: GlcNAc-1-phosphotransferase, ML II and III α/β, genotype-phenotype correlation.

Introduction

Mucolipidosis type II (MLII; MIM# 252500) and type III (MLIII; MIM# 252600) are rare autosomal recessive disorders of lysosomal hydrolases trafficking which are, respectively, caused by completely absent or reduced activity of the enzyme GlcNAc-1-phosphotransferase. This enzyme catalyzes the initial step of a two-step reaction that leads to the formation of the mannose 6-phosphate (M6P) recognition marker [Kornfeld and Sly, 2001). The mannose 6-phosphate marker is added exclusively to the N-linked oligosaccharides of lysosomal soluble hydrolases, as they pass through the *cis*-Golgi network and is responsible for their selective recognition by specific transmembrane receptors, present in the *trans*-Golgi (TGN): the cation-independent M6P receptor (CI-MPR) and/or the cation-dependent M6P receptor (CD-MPR). The sequential action of these proteins, which is triggered upon arrival in the Golgi apparatus of soluble lysosomal enzymes allowing their subsequent transport to the endosomal/lysosomal compartment, constitutes the M6P-dependent targeting pathway [Storch and Braulke, 2005]. Any failure in this process may stop an entire repertoire of degradative enzymes from reaching the lysosomes and ensuring the normal cellular processes. This is the reason why mutations in GlcNAc-1-phosphotransferase may have a tremendous deleterious effect.

GlcNAc-1-phosphotransferase is a heterohexameric enzyme composed of three subunits: $\alpha_2\beta_2\gamma_2$ [Bao *et al.*, 1996]. This enzyme is a product of two distinct genes: *GNPTAB* and *GNPTG* [Kornfeld and Sly, 2001]. The *GNPTAB* gene codes for the enzymatically inactive α/β-precursor [Tiede *et al.*, 2005a] and the *GNPTG* gene codes for the γ-subunit [Raas-Rothschild *et al.*, 2000]. The α/β-precursor contains the active site of the enzyme and consists of a 1256 amino acid pro-peptide in a hairpin orientation that spans the membrane twice in such way that both its N and C termini are located in the cytosol [Kudo *et al.*, 2005; Tiede *et al.*, 2005a]. This highly glycosylated precursor is latter cleaved into the individual α- and β-subunits by the site-1-protease (S1P), thus becoming catalytically active [Marschner *et al.*, 2010]. The γ-subunit represents a soluble glycoprotein of 305 amino acids containing two *in vivo N*-glycosylation sites at positions 88 and 115. It forms disulfide-linked homodimers and localizes mainly in the *cis*-Golgi apparatus. The function of the γ-subunit remains elusive but evidence points that it may facilitate substrate recognition and assembly of entire complex [Raas-Rothschild *et al.*, 2000; Pohl *et al.*, 2010; Encarnação *et al.*, 2011].

Mutations in *GNPTAB* result in ML II alpha/beta and ML III alpha/beta [Tiede *et al.*, 2005; Paik *et al.*, 2005; Kudo *et al.*, 2006], while mutations in *GNPTG* are only associated with ML III gamma [Raas-Rothschild *et al.*, 2000; Raas-Rothschild *et al.*, 2004; Tiede *et al.*, 2004]. ML II alpha/beta represents the most severe form of the disease with a fatal outcome usually between 5 and 8 years of age. Defects appear in early infancy or even prenatally, with patients typically presenting dwarfism, coarse facies, stiffness of the joints, several

skeletal abnormalities and severe mental retardation. ML III alpha/beta and ML III gamma represent attenuated forms of the disease, with indistinguishable clinical symptoms. These forms of disease are characterized by a later onset, at the age of 2 to 4 years, and a more slowly progressive course, permitting patient surviving into eighth decade, and mental retardation or learning disabilities affecting only 50 % of the patients [Dierks *et al.*, 2009]. The variability in clinical symptoms observed in mucolipidoses is due to differences in residual activities of the GlcNAc-1-phosphotransferase, which result in a wide spectrum of severity ranging from prenatally lethal ML II alpha/beta to mild adult onset forms of ML III alpha/beta and ML III gamma [Dierks *et al.*, 2009].

To date more than 100 different *GNPTAB* mutations have been described, causing either ML II alpha/beta or ML III alpha/beta [HGMD and references therein; Stenson *et al.*, 2009]. Early reports suggested that, while splicing and frameshift mutations were usually associated with more severe phenotypes (ML II alpha/beta), missense mutations were mainly associated with the milder ML III alpha/beta phenotype. Nevertheless, recent reports on ML II alpha/beta patients harboring *GNPTAB* missense mutations in homozygous state challenge this initial assumption. Here we present a biochemical analysis of several disease-causing *GNPTAB* mutations previously reported in patients with well described phenotypes in order to get a first insight on the factors that account for the pathogenical effect of different mutations. *GNPTAB* mutations were inserted into a *GNPTAB* miniconstruct, transfected into HEK and HeLa cell lines and analysed through Western blot (WB) and Immunofluorescence (IF). In general, the combined analysis of the mutants' expression levels and subcellular location perfectly explains the phenotype: mutants harboring the severe deletions are retained in the endoplasmic reticulum (ER), presenting only the non-cleaved α/β- and inactive precursor form; mild missense mutations are correctly located in the Golgi apparatus but present reduced levels of the mature α- and β-subunits, when compared to the wild-type and, finally, severe missense mutations were either retained in the ER on the non-cleaved and inactive precursor form or not expressed at all.

Material and Methods

Reagents

Penicillin/streptomycin (P/S), bovine serum albumin (BSA), 4´,6´-diamidino-2-phenylindol (DAPI), cycloheximide, protease inhibitor cocktail, and other common laboratory reagents were obtained from Sigma-Aldrich. Fetal calf serum (FCS) was from PAA Laboratories. Dulbecco´s modified Eagle´s medium (DMEM) and GlutaMAX™ were from Invitrogen. Transfection reagent JetPEI® was purchased from Peqlab. Phusion® polymerase, GeneJET Plasmid Miniprep Kit, GeneJET PCR Purification Kit, dNTPs and prestained protein ladder Pageruler™ were from Thermo Scientific. Oligonucleotides used for sequencing

and mutagenesis were synthesized by MWG Biotech. Peptide *N*-glycosidase F was from Roche Applied Sciences. Media for cultivating *Escherichia coli*, Mowiol® and Roti®quant Protein Assay were from Roth.

Antibodies

An antiserum against the human, native α/β-precursor protein of the GlcNAc-1-phosphotransferase was produced by cDNA immunization. The expression mini constructs wildtype and non-cleavable mutant R925A [Marschner *et al.*, 2010] were mixed and conjugated to 1-µm gold particles. Rats were immunized four times by ballistic DNA immunization. A serum sample was obtained 10 days after the last immunization [Koch-Nolte *et al.*, 2005]. The complete serum was used for immunofluorescence microscopy stainings at a 1:250 dilution.

Polyclonal rabbit antibodies against the human α- and β-subunit of the GlcNAc-1-phosphotransferase were described previously [Pohl *et al.*, 2010; Kollmann et al., 2012]. Monoclonal antibodies against protein-disulfide isomerase (PDI) and Golgi marker protein (GM130) were obtained from Biomol and BD Biosciences, respectively. The monoclonal antibody against β-tubulin was purchased from the NICHD Developmental Studies Hybridoma Bank (University of Iowa).

Secondary antibodies conjugated to horseradish peroxidase or Alexa Fluor® were purchased from Dianova and Invitrogen, respectively.

Generation of mutant *GNPTAB* cDNA constructs

The pathogenic *GNPTAB* mutations c.1196C>T, c.2956C>T, c.3707A>T, c.3503_3504delTC and c.242G>T (Figure 1) were inserted into the wildtype *GNPTAB* miniconstruct [Marschner et al., 2011] by site-directed mutagenesis using mutagenic primers and Phusion® polymerase. Mutagenic primers were designed individually according to the desired mutation using the web-based program PrimerX (www.bioinformatics.org/primerx). After PCR the methylated, nonmutated, parental DNAs were digested with D*pn* I and transformed into *E. coli*. The plasmid DNAs were isolated and commercially sequenced to confirm proper introduction of the mutations (SeqLab). The generation of the *GNPTAB* c.345insC mutant was described recently [Kollmann et al., 2012].

Cell culture and transiently transfection

Human embryonic kidney (HEK) and HeLa cells were maintained in DMEM supplemented with 10 % FCS, GlutaMAX™ and P/S at 37 °C and 5 % CO_2. Cultured cells were plated on 6-cm plates or on glass cover slips. The wildtype and each of the mutated cDNAs were transiently transfected independently in cells using JetPEI® reagent according to manufacturer instructions. 24 hours after transfection the cells were used for further analysis.

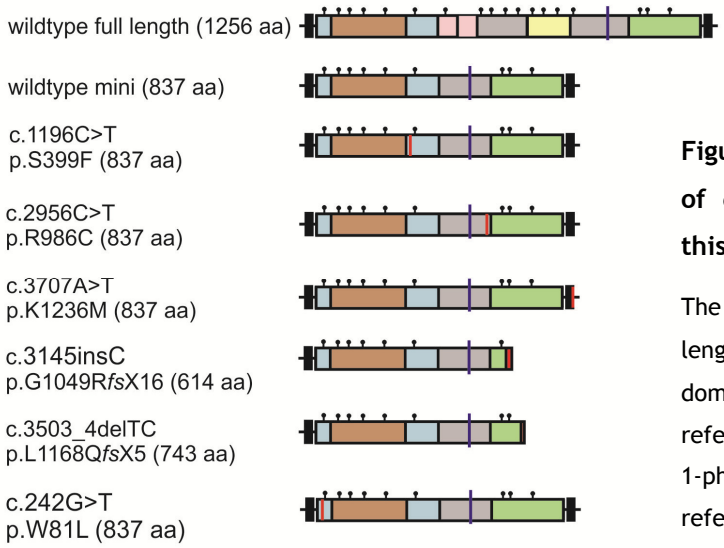


Figure 1: Schematic presentation of α/β-subunit constructs used in this study.

The first scheme shows the wildtype full length α/β-subunit precursor and its domain structure. The following schemes refer to wild type and mutated GlcNAc-1-phosphotransferase mini-constructs, as referenced in the image.

Western blot analysis

Transfected HEK cells were lysed for 30 minutes at 4 °C in lysis buffer containing 0.2 % Triton X-100 and protease inhibitors in phosphate-buffered saline (PBS). The protein content of cell extracts was measured by Bradford Protein Assay. Aliquots of cell extracts (75 µg protein) were incubated in the presence or absence of 1 U PNGaseF for 1 hour at 37 °C to remove *N*-glycans. Cell extracts were separated by SDS-PAGE (10 % polyacrylamide) and blotted onto nitrocellulose. After blocking with 5 % non-fat dry milk in PBS containg 0.1 % Tween-20, the membrane were incubated either with polyclonal antibodies anti α-subunit or anti β-subunit of the human GlcNAc-1-phosphotransferase. [Pohl *et al.*, 2010; Kollmann *et al.*, 2012]. β-Tubulin was used as loading control. The membrane was incubated with secondary HRP-conjugated goat anti-rabbit IgG antibody. The immunoreactive bands were visualized by enhanced chemoluminescence. Densitometric analyses were performed with the Image J software (http://rsbweb.nig.gov/ij) to quantify band intensities in Western blots.

Immunofluorescence microscopic analysis

Transfeted Hela cells were treated with 100 µg/ml cycloheximid for 40 min. Afterwards cells were fixed in 4 % paraformaldehyde for 20 minutes at room temperature and permeabilized with ice-cold 0.2 % Triton X-100 in PBS. Cells were incubated for 16 hours with primary antibodies and 1 hour with secondary antibodies conjugated to Alexa Fluor®546 and Alexa Fluor®488. After three washes the cells were embedded in Mowiol. Fluorescence was detected and images were obtained using a Leica DMIRE2 digital scanning confocal microscope and ADOBE PHOTOSHOP software, respectively.

Results

Expression analysis of α/β-subunit mutants of the GlcNAc-1-phosphotransferase

The human α/β-subunit precursor of the GlcNAc-1-phosphotransferase is a highly *N*-glycosylated type III membrane protein of 1265 amino acids. Cleavage by the site-1 protease gives origin to the enzymatically active α- and β-subunits. For expression analysis of pathogenic mutations in the α- and β-subunits we inserted the mutations in a mini construct of the α/β-subunit precursor, that lacks amino acids 431 to 819 (Figure 1) but allow its efficient expression [Marschner et al., 2011]. Protein extracts of non-transfected, wildtype and mutant α/β-subunit-overexpressing HEK cells were analysed by Western blotting using anti-serum against either the α-subunit (Figure 2 A) or the β-subunit (Figure 2 B) of the human GlcNAc-1-phosphotransferase.

In extracts of wildtype α/β-subunit-overexpressing cells, two immunoreactive forms of the protein were detected: the *N*-glycosylated ~120 kDa α/β-subunit precursor and the *N*-glycosylated cleaved/mature α-subunit of ~75 kDa and β-subunit of ~45 kDa (Figure 2 A and B, *lane 1 and 2*) that was not detectable in non-transfected cells (Figure 2 A and B, *lane 15 and 16*). The treatment with PNGase F, hydrolyzing all *N*-linked oligosaccharides, resulted in a shift of the electrophoretic mobility of the α/β-subunit precursor, cleaved α- and β-subunits to ~ 100 kDa, 55 kDa and 35 kDa, respectively (Figure 2 A and B, *lane 2*).

In protein extracts of cells overexpressing the missense mutations p.S399F the *N*-glycosylated and non-glycosylated α/β-subunit precursor were observed, together with small amounts of the cleaved α- and β-subunits were observed (Figure 2 A and B, *lane 3 and 4*). In extracts of cells overexpressing the p.R986C mutant, neither the α/β-subunit precursor nor the mature α-subunit could be detected. Same was registered when those cell extracts were incubated with PNGase F (Figure 2 A and B, *lane 5 and 6*). Protein extracts of cells overexpressing the missense mutation p.K1236M showed decreased levels of both α/β-subunit precursor α- and β-subunit (Figure 2 and B, *lane 7*) with the same molecular mass detected for cells overexpressing the wildtype α/β-subunit GlcNAc-1-phosphotransferase. After PNGase F treatment a shift to the non-glycosylated forms were observed (Figure 2 A and B, *lane 8)*. In protein extracts of cells overexpressing the frameshift mutation p.G1049R*fs*X16 or p.L1168*fs*X5 truncated *N*-glycosylated and non-glycosylated α/β-subunit precursor were observed. No cleaved α- and β-subunits were detectable (Figure 2 A and B, lane 9 to 12). In protein extracts of cells overexpressing the missense mutations p.W81L only the *N*-glycosylated and non-glycosylated α/β-subunit precursor form but no cleaved α- and β-subunits were observed (Figure 2 A and B, *lane 13 and 14*).

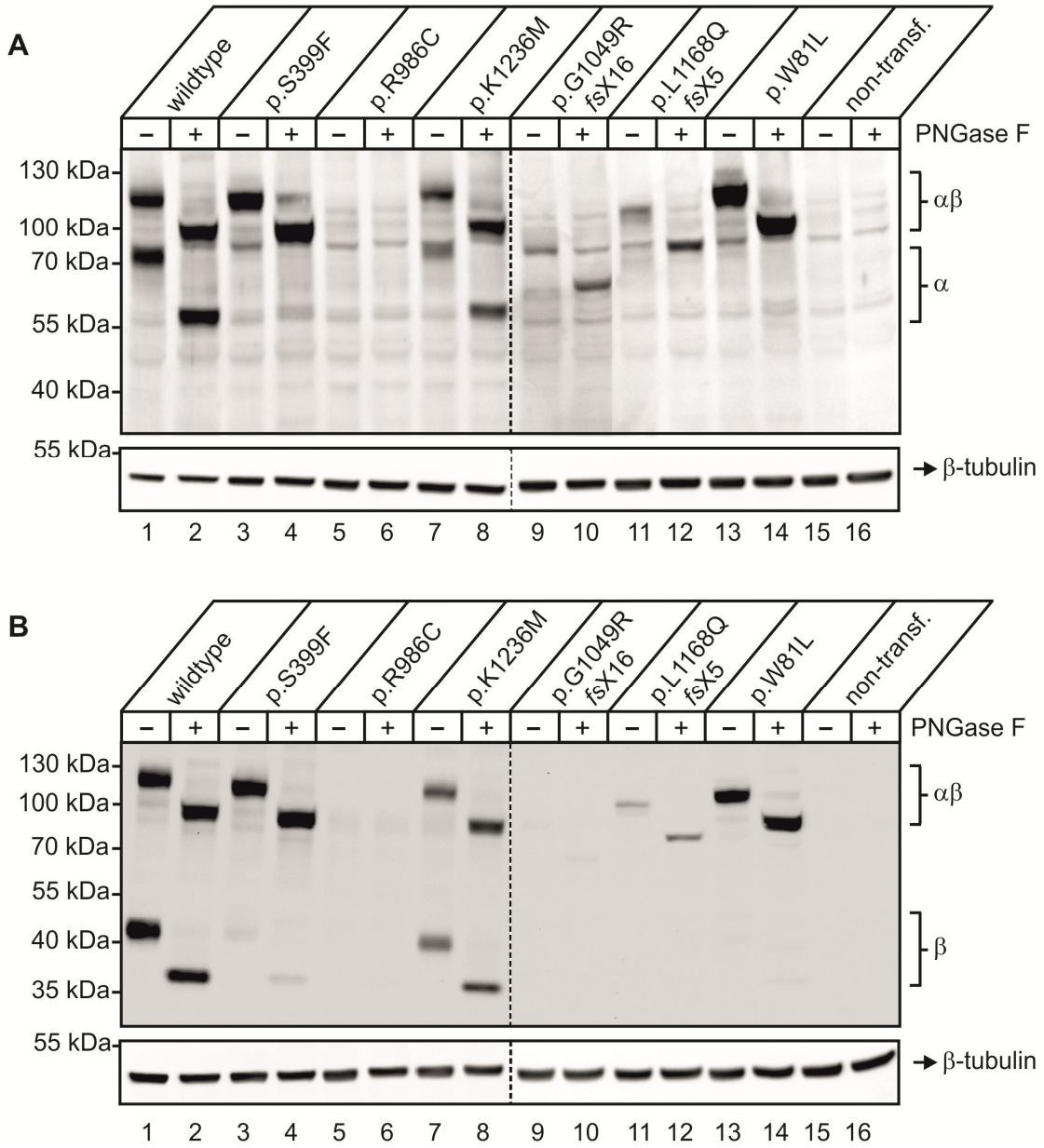


Figure 2: Western blot analysis of GlcNAc-1-phosphotransferase mutants in HEK cells.

HEK cells were transfected with cDNA encoding the *GNPTAB* wildtype miniconstruct and the mutants p.W81L, p.S399F, p.L1168Q*fs*X5, p.K1236M and p.G10489*fs*X16. The cells were incubated for 1 hour in the presence (+) or absence (-) of PNGase F. SDS-PAGE (10% polyacrylamide) was performed under reducing conditions followed by Western blot analysis against the human α-subunit (A) or β-subunit (B). β-Tubulin was used as a loading control. The positions of molecular mass marker proteins are indicated.

Intracellular localization of α/β-subunit mutants of the GlcNAc-1-phosphotransferase

The α/β-precursor of the GlcNAc-1-phosphotransferase is synthesized in the ER. The subsequent COPII-mediated vesicular transport from the ER to the Golgi apparatus requires a combinatorial sorting motif in the N- and C-terminal cytosolic domains [Franke et al., 2013]. In the Golgi apparatus the α/β-precursor is cleaved and activated by the site-

1 protease [Marschner et al., 2011].

To get an insight on the subcellular location of the different mutants here analysed, double immunofluorescence microscopy of HeLa cells overexpressing wildtype or mutant α/β-subunit of the GlcNAc-1-phosphotransferase was conducted with either marker proteins of the ER or the Golgi apparatus. For the staining of the GlcNAc-1-phosphotransferase a novel monoclonal antibody generated in rats was used. In cells transfected with the wildtype miniconstruct [Marschner et al., 2011], complete co-localization with the Golgi apparatus marker GM130 was observed, but not with the ER marker protein disulfide isomerase (PDI) (Figure 3). In cells overexpressing the mutant p.W81L, a co-localization, with the ER marker was observed. No mutant protein was present in the Golgi apparatus.

Cells overexpressing the p.S399F mutant, the protein co-localized mainly with the ER but also partially with the Golgi apparatus indicating that small amounts of the mutant reach the Golgi apparatus.

The mutant p.L1168Q*fs*X5 co-localized with the ER marker PDI but not with the Golgi apparatus marker protein GM130.

In cells overexpressing p.K1236M, on the other hand, a perfect co-localization with the Golgi apparatus was observed. No mutant protein was present in the ER.

Finally, for cells overexpressing the frameshift mutation p.G1049A*fs*X16, a complete co-localization with the ER marker PDI but not with the Golgi marker GM130 was observed.

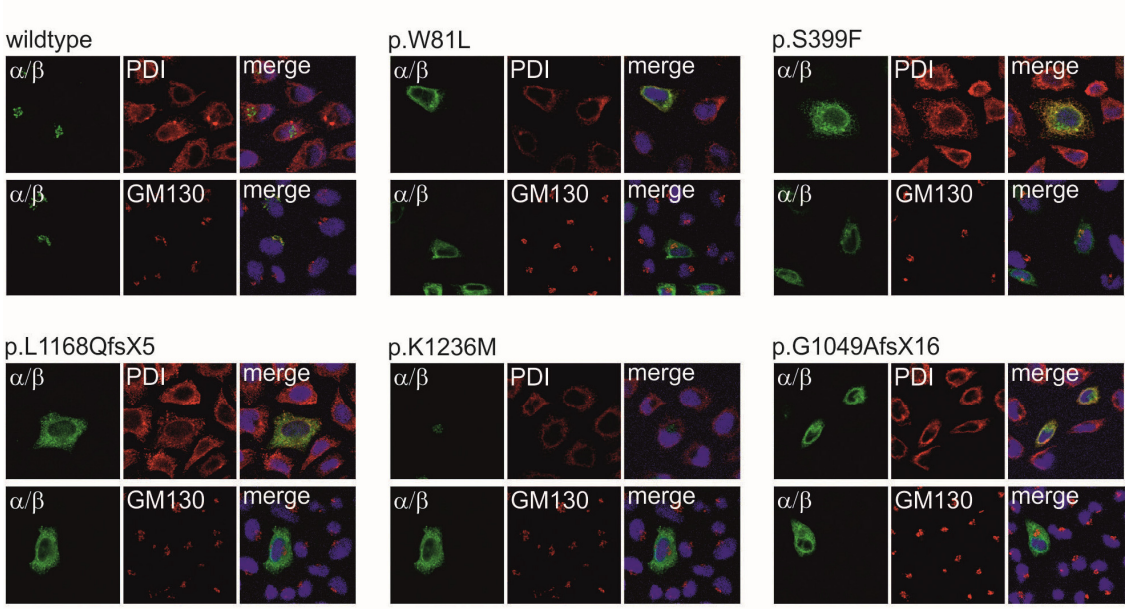


Figure 3: Analysis of subcellular location of GNPTAB mutant proteins in HeLa cell lines.

HeLa cells were transfected with cDNA encoding the wild-type *GNPTAB* miniconstruct (αβ-mini 1-430+849-1256 in pcDNA3.1), and the following mutants: p.W81L; p.S399F; p.L1168Q*fs*X5, p.K1236M and p.G10489*fs*X16. Cells were costained for either the *cis*-Golgi apparatus marker protein GM130 (*red*; 1:200 dilution) or the ER marker protein-disulfide isomerase, PDI (*red*; 1:400 dilution). Colocalization in merged images appears yellow.

Genotype-phenotype relation of analyzed α/β-subunit mutants of the GlcNAc-1-phosphotransferase

Table 1 – Summary classification of the known Mucolipidoses cases caused by the GNPTAB mutations under study according to zygosity and population of origin.

***GNPTAB* mutation** cDNA protein	**Clinical classification**	**Zygosity**	**Origin**	**Reference**
c.1196C>T p.S399F	MLIII alpha/beta	homozygous	Portuguese	Encarnacao et al., 2009
		heterozygous	French	Bargal et al., 2006
			Portuguese	Encarnacao et al., 2009
c.2956C>T p.R986C	MLII alpha/beta	homozygous	Indian	Coutinho et al., 2012
c.3707A>T p.K1236M	MLII alpha/beta (intermediate)	homozygous	German	Tiede et al., 2006
c.3145insC p.G1049R*fs*X16	MLII alpha/beta	homozygous	?	Tiede *et al.*, 2005
c.3503_3504delTC p.L1168*fs*X5	MLII alpha/beta	homozygous	Portuguese	Encarnacao *et al.*, 2009
			?	Kudo *et al.*, 2006
			Arab-muslim, Turkish, Irish	Bargal *et al.*, 2008
			French-Canadian	Plante *et al.*, 2008
			Italien, Argentinian, Bengalese	Tappino *et al.*, 2009
		heterozygous	Portuguese	Encarnacao *et al.*, 2009
			Italien, Hungarian, Bulgarian, Argentinian	Tappino *et al.*, 2009
	MLIII alpha/beta (intermediate)	heterozygous	?	Kudo *et al.*, 2006
			French, Irish	Bargal *et al.*, 2008
			Argentinian	Tappino *et al.*, 2009

c.242G>T p.W81L	MLII alpha/beta	homozygous	Portuguese	Encarnacao *et al.*, 2009

Mutations in the *GNPTAB* gene result in a broad phenotypic spectrum that, ultimately, may be divided into two different diseases: one with an early onset severe phenotype (ML II alpha/beta) and other with a latter onset milder presentation (ML III alpha/beta). Nevertheless, and as is described for most other lysosomal storage disorders, there is a phenotypic continuum. Consequently, clinical classification isn't always straightforward, with some cases being referenced/referred to as ML II/III intermediate.

The mutations analysed in this study have already been reported in different populations and zygosities (Table 1) and some genotype-phenotype correlations suggested but no expression studies were ever presented to support/justify them.

The missense mutation p.S399F was detected both in homozygosity and in heterozygosity and reported to cause the milder ML III alpha/beta disease. Patients harboring p.S399F were reported on the Portuguese [Encarnação *et al.*, 2009] and French population [Bargal *et al.*, 2006].

Mutation p.R986C, also missense, was detected in homozygous state in a patient of Indian origin with an early onset of symptoms and severe level of disability, therefore classified as ML II alpha/beta [Coutinho *et al.*, 2012].

The missense mutation p.K1236M, on the other hand, was reported in a single German patient, classified as ML II/III intermediate whose age of onset was higher and degree of disability milder than usually reported for typical ML II alpha/beta patients [Tiede *et al.*, 2006].

The frameshift mutation p.G10489*fs*X16 was reported in homozygous state in a patient suffering from severe ML II alpha/beta.

Finally, p.L1168*fs*X5, the most common disease-causing mutation reported in the *GNPTAB* gene, has been detected in virtually all studied populations, with the exception of the molecularly characterized patients from Japan. Its effect on the phenotype depends on zygozity: if present in homozygous state or in compound heterozygosity with another frameshift or nonsense mutation, underlies the more severe form of the disease (ML II alpha/beta) [Kudo *et al.*, 2006; Bargal *et al.*, 2008; Plante *et al.*, 2008; Encarnação *et al.*, 2009; Tappino *et al.*, 2009]; when combined with missense mutations, however, it has already been reported to cause a milder form of the disease [Kudo *et al.*, 2006; Bargal *et al.*, 2008; Tappino *et al.*, 2009].

Discussion

The first reports on the molecular characterization of large series of patients with mutations in the *GNPTAB* gene suggested that frameshift and nonsense mutations would be the ones causing the most severe ML II alpha/beta disease, while missense mutations either in homozygozity or in compound heterozygosity, would result in the milder phenotype, ML III alpha/beta [Tiede *et al.*, 2006]. Nevertheless, over the last few years, several studies have reported ML II alpha/beta patients carrying missense mutations [Encarnação *et al.*, 2009; Tappino *et al.*, 2009; Otomo *et al.*, 2009; Zarghooni and Dittakavi, 2009; Cathey *et al.*, 2010], this leading to change the initial paradigm. It became, then, mandatory to explore the factors which could account for the severity of a particular *GNPTAB* mutation and, even though biochemical analysis have already been published for some mutations, many others have only been reported on a genetic and clinical basis [Encarnação *et al.*, 2009; Guio *et al.*, 2011; Heo *et al.*, 2012]. In general, the phenotypic pathogenicity of missense mutations based strictly on genotype is difficult to predict, so expression of the mutant proteins is normally required to understand the effect of the amino acid change on folding, targeting, stability, and activity of the enzyme. So far, pathogenicity prediction of *GNPTAB* missense mutations has been mainly based on the nature of the affected amino acid, its evolutionary conservation, and position within the protein. Expression analyses were, then, required for a proper assessment of each mutant's severity.

In this work six pathogenic *GNPTAB* mutations were cloned and expressed into a *GNPTAB* miniconstruct ($\alpha\beta$-mini 1-430+849-1256 in pcDNA3.1): p.G1049A*fs*X16 (c.1581delC), p.L1168Q*fs*X5 (c.3503_3504delTC), p.W81L, p.S399F, p.R986C and p.K1236M. Plasmids containing wild-type and mutant GlcNAc-1-phosphotransferase proteins were transfected into HEK and HeLa cells and analysed through WB and IF, respectively. Protein expression levels were evaluated through WB analysis (Figure 2). The subcellular distribution pattern observed was addressed by IF (Figure 3).

In general, we observed a perfect correlation between the assessed features and the severity of the phenotype: mutants harboring the severe deletions are retained in the ER, presenting only the non-cleaved α/β- and inactive precursor form; mild missense mutations are correctly located in the Golgi apparatus but present reduced levels of the mature α- and β-subunits, when compared to the wild-type and, finally, severe missense mutations were either retained in the ER on the non-cleaved and inactive precursor form (p.W81L) or not expressed at all (p.R986C).

The frameshift mutation c.3503_3504delTC (p.L1168Q*fs*X5) is the most frequent ML II-causing mutation having been reported in virtually all studied populations [Bargal *et al.*, 2006; Kudo *et al.*, 2006; Plante *et al.*, 2008; Tappino *et al.*, 2009; Encarnação *et al.*, 2009]. When present in homozygous form, or in combination with other frameshift or nonsense

mutation, c.3503_3504delTC leads to a severe MLII alpha/beta phenotype. Patients harboring this mutation in homozygosity have been reported to present a severe phenotype with dysmorphic facial features, gingival hyperplasia, radiographic skeletal changes psychomotor retardation, all of them with early onset of symptoms (usually during the first year of life). At least one prenatal diagnosis was also reported [Encarnação *et al.*, 2009]. Previous reports had already demonstrated that the c.3503_3504delTC leads to a null or almost null GlcNAc-1-phosphotransferase activity. In fact, when a cDNA with this mutation was expressed, no activity could be demonstrated. Additionally, fibroblasts from a patient described in the original paper who was homozygous for this mutation, exhibited the very low value of <0.1% GlcNAc-1-phosphotransferase activity [Kudo *et al.*, 2006]. At protein level, such alteration was expected to result in a truncated protein (missing the last 84 amino acids) with a slightly lower molecular weight than the normal protein. In general, mutant transcripts harboring premature termination codons (PTC) such as this one trigger the nonsense mediated mRNA decay (NMD) mechanism, a cell quality control device which destroys abnormally small transcripts. Previous assessments of *GNPTAB* mRNA levels on fibroblasts harboring the c.3503_3504delTC in homozygosity showed a reduction of ~4.7 relatively to control [Encarnação *et al.*, 2009] suggesting that, even though the mutation may trigger NMD, this is not a fully efficient mechanism and some protein should still be produced. In the present study, protein expression levels were evaluated through Western blot and results showed a clear reduction on the GlcNAc-1-phosphotransferase expression levels for this mutant. Furthermore, only the non-cleaved precursor form could be detected. This is highly significant since only after cleavage does the GlcNAc-1-phosphotransferase become active. So, any mutant which does not suffer S1P cleavage in the Golgi, will necessarily present null activity. From the Western blot results alone it is also possible to infer that the p.L1168Q*fs*X5 (c.3503_3504delTC) mutant is retained in the ER, never reaching the Golgi apparatus. This was further confirmed through costaining with ER and *cis*-Golgi markers (Figure 3).

Similar WB and IF patterns were observed for the other severe deletion expressed on the *GNPTAB* miniconstruct, p.G10489*fs*X16 (c.1581delC): only the α/β-precursor could be detected and, from the IF images it was possible to verify it co-localized with the ER.

These results are in perfect agreement to those which would be expected for a frameshift mutation and with the phenotype reported for patients carrying it.

Concerning the missense mutations, different patterns were detected both through WB and IF. The two initial groups, composed one by the variants that were detected in patients suffering from a milder form of the disease (p.S399F and p.K1236M) and those identified in severely affected individuals (p.W81L and p.R986C), had completely different patterns, consistent with the observed phenotypes.

The mild missense mutation p.K1236M was correctly located in the Golgi apparatus

(Figure 3).

The mutant p.S399F, on the other hand, co-localized with the ER. Nevertheless, some of a slight co-localization with the Golgi apparatus was also observed (Figure 3). This is in accordance with the Western blot results for the same mutant that revealed major accumulation of the *N*-glycosylated ~120 kDa α/β-subunit precursor immunoreactive form together with minor bands which corresponded to the cleaved/mature α- and β-subunits of ~75 kDa and ~45 kDa, respectively (Figure 2). This implies that at least a small amount of the mutant α/β-precursor gets to be cleaved by S1P and correctly targeted to the Golgi. Clinical evidence seems to suggest that this small amount of mature α- and β-subunits may provide residual GlcNAc-1-phosphotransferase activity minimizing the phenotypic effect of the pathogenic mutation. The only reported p.S399F homozygous individual has Portuguese origin and a mild ML III alpha/beta phenotype, without neurological impairment. Patient is now 28 years old and has higher education, which further demonstrates that this mutation is associated to a mild phenotype.

The missense mutation c.242G>T (p.W81L) was retained in the ER on the non-cleaved and inactive precursor form, presenting a pattern similar to the one observed for frameshift mutations (Figure 2 and 3). This variant was first identified in homozygosity in a patient reported in 2009, who was diagnosed at the age of 4 months presenting clinical features of ML II alpha/beta such as gingival hypertrophy, hypertonia and delayed growth. Death occurred at 23 months of age by respiratory infection [Encarnação *et al.*, 2009]. Latter, we characterized a second patient of Portuguese origin (not published) who was diagnosed at the age of 3 months with the exact same genotype and presented with typical ML II alpha/beta features (dysostosis multiplex, gingival hypertrophy, hypothyroidism, respiratory difficulties and delayed growth) and slight hepatosplenomegaly. The fact that a second unrelated severely affected patient presented with the same genotype and a similar ML II alpha/beta phenotype is particularly relevant since it reinforces the causal relationship between the p.W81L missense mutation and the severe form of the disease. No fibroblast cell lines were available for the second patient but, the index case had already been analysed through qtRT-PCR showing a significant decrease when compared with the median control values [Encarnação *et al.*, 2009]. All computational predicted this mutation to be potentially damaging. Interspecific alignments of 54 homologous GlcNAc-phosphotransferases revealed that it occurred at an evolutionarily conserved amino acid residue.

Finally, we observed a complete absence of expression for the p.R986C mutant (Figure 2). The p.R986C variant was identified in homozygosity in a severely affected patient [Coutinho *et al.*, 2012]. The child, a boy of Indian origin, was only diagnosed at 5 years old, but first symptoms were reported one month after birth: coarse facies, depressed nasal bridge, epicanthic folds, severe gum hypertrophy, deformed chest, open mouth, widening

of wrists and contractures of fingers. The homozygous status of p.R986C was confirmed by sequencing genomic DNA from both parents. *In silico* predictions classified as potentially damaging and interspecific alignments revealed that it occurred at an evolutionarily conserved residue. Our present results support its association to a severe form of the disease, since no protein is present at all. Nevertheless, additional questions arise from the unusual pattern observed for this mutation. In theory, the absence of expression of the p.R986C GlcNAc-1-phosphotransferase mutant can be due to either mRNA instability or premature protein degradation but only additional analysis may bring some light on the effect of this variant in particular.

Conclusion

For the severe deletions c.440delC (p.G10489*fs*X16) and c.3503_3504delTC (p.L1168Q*fs*X5), phenotype may be explained by abnormal cellular location. Mutant α/β-precursor proteins are retained in the ER on their non-cleaved inactive form and accumulate on that compartment, never reaching the Golgi, where they should act.

For the milder missense mutations (p.S399F and p.K1236M) were observed decreased protein levels which are consistent with reduced enzymatic activity. Nevertheless, and unlike the severe deletions where only the precursor is detected, these mutants give rise to both cleaved and non-cleaved forms of the α/β-subunits which reach the Golgi, presenting a normal (p.K1236M) or partially normal (p.S399F) subcellular distribution pattern, coincident with the *cis*-Golgi complex and similar to the control wild-type protein.

For the atypically severe missense mutations (p.W81L and p.R986C), surprising observations were made: the p.W81L mutant is retained in the ER on the non-cleaved and inactive precursor form and the p.R986C does not seem to be expressed at all. The striking absence of expression observed for the p.R986C mutant raises the question of the role of residue R^{986} on the protein function and/or stability, prompting additional future studies on the topic.

In conclusion, the combined analysis of the mutants' expression levels and subcellular location perfectly explains the phenotype. Finally, being the identification of mutations in disease-affected genes the basis of genetic counseling, this study also highlights the need of a carefully and extensive analysis to provide a reliable genotype-phenotype correlation.

Acknowledgments

The authors would like to thank Joahnnes Brand for expert technical assistance. This work was supported by FCT - project PIC/IC/83252/2007 (http://alfa.fct.mctes.pt/). Coutinho MF is a grantee from the FCT (SFRH/BD/48103/2008).

References

Bao M, Booth JL, Elmendorf BJ, *et al*. Bovine UDP-N-acetylglucosamine:lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure. (1996) 271:31437-45.

Bargal R, Zeigler M, Abu-Libdeh B, *et al*. When Mucolipidosis III meets Mucolipidosis II: GNPTA gene mutations in 24 patients (2006) *Mol Genet Metab*. 88:359-363.

Cathey SS, Leroy JG, Wood T, *et al*. Phenotype and genotype in mucolipidoses II and III alpha/beta: a study of 61 probands (2010) *J Med Genet*. 47:38-48.

Coutinho MF, Da Silva Santos L, Girisha KM, *et al*. Mucolipidosis Type II α/ß with a Homozygous Missense Mutation in the *GNPTAB* Gene (2012) *Am J Med Genet A*. 158A:1225-8.

Koch-Nolte F, Glowacki G, Bannas P *et al*. Use of genetic immunization to raise antibodies recognizing toxin-related cell surface ADP-ribosyltransferases in native conformation (2005) *Cell Immunol*. 236:66-71.

Dierks T, Schlotawa L, Frese MA, *et al*. Molecular basis of multiple sulfatase deficiency, mucolipidosis II/III and Niemann-Pick C1 disease - Lysosomal storage disorders caused by defects of non-lysosomal proteins (2009) *Biochim Biophys Acta*. 1793:710-25.

Encarnação M, Lacerda L, Costa R, *et al*. Molecular analysis of the GNPTAB and GNPTG genes in 13 patients with mucolipidosis type II or type III - identification of eight novel mutations (2009) *Clin Genet*. 76:76-84.

Encarnação M, Kollmann K, Trusch M, *et al*. Post-translational modifications of the gamma-subunit affect intracellular trafficking and complex assembly of GlcNAc-1-phosphotransferase (2011) *J Biol Chem*. 286:5311-8.

Guio JA, Ramirez PA, Bermudez M, Obando FS. [Mucolipidosis type II-inclusion cell disease] (2012) *An Pediatr* 76:108-9.

Heo JS, Choi KY, Sohn SH *et al*. A case of mucolipidosis II presenting with prenatal skeletal dysplasia and severe secondary hyperparathyroidism at birth (2012) *Korean J Pediatr*. 55:438-44.

http://www.hgmd.org

Kornfeld S, Sly WS. I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localization. *In*: Scriver CR, Beaudet AL, Sly WS, Valle D., editors: The metabolic and molecular bases of inherited disease (2001) 8th ed, New York: MacGraw-Hill, Inc. 3421-3452.

Kollmann K, Damme M, Markmann S et al. Lysosomal dysfunction causes neurodegeneration in mucolipidosis II 'knock-in' mice (2012) *Brain*. 135: 2661-75.

Kudo M, Brem MS, Canfield WM. Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha / beta -subunits precursor gene (2006) *Am J Hum Genet*. 78:451-63.

Marschner K, Kollmann K, Schweizer M, *et al*. A key enzyme in the biogenesis of lysosomes is a protease that regulates cholesterol metabolism (2011) *Science*. 333:87-90.

Otomo T, Muramatsu T, Yorifuji T, *et al*. Mucolipidosis II and III alpha/beta: mutation analysis of 40 Japanese patients showed genotype-phenotype correlation (2009) *J Hum Genet*. 54:145-51.

Paik KH, Song SM, Ki CS, *et al*. Identification of mutations in the GNPTA (MGC4170) gene coding for GlcNAc-phosphotransferase alpha/beta subunits in Korean patients with mucolipidosis type II or type IIIA (2005) *Hum Mutat*. 26:308-14.

Plante M, Claveau S, Lepage P, *et al*. Mucolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTAB) in a French Canadian founder population (2008) *Clin Genet*. 73:236-244.

Pohl S, Tiede S, Marschner K, *et al*. Proteolytic processing of the gamma-subunit is associated with the failure to form GlcNAc-1-phosphotransferase complexes and mannose 6-phosphate residues on lysosomal enzymes in human macrophages (2010) *J Biol Chem*. 285:23936-44.

Raas-Rothschild A, Bargal R, Goldman O, *et al*. Genomic organisation of the UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) and its mutations in mucolipidosis III (2004) *J Med Genet*. 41:e52.

Raas-Rothschild A, Cormier-Daire V, Bao M, *et al*. Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC) (2000) *J Clin Invest*. 105:673-81.

Stenson PD, Mort M, Ball E, *et al*. The Human Gene Mutation Database: 2008 update. (2009) *Genome Med*. 1:13.

Storch S, Braulke T. Transport of lysosomal enzymes. *In*: Saftig P., editor. Lysosomes. Eurekah, Landes Bioscience, Springer & Business Media, New York, 2005:17-26.

Tappino B, Chuzhanova NA, Regis S, *et al*. Molecular characterization of 22 novel UDP-N-acetylglucosamine-1-phosphate transferase alpha- and beta-subunit (GNPTAB) gene mutations causing mucolipidosis types IIalpha/beta and IIIalpha/beta in 46 patients (2009) *Hum Mutat*. 30:E956-73.

Tiede S, Cantz M, A Raas-Rothchild, *et al*. A novel mutation in UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) in two siblings with Mucolipidosis type III alters a used glycosylation site (2004) *Hum Mutat*. 24:535.

Tiede S, Storch S, Lübke T, *et al*. Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase (2005a) *Nat Med*. 11:1109-12.

Tiede S, Cantz M, Spranger J, *et al*. Missense mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTA) in a patient with mucolipidosis II induces changes in the size and cellular distribution of GNPTG (2006) *Hum Mutat*. 27:830-1.

Zarghooni M, Dittakavi SS. Molecular analysis of cell lines from patients with mucolipidosis II and mucolipidosis III (2009) *Am J Med Genet A*. 149A:2753-61

3. Haplotypic analysis

Paper 6: Origin and spread of a common deletion causing Mucolipidosis type II: insights from patterns of haplotypic diversity

CLINICAL GENETICS An International Journal of Genetics, Molecular and Personalized Medicine

Clin Genet 2011: 80: 273–280
Printed in Singapore. All rights reserved

CLINICAL GENETICS
doi: 10.1111/j.1399-0004.2010.01539.x

Short Report

Origin and spread of a common deletion causing mucolipidosis type II: insights from patterns of haplotypic diversity

Coutinho MF, Encarnação M, Gomes R, da Silva Santos L, Martins S, Sirois-Gagnon D, Bargal R, Filocamo M, Raas-Rothschild A, Tappino B, Laprise C, Cury GK, Schwartz IV, Artigalás O, Prata MJ, Alves S. Origin and spread of a common deletion causing mucolipidosis type II: insights from patterns of haplotypic diversity.
Clin Genet 2011: 80: 273–280. © John Wiley & Sons A/S, 2010

Mucolipidosis II (ML II alpha/beta), or I-cell disease, is a rare genetic disease in which activity of the uridine diphosphate (UDP)-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase (GlcNAc-phosphotransferase) is absent. GlcNAc-phosphotransferase is a multimeric enzyme encoded by two genes, *GNPTAB* and *GNPTG*. A spectrum of mutations in *GNPTAB* has been recently reported to cause ML II alpha/beta. Most of these mutations were found to be private or rare. However, the mutation c.3503_3504delTC has been detected among Israeli and Palestinian Arab-Muslim, Turkish, Canadian, Italian, Portuguese, Irish traveller and US patients. We analysed 44 patients who were either homozygous or compound heterozygous for this deletion (22 Italians, 8 Arab-Muslims, 1 Turk, 3 Argentineans, 3 Brazilians, 2 Irish travellers and 5 Portuguese) and 16 carriers (15 Canadians and 1 Italian) for three intragenic polymorphisms: c.-41_-39delGGC, c.18G>A and c.1932A>G as well as two microsatellite markers flanking the *GNPTAB* gene (D12S1607 and D12S1727). We identified a common haplotype in all chromosomes bearing the c.3503_3504delTC mutation. In summary, we showed that patients carrying the c.3503_3504delTC deletion presented with a common haplotype, which implies a common origin of this mutation. Additionally, the level of diversity observed at the most distant locus indicates that the mutation is relatively ancient (around 2063 years old), and the geographical distribution further suggests that it probably arose in a peri-Mediterranean region.

Conflict of interest

None of the authors declare any conflict of interest. The study sponsors played no role in study design, collection, analysis and interpretation of data.

MF Coutinho[a,b], M Encarnação[a], R Gomes[a], L da Silva Santos[a], S Martins[c], D Sirois-Gagnon[d], R Bargal[e], M Filocamo[f], A Raas-Rothschild[e], B Tappino[f], C Laprise[d], GK Cury[g], IV Schwartz[g,h,i,j], O Artigalás[i,j], MJ Prata[b,c] and S Alves[a]

[a]Medical Genetics Center Jacinto Magalhães, Research and Development Unit, Department of Genetics, INSA, Portugal, [b]Department of Biology, Faculty of Sciences, Porto, Portugal, [c]IPATIMUP, Porto, Portugal, [d]Université du Québec à Chicoutimi, Département des Sciences Fondamentales, Saguenay, Québec, Canada, [e]Department of Human Genetics and Metabolic Diseases, Hadassah Hebrew University Medical Center, Jerusalem, Israel, [f]''Diagnosi Pre-Postnatale Malattie Metaboliche'' Laboratory-IRCCS G. Gaslini, Genova, Italy, [g]Post Graduation Program in Medicine: Medical Sciences, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [h]Medical Genetics Service, Clinical Hospital of Porto Alegre, Porto Alegre, Brazil, [i]Genetics Department, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, and [j]Post Graduation Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Key words: c.3503_3504delTC – common origin – haplotypic analysis – mucolipidosis

Corresponding author: Sandra Alves, Medical Genetics Center Jacinto Magalhães, INSA, Praça Pedro Nunes, 88, 4099-028 Porto, Portugal.
Tel.: +351 226070300;
fax: +351 226070399;

Coutinho et al.

e-mail:
sandra.alves@insa.min-saude.pt;
alvessandra@hotmail.com

Received 15 June 2010, revised and accepted for publication 31 August 2010

Mucolipidosis II (ML II) is an autosomal recessive disease of lysosomal hydrolase trafficking caused by total or near total deficiency of the uridine diphosphate (UDP)-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase (GlcNAc-phosphotransferase, EC 2.7.8.17). GlcNAc-phosphotransferase is a hexameric complex of $\alpha 2\beta 2\gamma 2$ whose α and β subunits are encoded by the *GNPTAB* gene (MIM# 607840). A spectrum of mutations in this gene has been reported to cause ML II alpha/beta, and most of the mutations have been found to be private or rare. However, the microdeletion mutation c.3503_3504delTC was found to be a common founder mutation in a French-Canadian population (1). Furthermore, it shows a remarkably wide geographical distribution, having been detected among Israeli and Palestinian Arab-Muslim, Turkish, Irish traveller (2), Italian (3), Portuguese (4) and US patients (5).

Up to now, c.3503_3504delTC has been considered the most common pathogenic mutation associated with ML II. Because the molecular defect involves a 2-bp deletion from a small repeat CTCT sequence, which appears to have originated through slippage mispairing (5), the question arises whether such an error occurred more than once, suggesting that the c.3503_3504delTC mutation has a recurrent origin.

To investigate whether the worldwide spread and frequent c.3503_3504delTC mutation was due to a unique founder molecular lesion or if the deletion arose more than once through a recurrent mutational event, we assessed the diversity at two microsatellite markers flanking the *GNPTAB* gene and three *GNPTAB* intragenic polymorphisms among 44 ML II patients whose previous molecular characterization revealed them to be homozygous or compound heterozygous for c.3503delTC (1–4). This sample corresponds to the majority of characterized patients harbouring this deletion and is comprised of individuals from very distinct populations: Portuguese, Italian, Israeli Arab-Muslim, Brazilian, Turkish, Argentinean and Irish travellers. The criterion for their selection was the availability of biological material for the analyses. Additionally, healthy individuals of French-Canadian and Italian origin known to be obligatory carriers of c.3503_3504delTC were also enrolled in the study. It is important to stress that this sample is highly representative because ML II is a rare disease, with an estimated birth prevalence in Portugal of 0.16 cases per 100,000 live births (6). This rate is in accordance with other values of birth prevalence calculated for different populations including 3 cases per 1,000,000 births in Australia (7) and 0.8 cases per 1,000,000 live births in the Netherlands (8).

Results from the haplotypic backgrounds clearly pointed to a common origin of the analysed c.3503delTC chromosomes. Furthermore, an age estimate of the mutation indicates that it is ancient enough to explain the present-day broad geographical distribution.

Materials and methods

Patients and controls

We analysed 44 patients who were either homozygous or compound heterozygous for the c.3503_3504delTC mutation (22 Italians, 8 Israeli Arab-Muslims, 5 Portuguese, 3 Argentineans, 3 Brazilians, 1 Turk and 2 Irish travellers). Additionally, 15 healthy carriers from Canada and 1 from Italy who were parents of children with ML II and known to be homozygous for c.3503_3504delTC were studied (see Table S1). Portuguese, Italian and Argentinean, Israeli Arab-Muslim and Canadian patients and/or carriers were previously characterized for the *GNPTAB* gene (1–4). A sample of 45 healthy Portuguese individuals was also analysed.

Polymorphisms and genotyping procedures

Three *GNPTAB* intragenic polymorphisms, c.-41_-39delGGC (rs76300806), c.18G>A (rs4764655) and c.1932A>G (rs10778148), located in the 5′ UTR region, exon 1 and exon 13, respectively, as well as two microsatellite markers, D12S1607 and D12S1727, flanking *GNPTAB* and 0.034 and 0.43 Mb distant from the gene, respectively, were used for the haplotypic analysis (see Fig. 1). D12S1607 and D12S1727 are dinucleotide repeats, whereas c.-41_-39delGGC is a 3-bp Indel. The three intragenic polymorphisms of c.-41_-39delGGC, c.18G>A and c.1932A>G

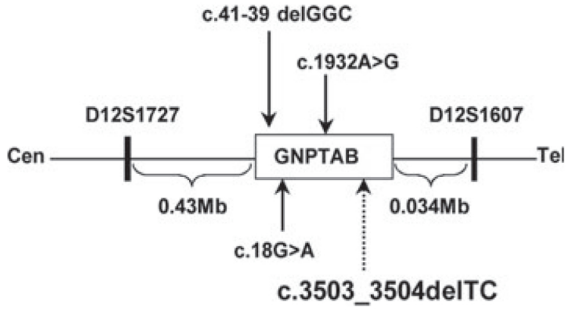


Fig. 1. Schematic representation of the polymorphisms studied and their relative distance to the c3503_3504delTC mutation. Cen: Centromere; Tel: Telomere.

were screened by direct sequencing as previously described in Encarnação et al. (4).

As for the two microsatellites, D12S1607 and D12S1727, the polymerase chain reaction (PCR) products were separated in an ABI PRISM *3130xl* Genetic Analyzer (Applied Biosystems, Foster City, CA) after amplification with fluorescently labelled primers. A multiplex system for the simultaneous detection of the two microsatellites was developed (primers and PCR conditions upon request).

Haplotype determination

In patients homozygous for the c.3503delTC deletion, the haplotype phase was directly inferred.

In patients who were compound heterozygous for the deletion or in healthy carriers, PHASE software version 2.0 (http://www.stat.washington.edu/stephens/software.html) was used to reconstruct the haplotypes from genetic data while taking into account allele frequencies and phase-known haplotypes. Only haplotype pairs with a probability greater than 0.6 were used for further analyses.

Allele frequency distributions in mutant and control Portuguese chromosomes were compared by evaluating the significance of the difference in proportions as previously described (9). The Z statistic was calculated using the Differ program of the PEPI package (http://sagebrushpress.com/pep.html).

Standard diversity indices were calculated according to Nei (10) using the ARLEQUIN program (11).

Estimation of age

To estimate the age of the deletion from the variation accumulated in their ancestral haplotypes, both recombination (c) and mutation (μ) rates were considered in the generation of variation as described in Ref. (12). The probability of change per generation was given by $\varepsilon = 1 - [(1 - c)(1 - \mu)]$, and the average of the mutation and recombination events (λ) equals εt, where t is the number of generations.

The recombination rate (c) was based on the physical distance between the two most distant markers D12S1727 and D12S1607 (540 Kb): $c = 0.0054$, when using a conversion factor of 1 Mb = 1 cM as reported for this genomic region (13). The estimated probability of mutation per generation and per polymorphism was 1.56×10^{-3} [as two dinucleotides and three single nucleotide polymorphisms (SNPs) were used]. The mutation rate for the dinucleotides was considered to be 7.8×10^{-4} (14). For the two SNPs and the Indel c.-41_-39delG, the assumed mutation rate was 2×10^{-9} (15). Moreover, a mean value of 25 years was presumed to correspond to each generation.

Results and discussion

Ancestral haplotype identification

Table 1 contains a list of the inferred or reconstructed haplotypes that were found as background in 61 chromosomes harbouring the c.3503-3504 delTC mutation. Haplotypes were defined by variation at the five markers of D12S1727, c.-41_-39delGGC, c.18G>A, c.1932A>G and D12S1607, whose order is indicated according to the chromosomal location presented in Fig. 1.

Without accounting for the population of origin, a total of nine different haplotypes were detected. One of them, H1-13:7:G:G:12, was clearly predominant and present in 29 out of the 61 mutant chromosomes (49%). This haplotype was shared by Italian, Portuguese and Israeli Arab-Muslim patients. Among the Italians, H1 reached the especially high frequency of 70% (21/30 chromosomes); among the Portuguese, it was also the most common haplotype occurring at a frequency of 37.5% (3/8 chromosomes). Among Israeli Arab-Muslims, 25% of the chromosomes were found to contain c.3503_3504delTC (4/16), despite it not being the most frequent haplotype.

It is noteworthy that 59 of the 61 disease-associated chromosomes (97%) shared a core haplotype covering four of the five polymorphic markers analysed. This core haplotype showed the allelic array of 7:G:G:12 at c.-41_-39delGGC, c.18G>A, c.1932A>G and D12S1607, respectively, which was found to be distinct in only two chromosomes from Italian patients: one harboured the 4-locus haplotype 8:G:G:12, and the other contained 7:G:G:18 according to the same locus order as the core haplotype. When compared to the core,

Table 1. Haplotypes present on the 3503_3504delTC chromosomes and age estimation

	Haplotype	D12S1727	c.41-39delGGC	c.18G>A	c.1932A>G	D12S1607	Chromosome No.	Mutation no.	Age (years)
All samples	H1	13	7	G	G	12	29	0	2063 ± 729
	H2	12	7	G	G	12	2	1	
	H3	14	7	G	G	12	8	1	
	H4	15	7	G	G	12	2	2	
	H5	16	7	G	G	12	12	1	
	H6	17	7	G	G	12	4	1	
	H7	11	7	G	G	12	2	2	
	H8	17	8	G	G	12	1	1	
	H9	13	7	G	G	18	1	1	
Italians	H1	13	7	G	G	12	21	0	1079 ± 531
	H2	12	7	G	G	12	2	1	
	H3	14	7	G	G	12	5	1	
	H8	17	8	G	G	12	1	1	
	H9	13	7	G	G	18	1	1	
Portuguese	H1	13	7	G	G	12	3	0	3147 ± 800
	H3	14	7	G	G	12	1	1	
	H4	15	7	G	G	12	2	2	
	H6	17	7	G	G	12	2	1	
Arab-Muslims	H1	13	7	G	G	12	4	0	2697 ± 2147
	H5	16	7	G	G	12	12	1	
Argentineans	H7	11	7	G	G	12	2	—	—
Brazilians	H6	17	7	G	G	12	1	—	—
Travellers	H1	13	7	G	G	12	1	—	—
	H3	14	7	G	G	12	2	—	
	H6	17	7	G	G	12	1	—	

each of the remaining two haplotypes differed by alleles at only one marker.

This finding indicates that the region encompassing the four polymorphic markers is in strong linkage disequilibrium with c.3503_3504delTC, which leads to the conclusion that this disease-causing mutation in *GNPTAB* occurred only once in a founder haplotype. Furthermore, the pattern of 7:G:G:12 distribution together with its wide population dissemination strongly support the hypothesis that this 4-locus haplotype was the ancestral configuration from which the molecular lesion c.3503_3504delTC first arose.

A network was constructed with the nine detected haplotypes (Fig. 2), assuming the most parsimonious relationships between them as well as taking into consideration the number of stepwise events required, the distance between the markers involved, the presence of intermediate haplotypes in the populations to which the patients belonged and data on allele and haplotype frequency in the controls when available.

The network illustrates the close molecular relationships between the distinct haplotypes and shows that relative to the most frequent haplotype H1 the majority of the remaining haplotypes are connected through simple stepwise events of mutation or recombination. Except for

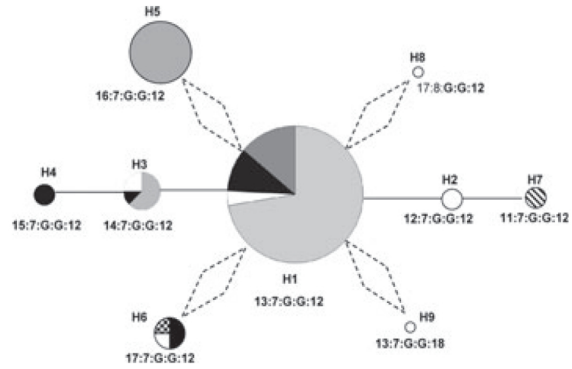


Fig. 2. Network showing the most parsimonious relationships among the nine detected haplotypes. The link between haplotypes (mutation or recombination) was determined while taking into consideration the number of stepwise events required, the distance between the genetic markers involved, the presence of intermediate haplotypes in the populations to which the patients belonged and data on the allele and haplotype frequencies in the controls (when available). Dashed diamonds represent the occurrence of recombination (haplotype H1 to haplotypes H5, H6, H8 and H9) and straight lines represent the occurrence of mutation (haplotype H1 to haplotypes H2, H3, H4 and H7). Circles are proportional to the absolute frequencies of the haplotypes. White: Italians; Black: Portuguese; Grey: Arab-Muslims; Squares: Argentineans.

the two molecular steps that underlie the origin of H8 and H9, which involved the markers c.-41_-39delGGC and D12S1607, respectively, all

other diversification events occurred at the locus D12S1727. This result is fairly understandable given that this marker is the farthest away from the target molecular lesion c.3503-3504delTC, which allows for more recombination opportunities. This observation explains why the second most frequent haplotype H5, which harbours the 5-allelic background 16:7:G:G:12, differs only from the most common haplotype at allele 16, which is present at D12S1727 instead of allele 13. H5 was found exclusively in Israeli Arab-Muslim patients, where it reached the elevated frequency of 75%, probably due to the effects of genetic drift.

To better evaluate the extent of linkage disequilibrium in mutation-bearing chromosomes, we compared the allele frequencies at the five polymorphic markers used to define haplotypes with the values estimated in healthy controls. Because our control sample contained only Portuguese subjects, we excluded the Israeli Arab-Muslims, Brazilians and Argentineans from the comparison. Consequently, the analysis was performed with only the haplotypes inferred for the European individuals from Table 1 (Italian and Portuguese) under the assumption that the Portuguese controls would work as a reasonable reference for healthy Europeans.

The graphical representation of allele frequencies in normal and mutation-bearing chromosomes is presented in Fig. 3. The five alleles defining the haplotype 13:7:G:G:12 had frequencies in chromosomes without the deletion that were significantly different from values in c.3503_3504delTC chromosomes. Extreme linkage disequilibrium was observed between the mutation and the polymorphic markers c.18G>A and c.1932A>G, but all five alleles were over-represented among the mutant chromosomes when compared to controls: allele 13 (D12S1727) 0.66 *vs* 0.43 ($p = 0.028$); allele 7 (c.-41_-39delGGC) 0.97 *vs* 0.56 ($p = 0$); allele G (c.18G>A) 1.0 *vs* 0.87 ($p = 0.046$); allele G (c.1932A>G) 1.0 *vs* 0.47 ($p = 0$); allele 12 (D12S1607) 0.97 *vs* 0.27 ($p = 0$). Phase-inferred 5-locus haplotypes were achieved for 84 normal chromosomes, among which 35 different haplotypes could be found. In the controls, the haplotype diversity reached the very high value of 0.954 as compared to 0.709, which was estimated for the overall sample of mutated chromosomes. Furthermore, the most common haplotypic background occurred at a moderate frequency of 14% for the extended haplotype 13:7:G:G:12 and 21% for the core 7:G:G:12, which sharply contrasted with 49% and 97%, respectively, in

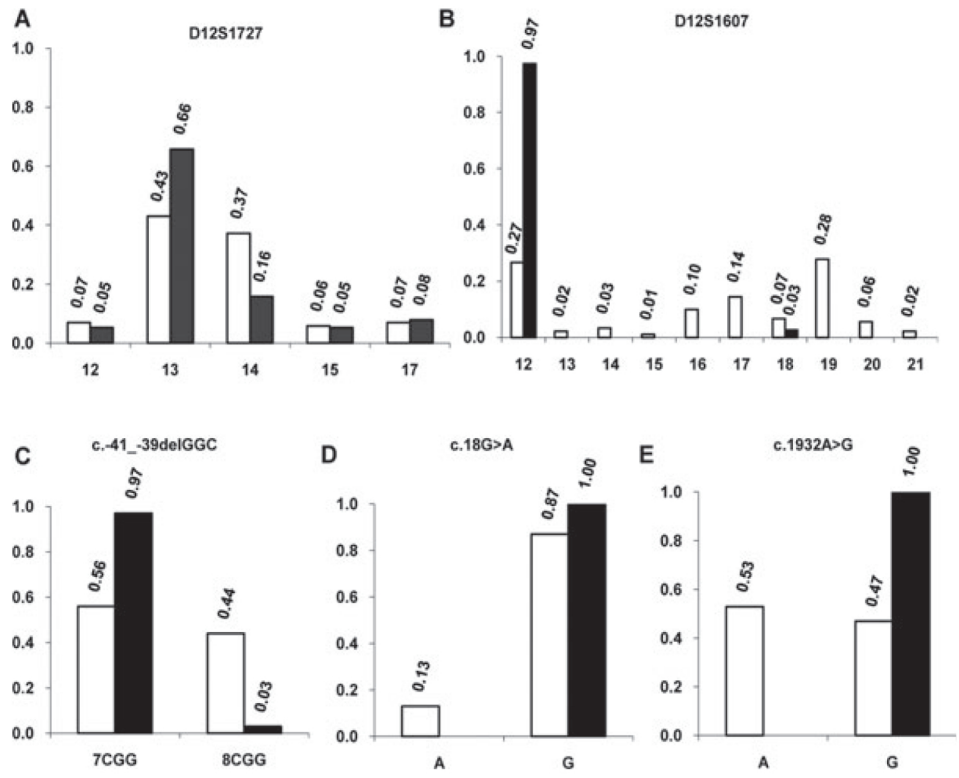


Fig. 3. Polymorphism allele frequency distributions within the wild-type Portuguese chromosomes and the 3503_3504delTC Italian and Portuguese chromosomes. (a) D12S1727; (b) D12S1607; (c) c.-41_−39delGGC; (d) c.18G>A; (e) c.1932A>G. White bars: wild type; black bars: 3503_3504delTC.

the mutated chromosomes. Therefore, it seems extremely unlikely that independent events leading to the same pathogenic deletion could have occurred in an identical haplotypic background, which reinforces the hypothesis of a single origin for the mutation under analysis.

In this study, 15 Canadian carriers of c.3503_3504delTC were also enrolled. They were healthy parents of confirmed homozygous patients for the deletion, but the PHASE program failed to determine their 5-locus haplotypes with enough certainty; however, this limitation could be overcome if the homozygous children were available. However, when haplotypes were inferred based upon the three polymorphic markers of c.-41_-39delGGC, c.18G>A, and D12S1607, the threshold of confidence was attained for the majority of chromosomes, and all shared the inferred haplotypic combination of 7:G:12. Furthermore, when looking at genotypes of the Canadian individuals (Table S1), it is possible to observe the presence of the common haplotype of 12:7:G:G:12 in all subjects, which is very likely in phase with the deletion. Because this haplotype is only one mutational step away from 13:7:G:G:12, this observation adds evidence of a common origin of the chromosomes carrying the molecular lesion.

The PHASE program was also unable to infer haplotypes with enough certainty in all the Argentinean, Brazilian and Turkish patients who carried the c.3503_3504delTC mutation together with a second pathogenic mutation, probably due to the very small number of studied patients. However, examination of their genotypes at the five polymorphic markers (Table S1) did not exclude the possibility that they harbour either 13:7:G:G:12 or other haplotypes with close molecular features.

c.3503-3504delTC age estimation

The presence of detectable linkage disequilibrium at markers flanking the c.3503_3504delTC mutation and the availability of defined map distances for each marker allowed for an estimation of the coalescent time of the mutated chromosomes.

Because 13:7:G:G:12 was the most widespread and well-represented haplotype, we first assumed that it was the most ancestral lineage when calculating the time needed for haplotypic diversification and assuming the events used to construct the network in Fig. 2 (also explicitly indicated in Table 1). Calculations were performed individually for chromosomes belonging to Italian, Israeli Arab-Muslim and Portuguese subjects and yielded ages of 1079, 2697 and 3147 years, respectively (Table 1). When the entire set of chromosomes was considered, the obtained age was 2063 years. Next, an alternative scenario for the founder haplotype was modelled. Only two haplotypes were detected among the Israeli Arab-Muslims, and we assumed that the most prevalent one, 16:7:G:G:12, was the founder from which the latter, 13:7:G:G:12, was derived through recombination. In this population, the derived age estimate was 899 $\pm$ 716 years, whereas it was 4186 $\pm$ 1504 years in the whole set of chromosomes (data not shown in Table 1). Because 16:7:G:G:12 was restricted to Israeli Arab-Muslim individuals, no other computations were performed.

Whichever founder was assumed, the range of ages was in agreement with the levels of haplotypic diversity within mutation-bearing chromosomes across the three populations, which were 0.492 $\pm$ 0.099 in Italians, 0.400 $\pm$ 0.114 in Israeli Arab-Muslims and 821 $\pm$ 0.100 in Portuguese individuals.

The greatest diversity found in Portuguese individuals does not necessarily mean that this population was the ancestral one where the mutation first arose. Low representation of Portuguese samples with respect to the total number of individuals may generate an ascertained bias. Patterns of diversity and, consequently, age estimates are highly influenced by selective pressures and demographic effects. While no reasons are envisioned for why selection over mutated chromosomes would have acted differently in Italian, Israeli Arab-Muslim and Portuguese individuals, reasons do exist to believe that demography might have played a role in the reduced level of diversity in Italians and Israeli Arab-Muslims. In fact, approximately 40% of the Italian patients studied here are from Sicilia, a region where extensive genetic drift effects have been documented (16), whereas among Israeli Arab-Muslims, consanguineous unions are a common tradition (17). If one admits that demographic effects very probably account for the reduced haplotypic diversity in Italians and Israeli Arab-Muslims, then the age estimates of the mutation in both population groups must be a serious underestimate.

To address questions about the region where c.3503_3504delTC initially originated and its subsequent dispersion route, it would be essential to have a more comprehensive picture of the worldwide haplotypic distribution within mutated chromosomes. Nevertheless, given the strong molecular evidence gathered here that this mutation arose only once and knowing that it is well represented both in Europe and the Middle East, the

answer certainly lies within the most important past migrations of the people into the two regions.

The timeframe of our age estimates either accommodates the demographic expansion subsequent to the Neolithic transition, which was initiated approximately 10,000 years ago in the Near East, or the movement associated with the expansion of the Phoenicians who spread their civilisation across the entire Mediterranean region during the more recent Bronze Age. Both of these movements of people were directed from the Eastern Mediterranean towards Europe, whereas no migration with an equivalent impact was recorded in the reverse direction.

For now, the best reconstruction for the origin and spread of the c.3503_3504delTC mutation is to assume that it probably arose in an Eastern peri-Mediterranean region of Europe, having later been introduced to the continent by migrants coming from the region encompassing the Middle/Near East. The remarkable frequency of the c.3503_3504delTC mutation among Arab-Muslims clearly favours this hypothesis, but once this population is examined, the deletion is anchored in two well-represented haplotypes, 13:7:G:G:12 and 16:7:G:G:12. Thus, some doubt persists about in which background the original mutational event occurred. Whatever the case, only 13:7:G:G:12 appears to have entered Europe or at least underwent successful dissemination throughout the continent.

In very recent historical times, the massive emigration of Europeans to the New World represented the travelling vehicle of the mutation to the American continent. All examined mutated chromosomes from individuals belonging to North or South American countries fell into the category of the European founder haplotype or its closest molecular derivates (for instance, in Quebec, Canada, the introduced haplotype seems to be 12:7:G:G:12).

In summary, the assessment of haplotypic diversity within a considerable worldwide sample of chromosomes bearing the pathogenic c.3503_3504delTC mutation allowed for the inference that they are descendents of a single founder ancestor. This result means that the molecular damage from which the most prevalent pathogenic mutation associated with ML II arose appears to have occurred only once, probably in a peri-Mediterranean region out of Europe. Its introduction in Europe is ancient enough to explain the present-day broad geographical distribution in the continent as well as in populations of recent European ancestry. Demography appears to have acted as a major determinant of the pattern of distribution of the c.3503_3504delTC mutation across distinct populations.

Supporting Information

The following Supporting information is available for this article:

Table S1. Samples analysed and genotypes.

Additional Supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Acknowledgements

This work was supported by FCT – project PIC/IC/83252/2007 (http://alfa.fct.mctes.pt/). C. M. F. and M. S. received grants from the FCT (SFRH/BD/48103/2008; SFRH/BPD/29225/2006). Italian samples were obtained from the 'Cell Line and DNA Biobank from Patients Affected by Genetic Diseases' (G. Gaslini Institute) – Telethon Genetic Biobank Network (Project No. GTB07001A). The Brazilian Research Agency (CNPq) financed the characterization of Brazilian patients. C. G. K. received a grant from CAPES/Brazil.

References

1. Plante M, Claveau S, Lepage P et al. Mucolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTAB) in a French Canadian founder population. Clin Genet 2008: 73: 236–244.
2. Bargal R, Zeigler M, Abu-Libdeh B et al. When Mucolipidosis III meets Mucolipidosis II: GNPTA gene mutations in 24 patients. Mol Genet Metab 2006: 88: 359–363.
3. Tappino B, Chuzhanova NA, Regis S et al. Molecular characterization of 22 novel UDP-N-acetylglucosamine-1-phosphate transferase alpha- and beta-subunit (GNPTAB) gene mutations causing mucolipidosis types IIalpha/beta and IIIalpha/beta in 46 patients. Hum Mutat 2009: 30: E956–E973.
4. Encarnação M, Lacerda L, Costa R et al. Molecular analysis of the GNPTAB and GNPTG genes in 13 patients with mucolipidosis type II or type III - identification of eight novel mutations. Clin Genet 2009: 76: 76–84.
5. Kudo M, Brem MS, Canfield WM. Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha/beta -subunits precursor gene. Am J Hum Genet 2006: 78: 451–463.
6. Pinto R, Caseiro C, Lemos M et al. Prevalence of lysosomal storage diseases in Portugal. Eur J Hum Genet 2004: 12: 87–92.
7. Meikle PJ, Hopwood JJ, Clague AE, Carey WF. Prevalence of lysosomal storage disorders. JAMA 1999: 281 (3): 249–254.
8. Poorthuis BJ, Wevers RA, Kleijer WJ et al. The frequency of lysosomal storage diseases in The Netherlands. Hum Genet 1999: 105: 151–156.
9. Durst R, Colombo R, Shpitzen S et al. Recent origin and spread of a common Lithuanian mutation, G197del LDLR, causing familial hypercholesterolemia: positive selection is not always necessary to account for disease incidence among Ashkenazi Jews. Am J Hum Genet 2001: 68: 1172–1188.
10. Nei M. Molecular evolutionary genetics. New York, NY: Columbia University Press, 1987.

Coutinho et al.

11. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinforms Online 2005: 1: 47–50.
12. Martins S, Calafell F, Gaspar C et al. Asian origin for the worldwide-spread mutational event in Machado-Joseph disease. Arch Neurol 2007: 64: 1502–1508.
13. Kong A, Gudbjartsson DF, Sainz J et al. A high-resolution recombination map of the human genome. Nat Genet 2002: 31 (3): 241–247.
14. Gyapay G, Morissette J, Vignal A et al. The 1993-94 Généthon human genetic linkage map. Nat Genet 1994: 7: 246–339.
15. Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. Genetics 2000: 156: 297–304.
16. Francalacci P, Morelli L, Underhill PA et al. Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. Am J Phys Anthropol 2003: 121: 270–279.
17. Zlotogora J, Habiballa H, Odatalla A, Barges S. Changing family structure in a modernizing society: a study of marriage patterns in a single Muslim village in Israel. Am J Hum Biol 2002: 14: 680–682.

Part 2

The M6P-independent pathway

Paper 7: Molecular and computational analyses of genes involved in mannose 6-phosphate independent trafficking

Paper 8: *SCARB2* mutations as modifiers in Gaucher disease: the wrong enzyme at the wrong place

Here we present a summary of the major results obtained in the studies focused on the molecular and biochemical analyses of genes involved in the mannose 6-phosphate independent trafficking pathway. For a detailed description see papers 7 and 8.

1: Molecular analyses of undiagnosed LSD suspects

In a set of 120 patients with clinical suspicion of LSD but without definitive biochemical and/or molecular diagnosis, no novel mutations were detected either on the *SCARB2* or on the *SORT1* gene. Patients were choosen on basis of their clinical manisfestations. Preference was given to individuals whose symptoms overlapped those already reported as associated to loss of function of GCase, saposins, GM2AP and/or ASM. In theory, such phenotypes, could result not only from defects of the specific enzyme(s) but also from failures in the M6P-independent pathways through which they reach the lysosome.
Despite from our study no evidence came that *SORT1* deficiencies may be associated to LSD phenotypes, enlarged sample sizes are needed to draw more reliable conclusions on the topic.

In addition to the molecular screening, and even though no novel variants were detected, a careful *in silico* analysis of the previously known *SORT1* coding variants was also performed. Assessement of these variants' patogenical potential had never been reported and only carrier individuals have been referenced. Currently, all these variants are classified as polymorphisms but, from our computational analysis, some of them seem to harbor the potential to damage protein structure and/or function. Final conclusions may only be drawn after *in vitro* analysis or when individuals carrying such variants either in homozygosity or in compound heterozygosity are described and their phenotype evaluated.

1: Molecular screening of the *SCARB2* gene the Portuguese Gaucher Disease patients

In a set of 91 unrelated Gaucher disease (GD) cases which constitute the whole population of Portuguese patients suffering from this disorder, one novel mutation affecting the *SCARB2* gene was identified. As far as it was possible to evaluate, the clinical phenotype presented by the patient in whom this variant was identified was somehow in accordance to what has been described for the *GBA* genotype (p.L444P/ p.L444P). Nevertheless, the child has biochemical evidences suggestive of poor response to treatment, which may be justified by the additional GCase carrier partial deficiency.
In general, though, results reinforce previous evidence that mutations in the gene that codes for the β-glucocerebrosidase (GCase) transporter can act as GD modifiers. However, it became also clear that, at least in the Portuguese population, genetic variability at S*CARB2* does not account much to the GD phenotypic spectrum.

The mutation detected, p.T398M is currently under evaluation for subsequent publication. Western blot and Immunofluorescence assays are being carried out in model

cells lines transfected with a minigene carrying the mutation and on the patient's fibroblasts to further investigate protein expression and subcellular localization of LIMP-2 and GCase.

ꕥ ꕥ

As a theoretical complement to our study on genes/proteins involved in the M6P-independent pathway, we published:

- A minireview on the current knowledge on the discovery, study, structural features and cellular function of LIMP-2 and sortilin, with special attention to their role as alternative receptors to lysosomal trafficking (Appendix 1, review paper 3: Coutinho *et al.*, 2012);
- A synthesis on the published works focusing a surprising association between *SORT1* levels and risk of cardiovascular disease that has recently been unveiled (Appendix 1, review paper 4: Coutinho *et al.*, *in press*).

Paper 8

SCARB2 mutations as modifiers in Gaucher disease: the wrong enzyme at the wrong place

Coutinho MF[1,2,3], Lacerda L[4], Gaspar A[5], Pinto E[4], Ribeiro I[4], Laranjeira F[4], Ribeiro H[4], Silva E[4], Ferreira C[4], Prata MJ[2,3], Alves S[1]

[1] *Research and Development Unit, Department of Human Genetics, INSA, Portugal*

[2] *IPATIMUP, Porto, Portugal*

[3] *Department of Biology, Faculty of Sciences, Porto, Portugal*

[4] *Biochemical Genetics Unit, CGMJM, Portugal*

[5] *Unit of Metabolic Diseases, General Hospital Santa Maria, Lisbon, Portugal*

(ongoing study)

Abstract

Unlike most lysosomal proteins, beta-glucocerebrosidase (GCase) – the hydrolase defective in Gaucher disease (GD) – is specifically delivered to the lysosome through interaction with the lysosomal integral membrane protein type 2 (LIMP-2). Recently, mutations in the LIMP-2 coding gene, *SCARB2*, were reported to affect the severity of Gaucher phenotype. To understand the role of variations in *SCARB2* in the broad phenotype spectrum observed for patients carrying similar *GBA* mutations, we have screened the gene in the Portuguese GD patients. We analyzed a total of 91 individuals, having identified 3 different *SCARB2* coding variants. Of those, 2 were known polymorphic variations, with high prevalence in the normal population (p.M159V and p.V396I) and the third was a novel coding variant, p.T398M, present in heterozygozity in one Gaucher patient. When analyzed both *in silico* and *in vitro*, this variation was predicted to be deleterious for protein function.

In summary, this study demonstrates that at least in the Portuguese population genetic variability at *SCARB2* does not account much to the GD phenotypic spectrum, even though reinforcing previous evidence that *SCARB2* mutations can act as GD modifiers.

Key Words: Gaucher disease, β-Glucocerebrosidase; LIMP-2, *SCARB2*.

Introduction

Unlike most lysosomal proteins, beta-glucocerebrosidase (GCase) – the hydrolase defective in Gaucher disease (GD) – is specifically delivered to the lysosome through its interaction with the transmembrane protein LIMP-2 [Reczek *et al.*, 2007; Griffiths, 2007]. Even though being an uncommon metabolic disorder, GD is the most prevalent lysosomal storage disorder with an estimated birth frequency of 1/50,000 in the Caucasian population [Meikle *et al.*, 1999]. It is an autosomal recessive disorder caused by mutations in *GBA* - the gene encoding for GCase, with nearly 300 mutations having been described.

GD is classically divided into three variants according to the absence or presence and progressivity of neuropathic disease. All variants have differing degrees of enlargement of liver and spleen, anemia, thrombocytopenia and skeletal disease. These can range from very severe to mild within each type, although the rate of progression generally is greater in younger patients. Also, the degrees of visceral organ involvement are not concordant in patients. For example, massive involvement of the liver and spleen is not necessarily accompanied by severe bone disease. The reverse is also true. In addition, this classification is not age dependent, but depends on the primary involvement of the CNS by GD in any age. Gaucher disease type 1 patients are free of primary CNS involvement. The variability of the phenotype of visceral manifestations ranges from severe fatal disease in the first two decades to essentially asymptomatic nonagenarians. Gaucher disease types 2 and 3 have primary CNS neuronopathic involvement. Types 2 and 3 represent a continuum of disease phenotypes that differ, primarily in their rates of CNS and visceral disease progression. This continuum encompasses phenotypes leading to death *in utero*, or in the first few days of life, to rapidly progressive CNS and visceral diseases that are fatal in the first years, to more slowly progressive (yet severe) CNS (with mild to severe visceral disease) deterioration over a period of 2–3 years to decades [Frederickson and Sloan, 1978; Dreborg *et al.*, 1980; Erikson *et al.*, 1987; Sidransky *et al.*, 1992; Lewis *et al.*, 1994; reviewed in Zhao and Grabowski, 2002].

Even though some specific genotype-phenotype correlations have been established for some of its causing mutations (i.e., p.N370S only encountered in patients with the attenuated type 1 GD; and p.L444P associated with type 3 GD patients with visceral manifestations), the pathological cascade leading to the great variety of phenotypes among GD patients remains elusive, being still unsolved why patients carrying identical GBA genotypes present with disparate phenotypes. Cases of twins, including monozygotic twins homozygous for the p.N370S mutation, who exhibit striking discordance in the clinical manifestations of GD [Cox and Schofield, 1997], illustrate some of the many reports available in the literature that highlight such variability. Clearly, restricting the molecular analysis to a single locus encoding an enzyme has failed to discern a great amount of the variability associated to GD phenotypes, giving reason for the current limited ability to

make prognostic predictions from genotypic data [Hruska *et al.*, 2008]. In addition, variation in factors that influence GCase function *in vivo*, such as saposin C, solely does not seem enough to explain the divergent expression of disease [Tamargo *et al.*, 2012]. Over the years, efforts were made to identify unknown genetic, epigenetic and/or environmental factors influencing GD phenotypic manifestations [reviewed in Cox, 2001 and Zhao and Grabowski, 2002], but despite that the steps taken forward were scarce. Recently, a mutation in the gene that codes for the GCase transporter, *SCARB2*, was described as one of those modifiers [Velayati *et al.*, 2011].

In order to understand if *SCARB2* mutations are a common cause for the broad phenotype spectrum observed for patients carrying similar *GBA* mutations, we have screened the *SCARB2* gene in the Portuguese GD patients. We screened 91 individuals and, apart from numerous recurrent intronic polymorphisms, only 3 *SCARB2* coding variants were identified. Of those, 2 were known polymorphic variations, with high prevalence in the normal population (p.M159V and p.V396I) and the third was a novel coding variant, p.T398M. This alteration was present in heterozygozity in one Gaucher case. Its presence was screened in 50 control individuals and in none was it present. Bioinformatic tools (Polyphen and PANTHER) predict it to be deleterious for protein function.

In general, this study reinforces previous evidence that *SCARB2* mutations can act as modifiers of GD, even though highlighting that those are neither the only nor the most frequent cause of GD phenotype variability, at least in the Portuguese population.

Material and Methods

Patients

A total of 91 GD patients was studied. No relationship could be detected between the different families. All GD diagnoses had previously been confirmed at clinical, biochemical and molecular levels.

DNA sequencing

Genomic DNA was isolated from cultured fibroblasts and/or peripheral blood according to standard procedures. PCR amplifications of all 12 exons of the *SCARB2* gene and their corresponding intronic flanking regions were done with specific primers (primer sequences and PCR conditions available on request). Fragments were purified with ExoSap-IT (GE Healthcare, United Kingdom) according to the manufacturer's instructions and sequenced using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on an ABI PRISM 3130xl Genetic Analyser (Applied Biosystems, Foster City, CA, USA). Results were analysed with the sequence analysis software FINCH TV

(Geospiza, Seattle, WA, USA) version 1.3.1.

Plasmids, Site-directed Mutagenesis and Transfection of wild-type and mutant *SCARB2* cDNA

A pCMV6-Neo vector expressing wild-type (WT) LIMP-2 protein was purchased from Origene Inc. (Rockville, MD). Using QuikChange Site-directed Mutagenesis kit (Stratagene, USA), the missense mutation p.T398M, was introduced into the cloned wild-type *SCARB2* cDNA sequence according to the manufacturer's protocol, and the presence of additional mutations, resulting from possible enzymatic misincorporation, was excluded by full sequencing of the construct. COS-7 cells were transfected with wild-type and mutant constructs using Lipofectamine 2000 Reagent (Invitrogen, USA) following Invitrogen guidelines.

Western blot Analysis

COS-7 cells were grown in 6-well plates and 24 h after transfection harvested and lysed for 1 hour at 4°C in 30 μL of RIPA buffer (50mM Tris-HCl pH 8.8, 0.1 % NP-40, 150mM NaCl, 2mM EDTA pH 7.5, 0.1 % SDS and 0.5 % Na-DOC). Proteins were quantified through a modified Lowry assay using Bio-Rad DC Protein Assay (Bio-Rad, USA), according to the manufacturer's instructions. Protein samples (75 μg) were separated in 8% SDS-PAGE gel and electroblotted to Hybond-C Extra nitrocellulose membranes (Amersham Biosciences, UK). The membranes were then incubated with the following primary antibodies: anti-LIMP-2 (1:3000, OriGen Inc., Rockville, MD) and mouse monoclonal anti-α-Tubulin antibody (1:10000, Sigma-Aldrich, USA). The secondary antibodies used were goat anti-mouse IgG (H+L) HRP conjugate (1:3000, Invitrogen, USA) and rat anti-mouse monoclonal (1:12000, Sigma-Aldrich, USA). The signal was developed with the ECL Western blotting Analysis Systems (GE Healthcare, UK) and detected on ECL Hyperfilms (GE Healthcare, UK), using the Fujifilm system (Fujifilm, Japan).

Bioinformatic Analysis

To evaluate the potential effect of the reported *SCARB2* coding variants, three bioinformatic tools were used: Sorting Intolerant From Tolerant (SIFT; http://blocks.fhcrc.org/sift/SIFT.html), PANTHER (http://www.pantherdb.org/), PolyPhen (http://coot.embl.de/PolyPhen/), and PROVEAN (http://provean.jcvi.org/index.php).

Results and Discussion

From the total sample of 91 GD patients sequenced for the *SCARB2* gene, only 3 *SCARB2* coding variants were detected, apart from numerous recurrent intronic polymorphisms. Of those, 2 were known polymorphic variations, with high prevalence in the normal population (p.M159V and p.V396I) and the third was a novel coding variant, p.T398M. This alteration was present in heterozygozity in one GD case. Its presence was screened in 50 control individuals and in none was it present. All bioinformatic tools used (Polyphen, PANTHER and PROVEAN) predict it to be deleterious for protein function.

To check if the p.T398M could be acting as a GD modifier, we started by carefully reviewing the clinical history of the patient in which the *SCARB2* variant was detected. Patient was the second child of young, healthy unrelated parents of Cape Verdean origin. The couple had two other children, both apparently healthy: a 6 year-old boy of and 1-year-old girl. The boy was born of normal delivery at 40 weeks, after a monitored pregnancy with no registered complications. At seven months, the child developed a multi-symptomatic clinical presentation of abdominal distension, diarrhea, anorexia, feeding difficulties, rough coughing and progressive weight loss, remaining hospitalized for two months in Cape Verde. During that period progressive degradation of his general condition occurred, with repeated infections and significant weight loss. The boy developed severe anemia, thrombocytopenia, cardiomegaly and marked splenomegaly. His aggravated condition led to the child's evacuation to Portugal for etiological clarification. He was received at the Hospital de Santa Maria (Lisbon) and reported mild hepatomegaly (3cm below the edge of the right ribs); exuberant splenomegaly (until the iliac crest); poor weight gain with cachexia; feeding difficulties and dysphagia; bilateral convergent strabismus; marked axial hypotonia; poor facial mimic; stridor; global psychomotor developmental delay; cardiomegaly with dilatation of the left cavities; interstitial lung disease with multiple recurrent infections including aspiration pneumonia. Three cardiac arrest events were also registered. Cerebral MRI showed supratentorial periventricular white matter alterations of tegmentum pontis and dentate nucleus, all aspects which are compatible with central nervous system (CNS) involvement in the context of GD. Initial investigations showed low β-glucocerebrosidase levels and the child was referenced for ERT, having started treatment. Nevertheless, subsequent therapeutic follow-up with assessment of chitotriosidase levels showed enzyme levels which were disparate from the ones expected for a GD patient under ERT treatment, suggesting lower response to treatment.

Recombinant GCase uptake is known to be dependent from the receptor density (mannose-6-phosphate receptors and LIMP-2). That is, indeed, one the major reason for the variable organ response to ERT [Desnick and Schuchman, 2012]. Taking this into account, it would be expectable that any alteration causing either dysfunction or reduction of

LIMP-2 lead to a decrease in the efficacy of ERT in GD patients, as observed in our case. Preliminary results of Western blot assays in COS7 cells transfected with a minigene carrying the mutation show a decrease of LIMP-2 levels when compared to wild-type protein (*data not shown*). Nevertheless, additional experiments are still needed to further evaluate the effect of p. T398M on GCase transport. A tempting approach would be to carry on Western blot and Immunofluorescence microscopy analysis on the patient's fibroblasts to investigate protein expression and subcellular localization of both LIMP-2 and GCase. Additional assays to measure activity and extra-cellular secretion of GCase on the patient cell line are also predicted.

Conclusion

To the best of our knowledge this is the first time that a whole GD population is screened for mutations in this gene. From our results on the Portuguese population it was possible to conclude that *SCARB2* mutations are neither the only nor the most frequent cause of GD phenotype variability. Nevertheless, with the identification of a novel mutation in one of the patients who present a severe GD phenotype and had a poor response to ERT standard treatment and its subsequent evaluation, our study reinforces previous evidence that *SCARB2* mutations do act as modifiers of GD.

Acknowledgments

This work was supported by FCT - project PTDC/SAU-GMG/102889/2008. Coutinho MF is a grantee from the FCT (SFRH/BD/48103/2008).

References

Cox TM. Gaucher disease: understanding the molecular pathogenesis of sphingolipidoses (2001) *J Inherit Metab Dis*. 24 Suppl 2:106-21.

Cox TM. Gaucher disease: understanding the molecular pathogenesis of sphingolipidoses (2001) *J Inherit Metab Dis*. 24:106-21.

Desnick RJ, Schuchman EH. Enzyme replacement therapy for lysosomal diseases: lessons from 20 years of experience and remaining challenges (2012) *Annu Rev Genomics Hum Genet*. 13:307-35.

Dreborg S, Erikson A, Hagberg B. Gaucher disease-Norrbottnian type. I. General clinical description (1980) *Eur J Pediatr*. 133:107-18.

Erikson A, Karlberg J, Skogman AL, *et al*. Gaucher disease (type III): intellectual profile (1987) *Pediatr Neurol*. 3:87-91.

Frederickson DS, Sloan HR. Glucosyl ceramide lipidoses: Gaucher's disease. In: Stanbury JB, Wyngaarden JB, Frederickson DS (eds) The metabolic basis of inherited disease. (1978) McGraw-Hill, New York, pp 730-759.

Griffiths GM. Gaucher disease: forging a new path to the lysosome (2007) *Cell*. 131:647-9.

Hruska KS, LaMarca ME, Scott CR, *et al*. Gaucher disease: mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA) (2008) *Hum Mutat*. 29:567-83.

Lewis BD, Nelson PV, Robertson EF, *et al*. Mutation analysis of 28 Gaucher disease patients: the Australasian experience (1994) *Am J Med Genet*. 49:218-23.

Meikle PJ, Hopwood JJ, Clague AE, *et al*. Prevalence of lysosomal storage disorders (1999) *JAMA*. 281:249-54.

Reczek D, Schwake M, Schröder J, *et al*. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase (2007) *Cell*. 131:770-83.

Sidransky E, Sherer DM, Ginns EI. Gaucher disease in the neonate: a distinct Gaucher phenotype is analogous to a mouse model created by targeted disruption of the glucocerebrosidase gene (1992) *Pediatr Res*. 32:494-8.

Tamargo RJ, Velayati A, Goldin E, *et al*. The role of saposin C in Gaucher disease (2012) *Mol Genet Metab*. 106:257-63.

Velayati A, DePaolo J, Gupta N, *et al*. A mutation in SCARB2 is a modifier in Gaucher disease (2011) *Hum Mutat*. 32:1232-8.

Zhao H, Grabowski GA. Gaucher disease: Perspectives on a prototype lysosomal disease (2002) *Cell Mol Life Sci*. 59:694-707.

Paper 7

Molecular and computational analyses of genes involved in mannose 6-phosphate independent trafficking

Coutinho MF[1,2,3], Lacerda L[4], Pinto E[4], Ribeiro H[4], Macedo-Ribeiro S[5], Prata MJ[2,3], Alves S[1]

[1] *Research and Development Unit, Department of Human Genetics, INSA, Portugal*

[2] *IPATIMUP, Porto, Portugal*

[3] *Department of Biology, Faculty of Sciences, Porto, Portugal*

[4] *Biochemical Genetics Unit, CGMJM, Portugal*

[5] *IBMC, Porto, Portugal*

(under preparation)

Abstract

Lysosomes are essential regulators of cell homeostasis, since they harbour a vast repertory of specialized enzymes responsible for degrading endocytosed and intracellular material. The newly-synthesized lysosomal enzymes travel first to the *trans*-Golgi network (TGN) and then must be driven to the acidic organelle in order to exert their function. Whist the best-known pathway for TGN-to-endosome transport is the delivery of soluble lysosomal hydrolases by the MPRs, additional pathways from the TGN to lysosomes do exist, as recently demonstrated by the identification of two alternative receptors: LIMP-2, shown to be implicated in the delivery of β-glucocerebrosidase to the lysosomes, and sortilin, proposed to be involved in the transport of several proteins including the sphingolipid activator proteins prosaposin and GM2AP, acid sphingomyelinase and cathepsins D and H.

Disruption of the intracellular transport and delivery pathways to the lysosomes may result in lysosomal dysfunction, predictably leading to a range of clinical manifestations suggestive of lysosomal storage diseases (LSDs). However, for a great percentage of patients presenting such manifestations, no condition is successfully diagnosed because their metabolic profile does not fit any known LSD. In order to analyse if, in this group of patients, the underlying causes of their phenotypes could be genetically determined by impairments in the two known M6P-independent receptors, we screened the genes *SCARB2* and *SORT1* that encode for, respectively, LIMP-2 and sortilin. No novel mutations were found on the *SCARB2* gene and no pathogenic mutations were identified on the *SORT1* gene. Other study approaches will be needed to clarify whether sortilin dysfunction may cause disease.

Key Words: M6P independent trafficking, LIMP-2, sortilin, lysosomal storage diseases

Introduction

Lysosomes are responsible for the degradation of different substrates in the cell. This major function can only be fulfilled if a number of specific degradative enzymes, synthesized in the ER, is properly transported from the Golgi to the lumen of the acidic organelle. The majority of those soluble acid hydrolases is modified with mannose 6-phosphate (M6P) residues that allow their recognition by specific M6P receptors in the Golgi complex, ensuring their transport to the endosomal/lysosomal system [Braukle and Bonifacino, 2008]. Even though M6P receptors play a major role in this intracellular transport, compelling evidence has emerged that alternative methods of lysosomal targeting do exist. Recently, two M6P alternative trafficking receptors were brought to light: lysosomal integral membrane protein LIMP-2 (encoded by the *SCARB2* gene) and sortilin (encoded by the *SORT1* gene). LIMP-2, a protein of the lysosomal membrane, was shown to be implicated in the delivery of β-glucocerebrosidase (GCase, the enzyme defective in patients with Gaucher disease) to the lysosomes [Reczek *et al.*, 2007; Berkovic *et al.*, 2008; Balreira *et al.*, 2008]. Sortilin, is a multiligand receptor thought to be involved in the lysosomal trafficking of sphingolipid activator proteins (SAPs), prosaposin (PSAP) and GM2 activator protein (GM2AP), acid sphingomyelinase (ASM) and cathepsins D and H [Ni *et al.*, 2006; Nielsen *et al.*, 1999; Lefrancois *et al.*, 2003; Canuel *et al.*, 2008].

Impairments of the M6P-dependent pathway are the basis of the rare genetic diseases Mucolipidosis II and III (ML II, OMIM# 252500 and ML III, OMIM# 252600), two conditions that currently are most of the times successfully diagnosed after routine screening tests. Similarly, impairments on the M6P-independent trafficking processes might also have hazardous effects to the organism, ultimately leading to overt disease. Actually, recent studies demonstrated that mutations in *SCARB2* may underlie a serious autosomal-recessive disorder presently known as action myoclonus-renal failure syndrome (AMRF) [Reczek *et al.*, 2007; Berkovic *et al.*, 2008; Balreira *et al.*, 2008]. So, having in mind that mutations in the genes coding for LIMP-2 and sortilin could result in missorting or mistargeting of some the above-referred proteins, we sought to perform the molecular screening of these two genes targeting five specific groups of individuals presenting clinical manifestations frequently associated to lysosomal storage diseases (LSDs) but still with undiagnosed disorders. The loss of activity of specific lysosomal hydrolases or, more rarely, the defective function of non-enzymatic lysosomal proteins is the hallmark of LSDs, a large group of rare genetic diseases. LSDs have been recognised as one of the major groups of genetic disorders affecting children and adults, including over 50 different disorders with a combined prevalence of around 1:4000-1:9000 live births. LSDs represent a major public health problem aside from placing an enormous burden on the individuals and families affected [Meikle *et al.*, 2006].

Not rarely, however, in patients with clinical suspicion of LDS, the definitive diagnosis

cannot be achieved with enzymatic or molecular testing. We concentrated in the uncharacterised patients, recruiting for the present study those with phenotypes overlapping manifestations that could predictably be due to loss of activities of GCase, saposins, GM2AP and/or ASM, to find out whether those features could be resultant from failures in M6P-independent pathways through which those proteins reach the lysosome. Patients were sorted into five different groups according to their phenotypic picture: (1) local glomerulosclerosis myoclonus epilepsy; (2) metachromatic leukodystrophy/ Gaucher-like phenotype; (3) GM2 gangliosidosis-like phenotype; (4) Niemann-Pick types A/B-like and (5) hydropsis fetalis. In each group, the option to study sorting or trafficking of lysosomal proteins through M6P-independent pathways was made according to the suggestive clinical manifestations: groups (1) and (5) were screened for *SCARB2*, groups (3) and (4) for *SORT1* and group (2) was screened for both genes.

Material and Methods

Patients

We started by reviewing the clinical histories of patients with samples still available in our institution who had already undergone diagnostic testing for LSDs without having obtained any positive result. Patients were spitted in 5 groups based on their particular clinical features. Group (1) was constituted by patients presenting clinical manifestations characteristic of individuals which suffer from AMRF (glomerulosclerosis and/or with glomerular collapse). This group, together with group (2), was screened for the *SCARB2* gene. Groups (2), (3) and (4), screened for the *SORT1* gene, were classified according to the following characteristics: (2) metachromatic leukodystrophy and/or Gaucher-like phenotype (acute neuronopathic form of the disorder with onset in infancy and premature death; additional features include hepatosplenomegaly, developmental regression, and growth arrest); (3) GM2 gangliosidosis-like phenotype (developmental retardation, followed by paralysis, dementia and blindness, with death in the first years of life; hallmark features include “cherry red spot”, balloon-shaped neurons in the central nervous system early and a persistent extension response to sound known as 'startle reaction') and (4) Niemann-Pick types A/B-like phenotype (ranging from severe and premature neurologic degeneration resulting in early death to later-onset nonneurologic manifestations affecting the spleen, the liver, and the lungs). Finally, group (5), also screened for the *SCARB2* gene, comprised all available samples of hydropsis fetalis cases. Whenever clinical descriptions were missing, we focused our attention on the initial physician requests referring the patients to our laboratory. So, in group (2) patients of unknown symptoms but whose initial physicians’ request asked for

Gaucher disease testing but whose final result was negative, were also included. Similarly, patients that had been tested for GM2 Gangliosidosis and Niemann-Pick disease but whose suspicions were not confirmed were additionally integrated in groups (3) and (4), respectively. The total sample comprehended 120 individuals.

Molecular screening of the *SCARB2* and *SORT1* genes

Genomic DNA was isolated from cultured fibroblasts and/or peripheral blood according to standard procedures. PCR amplifications of all 12 exons of the *SCARB2* gene and 21 exons of the *SORT1* gene and their corresponding intronic flanking regions were done with specific primers (primer sequences and PCR conditions available on request). Fragments were purified with ExoSap-IT (GE Healthcare, United Kingdom) according to the manufacturer's instructions and sequenced using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on an ABI PRISM 3130xl Genetic Analyser (Applied Biosystems, Foster City, CA, USA). Results were analysed with the sequence analysis software FINCH TV (Geospiza, Seattle, WA, USA) version 1.3.1.

In silico analysis

To evaluate the potential effect of the reported sortilin coding variants, three bioinformatic tools were used: Sorting Intolerant From Tolerant (SIFT; http://blocks.fhcrc.org/sift/SIFT.html), PANTHER (http://www.pantherdb.org/), PolyPhen (http://coot.embl.de/PolyPhen/), and PROVEAN (http://provean.jcvi.org/index.php).
Modelling of sortilin tertiary structure was performed using the structure of the Vps10p domain in complex with neurotensin (Protein Data Bank file 3F6K.pdb) as template. Modelling was carried out by Swiss Model Server and *in silico* mutagenesis visualization was performed with PyMOL (http://pymol.sourceforge.net). For LIMP-2, no tertiary model could be designed since there are no known homologous proteins whose structure has already been crystallized. Predictions of the presence of exonic splice enhancer or silencer sequences were made using ESEfinder (http://rulai.cshl.edu/cgi-bin/tools/ESE3) [Cartegni *et al.*, 2002] and SplicingRainbow (http://www.ebi.ac.uk/asd-srv/wb.cgi?method=8) [Stamm *et al.*, 2006]. Rescue-ESE (http://genes.mit.edu/burgelab/rescue-ese/) [Fairbrother *et al.*, 2002] and PESXs-Search (http://cubweb.biology.columbia.edu/pesx/) [Zhang and Chasin, 2004] were used to analyze the binding of several splicing factors.

Results

In total, 67 individuals were screened for the *SCARB2* gene. Among them, only non-pathogenic and frequent variations were detected: eight were already registered in

Ensembl (rs3733259; rs894248; rs2289512; rs143655258; rs2289513; rs35369082; rs77814624; rs35583533) while three were not registered. Two of the identified polymorphisms affected the coding region of the *SCARB2* gene: p.M159V (rs143655258) and p.V391I (rs77814624). Both were already registered in Ensembl as polymorphisms. In addition, all bioinformatic tools (PolyPhen, SIFT, PANTHER and PROVEAN) predicted them to be tolerated. These two variant alleles occurred at frequency of ≈2% and 3% in our patient sample and at 1% and 2% of our control sample (50 healthy Portuguese individuals), respectively. In the 1000 Genome project, frequencies of the two variations were 1% and 0%, respectively. Out of the three not registered alterations, two were intronic and relatively common in our samples. The third was exonic, even though silent (c.567T>C; p.H189H), being present in heterozygosity in only one of the patients. To understand if it could have a deleterious effect by altering any exonic splicing enhancer (ESE) or exonic splicing suppressor (ESS), which consequently could interfere with splicing, we used four bioinformatic tools, ESEfinder, SplicingRainbow, Rescue-ESE and PESXs-Search, with all but PESXs-Search yielding identical predictions for the normal and the altered sequence. However, since according to PESXs-Search the substitution 568T>C would give rise to the loss of a PESS motif ("ATGTTTTC"), we screened the patient's cDNA with specific primers but no differences were observed between control cDNAs and the patient's sample (data not shown). As no other pathogenic mutation was found in that patient, such silent heterozygous variant can very likely be excluded as the cause of the phenotype.

Concerning the *SORT1* gene, we screened 53 individuals and identified seven different variations, all of them non-pathogenic: five already registered in Ensembl as polymorphisms (rs11181665; rs11142; rs72646560; rs2228604; rs74584797) and two that hadn't yet been stored in the same database. Three of those variants occur in the coding region of the *SORT1* gene (rs11142; rs72646560 and rs2228604) but none of them causes aminoacidic changes. They are also frequent both on the patient sample (≈11, 12 and 17%, respectively) and on the general population (≈15, 6 and 30% of European alleles, respectively according to data from the 1000 Genome project on allele frequencies).

Discussion

LIMP-2 is a heavily *N*-glycoylated type III transmembrane protein [Fujita *et al.*, 1991], representing the sorting receptor of GCase that is the enzyme defective in Gaucher Disease, the most common lysosomal storage disorder [Beutler, 1991]. Mutations in LIMP-2 were recently reported to cause AMRF, a disease characterized by a remarkable combination of local glomerulosclerosis (frequently with glomerular collapse) and a series

of symptoms which are often associated to LSD, such as the presence of storage material in the brain [Berkovic *et al.*, 2008; Balreira *et al.*, 2008]. We admitted, therefore, that some of the patients presenting such symptoms but lacking positive results in previously performed diagnostic tests for LSDs, could carry mutations in *SCARB2*, the gene encoding LIMP-2.

Sortilin's role on disease, on the other hand, still remains elusive. To evaluate if some LSD-related symptoms could arise due to sortilin dysfunction, we focused attention on its putative targets, all of them lysosomal proteins of known biological relevance: SAPs (prosaposin and GM2 activator protein), ASM and cathepsins D and H [Ni *et al.*, 2006; Nielsen *et al.*, 1999; Lefrancois *et al.*, 2003; Canuel *et al.*, 2008]. Prosaposin (PSAP) is both a non-enzymatic precursor of sphingolipid activator proteins (SAPs) and a secreted neurotrophic and myelotrophic factor [Jolivalt, 2008]. This molecule is a precursor protein common to four homologous glycoproteins: saposins A, B, C and D. Such glycoproteins are required for the *in vivo* degradation of sphingolipids with short carbohydrate chains and, even though presenting remarkable structural similarities, their functions are specific, with some overlaps for individual sphingolipid hydrolases. Total deficiency of all saposins (OMIM#611721) and specific deficiencies of saposin A (OMIM#611722), B (OMIM#249900) or C (OMIM# 610539) were already described in humans [Matsuda *et al.*, 2001; Tamargo *et al.*, 2012]. Patients suffering from individual saposin deficiencies present atypical Krabbe disease, metachromatic leukodystrophy and Gaucher-like phenotypes, respectively [Christomanou *et al.*, 1986; Stevens *et al.*, 1981]. Combined saposin deficiency is a devastating disease with a complex multisystemic phenotype, demonstrating the essential role of the four saposins *in vivo* [Matsuda *et al.*, 2001]. It seemed, though, quite tempting to explore whether a failure in the lysosomal targeting of prosaposin could have deleterious consequences if leading to the absence of some saposin activites. So, we screened the *SORT1* and *SCARB2* genes in patients presenting a phenotypic picture coinciding with that characteristic of Gaucher disease but without mutations in both the *GBA* gene (which codes for glucocerebrosidase) and in the *PSAP* (which codes for prosaposin). We also screened the *SORT1* gene in unclassified patients with metachromatic leukodystrophy.

GM2 ganglioside activator protein (GM2AP) is an essential cofactor for the lysosomal degradation of ganglioside GM2 by β-hexosaminidase A (hex A). Functional deficiencies in this protein result in a fatal neurological storage disorder, the AB variant of GM2 gangliosidosis (OMIM #272750) [Wendeler *et al.*, 2006]. Once again, it appeared reasonable to admit that mutations in *SORT1* could cause a similar disease, by preventing the correct targeting of GM2AP to the lysosome and, consequently, its activator function.

Two other lysosomal proteins had still to be considered if we wanted to address all possible metabolic dysfunctions that could result from an impaired lysosomal targeting mediated by sortilin: cathepsins D and H [Canuel *et al.*, 2008]. Cathepsin D is a peptidase

belonging to the family of aspartic peptidases. Although it may assume some physiological effects such as hormone, antigen and mediator of neuropeptides processing, its major function is the digestion of proteins and peptides within the acidic compartment of the lysosome [Fusek *et al.*, 2005]. Cathepsin H, on the other hand, is a lysosomal glycoprotein and a member of the cysteine proteinase family, which may be easily distinguished from other members of this family by its unique aminopeptidase activity [Rojnik *et al.*, 2012]. Both cathepsin D and H seem to be cell biomarkers of specific cancers with distinct roles in pathological conditions [Mimae *et al.*, 2012; Schweiger *et al.*, 2004]. It is also well established that the proper distribution of cathepsins, not only cathepsins D and H but several other members of the family, is essential for efficient lysosomal and cellular function [Urbanelli *et al.*, 2008; Schweiger *et al.*, 2004]. Nevertheless, none of these enzymes has ever been related to any type of storage disorder and so, there are no known effects of their failure or absence of any of them in the cell.

In our patients sample, however, no genetic variants in *SORT1* were detected that could explain the observed phenotypes. This does not deny, *per se*, the possibility that impairments on sortilin function may cause disease, especially if we take in account, on the one hand, the small size of the tested sample and the low prevalence of LSDs in general, and on the other, the fact that the establishment of a phenotypic criterion to screen for mutations in any gene involved in lysosomal function is far from being straightforward. Actually, even though some phenotypic effects of protein loss/malfunction may be easily predicted (such being the case of deficiencies or dysfunctions of saposins, which result in LSDs that mimic deficiencies of the enzymes activated by each particular Sap), others exist in which the correlation between the missing protein and its effect at clinical level is not so obvious (such being the case of LIMP-2 deficiencies, which result in a phenotype different from the one caused by the dysfunction of GCase, the enzyme it transports).

In a complementary attempt to evaluate the possibility that *SORT1* mutations cause disease, we performed *in silico* analysis of the 12 non-synonymous coding variants that have been described in the scope of the "*1000GenomeProject*" and were classified as polymorphisms, although none of them was ever detected in homozygouzity. In order to evaluate how genetic variants could affect protein function and, ultimately, cause disease we created a 3D-model for sortilin and performed *in silico* mutagenesis. Then, 10 of the registered variants were tested on our 3D-model, cross-linking the results with PolyPhen and SIFT predictions, accounting for the size, cargo and structure of the altered residues. By doing so, we classified them into: (1) tolerated, (2) possibly and (3) probably damaging. The complete results of this *in silico* analysis are represented in Figure 1.

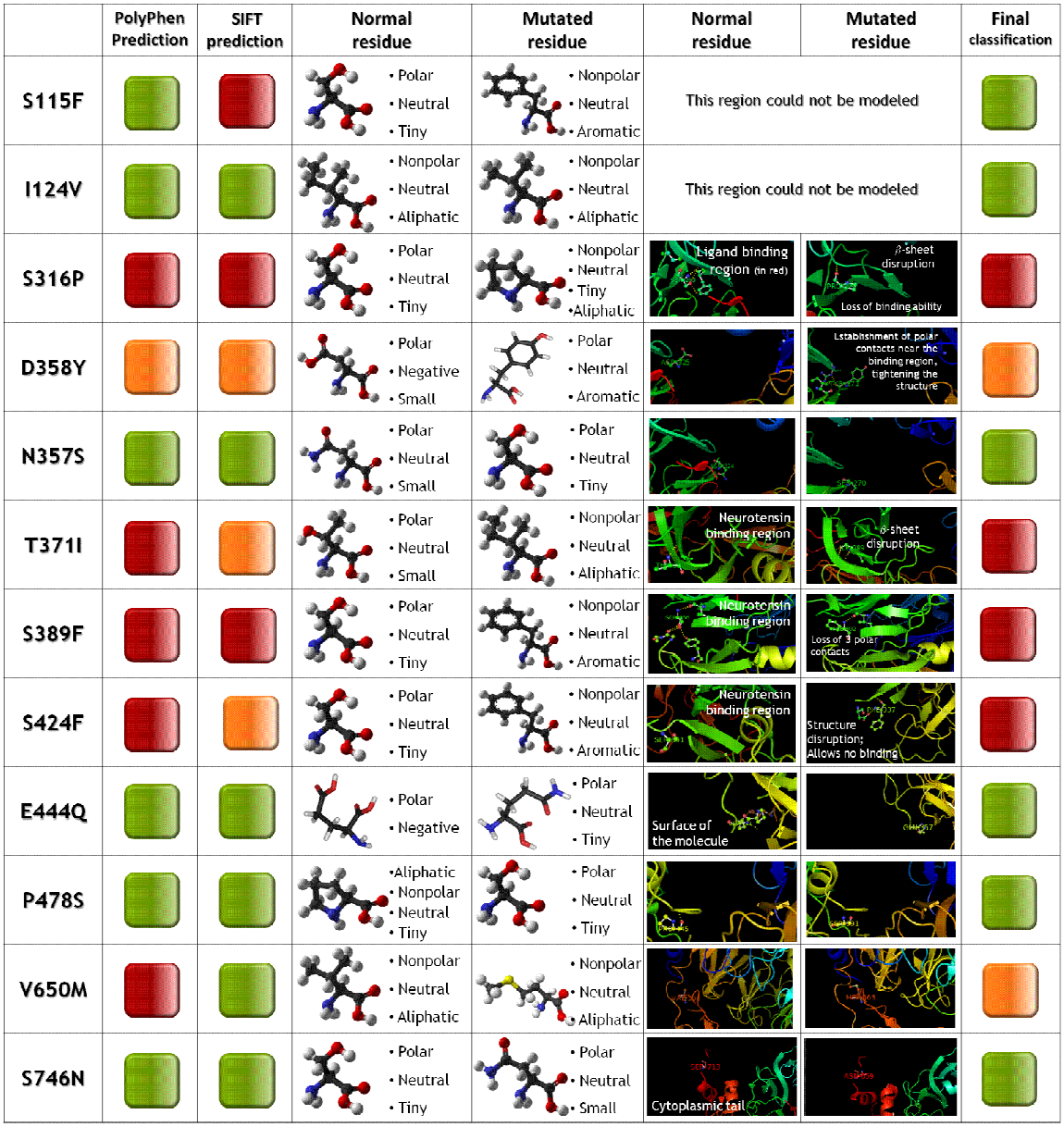


Figure 1: Complete results of the *in silico* analysis of *SORT1* coding variants.

A color code highlights the potential of each one to disrupt sortilin's structure, impair its biological function and, ultimately, cause lysosomal storage and disease. *Benign* - green; *possibly deleterious* - orange; *probably deleterious* - red.

Several parameters were taken in consideration: (a) conservation degree of the altered residue; (b) likelihood of the affected amino acid to change the secondary structure; (c) localization of the specific amino acid in the different protein domains, which together turned possible to predict which of the investigated 12 coding variants most likely affect the structure and, consequently, function of sortilin. Obviously, only additional *in vitro* and *in vivo* studies may confirm if the envisioned impairments will result indeed in disease.

Given that two of the variants, p.S115F and p.I124V, are located in a tip region of the

molecule that could not be modelled, for them only the bioinformatic predictions and interspecific alignments were considered to establish the final classification of severity/pathogenicity. All other variants were modelled and analysed with all available tools. Variants p.I124V, p.N357S, p.E444Q, p.P478S and p.S746N were predicted to be tolerated/benign by all the bioinformatic tools (data not shown for PANTHER and PROVEAN). In general, these variants involve aminoacid changes that do not affect regions of the molecule which are highly conserved among species. Furthermore, there are no significant differences between the properties of the "normal" and the "mutated" residues (similar polarities, sizes and/or acidity).

For p.S115F and p.V650M, discrepant results were yielded by PolyPhen and SIFT (and also PHANTER and PROVEAN) with half of them predicting them to be pathogenic and the other half to be tolerated. Concerning p.V650M, the similar properties of the involved aminoacids (See Figure 1, "normal" and "mutated" residues), could justify why a methionine in the position would be tolerated. Nevertheless, V^{650} is a highly conserved residue, indicating therefore that it may be crucial to guarantee protein functionality, even though no relevant sequence annotations were made for the loop region in the molecule where the residue is located. The situation is opposite for the missense alteration p.S115F, which despite implying a considerable alteration in aminoacid properties it changes S^{115} that is not particularly conserved among species. So, further approaches are still needed to illuminate the consequences of both variants on protein function. For the remaining variants, assessment of putative pathogenicity was less problematic, given the coincident or similar bioinformatic predictions. p.S316P, p.T371I, p.S389F and p.S424F, for example, are substitutions in which the properties of the involved aminoacids change drastically, not only in terms of polarity (polar $\rightarrow$ nonpolar) but also in size and chain properties. Whilst theoretically, any such substitutions is prone to disrupt protein structure and, consequently, its function, evidence that this might occur comes from the observation that the involved residue is evolutionary extremely conserved, as happens with p.S424F. The pathogenic nature of variants p.T371I, p.S389F and p.S424F can be further explained since they affect residues located on the surroundings of the neurotensin binding region, which is crucial for sortilin's function on the CNS. In the case of variant p.D358Y, it was classified as possibly deleterious, even if not implying a dramatic change in aminoacids properties, because a tyrosine at the highly conserved position 358 is capable of establishing new polar contacts which tighten the structure near the neurotensin binding site, potentially affecting binding.

A strong argument against the possibility that a lysosomal storage disorder could arise due to sortilin's loss of function came from studies on a sortilin knockout (KO) mouse that did not display any clinically relevant phenotype [Zeng *et al.*, 2009], although revealing a 80% reduction of levels of intraendosomal/lysosomal prosaposin, which thereby clearly

supported the role of sortilin as an alternative receptor. Actually, the results obtained for the sortilin KO mouse may be easily conciliated with the ones reported for mice deficient on the M6P receptors. In fact, when KO mice were created for the cation-dependent M6P-receptor (CD-MRP) and for the cation-independent M6P receptor (CI-MRP), none of them presented clinical phenotype, even though both had a partial missorting of phosphorylated lysosomal enzymes. Only when a double KO was created did the mouse present pathological changes, indicating that, when one of the receptors is not functional, the other partially replaces it. So, presuming that other receptors besides sortilin do exist able to transport lysosomal hydrolases, it is possible that, when sortilin is not functional, such receptors assume a reinforced role to avoid mistargeting and cellular malfunction. Sortilin, as well as most of the members of the Vps10p family of sorting proteins, appears to be evolutionary much older than the MPR and taking into account, as experimental results suggest [Marcusson *et al.*, 1994; Hampe *et al.*, 2001; Petersen *et al.*, 1997; Tauris *et al.*, 1998], that MPRs have the capacity of replacing one another when any is deficient, it seems even more probable that sortilin and its related proteins that share the same domains already present on yeast, do perform vacuolar transport. On the frontline of the most probable alternative receptors stands SorLA, the unique other known Vps10p family member containing motifs shown to be involved in trafficking from the Golgi to the endosome and *vice versa*. Although no studies have yet been conducted to evaluate its function in the M6P-independent transport of lysosomal hydrolases, the observation that KO mice in which the gene for SorLA was inactivated did not exhibit abnormal phenotypes, also pinpoint to the existence of a complementary receptor [Zhu *et al.*, 2004]. One final remark should also be made on animal models. There are differences in gene regulation between man and mouse respecting to sortilin, as demonstrated by the absence of a crucial C/EBP α binding site in mice [Musunuru *et al.*, 2010; Russell and Proctor, 2006], which inevitably may hinder extrapolation of mouse studies to human when dealing with sortilin.

In summary, unlike from what happens concerning LIMP-2, the role that sortilin plays on disease is still far from being unveiled. In general, there is still much to learn not only on sortilin but also on other Vps10p family receptors and their role on the M6P-independent trafficking pathway. Some steps forward would be taken by the generation of a double knockout for sortilin and SorLA. In addition, important insights on sortilin's role on disease, will probably come from finding out rare human patients harboring a large-effect mutation in *SORT1* (e.g., a loss of function mutation) and from a careful definition of their phenotype [Dubé *et al.*, 2011].

Conclusion

In this study we raised the question about whether some cases of patients suspected to have a LSD but remaining undiagnosed after routine tests could be due to genetic impairments in two recently discovered M6P-independent receptors LIMP-2 and sortilin. To address the issue we screened the encoding genes in subsamples of patients defined according to their phenotypic characteristics. No novel mutations were found on the *SCARB2* gene and no pathogenic mutations were identified on the *SORT1* gene. Other approaches will be needed to clarify whether sortilin dysfunction may cause disease.

The relevance of alternative receptors is demonstrated by their involvement in disease. *SCARB2* mutations were already shown to underlie the serious autosomal-recessive disorder AMRF [Berkovic *et al.*, 2008]. Recently, *SCARB2* mutations have also been demonstrated to act as modifiers in Gaucher disease (GD), contributing to a severe neuropathy even when present in heterozygous state [Velayati *et al.*, 2011]. So, a better understanding on the M6P-independent routes to the lysosome will contribute to a more accurate understanding not only of crucial cellular processes, but also of the pathophysiological bases of severe and disabling diseases. Finally, knowledge of the different trafficking mechanisms responsible for the sorting of lysosomal proteins may be an enormous help for the building of new therapeutic strategies for LSDs.

Acknowledgments

This work was supported by FCT - project PTDC/SAU-GMG/102889/2008 (http://alfa.fct.mctes.pt/). Coutinho MF is a grantee from the FCT (SFRH/BD/48103/2008).

References

Balreira A, Gaspar P, Caiola D *et al*. A nonsense mutation in the LIMP-2 gene associated with progressive myoclonic epilepsy and nephrotic syndrome (2008) *Hum Mol Genet*. 17:2238-43.

Berkovic SF, Dibbens LM, Oshlack A, *et al*. Array-based gene discovery with three unrelated subjects shows SCARB2/LIMP-2 deficiency causes myoclonus epilepsy and glomerulosclerosis (2008) *Am J Hum Genet*. 82:673-84.

Beutler E. Gaucher´s disease (1991) *N Engl J Med*. 325:1354-60.

Braukle T, Bonifacino JS, Sorting of lysosomal proteins (2008) *Biochim. Biophys, Acta*. 1793:605-614.

Canuel M, Korkidakis A, Konnyu K, *et al*. Sortilin mediates the lysosomal targeting of cathepsins D and H (2008) *Biochem Biophys Res Commun*. 373:292-7.

Canuel M, Libin Y, Morales CR, The interactomics of sortilin: an ancient lysosomal receptor evolving new functions (2009) *Histol Histopathol*. 24:481-92.

Christomanou H, Aignesberger A, Linke RP. Immunochemical characterization of two activator proteins stimulating enzymic sphingomyelin degradation in vitro. Absence of one of them in a human Gaucher disease variant (1986) *Biol Chem Hoppe Seyler*. 367:879-90.

Dubé JB, Johansen CT, Hegele RA. Sortilin: an unusual suspect in cholesterol metabolism: from GWAS identification to in vivo biochemical analyses, sortilin has been identified as a novel mediator of human lipoprotein metabolism (2011) *Bioessays*. 33:430-7.

Fairbrother WG, Yeh RF, Sharp PA, *et al*. Predictive identification of exonic splicing enhancers in human genes (2002) *Science*. 297:1007-13.

Fujita H, Tanaka Y, Noguchi Y, *et al*. Isolation and sequencing of a cDNA clone encoding rat liver lysosomal cathepsin D and the structure of three forms of mature enzymes (1991) *Biochem Biophys Res Commun*. 179:190-6.

Fusek M, Vetvicka V. Dual role of cathepsin D: ligand and protease (2005) *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 149:43-50.

Hampe W, Rezgaoui M, Hermans-Borgmeyer I *et al*. The genes for the human vps10 domain-containing receptors are large and contain many small exons (2001) *Hum Genet*. 108:529-536.

Jolivalt CG, Vu Y, Mizisin LM, *et al*. Impaired prosaposin secretion during nerve regeneration in diabetic rats and protection of nerve regeneration by a prosaposin-derived peptide (2008) *J Neuropathol Exp Neurol*. 67:702-10.

Lefrancois S, Zeng J, Hassan AJ, *et al*. The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin (2003) *EMBO J*. 22:6430-7. Erratum in: (2004) *EMBO J*. 23:1680.

Marcusson EG, Horazdovsky BF, Cereghino JL, *et al*. The sorting receptor for yeast vacuolar carboxypeptidase y is encoded by the vps10 gene (1994) *Cell*. 77:579-586.

Matsuda J, Vanier MT, Saito Y, *et al*. A mutation in the saposin A domain of the sphingolipid activator protein (prosaposin) gene results in a late-onset, chronic form of globoid cell leukodystrophy in the mouse (2001) *Hum Mol Genet*. 10:1191-9.

Meikle PJ, Ranieri E, Simonsen H, *et al*. Newborn screening for lysosomal storage disorders (2006) *Mol Genet Metab*. 88:307-14.

Mimae T, Tsuta K, Maeshima AM, *et al*. Cathepsin D as a potential prognostic marker for lung adenocarcinoma (2012) *Pathol Res Pract*. 208:534-40.

Musunuru K, Strong A, Frank-Kamenetsky M, *et al*. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus (2010) *Nature*. 466:714-9.

Ni X, Morales CR. The lysosomal trafficking of acid sphingomyelinase is mediated by sortilin and mannose 6-phosphate receptor (2006) *Traffic*. 7:889-902.

Nielsen MS, Gustafsen C, Madsen P, *et al*. Sorting by the cytoplasmic domain of the amyloid precursor protein binding receptor SorLA (2007) *Mol Cell Biol*. 27:6842-51.

Nielsen MS, Jacobsen C, Olivecrona G, *et al*. Sortilin/neurotensin receptor-3 binds and mediates degradation of lipoprotein lipase (1999) *J Biol Chem*. 274:8832-6.

Paiardini A, Caputo V. Insights into the interaction of sortilin with proneurotrophins: a computational approach (2008) *Neuropeptides*. 42:205-14.

Petersen CM, Nielsen MS, Nykjaer A, *et al*. Molecular identification of a novel candidate sorting receptor purified from human brain by receptor-associated protein affinity chromatography (1997) *J Biol Chem*. 272:3599-605.

Reczek D, Schwake M, Schröder J, *et al*. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase (2007) *Cell*. 131:770-83.

Rojnik M, Jevnikar ZR, Doljak B, *et al*. The influence of differential processing of procathepsin H on its aminopeptidase activity, secretion and subcellular localization in human cell lines (2012) *Eur J Cell Biol*. 91:757-64.

Russell JC, Proctor SD. Small animal models of cardiovascular disease: tools for the study of the roles of metabolic syndrome, dyslipidemia, and atherosclerosis (2006) *Cardiovasc Pathol*. 15:318-30.

Schweiger A, Christensen IJ, Nielsen HJ, *et al*. Serum cathepsin H as a potential prognostic marker in patients with colorectal cancer (2004) *Int J Biol Markers*. 19:289-94.

Schweiger A, Christensen IJ, Nielsen HJ, *et al*. Serum cathepsin H as a potential prognostic marker in patients with colorectal cancer (2004) *Int J Biol Markers*. 19:289-94.

Stevens RL, Fluharty AL, Kihara H, *et al*. Cerebroside sulfatase activator deficiency induced metachromatic leukodystrophy (1981) *Am J Hum Genet*. 33:900-6.

Tamargo RJ, Velayati A, Goldin E, *et al*. The role of saposin C in Gaucher disease (2012) *Mol Genet Metab*. 106:257-63.

Tauris J, Ellgaard L, Jacobsen C, *et al*. The carboxy-terminal domain of the receptor-associated protein binds to the vps10p domain of sortilin (1998) *FEBS Lett*. 429:27-30.

Urbanelli L, Emiliani C, Massini C, *et al*. Cathepsin D expression is decreased in Alzheimer's disease

fibroblasts. Neurobiol Aging. 2008;29:12-22.

Velayati A, DePaolo J, Gupta N, *et al*. A mutation in SCARB2 is a modifier in Gaucher disease (2011) *Hum Mutat*. 32:1232-8.

Wendeler M, Werth N, Maier T, *et al*. The enzyme-binding region of human GM2-activator protein (2006) *FEBS J*. 273:982-91.

Zeng J, Racicott J, Morales CR. The inactivation of the sortilin gene leads to a partial disruption of prosaposin trafficking to the lysosomes (2009) *Exp Cell Res*. 315:3112-24.

Zhang and Chasin. Computational definition of sequence motifs governing constitutive exon splicing (2004) *Genes and Development*. 18:1241-1250.

Zhu Y, Bujo H, Yamazaki H, *et al*. LR11, an LDL receptor gene family member, is a novel regulator of smooth muscle cell migration (2004) *Circ Res*. 94:752-8.

Chapter 4

Conclusions and Future Perspectives

Even though several questions were answered through our studies, many other novel and challenging questions have arisen.

In this last section are summarised the main conclusions we were able to draw from our analysis on both M6P dependent and independent transport pathways of lysosomal enzymes and their involvement in disease, and in addition to it is presented an overview of the next steps of this study as well as a summary of the novel lines of research we intent to follow supported on these foundations.

Part 1

Molecular genetics, pathophysiology and diagnosis of disorders related with malfunction of M6P pathway

In a set of 23 unrelated ML II and III cases, we have identified 18 different mutations: 16 in the *GNPTAB* gene and 2 in the *GNPTG* gene. Of those, 13 were novel: 11 in the *GNPTAB* gene and 2 in the *GNPTG* gene (Table 3.1).

All mutations identified were further analyzed with the most suitable approaches. Special attention was paid to the novel alterations:

- Real time PCR studies were performed whenever cDNA samples were available, which pinpointed the existence of feedback regulation mechanisms between the α/β- and the γ-subunits;
- Long-range PCR and extensive cDNA analysis were applied to characterize the novel deletion c.3435-386_3602+343del897, allowing the identification of the deletion breakpoints and the disentanglement of its effect at cDNA level that led to the creation of three abnormal transcripts;
- Cycloheximid treatment of cultured patients' cell lines was performed to inhibit mRNA degradation through NMD whenever nonsense and frameshift mutations detected through gDNA sequencing could not be confirmed by standard cDNA analysis, allowing their subsequent observation;
- Transient expression of wild-type and mutant *GNPTAB* gene sequences in HEK and HeLa cells revealed to be crucial to further analyze the impact of several *GNPTAB* mutations;
- Immunofluorescence and Western blot assays were used in a series of 7 GNPTAB mutants. Special attention was paid to the atypically severe missense mutations (p.W81L and p.R986C) whose results were cross-linked with the ones obtained for two severe deletions (c.1581delC and c.3503_4delTC) and three mild missense mutations (p.S399F, p.I403T and p.K1236M). From this study it became clear that the answer for the extreme pathogenicity of the two target mutations relied in ER-retention (abnormal subcellular location) of a non-cleaved precursor or in a total absence of the mutated protein, with consequent absence of phosphotransferase activity.

☙ Future Perspectives ❧

Presently, our future approaches concerning disorders related with malfunction of M6P pathway are mainly focused on specific causal mutations and may be divided into the following aspects:

☙ Evaluation of the p.R986C *GNPTAB* mutation effects at mRNA and protein levels

Given the unexpected results obtained for the *GNPTAB* mutation p.R986C, with a complete absence of GlcNAc-phosphotransferase expression on cells expressing it and with a extremely severe phenotype observed on the only known patient who harbours this variant in homozygosity, further analysis should be performed focusing residue R^{986} in particular.

In order to unveil whether the absence of expression of the p.R986C Glc-NAc-phosphotransferase mutant is due to mRNA instability or to premature protein degradation, different approaches may be followed. One of the possibilities is to perform quantitative real-time PCR to assess the *GNPTAB* mRNA levels of cells overexpressing p.R986C. Another tempting approach is to inhibit endoplasmic-reticulum-associated (ERAD) protein degradation with proteasome inhibitors and check whether the mutant protein can, then, be observed through standard Western blot assays. Also Immunofluorescence and co-immunofluorescence assays and pulse-chase experiments with ^{35}S-labelled mutant protein may be performed to verify when and where, at a subcellular level, is the mutant targeted for degradation.

Finally, knowledge of the phosphotransferase crystallographic structure could also provide a structural explanation for how distinct mutant residues affect the molecule function and/or stability including the assembly and interaction between the different subunits of this protein.

☙ Biochemical studies of the novel p.I941T *GNPTAB* mutation

Recent identification of the novel p.I941T *GNPTAB* mutation in a Turkish patient from whom no clinical data was available also prompts additional biochemical studies to disclosure possible effects of this alteration and predict its phenotypic effect.

Expression of p.I941T mutation and assessment of its associated protein expression levels and subcellular location, in a similar way as described on the Results and Discussion section (Paper 5) is already planned.

Part 2

The M6P independent trafficking pathways in LSDs

In a set of 120 individuals with clinical suspect of LSD but without definitive biochemical and/or molecular diagnosis, no novel mutations were detected either on the *SCARB2* or on the *SORT1* gene. From our experience so far, no evidence came that *SORT1* deficiencies may be associated to LSD phenotypes. Still, enlarged sample sizes are needed in order to draw more reliable conclusions on the topic.

In a set of 91 unrelated Gaucher disease (GD) cases, which altogether constitute the whole cohort of Portuguese patients suffering from this disorder, we have identified 1 novel mutation in the *SCARB2* gene, reinforcing previous indications that mutations in the gene that codes for the β-glucocerebrosidase transporter can act as GD modifier. From our study, however, it appears that genetic variability at S*CARB2* does not account much to the GD phenotypic spectrum.

☙ Future Perspectives ❧

In order to carry on the analysis on the contribution of genetic impairments of the M6P-independent receptors to LSD phenotypes, it would be important, not only to continue screening the *SORT1* and *SCARB2* genes in LDS suspects' and the study of the *SCARB2* mutation p.T398M as a GD modifier, but also to evaluate the potential of the Vps10p-family member SorLA as an alternative receptor for Golgi-to-lysosome transport. In summary, future approaches may be divided into the following main lines:

☙ Characterization of the M6P-independent pathway on patients with LSD-phenotype – molecular, *in silico* and *in vitro* studies

SORT1 and *SCARB2* molecular screenings should continue to be performed to obtain a better assessment of how both genes might contribute to LSD related phenotypes.
Every time a novel variant is identified in any of the genes, its effect on protein function will be bioinformatically predicted. Whenever possible (available immortalized fibroblast cell line), cDNA analysis will also be performed to confirm the presence of each mutation and study their effect upon mRNA processing. If cDNA is not available, expression constructs containing wild-type or mutant minigenes will be developed. The effect of each mutation will be further evaluated *in vitro* through functional studies. Western Blot

and Immunofluorescence assays will be performed to evaluate the effect of each variant on protein expression and cellular distribution, respectively.

ꕥ Biochemical studies of the novel p.T398M *SCARB2* mutation

The mutation detected, p.T398M is already under evaluation. Western Blot and Immunofluorescence assays are being carried out in COS7 cells transfected with a minigene carrying the mutation and on the patient's fibroblasts to further investigate protein expression and subcellular localization of LIMP-2 and β-glucocerebrosidase. Additional assays of β-glucocerebrosidase activity and extra-cellular secretion on the patient cell line are also predicted.

ꕥ Evaluation of the structural interaction between sortilin and its lysosomal ligands-*silico* studies

Light should be shed on the molecular mechanism underlying the interaction between sortilin and its lysosomal ligands: PSAP, GM2AP, ASM and cathepsins D and H. A tempting computational approach would comprise extensive sequence analysis, through multiple alignments, secondary structure/disordered regions prediction, threading and homology modeling.

ꕥ Evaluation of SorLA's potential to be an alternative receptor - molecular, *in silico* and *in vitro* studies

Taking into account that, apart from sortilin, SorLA is the only known Vps10p family member which contains Golgi-to-lysosome trafficking motifs, it's potential to be an alternative receptor should also be addressed. An interesting approach would be to develop a dominant negative SorLA vector and use it for co-immunoprecipitation and immunofluorescence assays. We propose cathepsins K and L as SorLA's possible ligands (i) for their long known ability to reach the lysosome in a M6P-independent fashion; (ii) for their structural and functional similarities to cathepsins D/H and (iii) for the already existing experimental evidences that sortilin isn't responsible for their Golgi-to-lysosome transport.

If the role of SorLA as an alternative receptor for cathepsins K/L is confirmed, their interaction could be as well modulated through a computational approach.

∞ Additional Remarks ∞

Apart from the specific data collected on impairments of the lysosomal travelling routes, several important and general progresses were achieved with this study:

∞ Importance of Molecular Genetic Testing (MGT)

First of all, it was possible to implement MGT techniques for three different diseases: ML II, ML III and AMRF. The establishment of MGT for LSDs is essential for the definitive diagnosis of these diseases, by confirming the enzymatic diagnosis. In the particular case of AMRF, only MGT may allow definitive diagnose. Also, for Mucolipidosis type III, only through MGT is it possible to classify the subtype of the disease (III alpha/beta or III gamma) and identify the phosphotransferase subunit which is affected in each patient. It is also the only suitable approach in post-mortem diagnoses, where the only specimens available are DNA samples. Such diagnoses are of major importance to the follow-up of affected families. In fact, knowing the genotype of LSD patients allows carrier detection in their family's members, which opens the possibility of offering them adequate genetic counseling. Once the molecular diagnosis of an LSD patient is ascertained, genetic counseling for at-risk couples includes prenatal testing on chorionic villi (at 11-12 weeks) or amniocytes (at 16 weeks). Finally, whenever genotype-phenotype correlations are reasonably strict, as was often observed in our study, prognosis of disease is much more reliable. Unfortunately, at present there are no available therapies for any of these diseases but still, in the future, the availability of MGT for MLII, MLIII and AMRF will be essential to support therapeutic choices that hopefully will diminish the burden of these devastating diseases.

∞ Relevance of cDNA analysis and alternative MGT methods for a correct diagnose

From the detection of a gross deletion involving the exon 19 of the *GNPTAB* gene, several important recommendations may be done for an accurate molecular diagnose of ML II disease. It is important to bear always in mind the possible existence of such large deletions when performing molecular screenings of the *GNPTAB* gene, being aware that when present in heterozygosity, they might easily escape detection through DNA genomic analyses. Without direct cDNA examination this kind of mutations may remain unidentified. Furthermore, our studies on two independent ML III cases in which only one of the mutant alleles could be detected through conventional cDNA screening due to the efficient action of the NMD mechanism, highlight the necessity of combining diverse analytical approaches on different samples to reach a precise diagnose.
In addition to the biochemical testing, both gDNA and cDNA analyses must be carried out

whenever suitable biological material is available. Furthermore, ML II α/β and ML III α/β cases with only one mutant allele detected by direct sequencing should be further examined for the presence of large heterozygous deletions to strengthen the accuracy of genetic counselling and prenatal diagnosis in the families at risk.

∞ Establishment and interpretation of genotype-phenotype correlations

The establishment of genotype-phenotype correlations is a major goal in Clinical Genetics. However, dealing with predictions on the severity of a given mutation is a delicate matter due to the influence they may have on genetic counseling.

From our molecular characterization of ML alpha/beta patients with subsequent identification of three cases in which homozygous missense mutations were underlying a severe phenotype, it became clear that, even when, for a certain gene, some genotype-phenotype correlations are already known, and clues do exist on the association between a certain type of mutation and the level of severity of its associated phenotype, care should be taken when classifying a novel variant as "mild" or "severe" respecting the putative phenotypic impact. Only extensive biochemical studies may allow a deep analysis of a mutation's effect.

The study on GD patients brought to light other limitations of genotype-phenotype analyses. In fact, although often saw as a prototype for LSDs and being the most common of these diseases, GD early turned out to be a good example of complexity in a monogenic disorder, with the observation of patients carrying the same genotype at *GBA* while presenting quite disparate phenotypes. Studies on different populations of GD patients over the years have been quite instructive in this regard, demonstrating that single GD genotypes, expressed in relatively similar environments, give rise to a wide range of phenotypes, from severe infantile neuropathic and visceral disease to milder variants with neuropathic disease onset in the 4^{th} to 6^{th} decades. From such analysis, became well established that the major variation in phenotype must have significant genetic basis, i.e. modifier genes. Abnormalities of GCase's transporter (LIMP-2) and activator (saposin C) proteins are examples of disease modifiers which explain why, frequently, the genotype and the phenotype do not seem to match. Nevertheless, other modifiers are thought to exist. Diseases like GD clearly illustrate that all manifestations encountered in a single gene disorder may not be thoroughly explained by the primary mutation in the gene of interest.

In conclusion, we should emphasise that the identification of novel variants per se, do not allow an accurate prediction of the phenotype. Ideally, studies on genotype-phenotype correlation should integrate the complexity of factors that underlie the range of phenotypic variability in a given disease, which is naturally critically hampered by the unsatisfactory knowledge on these factors and interaction between them.

Chapter 5

References

A

Alberts B, Johnson A, Lewis J, *et al*. Molecular Biology of the Cell, 4th edition, New York, 2002.

Andermann E, Andermann F, Carpenter S, *et al*. Action myoclonus-renal failure syndrome: a previously unrecognized neurological disorder unmasked by advances in nephrology (1986) *Adv Neurol*. 43:87-103.

Appelman Z, Blumberg BD, Golabi M, *et al*. Nonimmune hydrops fetalis may be associated with an elevated delta OD450 in the amniotic fluid (1988) *Obstet Gynecol*. 71:1005-8.

Applegarth DA, Toone JR, Wilson RD, *et al*. Morquio disease presenting as hydrops fetalis and enzyme analysis of chorionic villus tissue in a subsequent pregnancy (1987) *Pediatr Pathol*. 7:593-9.

Arvio M, Oksanen V, Autio S, *et al*. Epileptic seizures in aspartylglucosaminuria: a common disorder (1993) *Acta Neurol Scand*. 87:342-4.

B

Badhwar A, Berkovic SF, Dowling JP, *et al*. Action myoclonus-renal failure syndrome: characterization of a unique cerebro-renal disorder (2004) *Brain*. 127:2173-82.

Baldwin ET, Bhat TN, Gulnik S, *et al*. Crystal structures of native and inhibited forms of human cathepsin D: implications for lysosomal targeting and drug design (1993) *Proc Natl Acad Sci USA*. 90:6796-800.

Ballabio A and Gieselmann V. Lysosomal disorders: from storage to cellular damage (2009) *Biochim Biophys Acta*. 1793:684-96.

Balreira A, Gaspar P, Caiola D, *et al*. A nonsense mutation in the LIMP-2 gene associated with progressive myoclonic epilepsy and nephrotic syndrome (2008) *Hum Mol Genet*. 17:2238-43.

Bao M, Booth JL, Elmendorf BJ, *et al*. Bovine UDP-N-acetylglucosamine:lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure (1996) *J Biol Chem*. 271:31437-45.

Baranski TJ, Faust PL, Kornfeld S. Generation of a lysosomal enzyme targeting signal in the secretory protein pepsinogen (1990) *Cell*. 63:81-91.

Bargal R, Zeigler M, Abu-Libdeh B, *et al*. When Mucolipidosis III meets Mucolipidosis II: GNPTA gene mutations in 24 patients (2006) *Mol Genet Metab*. 88:359-363.

Beck M. Variable clinical presentation in lysosomal storage disorders (2001) *J Inherit Metab Dis*. 24:47-51.

Beck M. Emerging drugs for lysosomal storage diseases (2010) *Expert Opin Emerg Drugs*. 15:495-507.

Belzil VV, André-Guimont C, Atallah MR, *et al*. Analysis of the SORT1 gene in familial amyotrophic lateral sclerosis (2012) *Neurobiol Aging*. [Epub ahead of print]

Berkovic SF, Dibbens LM, Oshlack A, *et al*. Array-based gene discovery with three unrelated subjects shows SCARB2/LIMP-2 deficiency causes myoclonus epilepsy and glomerulosclerosis (2008) *Am J Hum Genet*. 82:673-84.

Beutler E and Grabowski G. (2001) 'Gaucher disease', in: Scriver, C.R. et al. (eds), The Metabolic and Molecular Bases of Inherited Disease (8th edn), McGraw-Hill, New York, NY, pp. 3635–3668.

Blanz J, Groth J, Zachos C, *et al*. Disease-causing mutations within the lysosomal integral membrane protein type 2 (LIMP-2) reveal the nature of binding to its ligand beta-glucocerebrosidase (2010) *Hum Mol Genet*. 19:563-72.

Boonen M, Vogel P, Platt KA, *et al*. Mice lacking mannose 6-phosphate uncovering enzyme activity have a milder phenotype than mice deficient for N-acetylglucosamine-1-phosphotransferase activity (2009) *Mol Biol Cell*. 20:4381-9.

Braulke T, Pohl S, Storch S. Molecular analysis of the GlcNac-1-phosphotransferase (2008) *J Inherit Metab Dis*. 31:253-7.

Braulke T and Bonifacino JS. Sorting of lysosomal proteins (2009) *Biochim Biophys Acta*. 1793:605-14.

Broustal O, Camuzat A, Guillot-Noel L, *et al*. French clinical and genetic research network on FTD/FTD-MND, FUS mutations in frontotemporal lobar degeneration with amyotrophic lateral sclerosis (2010) *J. Alzheimers Dis*. 22:765–769.

Calvo D, Dopazo J, Vega MA. The CD36, CLA-1 (CD36L1), and LIMPII (CD36L2) gene family: cellular distribution, chromosomal location, and genetic evolution (1995) *Genomics*. 25:100-106.

Cantor AB, Baranski TJ, Kornfeld S. Lysosomal enzyme phosphorylation. II. Protein recognition determinants in either lobe of procathepsin D are sufficient for phosphorylation of both the amino and carboxyl lobe oligosaccharides (1992) *J Biol Chem*. 267:23349-56.

Canuel M, Korkidakis A, Konnyu K, *et al*. Sortilin mediates the lysosomal targeting of cathepsins D and H (2008) *Biochem Biophys Res Commun*. 373:292-7.

Canuel M, Libin Y, Morales CR. The interactomics of sortilin: an ancient lysosomal receptor evolving new functions (2009) *Histol Histopathol*. 24:481-92.

Chavez CA, Bohnsack RN, Kudo M, *et al*. Domain 5 of the cation-independent mannose 6-phosphate receptor preferentially binds phosphodiesters (mannose 6-phosphate N-acetylglucosamine ester) (2007) *Biochemistry*. 46:12604-17.

Chien YH, Chiang SC, Zhang XK, *et al*. Early detection of Pompe disease by newborn screening is feasible: results from the Taiwan screening program (2008) *Pediatrics*. 22:39-45.

Costello DJ, Chiappa KH, Siao P. Progressive myoclonus epilepsy with demyelinating peripheral neuropathy and preserved intellect: a novel syndrome (2009) *Arch. Neurol*. 66:898-901.

Cuozzo JW, Tao K, Wu QL, *et al*. Lysine-based structure in the proregion of procathepsin L is the recognition site for mannose phosphorylation (1995) *J Biol Chem*. 270:15611-9.

Dahms NM and Kornfeld S. The cation-dependent mannose 6-phosphate receptor. Structural requirements for mannose 6-phosphate binding and oligomerization (1989) *J Biol Chem*. 264:11458-67.

Dahms NM, Lobel P, Breitmeyer J, *et al*. 46 kd mannose 6-phosphate receptor: cloning, expression, and homology to the 215 kd mannose 6-phosphate receptor (1987) *Cell*. 50:181-92.

Dahms NM, Rose PA, Molkentin JD, *et al*. The bovine mannose 6-phosphate/insulin-like growth factor II receptor. The role of arginine residues in mannose 6-phosphate binding (1993) *J Biol Chem*. 268:5457-63.

Dahms NM, Wick DA, Brzycki-Wessell MA. The bovine mannose 6-phosphate/insulin-like growth factor II receptor. Localization of the insulin-like growth factor II binding site to domains 5-11 (1994) *J Biol Chem*. 269:3802-9.

Dardis A, Filocamo M, Grossi S, *et al*. Biochemical and molecular findings in a patient with myoclonic epilepsy due to a mistarget of the beta-glucosidase enzyme (2009) *Mol Genet Metab*. 97:309-11.

de Duve C, Pressman BC, Gianetto R, *et al*. Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue (1955) *Biochem J*. 60:604-17.

de Duve C. The lysosome turns fifty (2005) *Nat Cell Biol*. 7:847-9.

Denis R, Wayenberg JL, Vermeulen M, Gorus *et al*. Hyperphosphatasemia in early diagnosed infantile GM1 gangliosidosis presenting as transient hydrops fetalis (1996) *Acta Clin Belg*. 51:320-7.

Dibbens LM, Michelucci R, Gambardella A, *et al*. SCARB2 mutations in progressive myoclonus epilepsy (PME) without renal failure (2009) *Ann. Neurol*. 66:532-536.

Dibbens LM, Karakis I, Bayly MA, *et al*. Mutation of SCARB2 in a patient with progressive myoclonus epilepsy and demyelinating peripheral neuropathy (2011) *Arch Neurol*. 68:812-3.

Do H, Lee WS, Ghosh P, *et al*. Human mannose 6-phosphate-uncovering enzyme is synthesized as a proenzyme that is activated by the endoprotease furin (2002) *J Biol Chem*. 277:29737-44.

Dubé JB, Johansen CT, Hegele RA. Sortilin: an unusual suspect in cholesterol metabolism: from GWAS identification to in vivo biochemical analyses, sortilin has been identified as a novel mediator of human lipoprotein metabolism (2011) *Bioessays*. 33:430-7.

Duffner PK, Caggana M, Orsini JJ, *et al*. Newborn screening for Krabbe disease: the New York State model (2009a) *Pediatr Neurol* 40:245–252.

Duffner PK, Caviness VS, Erbe RW, *et al*. The long-term outcomes of presymptomatic infants transplanted for Krabbe disease: report of the workshop held on July 11 and 12, 2008, Holiday Valley, New York (2009b) *Genet Med*. 11:450–454.

E

Encarnação M, Kollmann K, Trusch M, *et al*. Post-translational Modifications of the {gamma}-Subunit Affect Intracellular Trafficking and Complex Assembly of glcnac-1-phosphotransferase (2011) *J Biol Chem*. 286:5311-8.

Eyal N, Firon N, Wilder S, *et al*. Three unique base pair changes in a family with Gaucher disease (1991) *Hum Genet*. 87:328-32.

F

Filocamo M and Morrone A. Lysosomal storage disorders: molecular basis and laboratory testing (2011) *Hum Genomics*. 5:156-69.

Fuller M, Tucker JN, Lang DL, *et al*. Screening patients referred to a metabolic clinic for lysosomal storage disorders (2011) *J Med Genet*. 48:422-5.

Futerman AH and van Meer G. The cell biology of lysosomal storage disorders (2004) *Nat Rev Mol Cell Biol*. 5:554-65.

G

Gamp AC, Tanaka Y, Lüllmann-Rauch R, *et al*. LIMP-2/LGP85 deficiency causes ureteric pelvic junction obstruction, deafness and peripheral neuropathy in mice (2003) *Hum Mol Genet*. 12:631-46.

Gelfman CM, Vogel P, Issa TM, *et al*. Mice lacking alpha/beta subunits of GlcNAc-1-phosphotransferase exhibit growth retardation, retinal degeneration, and secretory cell lesions (2007) *Invest Ophthalmol Vis Sci*. 48:5221-8.

Ghosh P, Dahms NM, Kornfeld S. Mannose 6-phosphate receptors: new twists in the tale (2003) *Nat Rev Mol Cell Biol*. 4:202-12.

Giugliani R. Newborn screening for lysosomal diseases: current status and potential interface with population medical genetics in Latin America (2012) *J Inherit Metab Dis*. 35:871-7.

Gritti A. Gene therapy for lysosomal storage disorders (2011) *Expert Opin Biol Ther*. 11:1153-67.

H

Haltia M and Goebel HH. The neuronal ceroid-lipofuscinoses: A historical introduction (2012) *Biochim Biophys Acta*. [*Epub ahead of print*]

Hancock MK, Yammani RD, Dahms NM, Localization of the carbohydrate recognition sites of the insulin-like growth factor II/mannose 6-phosphate receptor to domains 3 and 9 of the extracytoplasmic region (2002) *J Biol Chem*. 277:47205-12.

Hasilik A, Wrocklage C, Schröder B. Intracellular trafficking of lysosomal proteins and lysosomes (2009) *Int J Clin Pharmacol Ther*. 47 Suppl 1:S18-33.

Haverkamp F, Jacobs D, Cantz M, *et al*. Nonimmune hydrops fetalis with galactosialidosis: consequences for family planning (1996) *Fetal Diagn Ther*. 11:114-9.

Helenius A and Aebi M. Intracellular functions of N-linked glycans (2001) *Science*. 291:2364-9.

Hopfner F, Schormair B, Knauf F, *et al*. Novel SCARB2 mutation in action myoclonus-renal failure syndrome and evaluation of SCARB2 mutations in isolated AMRF features (2011) *BMC Neurol*. 27:134.

Horoupian DS and Ross RT. Pigment variant of neuronal ceroid-lipofuscinosis (Kufs' disease) (1977) *Can J Neurol Sci*. 4:67-75.

Hu F, Padukkavidana T, Vægter CB, *et al*. Sortilin-mediated endocytosis determines levels of the frontotemporal dementia protein, progranulin (2010) *Neuron*. 68:654-67.

Huey ED, Ferrari R, Moreno JH, *et al*. FUS and TDP43 genetic variability in FTD and CBS (2012) *Neurobiol. Aging*. 33:1016.e9-17

Human Gene Mutation Database website [HGMD, http://www.hgmd.org]

Hwu WL, Chien YH, Lee NC, *et al*. Newborn screening for Fabry disease in Taiwan reveals a high incidence of the later-onset GLA mutation c.936+919G>A (IVS4+919G>A) (2009) *Hum Mutat*. 30:1397-405.

Igdoura SA, Rasky A, Morales CR. Trafficking of sulfated glycoprotein-1 (prosaposin) to lysosomes or to the extracellular space in rat Sertoli cells (1996) *Cell Tissue Res.* 283:385-94.

Kabashi E, Valdmanis PN, Dion P, *et al*. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis (2008) *Nat. Genet*. 40:572–574.

Kang C, Riazuddin S, Mundorff J *et al*. Mutations in the lysosomal enzyme-targeting pathway and persistent stuttering (2010) *N Engl J Med*. 362:677-85.

Kathiresan S, Melander O, Guiducci C, *et al*. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans (2008) *Nat Genet*. 40:189-97.

Kathiresan S, Willer CJ, Peloso GM, *et al*. Common variants at 30 loci contribute to polygenic dyslipidemia (2009) *Nat Genet*. 41:56-65.

Kattner E, Schäfer A, Harzer K. Hydrops fetalis: manifestation in lysosomal storage diseases including Farber disease (1997) *Eur J Pediatr*. 156:292-5.

Kjolby M, Andersen OM, Breiderhoff T, *et al*. Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export (2010) *Cell Metab*. 12:213-23.

Klinger SC, Glerup S, Raarup MK, *et al*. SorLA regulates the activity of lipoprotein lipase by intracellular trafficking (2011) *J Cell Sci*. 124:1095-105.

Knipper M, Claussen C, Rüttiger L, *et al*. Deafness in LIMP2-deficient mice due to early loss of the potassium channel KCNQ1/KCNE1 in marginal cells of the stria vascularis (2006) *J Physiol*. 576:73-86.

Kollmann K, Pohl S, Marschner K, *et al*. Mannose phosphorylation in health and disease (2010) *Eur J Cell Biol*. 89:117-23.

Kornfeld R and Kornfeld S. Assembly of asparagine-linked oligosaccharides (1985) *Annu Rev Biochem*. 54:631-64.

Kornfeld S and Mellman I. The biogenesis of lysosomes (1989) *Annu Rev Cell Biol*. 5:483-525.

Kornfeld S. Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors (1992) *Annu Rev Biochem*. 61:307-30

Kornfeld R, Bao M, Brewer K, *et al*. Molecular cloning and functional expression of two splice forms of human N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase (1999) *J Biol Chem*. 274:32778-85.

Kornfeld S and Sly WS. I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localisation. *In*: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. Childs B, Kinzler KW, Vogelstein B, assoc. eds. The metabolic and molecular bases of inherited disease, 8th edn (2001) New York: McGrawHill. 3421–3442.

Köster A, Saftig P, Matzner U, von *et al*. Targeted disruption of the M(r) 46,000 mannose 6-phosphate receptor gene in mice results in misrouting of lysosomal proteins (1993) *EMBO J*. 12:5219-23.

Kudo M, Bao M, D'Souza A, *et al*. The alpha- and beta-subunits of the human UDP-N-acetylglucosamine:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase [corrected] are encoded by a single cDNA (2005) *J Biol Chem*. 280:36141-9.

Kudo M and Canfield WM. Structural requirements for efficient processing and activation of recombinant human UDP-N-acetylglucosamine:lysosomal-enzyme-N-acetylglucosamine-1-phosphotransferase (2006a) *J Biol Chem*. 281:11761-8.

Kudo M, Brem MS, Canfield WM. Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha / beta -subunits precursor gene (2006b) *Am J Hum Genet*. 78:451-463.

Kulterer B, Friedl G, Jandrositz A, *et al*. Gene expression profiling of human mesenchymal stem cells derived from bone marrow during expansion and osteoblast differentiation (2007) *BMC Genomics*. 8:70.

Kwiatkowski TJ, Bosco DA, Leclerc AL, *et al*. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis (2009) *Science*. 323:1205–1208.

Lee WS, Rohrer J, Kornfeld R, *et al*. Multiple signals regulate trafficking of the mannose 6-phosphate-uncovering enzyme (2002) *J Biol Chem*. 277:3544-51.

Lee WS, Payne BJ, Gelfman CM, *et al*. Murine UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase lacking the gamma-subunit retains substantial activity toward acid hydrolases (2007) *J Biol Chem*. 282:27198-203.

Lee WS, Kang C, Drayna D *et al*. Analysis of mannose 6-phosphate uncovering enzyme mutations associated with persistent stuttering (2011) *J Biol Chem*. 286:39786-93.

Lefrancois S, Zeng J, Hassan AJ, *et al*. The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin (2003) *EMBO J*. 22:6430-7.

Leroy JG and Spranger JW. I-cell disease (1970) *N Engl J Med*. 283:598-9.

Lin HY, Chong KW, Hsu JH, *et al*. High incidence of the cardiac variant of Fabry disease revealed by newborn screening in the Taiwan Chinese population (2009) *Circ Cardiovasc Genet*. 2:450-6.

Linsel-Nitschke P, Samani NJ, Schunkert H. Sorting out cholesterol and coronary artery disease (2010) *N Engl J Med*. 363:2462-3.

Lobel P, Dahms NM, Kornfeld S. Cloning and sequence analysis of the cation-independent mannose 6-phosphate receptor (1988) *J Biol Chem*. 263:2563-70.

Lübke T, Lobel P, Sleat DE. Proteomics of the lysosome (2009) *Biochim Biophys Acta*. 1793:625-35.

Ludwig T, Ovitt CE, Bauer U, *et al*. Targeted disruption of the mouse cation-dependent mannose 6-phosphate receptor results in partial missorting of multiple lysosomal enzyme (1993) *EMBO J*. 12:5225-35.

Ludwig T, Munier-Lehmann H, Bauer U, *et al*. Differential sorting of lysosomal enzymes in mannose 6-phosphate receptor-deficient fibroblasts (1994) *EMBO J*. 13:3430-7.

Ludwig T, Le Borgne R, Hoflack B. Roles for mannose-6-phosphate receptors in lysosomal enzyme sorting, IGF-II binding and clathrin-coat assembly (1995) *Trends Cell Biol*. 5:202-6.

Maeda S, Nobukuni T, Shimo-Onoda K, *et al*. Sortilin is upregulated during osteoblastic differentiation of mesenchymal stem cells and promotes extracellular matrix mineralization (2002) *J Cell Physiol*. 193:73-9.

Marcusson EG, Horazdovsky BF, Cereghino JL, *et al*. The sorting receptor for yeast vacuolar carboxypeptidase Y is encoded by the VPS10 gene (1994) *Cell*. 77:579-86.

Marron-Terada PG, Hancock MK, Haskins DJ, *et al*. Recognition of Dictyostelium discoideum lysosomal enzymes is conferred by the amino-terminal carbohydrate binding site of the insulin-like growth factor II/mannose 6-phosphate receptor (2000) *Biochemistry*. 39:2243-53.

Marschner K, Kollmann K, Schweizer M, *et al*. A key enzyme in the biogenesis of lysosomes is a protease that regulates cholesterol metabolism (2011) *Science*. 333:87-90.

Meikle PJ and Hopwood JJ, Clague AE, *et al*.. Prevalence of lysosomal storage disorders (1999) *JAMA*. 281:249-54.

Meikle PJ, Hopwood JJ. Lysosomal storage disorders: emerging therapeutic options require early diagnosis (2003) *Eur J Pediatr*. 162:Suppl 1:S34-7.

Meizner I, Levy A, Carmi R, *et al*. Niemann-Pick disease associated with nonimmune hydrops fetalis (1990) *Am J Obstet Gynecol*. 163:128-9.

Michelakakis H, Dimitriou E, Tsagaraki S, *et al*. Lysosomal storage diseases in Greece (1995) *Genet Couns*. 6:43-7.

Mufson EJ, Wuu J, Counts SE, *et al*. Preservation of cortical sortilin protein levels in MCI and Alzheimer's disease (2010) *Neurosci Lett*. 471:129-33.

Munck-Petersen C, Nielsen MS, Jacobsen C, *et al*. Propeptide cleavage conditions sortilin/neurotensin receptor-3 for ligand binding (1999) *EMBO J*. 18:595-604.

Munier-Lehmann H, Mauxion F, Bauer U, *et al*. Re-expression of the mannose 6-phosphate receptors in receptor-deficient fibroblasts. Complementary function of the two mannose 6-phosphate receptors in lysosomal enzyme targeting (1996) *J Biol Chem*. 271:15166-74.

Musunuru K, Strong A, Frank-Kamenetsky M, *et al*. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus (2010) *Nature*. 466:714-9.

N

Ni X, Canuel M, Morales CR. The sorting and trafficking of lysosomal proteins (2006) *Histol Histopathol*. 21:899-913.

Ni X and Morales CR. The lysosomal trafficking of acid sphingomyelinase is mediated by sortilin and mannose 6-phosphate receptor (2006) *Traffic*. 7:889-902.

Nielsen MS, Jacobsen C, Olivecrona G, *et al*. Sortilin/neurotensin receptor-3 binds and mediates degradation of lipoprotein lipase (1999) *J Biol Chem*. 274:8832-6.

Nielsen MS, Madsen P, Christensen EI, Nykjaer *et al*. The sortilin cytoplasmic tail conveys Golgi-endosome transport and binds the VHS domain of the GGA2 sorting protein (2001) *EMBO J*. 20:2180-90.

Nielsen MS, Gustafsen C, Madsen P, *et al*. Sorting by the cytoplasmic domain of the amyloid precursor protein binding receptor SorLA (2007) *Mol Cell Biol*. 27:6842-51.

Nilsson SK, Christensen S, Raarup MK, *et al*. Endocytosis of apolipoprotein A-V by members of the low density lipoprotein receptor and the VPS10p domain receptor families (2008) *J Biol Chem*. 283: 283 25920-7.

Nykjaer A, Lee R, Teng KK, *et al*. Sortilin is essential for proNGF-induced neuronal cell death (2004) *Nature*. 427:843-8.

O

Olson LJ, Peterson FC, Castonguay A, *et al*. Structural basis for recognition of phosphodiester-containing lysosomal enzymes by the cation-independent mannose 6-phosphate receptor (2010) *Proc Natl Acad Sci USA*. 107:12493-8.

Oshima A, Nolan CM, Kyle JW, *et al*. The human cation-independent mannose 6-phosphate receptor. Cloning and sequence of the full-length cDNA and expression of functional receptor in COS cells (1988) *J Biol Chem*. 263:2553-62.

Otomo T, Muramatsu T, Yorifuji T, *et al*. Mucolipidosis II and III alpha/beta: mutation analysis of 40 Japanese patients showed genotype-phenotype correlation (2009) *J Hum Genet*. 54:145-51.

Ovali F, Samanci N, Güray A, *et al*. Congenital sialidosis (1998) *Turk J Pediatr*. 40:447-51.

P

Paiardini A and Caputo V. Insights into the interaction of sortilin with proneurotrophins: a computational approach (2008) *Neuropeptides*. 42:205-14.

Paik KH, Song SM, Ki CS, *et al*. Identification of mutations in the GNPTA (MGC4170) gene coding for GlcNAc-phosphotransferase alpha/beta subunits in Korean patients with mucolipidosis type II or type IIIA (2005) *Hum Mutat*. 26:308-14.

Parkinson-Lawrence EJ, Shandala T, Prodoehl M, *et al*. Lysosomal storage disease: revealing lysosomal function and physiology (2010) *Physiology (Bethesda)*. 25:102-15.

Petersen CM, Nielsen MS, Nykjaer A, *et al*. Molecular iden tification of a novel candidate sorting receptor purified from human brain by receptor-associated protein affinity chromatography (1997) *J Biol Chem*. 272:3599-605.

Pinto R, Caseiro C, Lemos M, *et al*. Prevalence of lysosomal storage diseases in Portugal (2004) *Eur J Hum Genet*. 12: 87-92.

Piraud M, Froissart R, Mandon G, *et al*. Amniotic fluid for screening of lysosomal storage diseases presenting in utero (mainly as non-immune hydrops fetalis) (1996) *Clin Chim Acta*. 248:143-55.

Plante M, Claveau S, Lepage P, *et al*. Mucolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTAB) in a French Canadian founder population (2008) *Clin Genet*. 73:236-244.

Pohl S, Tiede S, Castrichini M, *et al*. Compensatory expression of human N-acetylglucosaminyl-1-phosphotransferase subunits in mucolipidosis type III gamma (2009) *Biochim Biophys Acta*. 1792:221-5.

Pohl S, Tiede S, Marschner K, *et al*. Proteolytic processing of the gamma-subunit is associated with the

failure to form GlcNAc-1-phosphotransferase complexes and mannose 6-phosphate residues on lysosomal enzymes in human macrophages (2010) *J Biol Chem*. 285:23936-44.

Polymenidou M, Lagier-Tourenne C, Hutt KR, *et al*. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43 (2011) *Nat Neurosci*. 14:459–468.

Poorthuis BJ, Wevers RA, Kleijer WJ, *et al*. The frequency of lysosomal storage diseases in The Netherlands (1999) *Hum Genet*. 105:151-6.

Poupetová H, Ledvinová J, Berná L, *et al*. The birth prevalence of lysosomal storage disorders in the Czech Republic: comparison with data in different populations (2010) *J Inherit Metab Dis*. 33:387-96.

Pryor PRv and Luzio JP. Delivery of endocytosed membrane proteins to the lysosome (2009) *Biochim Biophys Acta*. 1793:615-24.

Q

Qian Y, Lee I, Lee WS, Qian M, Kudo *et al*. Functions of the alpha, beta, and gamma subunits of UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase (2010) *J Biol Chem*. 285:3360-70.

R

Raas-Rothschild A, Cormier-Daire V, Bao M, *et al*. Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC) (2000) *J Clin Invest*. 105:673-81.

Reczek D, Schwake M, Schröder J, *et al*. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase (2007) *Cell*. 131:770-83.

Reissner K, Tayebi N, Stubblefield BK, *et al*. Type 2 Gaucher disease with hydrops fetalis in an Ashkenazi Jewish family resulting from a novel recombinant allele and a rare splice junction mutation in the glucocerebrosidase locus (1998) *Mol Genet Metab*. 63:281-8.

Riaz N, Steinberg S, Ahmad J, *et al*. Genomewide significant linkage to stuttering on chromosome 12, *Am J Hum Genet*. 76 (2005) 647-51.

Rigat B, Wang W, Leung A, *et al*. Two mechanisms for the recapture of extracellular GM2 activator protein: evidence for a major secretory form of the protein (1997) *Biochemistry*. 36:8325-31.

Ron I and Horowitz M. Intracellular cholesterol modifies the ERAD of glucocerebrosidase in Gaucher disease patients (2008) *Mol Genet Metab*. 93:426-36.

Rothdach AJ, Dietl T, Kümpfel T, Gottschalk *et al*. [Familial myoclonus-renal failure syndrome] (2001) *Nervenarzt*. 72:636-40.

Rubboli G, Franceschetti S, Berkovic SF, *et al*. Clinical and neurophysiologic features of progressive myoclonus epilepsy without renal failure caused by SCARB2 mutations (2011) *Epilepsia*. 52:2356-63.

Ruddock LW and Molinari M. N-glycan processing in ER quality control (2006) *J Cell Sci*. 119:4373-80.

S

Saftig P and Klumperman J. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function (2009) *Nat Rev Mol Cell Biol*. 10:623-35

Samani NJ, Erdmann J, Hall AS, *et al*. Genomewide association analysis of coronary artery disease (2007) *N Engl J Med*. 357:443-53.

Santavuori P. Neuronal ceroid-lipofuscinoses in childhood (1988) *Brain Dev*. 10:80-3.

Schröder BA, Wrocklage C, Hasilik A, *et al*. The proteome of lysosomes (2010) *Proteomics*. 10:4053-76.

Schultz ML, Tecedor L, Chang M, *et al*. Clarifying lysosomal storage diseases (2011) *Trends Neurosci*. 34:401-10.

Shi J and Kandror KV. Sortilin is essential and sufficient for the formation of Glut4 storage vesicles in 3T3-L1 adipocytes (2005) *Dev Cell*. 9: 99-108.

Shi J and Kandror KV. The luminal Vps10p domain of sortilin plays the predominant role in targeting to insulin-responsive Glut4-containing vesicles (2007) *J Biol Chem*. 282:9008-16.

Shugart YY, Mundorff J, Kilshaw J, Doheny *et al*. Results of a genome-wide linkage scan for stuttering (2004) *Am J Med Genet A*. 124A:133-5.

Simonaro CM, D'Angelo M, Haskins ME, *et al*. Joint and bone disease in mucopolysaccharidoses VI and VII:

identification of new therapeutic targets and biomarkers using animal models (2005) *Pediatr Res.* 57:701-7.

Sohar I, Sleat D, Gong Liu C, *et al*. Mouse mutants lacking the cation-independent mannose 6-phosphate/insulin-like growth factor II receptor are impaired in lysosomal enzyme transport: comparison of cation-independent and cation-dependent mannose 6-phosphate receptor-deficient mice (1998) *Biochem J*. 330:903-8.

Sommerlade HJ, Selmer T, Ingendoh A, *et al*. Glycosylation and phosphorylation of arylsulfatase A (1994) *J Biol Chem*. 269:20977-81.

Spada M, Pagliardini S, Yasuda M, *et al*. High incidence of later-onset fabry disease revealed by newborn screening (2006) *Am J Hum Genet*. 79:31-40.

Spranger JW and Wiedemann HR. The genetic mucolipidoses. Diagnosis and differential diagnosis (1970) *Humangenetik*. 9:113-39.

Sreedharan J, Blair IP, Tripathi VB, *et al*. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis (2008) *Science*. 319:1668-72.

Suresh R, Ambrose N, Roe C, *et al*. New complexities in the genetics of stuttering: significant sex-specific linkage signals (2006) *Am J Hum Genet*. 78: 554-63.

T

Tall AR and Ai D. Sorting out sortilin (2011) *Circ Res.* 108:158-60.

Tappino B, Regis S, Corsolini F, *et al*. An Alu insertion in compound heterozygosity with a microduplication in GNPTAB gene underlies Mucolipidosis II (2008) *Mol Genet Metab*. 93:129-33.

Tappino B, Chuzhanova NA, Regis S, *et al*. Molecular characterization of 22 novel UDP-N-acetylglucosamine-1-phosphate transferase alpha- and beta-subunit (GNPTAB) gene mutations causing mucolipidosis types IIalpha/beta and IIIalpha/beta in 46 patients (2009) *Hum Mutat*. 30:E956-973.

Tauris J, Ellgaard L, Jacobsen C, *et al*. The carboxy-terminal domain of the receptor-associated protein binds to the Vps10p domain of sortilin (1998) *FEBS Lett*. 429:27-30.

Tiede S, Cantz M, Raas-Rothschild A, *et al*. A novel mutation in UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) in two siblings with mucolipidosis type III alters a used glycosylation site (2004) *Hum Mutat*. 24:535.

Tiede S, Storch S, Lübke T, *et al*. Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase (2005) *Nat Med*. 11:1109-12.

Tong PY and Kornfeld S. Ligand interactions of the cation-dependent mannose 6-phosphate receptor. Comparison with the cation-independent mannose 6-phosphate receptor (1989) *J Biol Chem*. 264:7970-5.

U

Umehara F, Matsumoto W, Kuriyama M, *et al*. Mucolipidosis III (pseudo-Hurler polydystrophy); clinical studies in aged patients in one family (1997) *J Neurol Sci*. 146:167-72.

V

Vadlamudi L, Vears DF, Hughes A, *et al*. Action myoclonus-renal failure syndrome: a cause for worsening tremor in young adults (2006) *Neurology.* 67:1310-1.

Vaegter CB, Jansen P, Fjorback AW, *et al*. Sortilin associates with Trk receptors to enhance anterograde transport and neurotrophin signaling (2011) *Nat Neurosci*. 14:54-61.

Vallance H, and Ford J. Carrier testing for autosomal-recessive disorders (2003) *Crit Rev Clin Lab Sci*. 40:473-97.

Van Dorpe J, Moerman P, Pecceu A, *et al*. Non-immune hydrops fetalis caused by beta-glucuronidase deficiency (mucopolysaccharidosis VII). Study of a family with 3 affected siblings (1996) *Genet Couns.* 7:105-12.

Vance C, Rogelj B, Hortobágyi T, *et al*. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6 (2009) *Science*. 323:1208-11.

Varki A and Kornfeld S. Purification and characterization of rat liver alpha-N-acetylglucosaminyl

phosphodiesterase (1981) *J Biol Chem*. 256:9937-43.

Velayati A, DePaolo J, Gupta N, *et al*. A mutation in SCARB2 is a modifier in Gaucher disease (2011) *Hum Mutat*. 32:1232-8.

Vellodi A. Lysosomal storage disorders (2005) *Br J Haematol*. 128:413-31.

Vitner EB, Platt FM, Futerman AH. Common and uncommon pathogenic cascades in lysosomal storage diseases (2010) *J Biol Chem*. 285:20423-7.

W

Waheed A and von Figura K. Rapid equilibrium between monomeric, dimeric and tetrameric forms of the 46-kDa mannose 6-phosphate receptor at 37 degrees C. Possible relation to the function of the receptor (1990) *Eur J Biochem*. 193:47-54.

Walkley SU. Pathogenic cascades in lysosomal disease-Why so complex? (2009) *J Inherit Metab Dis*. 32:181-9.

Walkley SU and Vanier MT. Secondary lipid accumulation in lysosomal disease (2009) *Biochim Biophys Acta*. 1793: 726-36.

Weiss N. Cross-talk between TRPML1 channel, lipids and lysosomal storage diseases (2012) *Commun Integr Biol*. 5:111-3.

Westlund B, Dahms NM, Kornfeld S. The bovine mannose 6-phosphate/insulin-like growth factor II receptor. Localization of mannose 6-phosphate binding sites to domains 1-3 and 7-11 of the extracytoplasmic region (1991) *J Biol Chem*. 266:23233-9.

Wilcox WR. Lysosomal storage disorders: the need for better pediatric recognition and comprehensive care (2004) *J Pediatr.* 144:S3-14.

Williamson WR and Hiesinger PR. On the role of v-ATPase V0a1-dependent degradation in Alzheimer disease (2010) *Commun Integr Biol*. 3:604-7.

Willnow TE, Petersen CM, Nykjaer A. VPS10P-domain receptors - regulators of neuronal viability and function (2008) *Nat Rev Neurosci*. 9:899-909:

Wittke-Thompson JK, Ambrose N, Yairi E, *et al*. Genetic studies of stuttering in a founder population (2007) *J Fluency Disord.* 32:33-50.

Wraith JE. Lysosomal disorders. *Semin Neonatol*. 7 (2002) 75-83.

Z

Zachos C, Blanz J, Saftig P, *et al*. A Critical Histidine Residue Within LIMP-2 Mediates pH Sensitive Binding to Its Ligand β-Glucocerebrosidase (2012) *Traffic*. [Epub ahead of print]

Appendix 1

Review papers

Review Paper 1: Glycosaminoglycan storage disorders: a review.

Review Paper 2: Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction.

Review Paper 3: A shortcut to the lysosome: the mannose-6-phosphate-independent pathway.

Review Paper 4: *Sortilina e Risco de Doença Cardiovascular*/ Sortilin and the risk of Cardiovascular Disease.

Hindawi Publishing Corporation
Biochemistry Research International
Volume 2012, Article ID 471325, 16 pages
doi:10.1155/2012/471325

Review Article

Glycosaminoglycan Storage Disorders: A Review

Maria Francisca Coutinho,[1] Lúcia Lacerda,[2] and Sandra Alves[1]

[1] *Research and Development Unit, Department of Genetics, CGMJM, INSA, Portugal*
[2] *Biochemical Genetics Unit, Department of Genetics, CGMJM, INSA, Portugal*

Correspondence should be addressed to Maria Francisca Coutinho, francisca_coutinho@yahoo.com

Received 2 June 2011; Accepted 9 August 2011

Academic Editor: Laura Alaniz

Impaired degradation of glycosaminoglycans (GAGs) with consequent intralysosomal accumulation of undegraded products causes a group of lysosomal storage disorders known as mucopolysaccharidoses (MPSs). Characteristically, MPSs are recognized by increased excretion in urine of partially degraded GAGs which ultimately result in progressive cell, tissue, and organ dysfunction. There are eleven different enzymes involved in the stepwise degradation of GAGs. Deficiencies in each of those enzymes result in seven different MPSs, all sharing a series of clinical features, though in variable degrees. Usually MPS are characterized by a chronic and progressive course, with different degrees of severity. Typical symptoms include organomegaly, dysostosis multiplex, and coarse facies. Central nervous system, hearing, vision, and cardiovascular function may also be affected. Here, we provide an overview of the molecular basis, enzymatic defects, clinical manifestations, and diagnosis of each MPS, focusing also on the available animal models and describing potential perspectives of therapy for each one.

1. Introduction

The mucopolysaccharidoses (MPSs) are a group of lysosomal storage disorders caused by deficiency of enzymes catalyzing the stepwise degradation of glycosaminoglycans (GAGs) and characterized by intralysosomal accumulation and increased excretion in urine of partially degraded GAGs, which ultimately results in cell, tissue, and organ dysfunction [1].

Glycosaminoglycans (previously called mucopolysaccharides), with the exception of hyaluronic acid, are the degradation products of proteoglycans that exist in the extracellular matrix and are proteolytic cleaved, giving origin to GAGs, which enter the lysosome for intracellular digestion. There are four different pathways of lysosomal degradation of GAGs, depending on the molecule to be degraded: dermatan sulfate, heparan sulfate, keratan sulfate, and chondroitin sulfate. The stepwise degradation of glycosaminoglycans requires 10 different enzymes: four glycosidases, five sulfatases, and one nonhydrolytic transferase, whose structure, biosynthesis, processing, and cDNA sequence have already been extensively documented. Deficiencies of each one of these enzymes have already been reported and result in seven different MPSs, all of them sharing a series of clinical features, even though in variable degrees (summarized in Table 1) [1, 2].

Usually, MPSs are characterized by a chronic and progressive course, with different velocities of progression depending on the severity of each one. The typical symptoms include organomegaly, dysostosis multiplex, and a characteristic abnormal facies. Hearing, vision, and cardiovascular function may also be affected. Additionally, joint mobility may also be compromised. The majority of symptoms may be explained by abnormal accumulation of undegraded substrates within the lysosomes. In fact, the continued presentation of GAGs to cell for degradation results in storage, which gives rise to an enlargement of lysosomes. As substrates accumulate, the lysosomes swell and occupy more and more of the cytoplasm. As a consequence of this increased number and size of lysosomes, other cellular organelles may be obscured, and the nuclear outline may be deformed. As the process continues, the enlarged cells lead to organomegaly. Abnormalities observed in heart cells and function may also be explained by GAGs accumulation. The increase of storage material within the cells of the heart valves causes an alteration of the cell's outline, changing

Table 1: Summary table of mucopolysaccharidoses.

Pathology	Subtype	Enzyme deficiency	Gene (localization)	Affected GAG	Clinical manifestations	Animal model
MPS I	Hurler (H)	α-L-iduronidase	*IDUA* 4p16.3	Dermatan and heparan sulfate	Corneal clouding; dysostosis multiplex; organomegaly; heart disease; mental retardation; death in childhood.	Feline [5]; canine [6]; knock-out mouse [7]
	Hurler -Scheie (H/S)	α-L-iduronidase		Dermatan and heparan sulfate	Intermediate phenotype, between MPS IH and MPS IS.	
	Scheie (S)	α-L-iduronidase		Dermatan and heparan sulfate	Corneal clouding; stiff joints; normal intelligence and life span.	
MPS II	Hunter	Irudonate sulfatase	*IDS* Xq28	Dermatan and heparan sulfate	Dysostosis multiplex; organomegaly; no corneal clouding; mental retardation; death before 15 years (severe); Short stature; normal intelligence; survival to 20s to 60s (mild)	Canine [8]; knock-out mouse [9]
MPS III	Sanfilippo A	Heparan-N-sulfatase	*SGSH* 17q25.3	Heparan sulfate	Relatively mild somatic manifestations; hyperactivity; profound mental deterioration.	Canine [10]; spontaneous mouse mutant [11]
	Sanfilippo B	α-N-Acetylglucosaminidase	*NAGLU* 17q21	Heparan sulfate	Phenotype similar to MPS IIIA.	Canine [12]
	Sanfilippo C	Heparan acetyl-CoA:α-glucosaminide N-acetyltransferase	*HGSNAT* 8p11.1	Heparan sulfate	Phenotype similar to MPS IIIA.	
	Sanfilippo D	N-Acetylglucosamine 6-sulfatase	*GNS* 12q14	Heparan sulfate	Phenotype similar to MPS IIIA.	Caprine [13]
MPS IV	Morquio A	Galactose 6-sulfatase	*GALNS* 16q24.3	Keratan and chondroitin sulfate	Distinctive skeletal abnormalities; corneal clouding; odontoid hypoplasia; milder forms known to exist.	Mouse [14]
	Morquio B	β-galactosidase	*GLB1* 3p21.33	Keratan sulfate	Phenotype similar to MPS IVA, with the same spectrum of severity.	
MPS V	** This designation is no longer used; the phenotype, which was first classified as MPS V, was found to be the milder form of MPS I (Scheie syndrome)**					
MPS VI (Maroteaux-Lamy)		Arylsulfatase B (N-acetylglucosamine 4-sulfatase)	*ARSB* 5q11-q13	dermatan sulfate	Dysostosis multiplex; corneal clouding; normal intelligence; survival to teens in severe form; milder forms also documented.	Feline [15]
MPS VII (Sly)		β-glucuronidase	*GUSB* 7q21.11	dermatan, keratan and chondroitin sulfate	Dysostosis multiplex; hepatomegaly; wide spectrum of severity including fetal hydrops and neonatal form.	Canine [16]; spontaneous mouse mutant [17]; mouse [18]
MPS VIII	*∗The designation MPS VIII was based on incorrect data and is no longer used.∗*					
MPS IX		Hyaluronidase 1	*HYAL* 3p21.3			Mouse [19]

them from fusiform to round. As a consequence, the valve leaflet and cordae tendinea become thickener and interfere with normal cardiac function, producing valvular stenosis. At corneal level, also, storage of undegraded GAGs results in reflection and refraction of light, leading to the cloudiness which is so typical of these pathologies. Also at the CNS level, swollen neurons and lysosomes may produce lesions that include the development of meganeurites and neurite sprouting (reviewed in [3, 4]).

Traditionally, MPSs are recognized through analysis of urinary GAGs. Several methods have been devised, to precise qualitative identification and quantitative measurements. These analyses of urinary GAGs allow discrimination between broad classes of MPSs but cannot distinguish subgroups. Definitive diagnosis is usually accessed through enzymatic assays of the defective enzyme in cultured fibroblasts, leukocytes, and serum or plasma (reviewed in [1]). During the last decade; however, dried blood spot technology was also introduced for enzymatic assays, allowing cheaper, easier, feasible diagnosis and opening the possibility for large population screenings (see Section 11 for more details).

In general, MPSs are transmitted in an autosomal recessive fashion, except for MPS II, which is X-linked.

This paper provides an overview of the molecular basis, enzymatic defects, clinical manifestations, and diagnosis of each glycosaminoglycan storage disease, focusing also on the respective animal models and describing potential perspectives of therapy which are being tested as well as the ones which are already available (summarized in Table 2).

2. Mucopolysaccharidosis I

Mucopolysaccharidosis I is caused by a deficiency of α-L-iduronidase (IDUA; EC 3.2.1.76) and can result in a wide range of phenotypic involvement with three major recognized clinical entities: Hurler (MPS IH; MIM#607014), Hurler-Scheie (MPS IH/S; MIM#607015), and Scheie (MPS IS) syndromes. Hurler and Scheie syndromes represent phenotypes at the severe and mild ends of the MPS I clinical spectrum; respectively, and the Hurler-Scheie syndrome is intermediate in phenotypic expression [20]. It is important to stress that, although MPS I may be subdivided into these three clinically diverse entities, the underlying enzymatic defect is common to all of them, being all caused by mutation in the gene encoding α-L-iduronidase (*IDUA*).

Functionally, α-L-iduronidase is essential to the correct metabolism of both dermatan sulfate and of heparan sulfate, hydrolyzing the terminal α-L-iduronic acid residues of the above-referred glycosaminoglycans [1].

In 1992, Scott and colleagues [21] were able to clone and purify the gene that encodes this enzyme, *IDUA*, demonstrating that it spans approximately 19 kb and contains 14 exons. The first 2 exons are separated by an intron of 566 bp, a large intron of approximately 13 kb follows, and the last 12 exons are clustered within 4.5 kb. Previously, this gene was mapped to 4p16.3, through unequivocal in situ hybridization and southern blot analysis of mouse-human cell hybrids [22].

There are, presently, several animal models known for MPS I.

In 1979, Haskins and colleagues [5] described α-L-iduronidase deficiency in a cat, and, few years later, Shull et al. [6] and Spellacy et al. [23] reported a similar deficiency in the dog. Subsequent studies lead to cloning and characterization of the canine IDUA gene as well as the mutation causing the observed phenotype [24, 25] and proved it to be a good model for study of human MPS I. So, in 1994, Shull and collaborators [26] published the first results of enzyme replacement therapy in the canine model. Through intravenous administration of recombinant human α-L-iduronidase, these authors managed to obtain a remarkable resolution of lysosomal storage in both hepatocytes and Kupffer's cells. In the same year, Grosson et al. [27] mapped the homologous IDUA locus in the mouse to chromosome 5. That knowledge was later used to create a knock-out mouse presenting the characteristic MPS I features [7, 28].

Currently, both hematopoietic stem cell transplantation (HSCT) and enzyme replacement therapy (ERT) using laronidase (recombinant human α-L-iduronidase, Aldurazyme) are available for MPS I. HSCT is the recommended treatment for patients with severe MPS I, before 2 years of age [29–32]. ERT is recommended for the other cases, and it has been shown to be effective in ameliorating some of the clinical manifestations of MPS disease. Among positive effects are decreased hepatosplenomegaly, improved respiratory and myocardial function and physical capacity [33–35] as well as improvement in active movement followed by enhanced self-care [36]. Recently, several reports have been published trying to evaluate long-term effect of ERT on the natural history of treated patients. From those studies, several conclusions have been reached. Concerning treated patients' growth pattern, it became clear that children with MPS I grow considerably slower than healthy individuals, and differences between healthy and affected children increase with age [37]. Other relevant evidences show that early treatment of attenuated MPS I may significantly delay or prevent the onset of the major clinical signs, substantially modifying the natural history of the disease [38]. Additional investigation is needed to clarify the mechanisms by which improvements are achieved in laronidase-treated patients. Such knowledge may support the development of ERT directly targeting the brain.

2.1. Hurler's Syndrome (MPS IH). Hurler's syndrome is the most severe form of MPS I and has been, over the last decades, the prototype description of MPS. Nevertheless, this may be misleading, since not all MPSs share the same features, and this pathology in particular is not representative of all of them, but only of the most severe end of a broad clinical spectrum (reviewed in [1]). Like all other MPSs, the clinical course of this disease is progressive, with multiple organ and tissue involvement. Hallmark clinical features of Hurler syndrome include coarse facies, corneal clouding, mental retardation, hernias, dysostosis multiplex, and hepatosplenomegaly. Children with Hurler's syndrome appear normal at birth and develop the characteristic appearance over the first years of life [39]. Length is often normal until about 2 years of age when growth stops; by age of 3 years, height is under the third percentile [40]. Cardiac

Table 2: Available therapeutic approaches for mucopolysaccharidoses.

Pathology	Subtype	Available therapeutic approaches
MPS I	Hurler (H)	HSCT (recommended before 2 years of age)
	Hurler -Scheie (H/S)	ERT with Aldurazyme (laronidase; recombinant human α-L-iduronidase)
	Scheie (S)	ERT with Aldurazyme (laronidase; recombinant human α-L-iduronidase)
MPS II	Hunter	ERT with Elaprase (idursulfase; recombinant human iduronate sulfatase)
MPS III	Sanfilippo A	*Not available*
	Sanfilippo B	*Not available*
	Sanfilippo C	*Not available*
	Sanfilippo D	*Not available*
MPS IV	Morquio A	*ERT: ongoing clinical trial (with recombinant human GALNS)*
	Morquio B	*Not available*
MPS VI (Maroteaux-Lamy)		ERT with Naglazyme (galsulfase; recombinant human arylsulfatase B)
MPS VII (Sly)		*Not available*
MPS IX		*Not available*

disease and respiratory complications are common. Acute cardiomyopathy associated with endocardial fibroelastosis has been a presenting condition in some infants with MPS I less than 1 year of age [41]. Upper and lower respiratory tract infections are also frequent [42]. Developmental delay is often apparent by 12 to 24 months of age, with a maximum functional age of 2 to 4 years followed by progressive deterioration. Most children develop limited language as a consequence of developmental delay, chronic hearing loss, and enlarged tongue [1]. Dermal melanocytosis may also be found in Hurler patients [43], as well as in patients suffering from other LSDs, such as GM1 gangliosidosis. Nevertheless, Hurler's syndrome is the most common lysosomal storage disease associated with dermal melanocytosis, as revealed by a literature analysis.

2.2. Hurler-Scheie's Syndrome (MPS IH/S). MPS IH/S corresponds to a clinical phenotype which is intermediate between the Hurler and the Scheie syndromes. It is characterized by progressive somatic involvement with dysostosis multiplex but little or no mental retardation. First symptoms usually occur between 3 and 8 years. Characteristic features of Hurler's syndrome, such as corneal clouding, joint stiffness, deafness, and valvular heart disease, can also appear in MPS IH/S patients. Nevertheless, the onset of these symptoms occurs much later than that in the severe MPS I type, beginning in the midteens and leading to significant impairment and loss of function. Other clinical features, such as micrognathism, pachymeningitis cervicalis, and compression of the cervical cord due to GAG accumulation in the dura, may also occur. Cardiac and respiratory complications may explain the high clinical mortality (reviewed in [1]).

2.3. Scheie's Syndrome (MPS IS). Scheie's syndrome was earlier thought to be a separate entity designated MPS V, instead of a phenotypical subtype of MPS I [44]. This pathology is characterized by a mild phenotype in which dysostosis multiplex can be present. Joint involvement is marked in the hand with a claw-hand deformity. Patients also have genu valgum, stiff, painful feet, and *pes cavus* [1]. Cardiac and respiratory complications are much milder than in the Hurler syndrome, with aortic and mitral valvular disease being a common feature [45]. At a respiratory level, Perks et al. [46] have reported two brothers with Scheie's syndrome suffering from sleep apnea, but no other complications are known. Intelligence is normal [1]. Pachymeningitis cervicalis (compression of the cervical cord secondary to glycosaminoglycan accumulation in the dura) may also occur.

3. Mucopolysaccharidosis II (Hunter's Syndrome)

Mucopolysaccharidosis II is the sole MPS transmitted in an X-linked manner and is caused by deficiency of the lysosomal enzyme iduronate sulfatase, which is crucial to the correct degradation of heparan and dermatan sulfate, by cleaving their O-linked sulfate. As a result, there is a progressive accumulation of glycosaminoglycans in nearly all cell types, tissues, and organs. Patients with MPS II excrete excessive amounts of dermatan sulfate and heparan sulfate in the urine [20, 47]. Hunter syndrome is caused by mutation in the gene encoding iduronate-2-sulfatase (*IDS*).

Although the disease is known since the early 1970s, being the first MPS to be defined clinically in humans, it was not until the 1990s that the IDS was cloned. In 1991, Wilson et al. [48] localized the gene to Xq28. Two years later, Flomen and coworkers [49] described the gene's structure as containing 9 exons and characterized the intron sequences surrounding them. In the same year, Wilson et al. [50] reported the complete sequence of the *IDS* gene, which spans approximately 24 kb. The potential promoter for *IDS* lacks a TATA box but contains GC box consensus sequences, which are consistent with its role as a housekeeping gene.

Curiously, a second *IDS* gene (*IDS2*) was identified by Bondeson et al. [51]. It is a pseudogene and is located within 90 kb telomeric region of the *IDS* gene and involved in a recombination event with the primary *IDS* gene in about 13% of patients with the Hunter syndrome.

Traditionally, the Hunter syndrome comprises 2 recognized clinical entities, according to the severity of symptoms: mild and severe. Although largely used, this nomenclature does have its difficulties, since the mild and severe forms represent the two ends of a wide and continuous spectrum of clinical severity. Also, in terms of iduronate deficiency, these forms cannot be distinguished since the enzyme's activity is equally deficient in both (reviewed in [1]). They are, though, separated almost exclusively on clinical grounds, although nowadays mutation analysis may help distinguish them.

This classification of MPS goes back to 1972, when McKusick distinguished between the severe form (which he called MPS IIA), with progressive mental retardation and physical disability and death before age 15 years in most cases, and the mild form (called MPS IIB) compatible with survival to adulthood and in which intellect is impaired minimally, if at all. He also pointed out the lack of corneal clouding in the X-linked form of MPS as opposed to the autosomal forms.

Presently, this classification has become obsolete since, in 2008, Wraith et al. [47] stated that MPS II should be regarded as a continuum between the two extremes (severe and attenuated). They noted that, although the clinical course for the more severely affected patients is relatively predictable, there is considerable variability in the clinical phenotype and progression of the more attenuated form of the disease and, so, it would not be correct to consider the milder form as a separate entity but, instead, look at Hunter's disease as a phenotypical continuum, with several possible degrees of severity.

In 1998, Wilkerson et al. [8] described Hunter's syndrome in a Labrador retriever, with the typical clinical features observed in humans: coarse facies, macrodactyly, corneal dystrophy, progressive CNS deterioration, and positive biochemical diagnosis for MPS through urine analysis.

After the successful results obtained in improving certain disease manifestations in patients with MPS I, including visceral manifestations and attenuation of neurologic disease progression [29, 52], hematopoietic stem cells transplantation (HSCT) has also been performed in several patients with MPS II. Unfortunately, although the transplantation of hematopoietic stem cells provides some enzymatic reconstruction in many target tissues with decreased excretion of GAGs in urine, decreased liver and spleen volumes, diminished facial coarsening, and improved respiratory function and joint mobility [53, 54], the results at neurological level were disappointing (reviewed in [55]). The additional risk of morbidity and mortality associated to this procedure led investigators to focus their attention in ERT for this pathology, with much better results, as discussed below.

A knock-out mouse model for MPS II was developed by replacing exon 4 and a portion of exon 5 of *IDS* with the neomycin-resistance gene [9, 56]. Affected mice exhibit a phenotype with notorious similarities to human disease, both at the biochemical and the clinical levels [9]. Several studies with this knock-out mouse model were done to assess the effect of ERT [56] as well as dose and various dosing regimens of idursulfase in urine and tissue GAG levels [57]. The results of these studies were quite promising, with a marked decrease in urinary GAGs as well as decreased GAG accumulation in several tissues [56] verified for several idursulfase doses and several dosing frequencies [57]. These studies have been used to support the first clinical trial of recombinant *IDS* in Hunter's syndrome patients. At the moment, both phase I/II [58] and phase II/III [59] clinical studies have proven not only the efficacy but also the safety of idursulfase replacement therapy. Consequently, ERT with recombinant human iduronate sulfatase (Elaprase, idursulfase, Shire Human Genetic Therapies Inc.) was approved in the US (July, 2006) and the European Union (January, 2007) for the treatment and the management of MPS II. The recommended dose is 0.5 mg/kg administered once weekly as an intravenous infusion (reviewed in [55]). As time goes by, additional evidence on the efficacy of ERT for MPS II patients is being published, as long-term treatments are successful. This is the case of a recently published report on the improvements observed in a 7 years and 10 months old child who began a 36 months' treatment with Elaprase at 4 years and 10 months. At the end of the treatment, the child presented normal excretion of GAGs in urine, normal-sized liver and spleen, and significant bone remodeling. Cardiac and neurological development, however, still progressively deteriorated [60]. This year, protective effects of ERT in MPS II patients were also reported for DNA damaging in leukocytes [61] and oxidative stress [62].

4. Mucopolysaccharidosis III (Sanfilippo's Syndrome)

The Sanfilippo syndrome, or mucopolysaccharidosis III, is caused by impaired degradation of heparan sulfate [1] and includes 4 subtypes, each due to the deficiency of a different enzyme: heparan N-sulfatase (type A; MIM no. 252900), α-N-acetylglucosaminidase (type B; MIM no. 252920), acetyl CoA: α-glucosaminide acetyltransferase (type C; MIM no. 252930), and N-acetylglucosamine-6-sulfatase (type D; MIM#252940). At a clinical level, the four subtypes are quite similar, with a characteristic severe central nervous system degeneration associated with mild somatic disease. Onset of clinical features usually occurs between 2 and 6 years, severe neurologic degeneration occurs in most patients between 6 and 10 years of age, and death occurs typically during the second or third decade of life. Type A has been reported to be the most severe, with earlier onset and rapid progression of symptoms and shorter survival [63].

4.1. Mucopolysaccharidosis IIIA (Sanfilippo A). General MPS IIIA clinical features include severe mental retardation with relatively mild somatic features (moderately severe claw hand and visceromegaly, little or no corneal clouding, little or no vertebral change). Usually, this pathology is characterized by marked overactivity, destructive tendencies, and other behavioral aberrations.

MPS IIIA phenotype is caused by mutations in the gene encoding N-sulfoglucosamine sulfohydrolase, also named heparan sulfate sulfatase (*SGSH*; 605270). This enzyme is specific for sulfate groups linked to the amino group of glucosamine.

In 1995, the gene encoding N-sulfoglucosamine sulfohydrolase, *SGSH*, was isolated, sequenced, and cloned [64]. Later, it was shown to contain 8 exons spanning approximately 11 kb [65].

There are two animal models known for MPS IIIA. The first to be discovered was the canine model when Fischer et al. [10] identified sulfaminidase deficiency in two adult wire-haired dachshund littermates. Subsequently, Aronovich et al. [66] determined the normal sequence of the canine heparan sulfate sulfatase gene and cDNA, through PCR-based approaches. Another model was described in 2001, when Bhattacharyya and collaborators [11] found a spontaneous mouse mutant of MPS IIIA resulting from a missense mutation (D31N) in the murine sulfatase gene. Affected mice die at about 10 months of age, exhibiting notorious visceromegaly, distended lysosomes and heparan sulfate accumulation in urine. Hemsley and Hopwood [67] found that these mice had severe brain involvement, with impaired open field locomotor activity and behavioral changes, suggesting axonal degeneration. Later, Settembre et al. [68] observed increased autophagosomes resulting from autophagosome-lysosome function in these mice. Similar findings were observed in another mouse model of another lysosomal storage disorder (multiple sulfatase deficiency; MSD; MIM no. 272200), reinforcing the recent idea that these diseases are disorders of autophagy, which may be a common mechanism for neurodegenerative lysosomal storage disorders.

MPS IIIA mice were recently tested for substrate deprivation therapy with both genistein and rhodamine B, two chemicals that inhibit GAG synthesis ([4, 69], reviewed in [70]). Encouraging results were obtained with both compounds, and this therapeutic approach started to be considered for several MPSs (see Section 11 for more details). Other interesting results were also obtained when siRNAs were used to reduce GAG synthesis in MPS IIIA mice. Last year, this approach was tested by Dziedzic et al. [71], who managed to reduce mRNA levels of four genes, *XYLT1, XYLT2, GALTI*, and *GALTII*, whose products are involved in GAG synthesis. This decrease of levels of transcripts corresponded to a decrease in levels of proteins encoded by them. Moreover, efficiency of GAG production in these fibroblasts was considerably reduced after treatment of the cells with siRNA. Either way, substrate deprivation therapy seems to be a promising approach for Sanfilippo's syndrome type A.

Gene therapy approaches are also being tested in MPS IIIA mice. Recently, promising results have been reported by Fraldi et al. [72], who performed experiments with intracerebral adeno-associated-virus- (AAV-) mediated delivery of *SGSH* gene, together with *SMUF1* gene, which exhibits an enhancing effect on sulfatase activity when coexpressed with sulfatases. They observed a visible reduction in lysosomal storage and inflammatory markers in transduced brain regions, together with an improvement in both motor and cognitive functions.

4.2. Mucopolysaccharidosis IIIB (Sanfilippo B). With a phenotype quite similar to MPS IIIA, the Sanfilippo syndrome B is characterized by deficiencies of α-N-acetylglucosaminidase, caused by mutations in the *NAGLU* gene that encodes this enzyme. α-N-Acetylglucosaminidase is required for the removal of the N-acetylglucosamine residues that exist in heparan sulfate or are generated during lysosomal degradation of this polymer by the action of heparan acetyl-CoA: α-glucosaminide N-acetyltransferase (reviewed in [1]).

The *NAGLU* gene was cloned in 1995 by Zhao and colleagues [73]. The deduced 743-amino acid protein has a 20- to 23-residue leader sequence, consistent with a signal peptide, and 6 potential N-glycosylation sites. It contains 6 exons and spans 8.3 kb on chromosome 17q21 [74].

Similarly to the above-referred MPS III syndrome, there is also a natural occurring mutant for Sanfilippo B. It was described by Ellinwood and coworkers, in 2003, in Schipperke's dogs [12].

During the last decade, Li et al. [75] created a laboratorial murine MPS IIIB was also constructed trough targeted disruption of the *NAGLU* gene [76]. With a phenotype quite similar to that of patients with MPS IIIB, this model began immediately to be used for therapeutic approaches as well as for pathogenesis studies. The first studies were done to evaluate the potential of ERT for this pathology [76]. The results, however, were quite disappointing since the recombinant NAGLU produced in Chinese hamster ovary (CHO) cells was not efficiently captured by MPS IIIB cells, either *in vitro* [77, 78] or *in vivo* [76]. This difficulty has turned the search for a treatment for MPS IIIB even more challenging. Presently, several therapies are under evaluation for this disease, including cell-mediated therapy, enzyme enhancement therapy, substrate deprivation therapy, and viral gene therapy (reviewed in [79]).

Promising results are being achieved through gene therapy approaches in MPS IIIB mice, namely, through direct microinjection into the brain of adeno-associated virus (AAV) vectors coding for NAGLU [80–82] and intravenous injections and intracranial gene delivery of lentiviral (LV) vector of NAGLU [83–85].

4.3. Mucopolysaccharidosis IIIC (Sanfilippo C). Sanfilippo syndrome C is, in general, characterized by the same clinical features described to MPS IIIA. Nevertheless, the enzyme deficiency in this pathology is different from the one causing the latter. Type C disease is caused by mutations in the gene encoding heparan acetyl-CoA: α-glucosaminide N-acetyltransferase (*HGSNAT*; 610453). This is the only known lysosomal enzyme that is not a hydrolase. It catalyzes the acetylation of the glucosamine amino groups that have become exposed by the action of heparan-N-sulfatase (reviewed in [1]).

The *HGSNAT* gene was cloned in parallel by two different groups, during the last decade: Fan et al. [86] and Hřebíček et al. [87]. The molecular defects underlying MPS IIIC remained unknown for almost three decades due to the low tissue content and the instability of *HGSNAT* [88].

To date, 54 *HGSNAT* sequence variants have been identified including 13 splice-site mutations, 11 insertions and deletions with consequent frameshifts and premature termination of translation, 8 nonsense, and 18 missense (reviewed in [89]).

Recently, two independent studies from Feldhammer et al. [88] and Fedele and Hopwood [90] have performed exhaustive functional analysis of the majority of the missense mutations already reported for the *HGSNAT* gene. Attention was focused in this particular type of mutations since there are several MPS IIIC patients carrying only missense mutations, either homozygous or heterozygous, who present an unexpected severe phenotype. In fact, although splicing and frameshift mutations are usually associated to that type of phenotype, since they give rise to premature termination codons and trigger nonsense-mediated mRNA decay (NMD); missense mutations are traditionally associated to milder disease. Nevertheless, this typical/general pattern is not observed for MPS IIIC. That is why these alterations were specifically cloned, expressed, and analyzed for their folding, targeting, and enzymatic activities. As a result, Fedele and Hopwood [90] have observed that the expression levels and enzymatic activity of most mutants were extremely low or even negligible. Feldhammer and colleagues [88], on the other hand, have observed that those mutations cause a misfolding of the enzyme, which is not correctly glycosylated. As a consequence, HGSNAT is not targeted to the lysosome but, instead, stays in the endoplasmic reticulum (ER). Thus, enzyme folding defects due to missense mutations, together with NMD seem to be the major molecular mechanisms underlying MPS IIIC. This makes MPS IIIC a good candidate for enzyme enhancement therapy, where active site-specific inhibitors are used as pharmacological chaperones to modify the conformation of the mutant lysosomal enzymes usually retained and degraded in the ER, in order to increase the level of the residual activity to a point which is sufficient to reverse the clinical phenotypes [88]. Together with inhibitors of heparan sulphate synthesis, pharmacological chaperones are currently being tested to reduce storage of this polymer in the CNS to levels sufficient to stop neuronal death and reverse inflammation.

4.4. Mucopolysaccharidosis IIID (Sanfilippo D). Like the previous MPS III subtypes, Sanfilippo's syndrome D presents a phenotype similar to MPS IIIA, with a singular enzyme deficiency underlying it: mutation in the gene encoding N-acetylglucosamine-6-sulfatase (GNS; 607664). The enzyme was originally described as specific for the 6-sulphated N-acetylglucosamine residues of heparan sulphate. However, the early data have been reinterpreted, and given that this sulfatase is in fact able to desulphate the 6-sulphated N-acetylglucosamine present in α- or in β-linkage or even as a free monosaccharide (reviewed in [1]).

N-Acetylglucosamine-6-sulfatase (EC 3.1.6.14) was purified and characterized by Freeman et al. [91], who identified 4 different forms of the enzyme in liver. Its catalytic properties were studied by Freeman and Hopwood [92]. Afterwards, Robertson et al. [93] assigned the glucosamine-6-sulfatase gene, which they symbolized *G6S*, to chromosome 12q14 by *in situ* hybridization of a tritium-labeled *G6S* cDNA probe. The localization was confirmed by using the cDNA clone in analyses of DNA from human/mouse hybrid cell lines. More recently, that information was completed by the work of Mok et al. [94], who amplified and sequenced the promoter and 14 exons of the *GNS* gene from a patient with MPS IIID. By analyzing that patient, it was also possible to identify a homozygous nonsense mutation in exon 9, predicted to result in premature termination at codon 355, as well as two common synonymous coding SNPs. At the same time, another group identified a 1-bp deletion in the *GNS* gene in another affected individual [95].

A naturally occurring large animal model was described by Thompson et al. [13], who reported type D Sanfilippo's syndrome in a Nubian goat. Later, caprine MPS IIID was used to evaluate the efficacy of ERT in this pathology. Recombinant caprine N-acetylglucosamine-6-sulfatase was administered intravenously to one MPS IIID goat at 2, 3, and 4 weeks of age. As a result, a marked reduction of lysosomal storage vacuoles was observed in hepatic cells, but no amelioration was noticed concerning the CNS lesions. No residual enzyme activity was observed either in brain or liver. Taking this preliminary results into account, it was considered that other treatment regimens will be necessary for MPS IIID [96].

5. Mucopolysaccharidosis IV

Mucopolysaccharidosis IV, or Morquio's syndrome, is caused by impaired degradation of keratan sulphate. Presently, there are two known enzyme deficiencies causing 2 different subtypes of Morquio's syndrome: deficiency in N-acetylglucosamine-6-sulfatase (causing Morquio's disease type A; MIM no. 253000) and deficiency in β-galactosidase (causing Morquio's disease type B; MIM no. 253010). Both MPS IV subtypes present a wide spectrum of clinical manifestations, but there are some characteristic common features: short trunk dwarfism, fine corneal deposits, spondyloepiphyseal dysplasia. Actually, the predominant clinical features of Morquio's syndrome are the ones related to the skeleton. Most of the times, this severe somatic disease is accompanied by a normal intelligence [1]. Patients with the severe phenotype do not normally survive past the second or third decade of life [97].

5.1. Morquio's Syndrome Type A. Morquio's syndrome A is caused by mutations in the gene encoding galactosamine-6-sulphate sulfatase (*GALNS*), which plays a crucial role on the degradation of both keratan sulphate and chondroitin sulphate.

The gene coding for human galactosamine-6-sulphate sulfatase (*GALNS*), was mapped to chromosome 16q24.3 trough fluorescence *in situ* hybridization assays [98]. Its structure was described at the same time by independent groups as comprising 14 exons and spanning approximately 40–50 kb [99, 100]. Curiously, the *GALNS* gene contains an *Alu* repeat in intron 5 and a VNTR-like sequence in intron 6 [100].

No natural occurring model is known for either type A or type B Morquio's syndrome. Nevertheless, a laboratorial murine model for type A syndrome was created from an induced disruption in exon 2 of the *GALNS* gene. Mutants presented no detectable enzyme activity and showed

increased GAG levels in urine. GAGs accumulation was also detected in several tissues including liver, kidney, spleen, heart, brain, and bone marrow [14]. These mice were later tested for enzyme replacement therapy and, after a 12-week long treatment with native GALNS or SUMF1-modified GALNS, showed manifest clinical improvement, demonstrated by a marked reduction of storage material in visceral organs, bone marrow, heart valves, ligaments, and connective tissue. The clearance of stored material in brain was dose dependent, and the keratan sulphate blood levels were reduced to normal [101].

Presently, there is no effective therapy for MPS IVA and care has been palliative, as in the majority of LSDs. Enzyme replacement therapy (ERT) and hematopoietic stem cells therapy (HSCT) have been considered as potential therapeutic approaches for MPS IVA (reviewed in [102]), ERT being, though, the most attractive candidate, since affected patients lack CNS involvement.

Recently, Rodríguez et al. [103] have produced a recombinant GALNS enzyme in *Escherichia coli* BL21. To produce sufficient amounts of purified GALNS enzyme, high level expression of GALNS in Chinese hamster ovary (CHO) cells has been established as a source of selectively secreted human recombinant enzyme. This recombinant enzyme has already been tested in the murine knock-out model, with consequent clearance of tissue and blood keratan sulphate [101]. These results provided important preclinical data for the design of GALNS ERT trials, which are now in course.

5.2. Morquio's Syndrome Type B. Although presenting overlapping clinical features, Morquio's syndrome B is genetically distinct from Morquio's syndrome A, being caused by impairments in another enzyme involved in the stepwise degradation of keratan sulphate: β-galactosidase, which is coded by the *GLB1* gene. Beta galactosidase hydrolases terminal β-linked galactose residues found in GM1 ganglioside, glycoproteins, and oligosaccharides, as well as in keratan sulphate (reviewed in [1]).

The *GLB1* gene spans 62.5 kb and contains 16 exons [104, 105] and maps to chromosome 3p21.33 [106]. The deduced 677-residue protein has a calculated molecular mass of 75 kD and contains a putative 23-residue signal sequence and 7 potential asparagine-linked glycosylation sites. It may be interesting to refer that the *GLB1* gene gives rise to 2 alternatively spliced mRNAs: a major 2.5-kb transcript that encodes the classic lysosomal form of the enzyme of 677 amino acids, and a minor 2.0-kb transcript that encodes a β-galactosidase-related protein (elastin-binding protein, EBP) of 546 amino acids with no enzymatic activity and a different subcellular localization. Exons 3, 4, and 6 are absent in the 2.0-kb mRNA as a consequence of alternative splicing of the pre-mRNA [107–109].

Presently, there are no known animal models for MPS IVB, either natural or engineered.

6. Mucopolysaccharidosis V

The designation MPS V is no longer used. In fact, the phenotype which was first classified as MPS V, was later found to be the milder form of MPS I (Scheie's syndrome), caused by deficiencies in α-L-iduronidase, with the typical stiff joints, clouding of the cornea most dense peripherally, survival to a late age with little if any impairment of intellect and aortic regurgitation [44].

7. Mucopolysaccharidosis VI (Maroteaux-Lamy Syndrome)

Mucopolysaccharidosis type VI is an autosomal recessive lysosomal storage disorder resulting from a deficiency of arylsulfatase B (N-acetylgalactosamine-4-sulfatase). Clinical features and severity are variable but usually include short stature, hepatosplenomegaly, dysostosis multiplex, stiff joints, corneal clouding, cardiac abnormalities, and facial dysmorphism. Intelligence is usually normal [110].

Arylsulfatase B is a lysosomal enzyme that removes the C4 sulphate ester group from the N-acetylgalactosamine sugar residue at the nonreducing terminus of dermatan sulphate and chondroitin sulphate, during lysosomal degradation [111]. The gene that codes for this enzyme was first mapped to chromosome 5q11-q13 [112] and is now known to contain 8 exons and span about 206 kb [111].

In 2002, a 3-year-old Siamese/short-haired European cat was referred for clinical disease characterized by dwarfism, facial dysmorphia, paralysis, small and curled ears, corneal clouding, and large areas of alopecia. X-ray examination showed multiple bone dysplasias. These features lead to suspect from a mucopolysaccharide storage disorder. Subsequent analysis proved it to be a natural occurring form of the Maroteaux-Lamy syndrome [15].

This MPS VI model has been extensively used over the last years to test ERT for this specific pathology. In 2003, Auclair and colleagues [113] have evaluated the cats' response to infusions of recombinant human N-acetylgalactosamine-4-sulfatase (rhASB) and observed an overall improvement in the disease condition at physical, neurological, and skeletal levels. Later, the same team has demonstrated that a high rate of immunotolerance towards rhASB can be achieved in MPS VI cats with a short-course tolerisation regimen [114], which may help the implementation of such procedures. Another interesting approach was designed, specifically to ameliorate joint disease in MPS IVA, through long-term articular administration of rhASB, leading to a notorious improvement in feline joint disease [115]. These successful results lead to the development of clinical trials in MPS VI patients, and three clinical studies including 56 patients have evaluated the efficacy and safety. As a consequence, enzyme replacement therapy (ERT) became available. The specific ERT for MPS VI, galsulfase (Naglazyme, Biomarin Pharmaceutical) was approved in 2005 by FDA and in 2006 by EMA. Long-term follow-up data with patients treated up to 5 years showed that ERT is well tolerated and associated with sustained improvements in the patients' clinical condition [2, 116].

Even though presently there is ERT available for these patients, other therapeutic approaches are being tested in animal models for MPS VI. In 2009, the first attempt of

successful gene therapy was performed through lentiviral-mediated gene transfer to joint tissues of the rat, with consequent correction of MPS VI cells [117]. This year, another study, involving intravascular administration of adeno-associated viral vectors in MPS VI cats, was published. After gene transfer the authors observed clearance of GAG storage, improvement of long bone length, reduction of heart valve thickness, and improvement in spontaneous mobility [118]. Either way, promising therapeutic strategies for MPS VI patients may be arising.

8. Mucopolysaccharidosis VII (Sly's Syndrome)

MPS VII, also known as Sly's syndrome, is characterized by the impossibility to degrade glucuronic acid-containing GAGs, due to impaired function of β-glucuronidase, which removes the glucuronic acid residues present in dermatan sulphate as well as in heparan and chonroitin sulphates (reviewed in [1]). Clinical features are highly variable, with phenotypes ranging from severe fetal hydrops to mild forms allowing survival into adulthood. Typical features include hepatomegaly, skeletal abnormalities, coarse facial features, and variable degrees of mental impairment [119].

MPS VII was first reported by Sly and collaborators in 1973, in a boy with skeletal changes consistent with MPS, hepatosplenomegaly, and granular inclusions in granulocytes. Additional features included hernias, unusual facies, protruding sternum, thoracolumbar gibbus, vertebral deformities, and mental deficiency. When β-glucuronidase activity was measured in fibroblasts, obtained values were less than 2% of control values. Both parents and several sibs of the mother showed an intermediate level of the enzyme [120].

In 1990, Miller et al. [121] reported that the gene encoding β-glucuronidase (*GUSB)* is 21 kb long, contains 12 exons, and gives rise to two different types of cDNAs, through an alternate splicing mechanism. Speleman et al. [122] used fluorescence in situ hybridization to map the *GUSB* gene to 7q11.21-q11.22. This map position was confirmed by dual-color hybridization of β-glucuronidase and another gene which had been mapped proximal to it: elastin (7q11.23).

Several pseudogenes, located on chromosomes 5, 6, 7, 20, 22, and Y, were also detected by Shipley et al. [123], when amplifying exons 2–4, 3, 6-7, and 11.

In 2009, Tomatsu et al. [124] provided a review of mutations in the *GUSB* gene that cause MPS type VII. Forty-nine different pathogenic mutations have been reported in the literature, with approximately 90% of them being missense mutations. Approximately 40% of the known *GUSB* mutations occur at CpG sites within the gene. The most common mutation is L176F, which has been found in several populations: American (Caucasian), Brazilian, British, Chilean, French, Mexican, Polish, Spanish, and Turkish ([125–127], reviewed in [124]). Genotype/phenotype analysis indicated that the most severe phenotype was associated with truncating mutations and with mutations affecting either the hydrophobic core or the modification of packing.

In 1984, mucopolysaccharidosis type VII (Sly syndrome) was described in a mixed-breed dog [16]. Since then, several other affected dogs have been studied, in the animal colony established at the University of Pennsylvania, the School of Veterinary Medicine [128] and, later, in a 12-week-old male German Shepherd dog studied in the same school [129]. All dogs shared the same missense mutation and developed similar phenotypes with skeletal deformities, corneal cloudiness, cytoplasmic granules in the neutrophils and lymphocytes of blood and CSF, and glycosaminoglycans in urine [129]. Another animal model was described as naturally occurring: the $gus^{mps/mps}$ mouse, which has a 1 bp deletion in exon 10 resulting in a progressive degenerative disease that reduces lifespan and causes facial dysmorphism, growth retardation, deafness, and behavioral defects [17]. Nevertheless, opportunities for experimental therapies were greatly expanded by the work of Tomatsu et al., in 2006 [18], who developed a new MPS VII mouse model, which is tolerant to both human and murine GUS, without the characteristic immune responses that complicated evaluation of the long-term benefits of enzyme replacement or gene therapy when the naturally occurring mice were used. Ever since, several therapeutic approaches have been attempted in MPS VII mice, and the results have been encouraging. That is the case of the works by Bosch and collaborators, who have been working on gene therapy for this pathology, in order to correct brain lesions. They have used both adeno-associated virus (AAV) [130] and lentivirus-mediated gene transfer [131] and observed that there was a significant correction of pathology in the brain of affected mice.

Other therapeutic approaches had already been attempted, but their results were not as promising. In fact, in 1998, allogenic bone marrow transplantation was reported in a 12-year-old Japanese girl with consequent improvement of motor function and daily life activities, decrease of upper respiratory and ear infections, but no improvement at all in cognitive function [132].

9. Mucopolysaccharidosis VIII

The clinical entity once known as MPS VIII was described in a single patient, in the late 1970s. The patient, a 5-year-old child, presented short stature, coarse hair, hepatomegaly, mild dysostosis multiplex, mental retardation, and no signs of corneal clouding. Biochemical analysis of the urine revealed increased excretion of keratan and heparan sulphate [133, 134]. The biochemical findings described by this group lead to suspect the existence of two hexosamine sulfatases and propose the existence of this novel MPS, caused by glucosamine-6-sulfatase [133].

Nevertheless, subsequent analysis on Diferrante's laboratory brought this idea down, and the designation MPS VIII was abandoned [135].

10. Mucopolysaccharidosis IX

Mucopolysaccharidosis IX, also known as hyaluronidase deficiency, is caused by mutations in the *HYAL1* gene.

This disease was first discovered by Natowicz et al. [136] in a 14-year-old girl with short stature and multiple periarticular soft-tissue masses. Radiographic analysis showed nodular synovia, acetabular erosions, and a popliteal cyst. Lysosomal storage of hyaluronan (HA) was evident within the macrophages and fibroblasts of biopsied soft-tissue masses, and serum concentrations were elevated 38–90-fold. She was proven to have a storage disease of hyaluronan (hyaluronic acid) due to a genetic deficiency of hyaluronidase. The descriptions of hyaluronidase deficiency in this family are consistent with autosomal recessive inheritance.

In order to determine the molecular basis of MPS IX, Triggs-Raine et al. [137] analyzed 2 different candidate genes tandemly distributed on chromosome 3p21.3, both encoding proteins with homology to a sperm enzyme with hyaluronidase activity. These genes, *HYAL1* and *HYAL2*, encode 2 distinct lysosomal hyaluronidases with different substrate specificities. When characterizing the patient with hyaluronidase deficiency originally reported in [136], they verified that he was a compound heterozygote for 2 mutations in the *HYAL1* gene: a missense mutation (c.1412G>A), which introduced a nonconservative amino acid substitution in a putative active site residue (p.Glu268Lys) and a complex intragenic rearrangement, 1361del37ins14, which resulted in a premature termination codon. Through this work, they have also showed that these 2 hyaluronidase genes, together with a third adjacent HYAL3 gene, had markedly different tissue expression patterns, consistent with differing roles in the metabolism of hyaluronan. These findings allowed this team not only to explain the unexpectedly mild phenotype of MPS IX but also to predict the existence of other hyaluronidase-deficiency disorders.

Presently, three other hyaluronidase-related genes (*HYAL4*, *HYALP1*, *SPAM1*) have been identified at 7q31.3 [138]. These genes are predicted to encode hyaluronidases, endoglycosidases that initiate the degradation of HA, a large negatively charged GAG found in the extracellular matrix (ECM) of all vertebrate cells [19].

Since there is only one patient reported to date, the development and characterization of a model of Hyal1 deficiency was the first logical step in understanding the main phenotypic symptoms associated with MPS IX. During this decade, a mouse model for MPS IX has become available and was fully characterized [19]. Overall, it was observed that the murine MPS IX model displays the key features of the human disease. Nevertheless, during the same year, another mutant mouse suffering from a hyaluronidase deficiency was described, this one deficient in HYAL2 [139]. Skeletal and hematological anomalies were described in this model, raising the possibility that a similar defect, defining a new MPS disorder, exists in humans [139].

11. Conclusion

The elucidation of enzyme deficiencies underlying mucopolysaccharidoses was crucial to unveil the normal pathways of glycosaminoglycan catabolism. In fact, only through the consequences of their absence became the role of several enzymes evident. The majority of these enzyme deficiencies were discovered during the 1970s. Over the last decades the enzyme deficiencies underlying each disease, and the molecular defects causing them have been identified and extensively analyzed and characterized. As a result, six different MPS are known, caused by deficiencies in one of the ten different enzymes necessary to intralysosomal degradation of GAGs through one of the four different degradation pathways.

Each disease has its own hallmark features. Nevertheless, a common pattern arouses: MPS are usually chronic, with a progressive course and different severity degrees. Organomegaly, dysostosis multiplex, and CNS involvement are common but not necessary features.

Over the years, several MPS have been recognized in animals as naturally occurring diseases, and others were created by knock-out technology. Most animal colonies have been established from single related heterozygous animals, in such a way that the affected offspring is homozygous for the same mutant allele. All these models present disease pathology similar to that seen in humans, making the animals extremely valuable for both investigation of disease pathogenesis and testing of therapies. Large animal homologues are similar to humans in natural genetic diversity, approaches to therapy and care, and possibility of evaluating long-term effects of treatment. Presently, therapeutic strategies for MPS include enzyme replacement therapy, heterologous bone marrow transplantation, and somatic cell gene transfer, all of which have been tested in animals with some success. During the 80s, transplantation of hematopoietic stem cells was tested for several MPSs. Theoretically, haematopoietic stem cells taken from a normal compatible donor and transplanted into an enzyme-deficient recipient can provide a safe, permanent, and self-replicating source of bone marrow-derived cells. By secreting active lysosomal enzymes, these cells cross-connect nonbone marrow-derived cells. Several animal models for GAGs storage diseases have already been subjected to/undergone BMT. From those experiments, along with human clinical trials already tried, it was possible to verify that there are important variations in therapeutic response among different pathologies with some diseases with CNS pathology which can be successfully treated by BMT (the severe form of MPS I, being the example for the GAG storage disorders; [140]) whereas others cannot (MPS II and III in which BMT was tested with few success; [3, 141]). These variations are usually attributed to the different capacities of secretion, stability, and uptake of each specific enzyme. Nevertheless, important conclusions could be drawn from the collective experience of postnatal transplantation including the idea that the earlier the transplants are performed, the better the clinical response. In the 90s; however, a novel approach started to be tested: ERT. Nowadays, it has been the most tested approach in animal models of GAG storage disorders. Until now, the obtained results have been highlighting the potential of administered recombinant enzyme to reduce GAG accumulation. ERTs are presently available for MPS I (since 2003), II (since 2005), and VI (since 2006). Clinical trials are also in course for MPS IVA treatment through ERT. Nevertheless, this approach is

ineffective for the brain since recombinant enzymes are not able to cross the blood-brain barrier (BBB). This is one of the reasons why other therapies are being tested for MPS with CNS involvement. ERT with direct administration of the recombinant enzyme into the brain (intrathecal injections) is also being considered in order to overcome that difficulty. Presently, such approaches are being considered for MPS IIIA and to overcome the cognitive deficit of MPS II and MPS I (reviewed in [142]).

Somatic cell gene transfer is another possible approach, but a long way needs yet to be travelled towards such a therapy is applicable to patients.

Finally, substrate deprivation therapy (or substrate reduction therapy) is also being considered for some MPSs. This approach is being tested with both genistein and rhodamine B (reviewed in [70]). Genistein, a chemical from the group of isoflavones, has been shown to inhibit the synthesis of GAGs in fibroblasts of patients with various forms of MPSs, namely, types I, II, IIIA, and IIIB [4, 143]. Similar results were obtained with rhodamine B, an inhibitor with an unknown mechanism of action. Remarkably, in MPS IIIA mice treated with rhodamine B, GAG storage decreased not only in somatic tissues, but also in brain, with improved behaviours of the animals [144, 145]. These encouraging results lead to the development of open-label pilot clinical studies with children suffering from Sanfilippo's syndrome types A and B in which a genistein-rich isoflavone extract (SE-2000, Biofarm, Poland) orally administered for 12 months. After one year of treatment, statistically significant improvement in all tested parameters was demonstrated (reviewed in [70]).

In order to better quantify and assess the efficacy of these therapeutic approaches, investigators have been trying to indentify suitable biomarkers for MPS, which allow the evaluation of short- and long-term treatment effects. This is also assuming particular importance since early detection of MPS is an important factor in treatment success. Recently discovered biomarkers include heparin cofactor II-thrombin complex and dermatan sulphate:chondroitin sulphate ratio [146]. Other biomarkers and/or therapeutic targets for MPS joint and bone disease recently identified through animal studies include several proinflammatory cytokines, nitric oxide, and matrix metalloproteinases (MMPs; [147]).

Another hot topic which is recently being discussed refers to the possibility of including some MPSs (particularly type I, IIIA, IIIB, and VI) in neonatal screening programs [148]. The ongoing development of enzyme replacement therapy and other treatments for several LSDs, including MPSs combined with the growing evidence that early commencement of therapy improves outcomes, has increased the pressure for the introduction of newborn screening programs, and a number of pilot studies are ongoing [148–152]. This is only possible thanks to the significant advances that were made in last decade since dried blood spot technology was introduced for enzymatic assays and lysosomal protein profile was developed.

Overall, there are encouraging results, with some therapeutic approaches already approved and others under development. Either way, it is important to stress that the management of MPS requires lifelong attention to the multisystemic involvement by a team of specialists experienced in dealing with these diseases, since none of the therapeutic options currently available result in complete resolution of morbidity.

References

[1] E. F. Neufeld and J. Muenzer, "The mucopolysaccharidoses," in *The Metabolic & Molecular Bases of Inherited Disease*, C. R. Scriver, A. L. Beaudet, W. S. Sly, and D. Valle, Eds., vol. 3, McGraw–Hill, New York, NY, USA, 8th edition, 2001.

[2] R. Giugliani, C. G. Carvalho, S. Herber, and L. L. de Camargo Pinto, "Recent advances in treatment approaches of mucopolysaccharidosis VI," *Current Pharmaceutical Biotechnology*, vol. 12, no. 6, pp. 956–962, 2011.

[3] M. Haskins, M. Casal, N. M. Ellinwood, J. Melniczek, H. Mazrier, and U. Giger, "Animal models for mucopolysaccharidoses and their clinical relevance," *Acta Paediatrica*, vol. 91, no. 439, supplement, pp. 88–97, 2002.

[4] E. Piotrowska, J. Jakóbkiewicz-Banecka, S. Barańska et al., "Genistein-mediated inhibition of glycosaminoglycan synthesis as a basis for gene expression-targeted isoflavone therapy for mucopolysaccharidoses," *European Journal of Human Genetics*, vol. 14, no. 7, pp. 846–852, 2006.

[5] M. E. Haskins, P. F. Jezyk, R. J. Desnick, S. K. McDonaugh, and D. F. Patterson, "Alpha L iduronidase deficiency in a cat: a model of mucopolysaccharidosis I," *Pediatric Research*, vol. 13, no. 11, pp. 1294–1297, 1979.

[6] R. M. Shull, R. J. Munger, and E. Spellacy, "Canine α-L-iduronidase deficiency. A model of mucopolysaccharidosis I," *American Journal of Pathology*, vol. 109, no. 2, pp. 244–248, 1982.

[7] L. A. Clarke, C. S. Russell, S. Pownall et al., "Murine mucopolysaccharidosis type I: targeted disruption of the murine α-L-iduronidase gene," *Human Molecular Genetics*, vol. 6, no. 4, pp. 503–511, 1997.

[8] M. J. Wilkerson, D. C. Lewis, S. L. Marks, and D. J. Prieur, "Clinical and morphologic features of mucopolysaccharidosis type II in a dog: naturally occurring model of Hunter syndrome," *Veterinary Pathology*, vol. 35, no. 3, pp. 230–233, 1998.

[9] A. R. Garcia, J. Pan, J. C. Lamsa, and J. Muenzer, "The characterization of a murine model of mucopolysaccharidosis II (Hunter syndrome)," *Journal of Inherited Metabolic Disease*, vol. 30, no. 6, pp. 924–934, 2007.

[10] A. Fischer, K. P. Carmichael, J. F. Munnell et al., "Sulfamidase deficiency in a family of dachshunds: a canine model of mucopolysaccharidosis IIIA (Sanfilippo A)," *Pediatric Research*, vol. 44, no. 1, pp. 74–82, 1998.

[11] R. Bhattacharyya, B. Gliddon, T. Beccari, J. J. Hopwood, and P. Stanley, "A novel missense mutation in lysosomal sulfamidase is the basis of MPS III A in a spontaneous mouse mutant," *Glycobiology*, vol. 11, no. 1, pp. 99–103, 2001.

[12] N. M. Ellinwood, P. Wang, T. Skeen et al., "A model of mucopolysaccharidosis IIIB (Sanfilippo syndrome type IIIB): N-acetyl-α-D-glucosaminidase deficiency in Schipperke dogs," *Journal of Inherited Metabolic Disease*, vol. 26, no. 5, pp. 489–504, 2003.

[13] J. N. Thompson, M. Z. Jones, G. Dawson, and P. S. Huffman, "N-acetylglucosamine 6-sulphatase deficiency in a Nubian

goat: a model of Sanfilippo syndrome type D (mucopolysaccharidosis IIID)," *Journal of Inherited Metabolic Disease*, vol. 15, no. 5, pp. 760–768, 1992.

[14] S. Tomatsu, K. O. Orii, C. Vogler et al., "Mouse model on N-acetylgalactosamine-6-sulfate sulfatase deficiency (Galns-/-) produced by targeted disruption of the gene defective in Morquio A disease," *Human Molecular Genetics*, vol. 12, no. 24, pp. 3349–3358, 2003.

[15] B. MacRì, F. Marino, G. Mazzullo et al., "Mucopolysaccharidosis VI in a Siamese/short-haired European cat," *Journal of Veterinary Medicine A*, vol. 49, no. 8, pp. 438–442, 2002.

[16] M. E. Haskins, R. J. Desnick, and N. DiFerrante, "β-glucuronidase deficiency in a dog: a model of human mucopolysaccharidosis VII," *Pediatric Research*, vol. 18, no. 10, pp. 980–984, 1984.

[17] E. H. Birkenmeier, M. T. Davisson, W. G. Beamer et al., "Murine mucopolysaccharidosis type VII. Characterization of a mouse with β-glucuronidase deficiency," *Journal of Clinical Investigation*, vol. 83, no. 4, pp. 1258–1266, 1989.

[18] S. Tomatsu, K. Sukegawa, G. G. Trandafirescu et al., "Differences in methylation patterns in the methylation boundary region of IDS gene in hunter syndrome patients: implications for CpG hot spot mutations," *European Journal of Human Genetics*, vol. 14, no. 7, pp. 838–845, 2006.

[19] D. C. Martin, V. Atmuri, R. J. Hemming et al., "A mouse model of human mucopolysaccharidosis IX exhibits osteoarthritis," *Human Molecular Genetics*, vol. 17, no. 13, pp. 1904–1915, 2008.

[20] V. A. McKusick, *Heritable Disorders of Connective Tissue*, Mosby, St. Louis, Mo, USA, 4th edition, 1972.

[21] H. S. Scott, X. H. Guo, J. J. Hopwood, and C. P. Morris, "Structure and sequence of the human α-L-iduronidase gene," *Genomics*, vol. 13, no. 4, pp. 1311–1313, 1992.

[22] H. S. Scott, L. J. Ashton, H. J. Eyre et al., "Chromosomal localization of the human α-L-iduronidase gene (IDUA) to 4p16.3," *American Journal of Human Genetics*, vol. 47, no. 5, pp. 802–807, 1990.

[23] E. Spellacy, R. M. Shull, G. Constantopoulos, and E. F. Neufeld, "A canine model of human α-L-iduronidase deficiency," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 80, no. 19, pp. 6091–6095, 1983.

[24] L. J. Stoltzfus, B. Sosa-Pineda, S. M. Moskowitz et al., "Cloning and characterization of cDNA encoding canine α-L-iduronidase. mRNA deficiency in mucopolysaccharidosis I dog," *Journal of Biological Chemistry*, vol. 267, no. 10, pp. 6570–6575, 1992.

[25] K. P. Menon, P. T. Tieu, and E. F. Neufeld, "Architecture of the canine IDUA gene and mutation underlying canine mucopolysaccharidosis I," *Genomics*, vol. 14, no. 3, pp. 763–768, 1992.

[26] R. M. Shull, E. D. Kakkis, M. F. McEntee, S. A. Kania, A. J. Jonas, and E. F. Neufeld, "Enzyme replacement in a canine model of Hurler syndrome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 26, pp. 12937–12941, 1994.

[27] C. L. S. Grosson, M. E. MacDonald, M. P. Duyao, C. M. Ambrose, S. Roffler Tarlov, and J. F. Gusella, "Synteny conservation of the Huntington's disease gene and surrounding loci on mouse chromosome 5," *Mammalian Genome*, vol. 5, no. 7, pp. 424–428, 1994.

[28] C. Russell, G. Hendson, G. Jevon et al., "Murine MPS I: insights into the pathogenesis of Hurler syndrome," *Clinical Genetics*, vol. 53, no. 5, pp. 349–361, 1998.

[29] C. Peters, E. G. Shapiro, and W. Krivit, "Neuropsychological development in children with Hurler syndrome following hematopoietic stem cell transplantation," *Pediatric Transplantation*, vol. 2, no. 4, pp. 250–253, 1998.

[30] C. Peters, E. G. Shapiro, and W. Krivit, "Hurler syndrome: past, present, and future," *Journal of Pediatrics*, vol. 133, no. 1, pp. 7–9, 1998.

[31] S. L. Staba, M. L. Escolar, M. Poe et al., "Cord-blood transplants from unrelated donors in patients with Hurler's syndrome," *New England Journal of Medicine*, vol. 350, no. 19, pp. 1960–1969, 2004.

[32] J. J. Boelens, R. F. Wynn, A. O'Meara et al., "Outcomes of hematopoietic stem cell transplantation for Hurler's syndrome in Europe: a risk factor analysis for graft failure," *Bone Marrow Transplantation*, vol. 40, no. 3, pp. 225–233, 2007.

[33] L. A. Clarke, J. E. Wraith, M. Beck et al., "Long-term efficacy and safety of laronidase in the treatment of mucopolysaccharidosis I," *Pediatrics*, vol. 123, no. 1, pp. 229–240, 2009.

[34] J. E. Wraith, L. A. Clarke, M. Beck et al., "Enzyme replacement therapy for mucopolysaccharidosis I: a randomized, double-blinded, placebo-controlled, multinational study of recombinant human α-L-iduronidase (laronidase)," *Journal of Pediatrics*, vol. 144, no. 5, pp. 581–588, 2004.

[35] H. Harada, H. Uchiwa, M. Nakamura et al., "Laronidase replacement therapy improves myocardial function in mucopolysaccharidosis I," *Molecular Genetics and Metabolism*, vol. 103, no. 3, pp. 215–219, 2011.

[36] A. Tylki-Szymanska, J. Marucha, A. Jurecka, M. Syczewska, and B. Czartoryska, "Efficacy of recombinant human α-L-iduronidase (laronidase) on restricted range of motion of upper extremities in mucopolysaccharidosis type i patients," *Journal of Inherited Metabolic Disease*, vol. 33, no. 2, pp. 151–157, 2010.

[37] A. Tylki-Szymanska, A. Rozdzynska, A. Jurecka, J. Marucha, and B. Czartoryska, "Anthropometric data of 14 patients with mucopolysaccharidosis I: retrospective analysis and efficacy of recombinant human α-l-iduronidase (laronidase)," *Molecular Genetics and Metabolism*, vol. 99, no. 1, pp. 10–17, 2010.

[38] O. Gabrielli, L. A. Clarke, S. Bruni, and G. V. Coppa, "Enzyme-replacement therapy in a 5-month-old boy with attenuated presymptomatic MPS I: 5-year follow-up," *Pediatrics*, vol. 125, no. 1, pp. e183–e187, 2010.

[39] J. E. Wraith, J. G. Rogers, and D. M. Danks, "The mucopolysaccharidoses," *Australian Paediatric Journal*, vol. 23, no. 6, pp. 329–334, 1987.

[40] R. J. Gorlin, M. M. Cohen Jr., and R. C. M. Hennekam, *Syndromes of the Head and Neck*, Oxford University Press, New York, NY, USA, 4th edition, 2001.

[41] M. D. C. Donaldson, C. A. Pennock, P. J. Berry, A. W. Duncan, J. E. Cawdery, and J. V. Leonard, "Hurler syndrome with cardiomyopathy in infancy," *Journal of Pediatrics*, vol. 114, no. 3, pp. 430–432, 1989.

[42] J. Shapiro, M. Strome, and A. C. Crocker, "Airway obstruction and sleep apnea in Hurler and Hunter syndromes," *Annals of Otology, Rhinology and Laryngology*, vol. 94, no. 5 I, pp. 458–461, 1985.

[43] M. Hanson, J. R. Lupski, J. Hicks, and D. Metry, "Association of dermal melanocytosis with lysosomal storage disease: clinical features and hypotheses regarding pathogenesis," *Archives of Dermatology*, vol. 139, no. 7, pp. 916–920, 2003.

[44] V. A. McKusick, D. Kaplan, D. Wise et al., "The genetic mucopolysaccharidoses," *Medicine*, vol. 44, no. 6, pp. 445–483, 1965.

[45] C. B. Whitley, "The mucopolysaccharidoses," in *McKusick's Heritable Disorders of Connective Tissue*, P. Beighton, Ed., Mosby, St. Louis, Mo, USA, 5th edition, 1993.

[46] W. H. Perks, R. A. Cooper, and S. Bradbury, "Sleep apnoea in Scheie's syndrome," *Thorax*, vol. 35, no. 2, pp. 85–91, 1980.

[47] J. E. Wraith, M. Scarpa, M. Beck et al., "Mucopolysaccharidosis type II (Hunter syndrome): a clinical review and recommendations for treatment in the era of enzyme replacement therapy," *European Journal of Pediatrics*, vol. 167, no. 3, pp. 267–277, 2008.

[48] P. J. Wilson, G. K. Suthers, D. F. Callen et al., "Frequent deletions at Xq28 indicate genetic heterogeneity in Hunter syndrome," *Human Genetics*, vol. 86, no. 5, pp. 505–508, 1991.

[49] R. H. Flomen, E. P. Green, P. M. Green, D. R. Bentley, and F. Giannelli, "Determination of the organisation of coding sequences within the iduronate sulphate sulphatase (IDS) gene," *Human Molecular Genetics*, vol. 2, no. 1, pp. 5–10, 1993.

[50] P. J. Wilson, C. A. Meaney, J. J. Hopwood, and C. P. Morris, "Sequence of the human iduronate 2-sulfatase (IDS) gene," *Genomics*, vol. 17, no. 3, pp. 773–775, 1993.

[51] M. L. Bondeson, H. Malmgren, N. Dahl, B. M. Carlberg, and U. Pettersson, "Presence of an IDS-related locus (IDS2) and Xq28 complicates the mutational analysis of Hunter syndrome," *European Journal of Human Genetics*, vol. 3, no. 4, pp. 219–227, 1995.

[52] C. Peters, M. Balthazor, E. G. Shapiro et al., "Outcome of unrelated donor bone marrow transplantation in 40 children with Hurler syndrome," *Blood*, vol. 87, no. 11, pp. 4894–4902, 1996.

[53] G. V. Coppa, O. Gabrielli, L. Zampini et al., "Bone marrow transplantation in Hunter syndrome (mucopolysaccharidosis type II): two year follow-up of the first Italian patient and review of the literature," *Pediatria Medica e Chirurgica*, vol. 17, no. 3, pp. 227–235, 1995.

[54] E. J. R. McKinnis, S. Sulzbacher, J. C. Rutledge, J. Sanders, and C. R. Scott, "Bone marrow transplantation in Hunter syndrome," *Journal of Pediatrics*, vol. 129, no. 1, pp. 145–148, 1996.

[55] T. A. Burrow and N. D. Leslie, "Review of the use of idursulfase in the treatment of mucopolysaccharidosis II," *Biologics*, vol. 2, pp. 311–320, 2008.

[56] J. Muenzer, J. C. Lamsa, A. Garcia, J. Dacosta, J. Garcia, and D. A. Treco, "Enzyme replacement therapy in mucopolysaccharidosis type II (Hunter syndrome): a preliminary report," *Acta Paediatrica*, vol. 91, no. 439, supplement, pp. 98–99, 2002.

[57] A. R. Garcia, J. M. DaCosta, J. Pan, J. Muenzer, and J. C. Lamsa, "Preclinical dose ranging studies for enzyme replacement therapy with idursulfase in a knock-out mouse model of MPS II," *Molecular Genetics and Metabolism*, vol. 91, no. 2, pp. 183–190, 2007.

[58] J. Muenzer, M. Gucsavas-Calikoglu, S. E. McCandless, T. J. Schuetz, and A. Kimura, "A phase I/II clinical trial of enzyme replacement therapy in mucopolysaccharidosis II (Hunter syndrome)," *Molecular Genetics and Metabolism*, vol. 90, no. 3, pp. 329–337, 2007.

[59] J. Muenzer, J. E. Wraith, M. Beck et al., "A phase II/III clinical study of enzyme replacement therapy with idursulfase in mucopolysaccharidosis II (Hunter syndrome)," *Genetics in Medicine*, vol. 8, no. 8, pp. 465–473, 2006.

[60] F. Papadia, M. S. Lozupone, A. Gaeta, D. Capodiferro, and G. Lacalendola, "Long-term enzyme replacement therapy in a severe case of mucopolysaccharidosis type II (Hunter syndrome)," *European Review for Medical and Pharmacological Sciences*, vol. 15, no. 3, pp. 253–258, 2011.

[61] L. Filippon, C. A. Y. Wayhs, D. M. Atik et al., "DNA damage in leukocytes from pretreatment mucopolysaccharidosis type II patients; protective effect of enzyme replacement therapy," *Mutation Research*, vol. 721, no. 2, pp. 206–210, 2011.

[62] L. Filippon, C. S. Vanzin, G. B. Biancini et al., "Oxidative stress in patients with mucopolysaccharidosis type II before and during enzyme replacement therapy," *Molecular Genetics and Metabolism*, vol. 103, no. 2, pp. 121–127, 2011.

[63] J. J. P. Van De Kamp, M. F. Niermeijer, K. Von Figura, and M. A. H. Giesberts, "Genetic heterogeneity and clinical variability in the Sanfilippo syndrome (types A, B, and C)," *Clinical Genetics*, vol. 20, no. 2, pp. 152–160, 1981.

[64] H. S. Scott, L. Blanch, X. H. Guo et al., "Cloning of the sulphamidase gene and identification of mutations in Sanfilippo A syndrome," *Nature Genetics*, vol. 11, no. 4, pp. 465–467, 1995.

[65] L. E. Karageorgos, X. H. Guo, L. Blanch et al., "Structure and sequence of the human sulphamidase gene," *DNA Research*, vol. 3, no. 4, pp. 269–271, 1996.

[66] E. L. Aronovich, K. P. Carmichael, H. Morizono et al., "Canine heparan sulfate sulfamidase and the molecular pathology underlying Sanfilippo syndrome type A in Dachshunds," *Genomics*, vol. 68, no. 1, pp. 80–84, 2000.

[67] K. M. Hemsley and J. J. Hopwood, "Development of motor deficits in a murine model of mucopolysaccharidosis type IIIA (MPS-IIIA)," *Behavioural Brain Research*, vol. 158, no. 2, pp. 191–199, 2005.

[68] C. Settembre, A. Fraldi, L. Jahreiss et al., "A block of autophagy in lysosomal storage disorders," *Human Molecular Genetics*, vol. 17, no. 1, pp. 119–129, 2008.

[69] A. L. K. Roberts, B. J. Thomas, A. S. Wilkinson, J. M. Fletcher, and S. Byers, "Inhibition of glycosaminoglycan synthesis using rhodamine B in a mouse model of mucopolysaccharidosis type IIIA," *Pediatric Research*, vol. 60, no. 3, pp. 309–314, 2006.

[70] J. Jakóbkiewicz-Banecka, A. Wegrzyn, and G. Wegrzyn, "Substrate deprivation therapy: a new hope for patients suffering from neuronopathic forms of inherited lysosomal storage diseases," *Journal of Applied Genetics*, vol. 48, no. 4, pp. 383–388, 2007.

[71] D. Dziedzic, G. Wgrzyn, and J. Jakóbkiewicz-Banecka, "Impairment of glycosaminoglycan synthesis in mucopolysaccharidosis type IIIA cells by using siRNA: a potential therapeutic approach for Sanfilippo disease," *European Journal of Human Genetics*, vol. 18, no. 2, pp. 200–205, 2010.

[72] A. Fraldi, K. Hemsley, A. Crawley et al., "Functional correction of CNS lesions in an MPS-IIIA mouse model by intracerebral AAV-mediated delivery of sulfamidase and SUMF1 genes," *Human Molecular Genetics*, vol. 16, no. 22, pp. 2693–2702, 2007.

[73] H. G. Zhao, H. H. Li, A. Schmidtchen, G. Bach, and E. F. Neufeld, "The gene encoding alpha-N-acetylglucosaminidase and mutations underlying Sanfilippo B syndrome," *American Journal of Human Genetics*, vol. 57, p. A185, 1995.

[74] H. G. Zhao, H. H. Li, G. Bach, A. Schmidtchen, and E. F. Neufeld, "The molecular basis of Sanfilippo syndrome type B," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 12, pp. 6101–6105, 1996.

[75] H. H. Li, W. H. Yu, N. Rozengurt et al., "Mouse model of Sanfilippo syndrome type B produced by targeted disruption of the gene encoding alpha-N-acetylglucosaminidase," *Proc Natl Acad Sci USA*, vol. 96, no. 5, pp. 14505–14510, 1999.

[76] W. H. Yu, K. W. Zhao, S. Ryazantsev, N. Rozengurt, and E. F. Neufeld, "Short-term enzyme replacement in the murine model of sanfilippo syndrome type B," *Molecular Genetics and Metabolism*, vol. 71, no. 4, pp. 573–580, 2000.

[77] K. W. Zhao and E. F. Neufeld, "Purification and characterization of recombinant human α-N-acetylglucosaminidase secreted by Chinese hamster ovary cells," *Protein Expression and Purification*, vol. 19, no. 1, pp. 202–211, 2000.

[78] B. Weber, J. J. Hopwood, and G. Yogalingam, "Expression and characterization of human recombinant and α-N-actylglucosaminidase," *Protein Expression and Purification*, vol. 21, no. 2, pp. 251–259, 2001.

[79] M. Beck, "New therapeutic options for lysosomal storage disorders: enzyme replacement, small molecules and gene therapy," *Human Genetics*, vol. 121, no. 1, pp. 1–22, 2007.

[80] H. Fu, R. J. Samulski, T. J. McCown, Y. J. Picornell, D. Fletcher, and J. Muenzer, "Neurological correction of lysosomal storage in a mucopolysaccharidosis IIIB mouse model by adeno-associated virus-mediated gene delivery," *Molecular Therapy*, vol. 5, no. 1, pp. 42–49, 2002.

[81] A. Cressant, N. Desmaris, L. Verot et al., "Improved behavior and neuropathology in the mouse model of sanfilippo type IIIB disease after adeno-associated virus-mediated gene transfer in the striatum," *Journal of Neuroscience*, vol. 24, no. 45, pp. 10229–10239, 2004.

[82] C. D. Heldermon, K. K. Ohlemiller, E. D. Herzog et al., "Therapeutic efficacy of bone marrow transplant, intracranial AAV-mediated gene therapy, or both in the mouse model of MPS IIIB," *Molecular Therapy*, vol. 18, no. 5, pp. 873–880, 2010.

[83] P. Di Natale, C. Di Domenico, N. Gargiulo et al., "Treatment of the mouse model of mucopolysaccharidosis type IIIB with lentiviral-NAGLU vector," *Biochemical Journal*, vol. 388, no. 2, pp. 639–646, 2005.

[84] G. R. D. Villani, A. Follenzi, B. Vanacore, C. Di Domenico, L. Naldini, and P. Di Natale, "Correction of mucopolysaccharidosis type IIIB fibroblasts by lentiviral vector-mediated gene transfer," *Biochemical Journal*, vol. 364, no. 3, pp. 747–753, 2002.

[85] C. Di Domenico, G. R. D. Villani, D. Di Napoli et al., "Intracranial gene delivery of LV-NAGLU vector corrects neuropathology in murine MPS IIIB," *American Journal of Medical Genetics A*, vol. 149, no. 6, pp. 1209–1218, 2009.

[86] X. Fan, H. Zhang, S. Zhang et al., "Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C)," *American Journal of Human Genetics*, vol. 79, no. 4, pp. 738–744, 2006.

[87] M. Hřebíček, L. Mrázová, V. Seyrantepe et al., "Mutations in TMEM76* cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome)," *American Journal of Human Genetics*, vol. 79, no. 5, pp. 807–819, 2006.

[88] M. Feldhammer, S. Durand, and A. V. Pshezhetsky, "Protein misfolding as an underlying molecular defect in mucopolysaccharidosis III type C," *PloS one*, vol. 4, no. 10, article e7434, 2009.

[89] S. Durand, M. Feldhammer, E. Bonneil, P. Thibault, and A. V. Pshezhetsky, "Analysis of the biogenesis of heparan sulfate acetyl-CoA: α- glucosaminide N-acetyltransferase provides insights into the mechanism underlying its complete deficiency in mucopolysaccharidosis IIIC," *Journal of Biological Chemistry*, vol. 285, no. 41, pp. 31233–31242, 2010.

[90] A. O. Fedele and J. J. Hopwood, "Functional analysis of the HGSNAT gene in patients with mucopolysaccharidosis IIIC (Sanfilippo C syndrome)," *Human Mutation*, vol. 31, no. 7, pp. E1574–E1586, 2010.

[91] C. Freeman, P. R. Clements, and J. J. Hopwood, "Human liver N-acetylglucosamine-6-sulphate sulphatase. Purification and characterization," *Biochemical Journal*, vol. 246, no. 2, pp. 347–354, 1987.

[92] C. Freeman and J. J. Hopwood, "Human liver N-acetylglucosamine-6-phosphate sulphatase. Catalytic properties," *Biochemical Journal*, vol. 246, no. 2, pp. 355–365, 1987.

[93] D. A. Robertson, D. F. Callen, E. G. Baker, C. P. Morris, and J. J. Hopwood, "Chromosomal localization of the gene for human glucosamine-6-sulphatase to 12q14," *Human Genetics*, vol. 79, no. 2, pp. 175–178, 1988.

[94] A. Mok, H. Cao, and R. A. Hegele, "Genomic basis of mucopolysaccharidosis type IIID (MIM 252940) revealed by sequencing of GNS encoding N-acetylglucosamine-6-sulfatase," *Genomics*, vol. 81, no. 1, pp. 1–5, 2003.

[95] C. E. Beesley, D. Burke, M. Jackson, A. Vellodi, B. G. Winchester, and E. P. Young, "Sanfilippo syndrome type D: identification of the first mutation in the N-acetylglucosamine-6-sulphatase gene," *Journal of Medical Genetics*, vol. 40, no. 3, pp. 192–194, 2003.

[96] E. Downs-Kelly, M. Z. Jones, J. Alroy et al., "Caprine mucopolysaccharidosis IIID: a preliminary trial of enzyme replacement therapy," *Journal of Molecular Neuroscience*, vol. 15, no. 3, pp. 251–262, 2000.

[97] A. M. Montaño, S. Tomatsu, A. Brusius, M. Smith, and T. Orii, "Growth charts for patients affected with Morquio A disease," *American Journal of Medical Genetics A*, vol. 146, no. 10, pp. 1286–1295, 2008.

[98] S. Tomatsu, S. Fukuda, M. Masue, K. Sukegawa, M. Masuno, and T. Orii, "Mucopolysaccharidosis type IVA: characterization and chromosomal localization of N-acetylgalactosamine-6-sulfate sulfatase gene and genetic heterogeneity," *American Journal of Human Genetics*, vol. 51, p. A178, 1992.

[99] Y. Nakashima, S. Tomatsu, T. Hori et al., "Mucopolysaccharidosis IV A: molecular cloning of the human N-acetylgalactosamine-6-sulfatase gene (GALNS) and analysis of the 5'-flanking region," *Genomics*, vol. 20, no. 1, pp. 99–104, 1994.

[100] C. P. Morris, X. H. Guo, S. Apostolou, J. J. Hopwood, and H. S. Scott, "Morquio A syndrome: cloning, sequence, and structure of the human N- acetylgalactosamine 6-sulfatase (GALNS) gene," *Genomics*, vol. 22, no. 3, pp. 652–654, 1994.

[101] S. Tomatsu, A. M. Montaño, A. Ohashi et al., "Enzyme replacement therapy in a murine model of Morquio A syndrome," *Human Molecular Genetics*, vol. 17, no. 6, pp. 815–824, 2008.

[102] S. Tomatsu, A. M. Montaño, H. Oikawa et al., "Mucopolysaccharidosis type IVA (morquio a disease): clinical review and current treatment: a special review," *Current Pharmaceutical Biotechnology*, vol. 12, no. 6, pp. 931–945, 2011.

[103] A. Rodríguez, A. J. Espejo, A. Hernández et al., "Enzyme replacement therapy for Morquio A: an active recombinant N-acetylgalactosamine-6-sulfate sulfatase produced in

Escherichia coli BL21," *Journal of Industrial Microbiology and Biotechnology*, vol. 37, no. 11, pp. 1193–1201, 2010.

[104] A. Oshima, A. Tsuji, Y. Nagao, H. Sakuraba, and Y. Suzuki, "Cloning, sequencing, and expression of cDNA for human β-galactosidase," *Biochemical and Biophysical Research Communications*, vol. 157, no. 1, pp. 238–244, 1988.

[105] R. Santamaria, M. Blanco, A. Chabas, D. Grinberg, and L. Vilageliu, "Identification of 14 novel GLB1 mutations, including five deletions, in 19 patients with GM1 gangliosidosis from South America," *Clinical Genetics*, vol. 71, no. 3, pp. 273–279, 2007.

[106] T. Takano and Y. Yamanouchi, "Assignment of human β-galactosidase-A gene to 3p21.33 by fluorescence in situ hybridization," *Human Genetics*, vol. 92, no. 4, pp. 403–404, 1993.

[107] H. Morreau, N. J. Galjart, N. Gillemans, R. Willemsen, G. T. J. Van Der Horst, and A. D'Azzo, "Alternative splicing of β-galactosidase mRNA generates the classic lysosomal enzyme and a β-galactosidase-related protein," *Journal of Biological Chemistry*, vol. 264, no. 34, pp. 20655–20663, 1989.

[108] E. Hinek, "Biological roles of the non-integrin elastin/laminin receptor," *Biological Chemistry*, vol. 377, no. 7-8, pp. 471–480, 1996.

[109] S. Privitera, C. A. Prody, J. W. Callahan, and A. Hinek, "The 67-kDa enzymatically inactive alternatively spliced variant of β- galactosidase is identical to the elastin/laminin-binding protein," *Journal of Biological Chemistry*, vol. 273, no. 11, pp. 6319–6326, 1998.

[110] A. C. M. M. Azevedo, I. V. Schwartz, L. Kalakun et al., "Clinical and biochemical study of 28 patients with mucopolysaccharides type VI," *Clinical Genetics*, vol. 66, no. 3, pp. 208–213, 2004.

[111] L. Karageorgos, D. A. Brooks, A. Pollard et al., "Mutational analysis of 105 mucopolysaccharidosis type VI patients," *Human Mutation*, vol. 28, no. 9, pp. 897–903, 2007.

[112] E. Fidzianska, T. Abramowicz, and B. Czartoryska, "Assignment of the gene for human arylsulfatase B, ARSB, to chromosome region 5p11 → 5qter," *Cytogenetics and Cell Genetics*, vol. 38, no. 2, pp. 150–151, 1984.

[113] D. Auclair, J. J. Hopwood, D. A. Brooks, J. F. Lemontt, and A. C. Crawley, "Replacement therapy in Mucopolysaccharidosis type VI: advantages of early onset of therapy," *Molecular Genetics and Metabolism*, vol. 78, no. 3, pp. 163–174, 2003.

[114] D. Auclair, J. Finnie, J. White et al., "Repeated intrathecal injections of recombinant human 4-sulphatase remove dural storage in mature mucopolysaccharidosis VI cats primed with a short-course tolerisation regimen," *Molecular Genetics and Metabolism*, vol. 99, no. 2, pp. 132–141, 2010.

[115] D. Auclair, J. J. Hopwood, J. F. Lemontt, L. Chen, and S. Byers, "Long-term intra-articular administration of recombinant human N-acetylgalactosamine-4-sulfatase in feline mucopolysaccharidosis VI," *Molecular Genetics and Metabolism*, vol. 91, no. 4, pp. 352–361, 2007.

[116] P. Harmatz, R. Giugliani, I. V. D. Schwartz et al., "Long-term follow-up of endurance and safety outcomes during enzyme replacement therapy for mucopolysaccharidosis VI: final results of three clinical studies of recombinant human N-acetylgalactosamine 4-sulfatase," *Molecular Genetics and Metabolism*, vol. 94, no. 4, pp. 469–475, 2008.

[117] S. Byers, M. Rothe, J. Lalic, R. Koldej, and D. S. Anson, "Lentiviral-mediated correction of MPS VI cells and gene transfer to joint tissues," *Molecular Genetics and Metabolism*, vol. 97, no. 2, pp. 102–108, 2009.

[118] G. Cotugno, P. Annunziata, A. Tessitore et al., "Long-term amelioration of feline mucopolysaccharidosis VI after AAV-mediated liver gene transfer," *Molecular Therapy*, vol. 19, pp. 461–469, 2011.

[119] J. M. Shipley, M. Klinkenberg, B. M. Wu, D. R. Bachinsky, J. H. Grubb, and W. S. Sly, "Mutational analysis of a patient with mucopolysaccharidosis type VII, and identification of pseudogenes," *American Journal of Human Genetics*, vol. 52, no. 3, pp. 517–526, 1993.

[120] W. S. Sly, B. A. Quinton, W. H. McAlister, and D. L. Rimoin, "Beta glucuronidase deficiency: report of clinical, radiologic, and biochemical features of a new mucopolysaccharidosis," *The Journal of Pediatrics*, vol. 82, no. 2, pp. 249–257, 1973.

[121] R. D. Miller, J. W. Hoffmann, P. P. Powell et al., "Cloning and characterization of the human β-glucuronidase gene," *Genomics*, vol. 7, no. 2, pp. 280–283, 1990.

[122] F. Speleman, R. Vervoort, N. Van Roy, I. Liebaers, W. S. Sly, and W. Lissens, "Localization by fluorescence in situ hybridization of the human functional β-glucuronidase gene (GUSB) to 7q11.21 → q11.22 and two pseudogenes to 5p13 and 5q13," *Cytogenetics and Cell Genetics*, vol. 72, no. 1, pp. 53–55, 1996.

[123] J. M. Shipley, M. Klinkenberg, B. M. Wu, D. R. Bachinsky, J. H. Grubb, and W. S. Sly, "Mutational analysis of a patient with mucopolysaccharidosis type VII, and identification of pseudogenes," *American Journal of Human Genetics*, vol. 52, no. 3, pp. 517–526, 1993.

[124] S. Tomatsu, A. M. Montano, V. C. Dung, J. H. Grubb, and W. S. Sly, "Mutations and polymorphisms in GUSB gene in mucopolysaccharidosis VII (sly syndrome)," *Human Mutation*, vol. 30, no. 4, pp. 511–519, 2009.

[125] B. M. Wu, S. Tomatsu, S. Fukuda, K. Sukegawa, T. Orii, and W. S. Sly, "Overexpression rescues the mutant phenotype of L176F mutation causing β- glucuronidase deficiency mucopolysaccharidosis in two Mennonite siblings," *Journal of Biological Chemistry*, vol. 269, no. 38, pp. 23681–23688, 1994.

[126] R. Vervoort, M. R. Islam, W. S. Sly et al., "Molecular analysis of patients with β-glucuronidase deficiency presenting as hydrops fetalis or as early mucopolysaccharidosis VII," *American Journal of Human Genetics*, vol. 58, no. 3, pp. 457–471, 1996.

[127] I. Schwartz, L. R. Silva, S. Leistner et al., "Mucopolysaccharidosis VII: clinical, biochemical and molecular investigation of a Brazilian family," *Clinical Genetics*, vol. 64, no. 2, pp. 172–175, 2003.

[128] M. E. Haskins, G. D. Aguirre, P. F. Jezyk, E. H. Schuchman, R. J. Desnick, and D. F. Patterson, "Mucopolysaccharidosis type VII (Sly syndrome): beta-glucuronidase-deficient mucopolysaccharidosis in the dog," *American Journal of Pathology*, vol. 138, no. 6, pp. 1553–1555, 1991.

[129] D. C. S. Dombrowski, K. P. Carmichael, P. Wang, T. M. O'Malley, M. E. Haskins, and U. Giger, "Mucopolysaccharidosis type VII in a German shepherd dog," *Journal of the American Veterinary Medical Association*, vol. 224, no. 4, pp. 553–532, 2004.

[130] A. Bosch, E. Perret, N. Desmaris, and J. M. Heard, "Long-term and significant correction of brain lesions in adult mucopolysaccharidosis type VII mice using recombinant AAV vectors," *Molecular Therapy*, vol. 1, no. 1, pp. 63–70, 2000.

[131] A. Bosch, E. Perret, N. Desmaris, D. Trono, and J. M. Heard, "Reversal of pathology in the entire brain of mucopolysaccharidosis type VII mice after lentivirus-mediated gene

transfer," *Human Gene Therapy*, vol. 11, no. 8, pp. 1139–1150, 2000.

[132] Y. Yamada, K. Kato, K. Sukegawa et al., "Treatment of MPS VII (Sly disease) by allogeneic BMT in a female with homozygous A619V mutation," *Bone Marrow Transplantation*, vol. 21, no. 6, pp. 629–634, 1998.

[133] L. C. Ginsburg, D. T. DiFerrante, C. T. Caskey, and N. M. DiFerrante, "Glucosamine-6-SO_4 sulfatase deficiency: a new mucopolysaccharidosis," *Clinical Research*, vol. 25, p. 471A, 1977.

[134] L. C. Ginsberg, P. V. Donnelly, and D. T. Di Ferrante, "N-acetylglucosamine-6-sulfate sulfatase in man: deficiency of the enzyme in a new mucopolysaccharidosis," *Pediatric Research*, vol. 12, no. 7, pp. 805–809, 1978.

[135] N. Di Ferrante, "N-acetylglucosamine-6-sulfate sulfatase deficiency reconsidered," *Science*, vol. 210, no. 4468, p. 448, 1980.

[136] M. R. Natowicz, M. P. Short, Y. Wang et al., "Brief report: clinical and biochemical manifestations of hyaluronidase deficiency," *New England Journal of Medicine*, vol. 335, no. 14, pp. 1029–1033, 1996.

[137] B. Triggs-Raine, T. J. Salo, H. Zhang, B. A. Wicklow, and M. R. Natowicz, "Mutations in HYAL1, a member of a tandemly distributed multigene family encoding disparate hyaluronidase activities, cause a newly described lysosomal disorder, mucopolysaccharidosis IX," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 11, pp. 6296–6300, 1999.

[138] A. B. Csoka, G. I. Frost, and R. Stern, "The six hyaluronidase-like genes in the human and mouse genomes," *Matrix Biology*, vol. 20, no. 8, pp. 499–508, 2001.

[139] L. Jadin, X. Wu, H. Ding et al., "Skeletal and hematological anomalies in HYAL2-deficient mice: a second type of mucopolysaccharidosis IX?" *FASEB Journal*, vol. 22, no. 12, pp. 4316–4326, 2008.

[140] W. Krivit, "Allogeneic stem cell transplantation for the treatment of lysosomal and peroxisomal metabolic diseases," *Springer Seminars in Immunopathology*, vol. 26, no. 1-2, pp. 119–132, 2004.

[141] R. Lachmann, "Treatments for lysosomal storage disorders," *Biochemical Society Transactions*, vol. 38, no. 6, pp. 1465–1468, 2010.

[142] P. I. Dickson and A. H. Chen, "Intrathecal enzyme replacement therapy for mucopolysaccharidosis I: translating success in animal models to patients," *Current Pharmaceutical Biotechnology*, vol. 12, no. 6, pp. 946–955, 2011.

[143] A. Friso, R. Tomanin, M. Salvalaio, and M. Scarpa, "Genistein reduces glycosaminoglycan levels in a mouse model of mucopolysaccharidosis type II," *British Journal of Pharmacology*, vol. 159, no. 5, pp. 1082–1091, 2010.

[144] A. L. K. Roberts, B. J. Thomas, A. S. Wilkinson, J. M. Fletcher, and S. Byers, "Inhibition of glycosaminoglycan synthesis using rhodamine B in a mouse model of mucopolysaccharidosis type IIIA," *Pediatric Research*, vol. 60, no. 3, pp. 309–314, 2006.

[145] A. L. K. Roberts, M. H. Rees, S. Klebe, J. M. Fletcher, and S. Byers, "Improvement in behaviour after substrate deprivation therapy with rhodamine B in a mouse model of MPS IIIA," *Molecular Genetics and Metabolism*, vol. 92, no. 1-2, pp. 115–121, 2007.

[146] K. J. Langford-Smith, J. Mercer, J. Petty et al., "Heparin cofactor II-thrombin complex and dermatan sulphate: chondroitin sulphate ratio are biomarkers of short- and long-term treatment effects in mucopolysaccharide diseases," *Journal of Inherited Metabolic Disease*, vol. 34, no. 2, pp. 499–508, 2011.

[147] C. M. Simonaro, M. D'Angelo, M. E. Haskins, and E. H. Schuchman, "Joint and bone disease in mucopolysaccharidoses VI and VII: identification of new therapeutic targets and BioMarkers using animal models," *Pediatric Research*, vol. 57, no. 5 I, pp. 701–707, 2005.

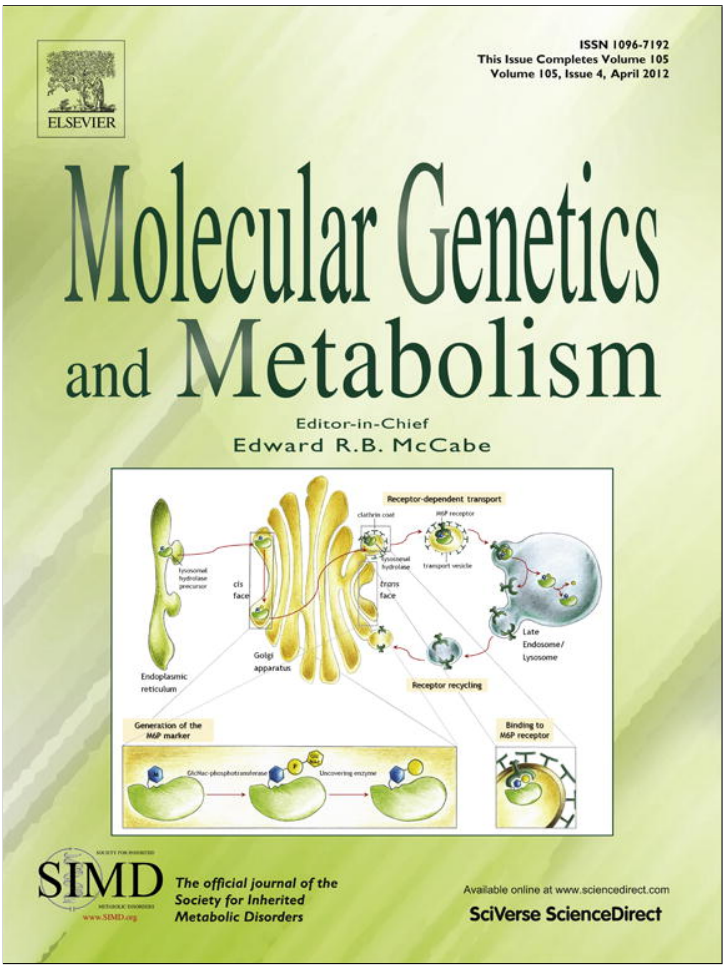
[148] M. Fuller, J. N. Tucker, D. L. Lang et al., "Screening patients referred to a metabolic clinic for lysosomal storage disorders," *Journal of Medical Genetics*, vol. 48, no. 6, pp. 422–425, 2011.

[149] M. Spada, S. Pagliardini, M. Yasuda et al., "High incidence of later-onset Fabry disease revealed by newborn screening," *American Journal of Human Genetics*, vol. 79, no. 1, pp. 31–40, 2006.

[150] H. Y. Lin, K. W. Chong, J. H. Hsu et al., "High incidence of the cardiac variant of fabry disease revealed by newborn screening in the Taiwan Chinese population," *Circulation*, vol. 2, no. 5, pp. 450–456, 2009.

[151] Y. H. Chien, S. C. Chiang, X. K. Zhang et al., "Early detection of pompe disease by newborn screening is feasible: results from the Taiwan screening program," *Pediatrics*, vol. 122, no. 1, pp. e39–e45, 2008.

[152] P. K. Duffner, M. Caggana, J. J. Orsini et al., "Newborn screening for Krabbe disease: the New York state model," *Pediatric Neurology*, vol. 40, no. 4, pp. 245–252, 2009.



Molecular Genetics and Metabolism 105 (2012) 542–550

Contents lists available at SciVerse ScienceDirect

Molecular Genetics and Metabolism

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

Minireview

Mannose-6-phosphate pathway: A review on its role in lysosomal function and dysfunction

Maria Francisca Coutinho [a,b,c,*], Maria João Prata [b,c], Sandra Alves [a]

[a] Research and Development Unit, Department of Genetics, CGMJM, INSA Portugal
[b] IPATIMUP, Porto, Portugal
[c] Department of Biology, Faculty of Sciences, Porto, Portugal

ARTICLE INFO

Article history:
Received 11 October 2011
Received in revised form 15 December 2011
Accepted 15 December 2011
Available online 23 December 2011

Keywords:
Mannose-6-phosphate-dependent pathway
GlcNAc-1-phosphotransferase
Uncovering enzyme (UCE)
Mannose-6-phosphate receptors
Mucolipidosis type II (I-cell disease)
Mucolipidosis type III
(Pseudo-Hurler disease)

ABSTRACT

Lysosomal hydrolases are synthesized in the rough endoplasmic reticulum and specifically transported through the Golgi apparatus to the *trans*-Golgi network, from which transport vesicles bud to deliver them to the endosomal/lysosomal compartment.

The explanation of how are the lysosomal enzymes accurately recognized and selected over many other proteins in the *trans*-Golgi network relies on being tagged with a unique marker: the mannose-6-phosphate (M6P) group, which is added exclusively to the N-linked oligosaccharides of lysosomal soluble hydrolases, as they pass through the *cis*-Golgi network. Generation of the M6P recognition marker depends on a reaction involving two different enzymes: UDP-N-acetylglucosamine 1-phosphotransferase and α-N-acetylglucosamine-1-phosphodiester α-N-acetylglucosaminidase.

The M6P groups are then recognized by two independent transmembrane M6P receptors, present in the *trans*-Golgi network: the cation-independent M6P receptor and/or the cation-dependent M6P receptor. These proteins bind to lysosomal hydrolases on the lumenal side of the membrane and to adaptins in assembling clathrin coats on the cytosolic side. In this way, the M6P receptors help package the hydrolases into vesicles that bud from the *trans*-Golgi network to deliver their contents to endosomes that ultimately will develop into mature lysosomes, where recently-delivered hydrolases may start digesting the endocyted material.

The above described process is known as the M6P-dependent pathway and is responsible for transporting most lysosomal enzymes.

This review synthesizes the current knowledge on each of the major proteins involved in the M6P-dependent pathway. Impairments in this pathway will also be addressed, highlighting the lysosomal storage disorders associated to GlcNAc-1-phosphotransferase loss of function: mucolipidosis type II and III.

Contents

Abbreviations: M6P, Mannose-6-phosphate; (LSD), Lysosomal storage disorders; (UCE), N-acetylglucosamine-1-phosphodiester α-N-acetylglucosaminidase; (MPRs), Mannose-6-phosphate receptors; (CI-MPR), Cation-independent mannose-6-phosphate receptor; (CD-MPR), Cation-dependent mannose-6-phosphate receptor; (ERT), Enzyme replacement therapies; (ML II), Mucolipidosis II; (ML III), Mucolipidosis III.

* Corresponding author at: Medical Genetics Center Jacinto Magalhães, INSA, Praça Pedro Nunes, 88, 4099-028 Porto, Portugal. Fax: +351 226070399.
E-mail addresses: francisca_coutinho@yahoo.com, francisca.coutinho@insa.min-saude.pt (M.F. Coutinho).

1096-7192/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2011.12.012

1. Introduction

Lysosomes are membrane delimited organelles with a characteristic acidic pH that are responsible for the degradation of many different substrates in the cell. This catabolic process is carried out by over 60 soluble enzymes that are contained inside the organelle. The majority of soluble acid hydrolases are modified with mannose-6-phosphate (M6P) residues, allowing their recognition by M6P receptors in the Golgi complex and ensuring their transport to the endosomal/lysosomal system [1].

Like several other proteins, lysosomal hydrolases are synthesized in the endoplasmic reticulum and cotranslationally glycosylated on selected asparagine residues. When moving to the secretory pathway, these proteins are selectively recognized by a phosphotransferase, UDP-N-acetylglucosamine 1-phosphotransferase (GlcNAc-1-phosphotransferase), which initiates the two-step reaction that results in the generation of the M6P marker on specific N-linked oligosaccharides. The enzyme catalyzes the transfer of a GlcNAc-1-phosphate residue from UDP-GlcNAc to C6 positions of selected mannoses in high-manose type oligosaccharides of the hydrolases. The second step involves the removal of the terminal GlcNAc by an N-acetylglucosamine-1-phosphodiester α-N-acetyl-glucosaminidase, also known as "uncovering enzyme" (UCE), exposing the M6P recognition signal. Next, the modified proteins are recognized by two independent receptors that bind the M6P residue of the newly synthesized lysosomal hydrolases in the *trans*-Golgi network. Finally, the ligand-receptor complex is packaged into clathrin-coated transport vesicles for delivery to endosomes and lysosomes (Fig. 1) [1].

The sequential action of these proteins, which is triggered upon arrival in the Golgi apparatus of soluble lysosomal enzymes allowing their subsequent transport to the endosomal/lysosomal compartment, constitutes the M6P-dependent targeting pathway. The importance of providing lysosomes with an entire repertoire of degradative enzymes for normal cellular processes is illustrated by the existence of over 50 different monogenic human lysosomal storage diseases that are caused by a deficiency of a catabolic enzyme [2] that as a whole are estimated to affect 1 in 5000 live births [3].

This review aims to synthesize the current knowledge on each of these proteins and their cellular function, passing through lysosomal storage disorders (LSD) associated to the loss of function of some of them. Recent studies on secondary impaired uptake of lysosomal enzymes and mutations in the lysosomal enzyme targeting pathway

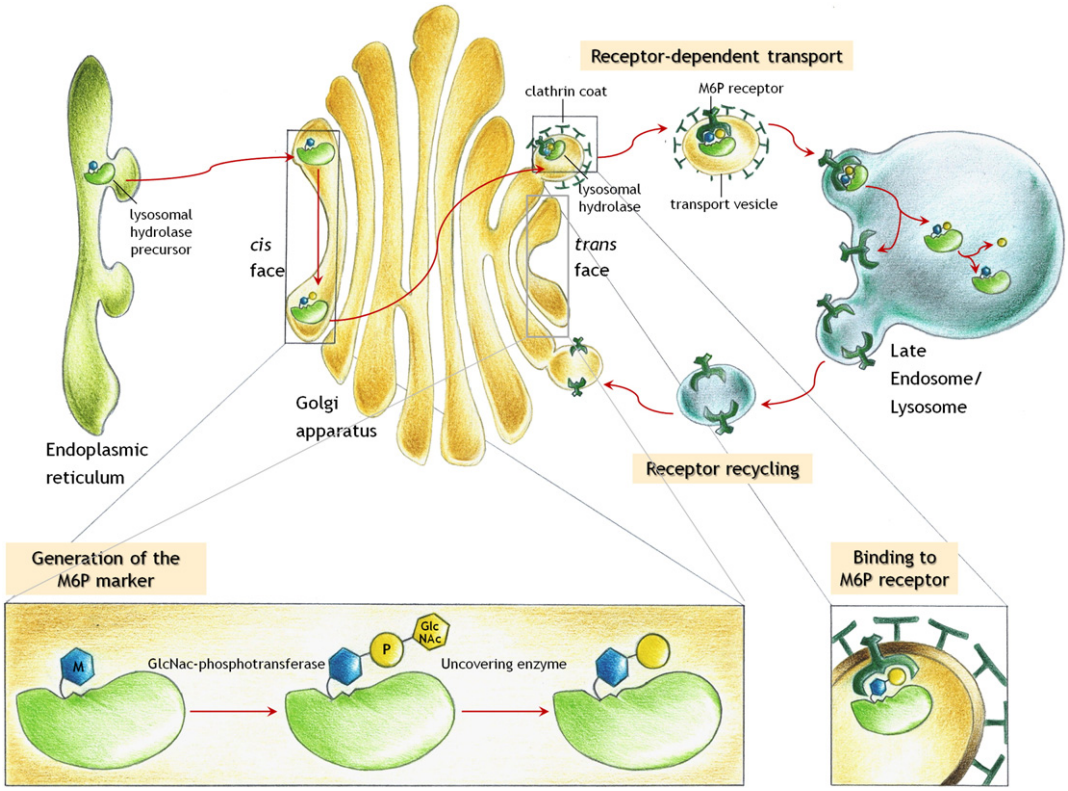


Fig. 1. Transport of newly synthesized lysosomal hydrolases to lysosomes.Lysosomal hydrolases are synthesized in the endoplasmic reticulum and move to the *cis* Golgi network, were they are covalently modified by the addition of mannose-6-phosphate (M6P) groups. The formation of this marker depends on the sequential effect of two lysosomal enzymes: UDP-N-acetylglucosamine 1-phosphotransferase (GlcNac-phosphotransferase) and N-acetylglucosamine-1-phosphodiester α-N-acetyl-glucosaminidase (uncovering enzyme). GlcNac-phosphotransferase catalyzes the transfer of a GlcNAc-1-phosphate residue from UDP-GlcNAc to C6 positions of selected mannoses in high-manose type oligosaccharides of the hydrolases. Then, the uncovering enzyme removes the terminal GlcNAc, exposing the M6P recognition signal. At the *trans* Golgi network, the M6P signal allows the segregation of lysosomal hydrolases from all other types of proteins through selective binding to the M6P receptors. The clathrin-coated vesicles produced bud off from the trans-Golgi network and fuse with late endosomes. At the low pH of the late endosome, the hydrolases dissociate from the M6P receptors and the empty receptors are recycled to the Golgi apparatus for further rounds of transport.

causing stuttering, will also be addressed, together with the variable efficacy of the some therapeutic approaches depending on the expression levels of the M6P pathway functional components.

2. The mannose-6-phosphate pathway: From amoebas to humans

The mannose-6-phosphate dependent mechanism for lysosomal enzyme routing was first described in mammalian cells as involving the sequential action of the above referred proteins: GlcNAc-1-phosphotransferase, N-acetylglucosamine-1-phosphodiester α-N-acetyl-glucosaminidase and the two mannose-6-phosphate receptors. An identical mechanism was then found to occur in avian (chicken) and amphibian (*Xenopus*) cells [4], suggesting that the mannose-6-phoshate pathway was likely shared by all vertebrates. In 1994, a UDP-GlcNAc: lysosomal enzyme GlcNAc-1-phosphotransferase was detected in the fruit fly *Ceratitis capitata* with specificity and catalytic properties resembling those found in mammalian cells [5]. The question of whether a homologous mannose-6-phosphate-dependent mechanism for lysosomal enzyme targeting was also present in invertebrates, was only answered two years later when Alvarez and collaborators firstly reported the mechanism in the crab *Chasmagnatus granulata* [6]. These authors purified an acid hydrolase with M6P residues, demonstrated the existence of a GlcNAc-1-phosphodiester α-N-acetyl-glucosaminidase in microsomal membranes of the crab's hepatopancreas and proved that two mannose-6-phosphate receptors (see information about mannose-6-phosphate receptors in Section 3.3 of this review), recognized by specific antisera raised against each bovine receptor, were also present in the microsomes of the crab [6].

Although GlcNAc-1-phosphotransferases have been described in lower eukaryotes, Lang et al. [7] have demonstrated that both the soil amoeba *Acanthamoeba castellanii* and the vegetative amoeba of the cellular slime mold *Dictyostelium discoideum* contain an enzyme that transfers GlcNAc-P from UDP-GlcNAc to mannose residues on high mannose oligosaccharides. The enzyme of *A. castellani* is similar to the mammalian enzyme in that it is able to recognize the common protein determinant of lysosomal mammalian enzymes. By contrast, the N-acetylglucosaminylphosphotransferase of *D. discoideum* lacks this capability demonstrating that functional differences exist between GlcNAc-1-phosphotransferases in lower eukaryotes and the mammalian ones. Recently, Qian and colleagues [8] identified the gene (*gpt1*) that encodes the *D. discoideum* GlcNAc-1-phosphotransferase. This enzyme presents sequence similarity to the human GlcNAc-1-phosphotransferase, differing from this one in by the lack of a DMAP binding domain on the αsubunit and by the absence of the γ subunit. Nevertheless, the lysosomal hydrolases of wild type *D. discoideum* are well phosphorylated and, so, the γ subunit seems to be dispensable in this species. [8].

Also in the unicellular organism *A. castellanii* referred above, despite the presence of an enzyme with homology, in terms of specificity, to mammalian GlcNAc-1-phosphotransferases, the additional components of the mammalian routing mechanism, such as GlcNAc-1-phosphodiester α-N-acetyl-glucosaminidase or mannose-6-phosphate receptors, were not detected, suggesting that in this ameboid microorganism lysosomal targeting is not mediated by M6P markers [9].

3. Functional components of the mannose-6-phosphate pathway

3.1. GlcNAc-1-phosphotransferase

GlcNac-phosphotransferase is a Golgi-resident 540-kDa hexameric transmembrane enzyme composed by three subunits, $\alpha_2\beta_2\gamma_2$, which catalyzes the first step in the formation of the M6P recognition marker on lysosomal enzymes (Fig. 1). This enzyme is coded by two different genes: *GNPTAB* and *GNPTG*, encoding for the α/β subunits and the α subunit, respectively. The *GNPTAB* gene contains 21 exons and spans 85 kb on chromosome 12q23.3. It encodes a protein of 1256 amino acids with a predicted molecular mass of 144 kDa (α/β precursor). Hydrophobicity analysis showed two transmembrane domains and 19 potential N-glycosylation sites (17 in the α subunit and 3 in the β subunit) [10].

Proteolytic processing of the α/β precursor between Lys928 and Asp929 generates the individual α and β subunits and is a prerequisite for the enzymatic activity of the GlcNAc-1-phosphotransferase complex [10]. Recently Marschner et al. [11] provided evidence that the α/β-subunit precursor of GlcNAc-1-phosphotransferase is cleaved by the site-1 protease (S1P; also known as subtilisin kexin isoenzyme–1, SKI-1). This enzyme is encoded by the *MBTPS1* gene and is a membrane bound serine protease. The prototypical membrane-bound S1P substrates are the sterol regulatory element-binding proteins SREBP1 and 2, which play a major role in lipid metabolism and cholesterol homeostasis [referred in 11]. These authors verified that S1P-deficient cells failed to activate the α/β-subunit precursor and exhibited a mucolipidosis II-like phenotype. This implies that these findings may also have consequences in terms of diagnosis, since mutations in the gene that codes for this protease may be present in individuals with genetically undefined mucolipidosis II-like phenotypes such as Pacman dysplasia.

The *GNPTG* gene, spanning 11 exons on chromosome 16p13.3, encodes the γ subunit, which is a soluble glycoprotein of 305 amino acids, with a molecular weight of 97 kDa capable to form disulfide-linked dimers [12]. The initial 305 aminoacid polypeptide is processed into a mature form by proteolytic cleavage of a signal peptide with 24 amino acids; then the mature protein can form disulfide bonds [13]. The γ subunit contains two potential N-glycosylation sites at N88 and N115 [14].

Functionally, the α/β subunits constitute the catalytic portion of the enzyme and carry the substrate binding sites for UDP-GlcNAc and high mannose-type lysosomal enzymes [15,16]. In the α subunit are located the UDP-GlcNAc binding site and homology region to the bacterial capsular polymerases [10,15,17]. Nevertheless, the β subunit also appears to be necessary, as demonstrated by the failure of transfection of the α subunit alone to generate GlcNAc-1-phosphotransferase activity [18] plus the genetic evidence that truncation of the β subunit results in GlcNAc-1-phosphotransferase deficiency [13,18,19]. Moreover, the transmembrane domains on the α and β subunits are required for natural processing of the GlcNAc-1-phosphotransferase precursor [18].

Analysis of a library of cathepsin D/pepsinogen chimeric proteins along with mutant cathepsin L has shown that GlcNAc-1-phosphotransferase recognizes common conformation-dependent protein structures of lysosomal enzymes in which lysine residues are the major determinants [20–23], as was also proven through antibody inhibition experiments with arylsulfatase A [24]. In addition, the interaction of both subunits with the protein determinant of acid hydrolases stimulates the catalytic function of the transferase [23].

In 2007, Gelfman and colleagues [25] generated a mouse model for GNPTAB deficiencies by microinjection of embryonic stem (ES) cell clones *GNPTAB*$^{-/-}$ into host blastocysts. Homozygous mice lacking α/β subunits of GlcNAc-1-phosphotransferase presented with growth retardation, retinal degeneration and secretory cell lesions that do not exactly mimic the human disease. Nevertheless, the less severe phenotype of the *GNPTAB*-knockout mouse enables studies designed to understand alternative molecular mechanisms involved in the trafficking of lysosomal enzymes to lysosomes and provides a valuable tool for probing the role of proper lysosome function in the maintenance of the retina and the secretory cells of exocrine glands [25]. The role of the γ subunits of the GlcNAc-1-phosphotransferase remained elusive [10,18] up to recently, when studies become to appear demonstrating that the γ chains are important to facilitate the proper folding of the full set of subunits in the GlcNAc-1-phosphotransferase and to maintain them in the right conformation,

turning the enzyme competent for substrate recognition and binding. The γ chains are also necessary to regulate the activity and expression of the α/β subunits [16,23,26,27]. Pohl and colleagues [26], after having identified, in a boy with MLIII γ a novel homozygous mutation in the *GNPTG* gene resulting in a truncated but stable γ subunit protein, performed in vitro and in vivo experiments to characterize the mutation, obtained compelling support that the γ subunit was involved in the regulation of the GlcNAc-1-phosphotransferase activity. Latter, the same group [27] showed that in human macrophages the proteolytic cleavage of the γ subunit of the GlcNAc-1-phosphotransferase leads to the formation of unique oligomers that fail to associate with the other GlcNAs-1-phosphotransferase subunits in higher molecular mass complexes, which might explain why almost all proteins in human macrophages lack M6P residues. The limited proteolysis of the γ subunits seems to be the regulator mechanism that controls the GlcNAc-1-phosphotransferase activity and subsequent sorting efficiency of lysosomal enzymes. In mice lacking the γ subunit, the extent of acid hydrolases phosphorylation relative to wild-type animals was investigated revealing that the γ subunit increases the activity of the α/β subunits toward protein acceptors to variable extents [23]. The same study afforded evidence that the γ subunit serves at least two roles. One is to interact with the high mannose oligosaccharides of the acceptor hydrolase facilitating the addiction of the second GlcNAc-P to the molecule. This GlcNAc-P is transferred to a specific mannose residue on the 3′ arm of the high mannose unit, whereas the first GlcNAc-P almost always is added to a mannose on the 6′ arm of the glycan. In addition, the authors also postulate that a specific domain of the gamma subunit, the mannose 6-phosphate receptor homology (MRH) domain binds the high mannose oligosaccharide in such a way that the transfer of the second GlcNAc-P is enhanced being the other role of the γ subunit to enhance the overall phosphorylation in a subset of the acid hydrolases [23]. This role is in line with previous data from Lee et al. [16], who having analyzed brain tissue of knock-out mice for the γ gene, observed that, for example, α-mannosidase and β-glucuronidase are similarly phosphorylated as in the wild-type brain, whereas β-hexosaminidase and β-galactosidase are not or are weakly phosphorylated [16].

3.2. The uncovering enzyme: N-acetylglucosamine-1-phosphodiester α-N-acetyl-glucosaminidase

N-acetylglucosamine-1-phosphodiester α-N-acetylglucosaminidase, also known as the uncovering enzyme (UCE) excises the covering GlcNAc from the GlcNAc-P-Man diesters, to form the M6P monoester recognition signal on lysosomal acid hydrolases (Fig. 1). UCE is a type I transmembrane glycoprotein of 515 amino acids which exists as a tetramer that cycles between the *trans*-Golgi network and the plasma membrane [28–30]. This enzyme is coded by the *NAGPA* gene, that contains 10 exons and is located on chromosome 16 [29]. UCE presents six potential N-glycosylation sites [29], being synthesized as an inactive precursor that is activated upon proteolytic cleavage of a 24-aminoacid propiece by the endoprotease furin. As furin is localized in the *trans*-Golgi network, newly synthesized UCE is inactive until reaching the terminal Golgi compartment [31].

To date no pathological conditions have been associated with the loss of UCE activity. Speculation on the existence or inexistence of such pathologies was maintained over years, with some authors arguing that alternative mechanism of conversion of the M6P diesters to monoesters should exist able to correct malformation of the recognition signal induced by UCE deficiency, while others argued that inactivating mutations in the UCE gene should be so rare events that could explain the non-observation of UCE related phenotypes. However, a few years ago, Chavez and collaborators [32], shed some light on why UCE loss of function might not be associated with overt pathological condition. They demonstrated that the fifth repeat domain, of a total of 15 that constitute the cation-independent mannose-6-phosphate receptor (CI-MPR), is able to bind GlcNAc-P-Man diesters. This binding occurred with much lower affinity than the binding to M6P monoesters, which take place in domains 1–3 and 9 of the same receptor. Probably, then, in the absence of UCE activity, acid hydrolases bearing M6P diesters might be able to bind sufficiently to the CI-MPR, enabling trafficking to lysosome enough to prevent the clinical manifestations seen when GlcNAc-1-phosphotransferase activity is absent/deficient [32]. These first evidences came from in vitro studies but next were corroborated by in vivo experiments with mice lacking UCE, created through insertional mutagenesis in the gene that codes for UCE [33]. Despite UCE$^{-/-}$ mice were viable, grew normally and lacked detectable histological abnormalities, the plasma levels of six monitored lysosomal hydrolases were 1.6 to 5.4-fold elevated over wild type levels. The secreted hydrolases contained GlcNAc-P-Man diesters, exhibited a decreased affinity for the CI-MPR receptor and no affinity at all to the CD-MPR. On the grounds of these results, Boonen and colleagues [33] proposed that, in the absence of UCE, the weak binding of the lysosomal hydrolases to the CI-MPR allows enough sorting to lysosomes to prevent the tissue abnormalities seen in GlcNAc-1-phosphotransferase deficiency. The ability of the CI-MPR to recognize the product of GlcNAc-1-phosphotransferase implies that deficiencies in the uncovering enzyme probably will not be lethal [33]. Furthermore, it allows delivery to the lysosome of those acid hydrolases that may be poor substrates for the uncovering enzyme in quantities sufficient to prevent the severe phenotypes usually associated with lysosomal storage diseases [34].

3.3. Mannose-6-phosphate receptors

Once anchored in the N-linked oligosaccharides of newly synthesized hydrolases, the M6P recognition marker has to be recognized by specific receptors, so that the labeled hydrolases are correctly transported to the lysosome (Fig. 1) [reviewed in 35]. This role is carried out by the mannose-6-phosphate receptors (MPRs), which are type I transmembrane glycoproteins that bind their specific oligosaccharide at pH 6,5-6,7, in the *trans*-Golgi network and release it at pH 6, the typical pH inside late endosomes. Inside these compartments, lysosomal hydrolases dissociate from the MPRs, and when the pH drops further, during endosomal maturation, and reaches 5 — the typical acidic pH of lysosomes which is maintained by a membrane ATP-driven H^+ pump, hydrolases begin to digest the endocytosed material delivered from early endosomes [reviewed in 36].

Sorting of MPRs at the TGN was long thought to predominantly depend on binding to the heterotetrameric adaptor-protein complex (AP)-1, which also mediates the recruitment of clathrin. The discovery of the Golgi localized, γ-ear-containing, Arf-binding family of proteins (GGA) as clathrin adaptors changed this view. GGAs are monomeric proteins and three different forms are known in mammals: GGA1, GGA2 and GGA3. Presently, GGAs and AP-1 are thought to function in parallel to generate distinct MPR-containing vesicles at the TGN, thereby allowing delivery to different endosomal compartments. Alternatively, or additionally, GGAs may facilitate entry of MPR into clathrin-coated vesicles by interacting with AP-1. Moreover, it is possible that different adaptor proteins can also act separately to a certain extent providing a means to establish the distinct pathways [reviewed in 37].

A multiprotein complex known as retromer has been shown to mediate retrieval of MPR to the TGN. It is likely that MPR retrieval to the TGN involves retromer in the early part and Rab9-TIP47 in the late part of the pathway. Retromer has been localized to membrane tubules that extend from the vacuolar part of endosomes and may participate in retrograde transport to the TGN. In contrast, Rab9 is associated with vesicular intermediates that bud from late endosomes and merge with the TGN. The existence of these two

retrieval mechanisms likely ensures efficient recycling of MPR during endosomal maturation. In addition, retrieval from endosomes to the TGN also depends on specific signals in the cytosolic tails of the MPR [reviewed in 1].

There are two different M6P receptors, both of them extensively characterized over the last decades [38,39 reviewed in 40]. The first is the cation-independent M6P receptor (CI-MPR or MPR300), a transmembrane glycoprotein of approximately 300-kDa that also binds with high affinity other ligands such as the insulin-like growth factor II (IGF II) that is involved in early grow and development. This is why the CI-MPR is also called the insulin-like growth factor II receptor. The extracytoplasmic domain of this receptor contains fifteen homologous repeating units of around 150 amino acids [41,42]. The third and the ninth of them provide two M6P binding sites, whereas the eleventh element provides a binding site to IGF II [43–45]. Additionally, CI-MPR also contains one phosphodiester (Man6P-GlcNAc) binding site [46,47]. The other receptor is the cation-dependent M6P receptor (CD-MPR or MPR46). The extracytoplasmic domain of this 46-kDa transmembrane glycoprotein is homologous to each of the fifteen repeats found in the extracytoplasmic domain of CI-MPR [48] and contains one M6P binding site [49]. The detergent-solubilized CD-MPR is a homodimer which may form tetramers in Golgi membranes [50,51].

Although related to one another, and even having similar domains, both MPRs are essential and non-synonymous components of the M6P-dependent targeting system. This was proven by gene disruption experiments in mice. Animals with disrupted CI-MPRs or CD-MPRs genes were partially impaired in intracellular lysosomal enzyme sorting [52,53]. Mutant mice lacking the CD-MPR are viable, while mice lacking the CI-MPR accumulate high levels of IGF-II and usually die perinatally, whereas mice that lack both IGF-II and CI-MPR are viable [54]. A partial missorting of phosphorylated lysosomal enzymes was also observed in embryonic fibroblasts derived from mouse embryos lacking either the CI-MPR or the CD-MPR [55]. Further experiments enlightened the fact that each mutant cell line secretes, in a large part, different phosphorylated ligands and that high levels of one MPR do not fully compensate for the absence of the other, demonstrating that the two MPRs have complementary targeting functions. One possible explanation is that each MPR recognizes different features of the lysosomal enzymes [55]. Furthermore, primary fibroblasts from embryos lacking both MPRs are almost totally impaired in lysosomal enzyme sorting. As a consequence, cells accumulate undigested material in their endosomes/lysosomes [56] presenting a phenotype quite similar to the one characteristic of I-cell disease fibroblasts, which have a lack of synthesis of the M6P signal due to defects in GlcNAc-1-phosphotransferase.

The functional mechanism through which MPRs work in vivo to ensure hydrolases transport to lysosomes has largely remained unclear. However, from the studies of receptor deficient fibroblasts, three different subgroups of M6P containing proteins could be distinguished: (i) those which are preferentially secreted in the absence of the CD-MPR and, therefore, do not interact well with the remaining CI-MPR; (ii) those that are preferentially secreted in the absence of the CI-MPR and, therefore, do not interact well with CD-MPR and, finally, (iii) those which are only secreted in the absence of both MPRs, thus interacting equally well with either MPR [55]. This reinforces the idea that the MPRs are complementary in function.

In terms of localization, the MPRs are distributed over several intracellular compartments including the *trans*-Golgi network, the plasma membrane and the endosomes. At steady level, the majority of MPRs reside in the endosomes [35].

When located in the *trans*-Golgi network, the MPRs divert the newly synthesized lysosomal hydrolases from the secretory pathway to the endocytic pathway, after which they unload their bound ligands.

However, not all the hydrolases that are tagged for delivery to the endosome/lysosome compartment arrive to their proper destination. Actually, some escape binding to MPR in the *trans*-Golgi network being directed, by default, to the cell surface, where they are secreted to the extracellular fluid. When this happens some MPRs can take a detour back to the plasma membrane, where they recapture the escaped lysosomal enzymes returning them to lysosomes by receptor-mediated endocytosis via early and late endosomes [in 36]. The CI-MPR localized at the plasma membrane may also contribute to the partial rescue of those hydrolases [45,47]. Notwithstanding, as the majority of lysosomal hydrolases require an acidic milieu to work, they can do not any harm in the extracellular fluid where the pH is usually neutral.

The additional presence of mannose-6-phosphate receptors (CI-MPR) in the plasma membrane allows internalization of glycoproteins containing the M6P recognition mark, which can bind to the cell surface and then be targeted to the lysosome. Taking advantage of this possibility, quite successful enzyme replacement therapies (ERT) were designed for several lysosomal storage disorders. Before administration in patients, the exogenous missing enzyme must be modified with sufficient M6P to bind to the receptor with the appropriate affinity. There are now available licensed products to replacement therapy for Fabry disease, mucopolysacharidosis I, II and VI and Pompe disease. Nevertheless, these approaches still present some limitations since many manifestations of established disease do not respond do ERT and, in some cases (e.g. renal impairment in Fabry disease), disease progression continues despite therapy. Furthermore, such therapeutic options are only suitable for lysosomal storage disorders without neurological manifestations or with low neurological involvement, because recombinant enzymes are not capable of crossing the blood–brain barrier, which would be necessary to neutralize the neurological manifestations of the diseases [reviewed in 57].

4. Impairments in the mannose-6-phosphate pathway and disease

4.1. Defective GlcNAc-1-phosphotransferase causing mucolipidosis II and III

Defective GlcNAc-1-phosphotransferase causes two distinct human lysosomal storage diseases, Mucolipidosis II (ML II) and Mucolipidosis III (ML III), which are among the few lysosomal storage disorders related to defects in non-lysosomal proteins. ML II, also known as I-cell disease, is characterized by a total loss of GlcNAc-1-phosphotransferase activity whether ML III, often referred to as pseudo-Hurler polydystrophy, manifests when enzymatic activity is reduced [58,59]. In both ML II and ML III patients, newly synthesized lysosomal enzymes fail to be correctly sorted to the endosome/lysosome compartment due to the absence or weak equipment in M6P residues. As a consequence, lysosomal dysfunction develops leading to accumulation of non-degraded material, the hallmark of this group of diseases. Unlike the majority of lysosomal storage disorders, which involve single enzymes acting in a catabolic pathway, MLII and III results from impaired sorting of multiple enzymes to lysosomes that instead are over-secreted from cells. The excessive accumulation of non-degraded substrates results in the subsequent formation of large inclusion bodies.

While ML II and ML III share similar clinical features, including skeletal abnormalities, ML II is the more severe in terms of phenotype [58]. In this pathology the skeletal system is severely affected, with abnormalities in both cartilage and bone. Linear growth decelerates during the first year of life almost stopping during the second year and death usually occurs between 5 and 8 years of age.

ML III is a much milder disorder, being characterized by latter onset of clinical symptoms and slower progressive course, which may allow the survival into the eighth decade. Usual clinical findings in ML III patients include restricted joint mobility, short stature and mild Hurler-like dysmorphism, among other less severe features

[59,60]. Only 50% of the patients present with mental retardation [60].

Once GlcNAc-1-phosphotransferase is an hexameric complex whose protein subunits are encoded by two genes, depending on which of them harbors the causal mutation(s) and simultaneously on the severity/clinical course of the disease, the associated pathologies are classified as ML II α/β (OMIM: 252500) and ML III α/β (OMIM: 252600) if mutations are present in the *GNPTAB* gene, or ML III γ (OMIM: 252605) if mutations occur in the *GNPTG* gene [10,12].

To date, more than 100 different *GNPTAB* mutations have been reported, causing either ML II α/β or ML III α/β, including 30 missense, 20 nonsense, 32 small deletions, 25 small insertions, 2 small indels and 14 splice site mutations [61 and references therein, reviewed in 62]. Large genomic rearrangements appear to be rare (1.6%) although two gross insertions [63,64] and one large deletion have already been detected [65]. Most of these mutations are private or rare. Uncommonly, however, the microdeletion c.3503_3504delTC, presents a remarkably wide geographical distribution, having been detected among Israeli and Palestinian Arab-Muslim, Turkish, Irish traveler [66], Italian [67], Portuguese [68] and U.S. patients [69]. That deletion was found to be a frequent mutation in a French Canadian founder population [70] and, later, haplotypic analysis performed on the majority of c.3503_3504delTC chromosomes up to now detected, clearly pointed to the common origin of all of them, indicating that the mutation is ancient enough to explain its present-day broad geographical distribution [71].

Concerning the *GNPTG* gene, the mutations until now reported associated to Mucolipidosis type III include 4 missense, 2 nonsense, 5 small deletions, 4 small insertions, 4 splice site mutations and 2 gross deletions [61 and references therein, reviewed in 62].

4.2. Mutations in the lysosomal enzyme-targeting pathway causing stuttering

Recently, genomewide scans have unveiled a curious and unexpected relation between mutations in the lysosomal enzyme-targeting pathway and persistent stuttering.

This innovative technology is becoming widely applied when complex, multifactorial and heterogeneous diseases are under investigation since it may yield suggestive evidence of linkage at several chromosomal sites that influence phenotypes. Stuttering, is a disorder in which speech fluency can be severely compromised, but its primary causes, the high rates of spontaneous recovery and the etiological differences between persistent and resolved forms have resisted explanation. In the last years, genomewide scans provided signals of linkage at multiple chromosomal sites with stuttering, [72–75] but the stronger evidence was with a locus on chromosome 12q, which was found in a study of stuttering in 46 consanguineous families from Pakistan [73]. Afterwards, when Kang and colleagues [76], focusing on the largest of those families, carried out a search for mutations across the chromosomal interval showing the strongest linkage, ended up identifying a single nucleotide change in the *GNPTAB* gene (G3598A), coding for both α and β subunits of GlcNAc-1-phosphotransferase. Then, the authors went on to sequence *GNPTG*, as well as *NAGPA*, in which mutations causing changes at highly conserved amino acid residues were also identified, as occurred in *GNPTAB*. A strong association was established between persistent stuttering and four mutations in the *GNPTAB* gene (S321G, A455S, F624L and E1200K); three mutations in the *GNPTG* gene (L5_R7dup; A25E and L230V) and three mutations in the *NAGPA* gene (H84Q; R328C and F523SfsX113).

This elegant study unveiled an unsuspected culprit to explain stuttering in some cases, due to the nature of the implicated biologic pathway whose impairment had been essentially connected to lysosomal storage diseases. Yet, the affected subjects studied by Kang et al. [76] stuttered but were otherwise normal, without presenting any of the typical symptoms of lysosomal malfunction, even being homozygous for some of the identified mutations on the *GNPTAB* and *GNPTG* genes. Eventually, this can be explained admitting that the efficiency of lysosomal targeting was only partially reduced in the presence of the detected mutations. However, the effect of the mutations on the function of the GlcNAc-1-phosphotransferase or UCE was not examined, leaving open the possibility that the mutations were not detrimental. Recently, Lee et al. performed a biochemical evaluation of the effect of the three *NAGPA* mutations which were associated with persistent stuttering and found that each mutation leads to lower cellular UCE activity [77].

The most interesting finding was the strong connection between stuttering and three enzymes that generate the mannose 6 phosphate signal. Expression levels of the encoding genes in human brain remain largely unknown, but in the mouse *GNPTG* has the most localized expression in the brain, with high levels of expression in the hippocampus, hippocampal formation and cerebellum (according to the Allen Brain Atlas and as postulated in [76]), well-known structures associated with emotion and motor function. Taking this into account, it seems plausible that mutations in such gene may led to stuttering since a person's emotional state can be a major conditioner of the severity of stuttering. Furthermore, stuttering does not affect the ability to conceptualize words and/or sentences, but only the fluency of speech, which is in accordance with a disruption of the motor functions required for speech [76].

Further studies in this area may contribute to a better understanding of the neural structures and functions within the brain that generate human speech, which are still poorly elucidated and to a better comprehension on the role of the mannose-6-phosphate-dependent pathway for lysosomal enzyme routing in this area.

4.3. Secondary impaired uptake of lysosomal enzymes: A 'symptom' instead of a 'cause'

Up to this point, we have only addressed primary impairments in the mannose-6-phosphate pathway, which are damages caused by mutations in genes coding for the functional components of the mannose-6-phosphate-dependent pathway that lead to specific clinical phenotypes. However, recent studies revealed that expression levels of these functional components may also be secondarily affected by other disease conditions.

Secondary impairments include, for example, increased levels of the cation-independent mannose-6-phosphate receptor (CI-MPR) localized at the plasma membrane in cells from patients suffering from metachromatic leukodystrophy, a lysosomal storage disorder, caused by the deficiency of arylsulfatase A (ARSA) with consequent storage of the sphingolipid 3-O-sulfogalactosylceramide. This compound accumulates in several cell types, such as renal tubular cells. It was demonstrated that in sulfatide-storing kidney cells the uptake of lysosomal enzymes is more than two-fold increased. Uptake rates were inferred trough analysis of CI-MPR expression levels (1.5-fold increased) and internalization rate (1.5-fold increased), with the two alterations accounting for the global 2.2-fold increase in M6P-dependent endocytosis of recombinant ARSA in sulfatide-storing cells. Recycling of CI-MPR was also found to be decreased. Similarly, altered expression levels were observed at the transferrin receptor, indicating that ARSA deficient cells present a general alteration of the endocytotic pathway [78].

4.4. Influence of expression levels of M6P pathway functional components on the efficacy of some therapeutic approaches

Evidence is emerging that the expression levels of several mannose-6-phosphate pathway functional components may affect the efficacy of different therapeutic approaches in LSDs.

In a recent study, Sun and colleagues [78] have evaluated the efficacy of gene replacement therapy in Pompe disease under different conditions. They used acid α-glucosidase (GAA) knockout mice, immune tolerant GAA KO mice and mannose-6-phophate deficient GAA-KO mice, in order to assess the impact of the following three factors in muscle-targeted gene therapy: antibody formation, age and mannose-6-phosphate receptor availability. Overall, antibody formation had a subtle effect upon biochemical efficacy from GAA expression, whereas the absence of mannose-6-phosphate receptors markedly impaired muscle-targeted gene therapy in murine Pompe disease. These results indicate that reduced mannose-6-phosphate receptor availability in type II myofibers represents an obstacle to a therapeutic strategy in Pompe disease that otherwise could be advantageous [79].

5. Conclusion and future perspectives

Over the last decades, much was learned about the mannose-6-phosphate-dependent pathway, the main cellular passageway for transporting soluble hydrolases to lysosomes. Its major functional components are now quite well characterized both at biochemical and genetic levels. Nevertheless, several important questions remain unanswered. Although the key enzymes necessary to the formation of the mannose-6-phosphate recognition signal have been identified, the protease that cleaves and thereby activates the α/β subunit precursor of GlcNAc-1-phosphotransferase is still unknown. Further studies on the β and γ subunits of GlcNAc-1-phosphotransferase are also needed, in order to elucidate better the role of β subunit and what is the functional relevance of γ subunit.

Secondary impairments of the mannose-6-phosphate pathway were only very recently brought to light, as well were signals of how expression levels of its functional components might affect therapeutic approaches in LSDs. The next few years, certainly will bring new data allowing for a better understanding of the biological impact of this pathway, which will be pivotal to design more efficient therapeutic strategies to overcome M6P related disruptions of enzyme lysosomal sorting.

Another issue that importantly needs additional insights concerns alternative mechanisms to the M6P pathway for enzyme lysosomal targeting that mammalian cells possess. Some years ago, studies with cells from patients with I-cell disease demonstrated that in some cell types, such as hepatocytes, nearly normal levels of lysosomal enzymes were present, despite the absence of M6P signals due to GlcNac-1-phosphotransferase deficiency. This finding clearly pointed to the existence of M6P-independent pathway(s) for lysosomal sorting and trafficking that could be used in some cell types but not in others. Accordingly, two alternative receptors have been discovered: LIMP-2 and sortilin. LIMP-2 was shown to be implicated in the delivery of β-glucocerebrosidase to the lysosomes [80–82], while sortilin was acknowledged to be a multifunctional receptor capable of binding several ligands including neurotensin, RAP (receptor-associated protein) and to ensure the lysosomal trafficking of several proteins, both hydrolytic or non-hydrolytic such as the sphingolipid activator proteins prosaposin and GM2 activator protein, acid sphingomyelinase and cathepsins D and H [83–88]. But there is still much to be learnt on these and other alternative pathways, as well on how M6P-dependent and independent routes possibly cooperates targeting to lysosomes the full set of enzymes necessary to accomplish the organelle function.

Acknowledgments

This work was supported by FCT — project PIC/IC/83252/2007 (http://alfa.fct.mctes.pt/). Coutinho MF is a grantee from the FCT (SFRH/BD/48103/2008).

References

[1] T. Braulke, J.S. Bonifacino, Sorting of lysosomal proteins, Biochim. Biophys. Acta 1793 (2009) 605–614.
[2] T. Lübke, P. Lobel, D.E. Sleat, Proteomics of the lysosome, Biochim. Biophys. Acta 1793 (2009) 625–635.
[3] P.J. Meikle, J.J. Hopwood, Lysosomal storage disorders: emerging therapeutic options require early diagnosis, Eur. J. Pediatr. 162 (Suppl. 1) (2003) S34–S37.
[4] K.B. Clairmont, M.P. Czech, Chicken and *Xenopus* mannose 6-phosphate receptors fail to bind insulin-like growth factor II, J. Biol. Chem. 264 (1989) 16390–16392.
[5] E. Guillen, L.A. Quesada-Allue, R.O. Couso, UDP-N-acetylglucosamine: glycoprotein N-acetylglucosamine-1-phosphotransferase activity in pupae of the Mediterranean fruit fly *Ceratitis capitata*, Insect Biochem. Mol. Biol. 24 (1994) 213–219.
[6] V. Alvarez, A.J. Parodi, R. Couso, Characterization of the mannose 6-phosphate-dependent pathway of lysosomal enzyme routing in an invertebrate, Biochem. J. 310 (1995) 589–595.
[7] L. Lang, R. Couso, S. Kornfeld, Glycoprotein phosphorylation in simple eucaryotic organisms. Identification of UDP-GlcNAc:glycoprotein N-acetylglucosamine-1-phosphotransferase activity and analysis of substrate specificity, J. Biol. Chem. 261 (1986) 6320–6325.
[8] Y. Qian, C.M. West, S. Kornfeld, UDP-GlcNAc:Glycoprotein N-acetylglucosamine-1-phosphotransferase mediates the initial step in the formation of the methylphosphomannosyl residues on the high mannose oligosaccharides of *Dictyostelium discoideum* glycoproteins, Biochem. Biophys. Res. Commun. 393 (2010) 678–681.
[9] C.M. Ketcham, S. Kornfeld, Characterization of UDP-N-acetylglucosamine:glycoprotein N-acetylglucosamine-1-phosphotransferase from *Acanthamoeba castellanii*, J. Biol. Chem. 267 (1992) 11654–11659.
[10] S. Tiede, S. Storch, T. Lübke, B. Henrissat, R. Bargal, A. Raas-Rothschild, T. Braulke, Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase, Nat. Med. 11 (10) (2005) 1109–1112.
[11] K. Marschner, K. Kollmann, M. Schweizer, T. Braulke, S. Pohl, A key enzyme in the biogenesis of lysosomes is a protease that regulates cholesterol metabolism, Science 333 (6038) (2011) 87–90.
[12] A. Raas-Rothschild, V. Cormier-Daire, M. Bao, E. Genin, R. Salomon, K. Brewer, M. Zeigler, H. Mandel, S. Toth, B. Roe, A. Munnich, W.M. Canfield, Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC), J. Clin. Invest. 105 (2000) 673–681.
[13] S. Tiede, M. Cantz, A. Raas-Rothschild, N. Muschol, F. Bürger, K. Ullrich, T. Braulke, A novel mutation in UDP-N-acetylglucosamine-1-phosphotransferase gamma subunit (GNPTAG) in two siblings with mucolipidosis type III alters a used glycosylation site, Hum. Mutat. 24 (2004) 535.
[14] M. Encarnação, K. Kollmann, M. Trusch, T. Braulke, S. Pohl, Post-translational modifications of the {gamma}-subunit affect intracellular trafficking and complex assembly of glcnac-1-phosphotransferase, J. Biol. Chem. 286 (2011) 5311–5318.
[15] M. Kudo, M. Bao, A. D'Souza, F. Ying, H. Pan, B.A. Roe, W.M. Canfield, The alpha- and beta-subunits of the human UDP-N-acetylglucosamine:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase [corrected] are encoded by a single cDNA, J. Biol. Chem. 280 (2005) 36141–36149.
[16] W.S. Lee, B.J. Payne, C.M. Gelfman, P. Vogel, S. Kornfeld, Murine UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase lacking the gamma-subunit retains substantial activity toward acid hydrolases, J. Biol. Chem. 282 (2007) 27198–27203.
[17] M. Bao, J.L. Booth, B.J. Elmendorf, W.M. Canfield, Bovine UDP-N-acetylglucosamine: lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure, J. Biol. Chem. 271 (1996) 31437–31445.
[18] M. Kudo, W.M. Canfield, Structural requirements for efficient processing and activation of recombinant human UDP-N-acetylglucosamine:lysosomal-enzyme-N-acetylglucosamine-1-phosphotransferase, J. Biol. Chem. 281 (2006) 11761–11768.
[19] K.H. Paik, S.M. Song, C.S. Ki, H.W. Yu, J.S. Kim, K.H. Min, S.H. Chang, E.J. Yoo, I.J. Lee, E.K. Kwan, S.J. Han, D.K. Jin, Identification of mutations in the GNPTA (MGC4170) gene coding for GlcNAc-phosphotransferase alpha/beta subunits in Korean patients with mucolipidosis type II or type IIIA, Hum. Mutat. 26 (2005) 308–314.
[20] T.J. Baranski, P.L. Faust, S. Kornfeld, Generation of a lysosomal enzyme targeting signal in the secretory protein pepsinogen, Cell 63 (1990) 81–91.
[21] A.B. Cantor, T.J. Baranski, S. Kornfeld, Lysosomal enzyme phosphorylation. II. Protein recognition determinants in either lobe of procathepsin D are sufficient for phosphorylation of both the amino and carboxyl lobe oligosaccharides, J. Biol. Chem. 267 (1992) 23349–23356.
[22] J.W. Cuozzo, K. Tao, Q.L. Wu, W. Young, G.G. Sahagian, Lysine-based structure in the proregion of procathepsin L is the recognition site for mannose phosphorylation, J. Biol. Chem. 270 (1995) 15611–15619.
[23] Y. Qian, I. Lee, W.S. Lee, M. Qian, M. Kudo, W.M. Canfield, P. Lobel, S. Kornfeld, Functions of the alpha, beta, and gamma subunits of UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase, J. Biol. Chem. 285 (2010) 3360–3370.
[24] H.J. Sommerlade, T. Selmer, A. Ingendoh, V. Gieselmann, K. von Figura, K. Neifer, B. Schmidt, Glycosylation and phosphorylation of arylsulfatase A, J. Biol. Chem. 269 (1994) 20977–20981.
[25] C.M. Gelfman, P. Vogel, T.M. Issa, C.A. Turner, W.S. Lee, S. Kornfeld, D.S. Rice, Mice lacking alpha/beta subunits of GlcNAc-1-phosphotransferase exhibit growth retardation, retinal degeneration, and secretory cell lesions, Invest. Ophthalmol. Vis. Sci. 48 (11) (2007) 5221–5228.
[26] S. Pohl, S. Tiede, M. Castrichini, M. Cantz, V. Gieselmann, T. Braulke, Compensatory expression of human N-acetylglucosaminyl-1-phosphotransferase subunits in mucolipidosis type III gamma, Biochim. Biophys. Acta 1792 (2009) 221–225.

[27] S. Pohl, S. Tiede, K. Marschner, M. Encarnação, M. Castrichini, K. Kollmann, N. Muschol, K. Ullrich, S. Müller-Loennies, T. Braulke, Proteolytic processing of the gamma-subunit is associated with the failure to form GlcNAc-1-phosphotransferase complexes and mannose 6-phosphate residues on lysosomal enzymes in human macrophages, J. Biol. Chem. 285 (2010) 23936–23944.

[28] A. Varki, S. Kornfeld, Purification and characterization of rat liver alpha-N-acetylglucosaminyl phosphodiesterase, J. Biol. Chem. 256 (1981) 9937–9943.

[29] R. Kornfeld, M. Bao, K. Brewer, C. Noll, W. Canfield, Molecular cloning and functional expression of two splice forms of human N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase, J. Biol. Chem. 274 (1999) 32778–32785.

[30] W.S. Lee, J. Rohrer, R. Kornfeld, S. Kornfeld, Multiple signals regulate trafficking of the mannose 6-phosphate-uncovering enzyme, J. Biol. Chem. 277 (2002) 3544–3551.

[31] H. Do, W.S. Lee, P. Ghosh, T. Hollowell, W. Canfield, S. Kornfeld, Human mannose 6-phosphate-uncovering enzyme is synthesized as a proenzyme that is activated by the endoprotease furin, J. Biol. Chem. 277 (2002) 29737–29744.

[32] C.A. Chavez, R.N. Bohnsack, M. Kudo, R.R. Gotschall, W.M. Canfield, N.M. Dahms, Domain 5 of the cation-independent mannose 6-phosphate receptor preferentially binds phosphodiesters (mannose 6-phosphate N-acetylglucosamine ester), Biochemistry 46 (2007) 12604–12617.

[33] M. Boonen, P. Vogel, K.A. Platt, N. Dahms, S. Kornfeld, Mice lacking mannose 6-phosphate uncovering enzyme activity have a milder phenotype than mice deficient for N-acetylglucosamine-1-phosphotransferase activity, Mol. Biol. Cell 20 (2009) 4381–4389.

[34] L.J. Olson, F.C. Peterson, A. Castonguay, R.N. Bohnsack, M. Kudo, R.R. Gotschall, W.M. Canfield, B.F. Volkman, N.M. Dahms, Structural basis for recognition of phosphodiester-containing lysosomal enzymes by the cation-independent mannose 6-phosphate receptor, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 12493–12498.

[35] S. Kornfeld, I. Mellman, The biogenesis of lysosomes, Annu. Rev. Cell Biol. 5 (1989) 483–525.

[36] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, Molecular Biology of the Cell, 4th ed., 2002 (New York).

[37] E. van Meel, J. Klumperman, Imaging and imagination: understanding the endo-lysosomal system, Histochem. Cell Biol. 129 (3) (2008) 253–266.

[38] S. Kornfeld, Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors, Annu. Rev. Biochem. 61 (1992) 307–330.

[39] T. Ludwig, R. Le Borgne, B. Hoflack, Roles for mannose-6-phosphate receptors in lysosomal enzyme sorting, IGF-II binding and clathrin-coat assembly, Trends Cell Biol. 5 (1995) 202–206.

[40] P. Ghosh, N.M. Dahms, S. Kornfeld, Mannose 6-phosphate receptors: new twists in the tale, Nat. Rev. Mol. Cell Biol. 4 (2003) 202–212.

[41] P. Lobel, N.M. Dahms, S. Kornfeld, Cloning and sequence analysis of the cation-independent mannose 6-phosphate receptor, J. Biol. Chem. 263 (1988) 2563–2570.

[42] A. Oshima, C.M. Nolan, J.W. Kyle, J.H. Grubb, W.S. Sly, The human cation-independent mannose 6-phosphate receptor. Cloning and sequence of the full-length cDNA and expression of functional receptor in COS cells, J. Biol. Chem. 263 (1988) 2553–2562.

[43] B. Westlund, N.M. Dahms, S. Kornfeld, The bovine mannose 6-phosphate/insulin-like growth factor II receptor. Localization of mannose 6-phosphate binding sites to domains 1–3 and 7–11 of the extracytoplasmic region, J. Biol. Chem. 266 (1991) 23233–23239.

[44] N.M. Dahms, P.A. Rose, J.D. Molkentin, Y. Zhang, M.A. Brzycki, The bovine mannose 6-phosphate/insulin-like growth factor II receptor. The role of arginine residues in mannose 6-phosphate binding, J. Biol. Chem. 268 (1993) 5457–5463.

[45] N.M. Dahms, D.A. Wick, M.A. Brzycki-Wessell, The bovine mannose 6-phosphate/insulin-like growth factor II receptor. Localization of the insulin-like growth factor II binding site to domains 5–11, J. Biol. Chem. 269 (1994) 3802–3809.

[46] P.G. Marron-Terada, M.K. Hancock, D.J. Haskins, N.M. Dahms, Recognition of Dictyostelium discoideum lysosomal enzymes is conferred by the amino-terminal carbohydrate binding site of the insulin-like growth factor II/mannose 6-phosphate receptor, Biochemistry 39 (2000) 2243–2253.

[47] M.K. Hancock, R.D. Yammani, N.M. Dahms, Localization of the carbohydrate recognition sites of the insulin-like growth factor II/mannose 6-phosphate receptor to domains 3 and 9 of the extracytoplasmic region, J. Biol. Chem. 277 (2002) 47205–47212.

[48] N.M. Dahms, P. Lobel, J. Breitmeyer, J.M. Chirgwin, S. Kornfeld, 46 kd mannose 6-phosphate receptor: cloning, expression, and homology to the 215 kd mannose 6-phosphate receptor, Cell 50 (1987) 181–192.

[49] P.Y. Tong, S. Kornfeld, Ligand interactions of the cation-dependent mannose 6-phosphate receptor. Comparison with the cation-independent mannose 6-phosphate receptor, J. Biol. Chem. 264 (1989) 7970–7975.

[50] N.M. Dahms, S. Kornfeld, The cation-dependent mannose 6-phosphate receptor. Structural requirements for mannose 6-phosphate binding and oligomerization, J. Biol. Chem. 264 (1989) 11458–11467.

[51] A. Waheed, K. von Figura, Rapid equilibrium between monomeric, dimeric and tetrameric forms of the 46-kDa mannose 6-phosphate receptor at 37 degrees C. Possible relation to the function of the receptor, Eur. J. Biochem. 193 (1990) 47–54.

[52] A. Köster, P. Saftig, U. Matzner, K. von Figura, C. Peters, R. Pohlmann, Targeted disruption of the M(r) 46,000 mannose 6-phosphate receptor gene in mice results in misrouting of lysosomal proteins, EMBO J. 12 (1993) 5219–5223.

[53] T. Ludwig, C.E. Ovitt, U. Bauer, M. Hollinshead, J. Remmler, P. Lobel, U. Rüther, B. Hoflack, Targeted disruption of the mouse cation-dependent mannose 6-phosphate receptor results in partial missorting of multiple lysosomal enzyme, EMBO J. 12 (1993) 5225–5235.

[54] I. Sohar, D. Sleat, C. Gong Liu, T. Ludwig, P. Lobel, Mouse mutants lacking the cation-independent mannose 6-phosphate/insulin-like growth factor II receptor are impaired in lysosomal enzyme transport: comparison of cation-independent and cation-dependent mannose 6-phosphate receptor-deficient mice, Biochem. J. 330 (1998) 903–908.

[55] H. Munier-Lehmann, F. Mauxion, U. Bauer, P. Lobel, B. Hoflack, Re-expression of the mannose 6-phosphate receptors in receptor-deficient fibroblasts. Complementary function of the two mannose 6-phosphate receptors in lysosomal enzyme targeting, J. Biol. Chem. 271 (1996) 15166–15174.

[56] T. Ludwig, H. Munier-Lehmann, U. Bauer, M. Hollinshead, C. Ovitt, P. Lobel, B. Hoflack, Differential sorting of lysosomal enzymes in mannose 6-phosphate receptor-deficient fibroblasts, EMBO J. 13 (1994) 3430–3437.

[57] R. Lachmann, Treatments for lysosomal storage disorders, Biochem. Soc. Trans. 38 (2010) 1465–1468.

[58] J.G. Leroy, J.W. Spranger, I-cell disease, N. Engl. J. Med. 283 (1970) 598–599.

[59] J.W. Spranger, H.R. Wiedemann, The genetic mucolipidoses. Diagnosis and differential diagnosis, Humangenetik 9 (1970) 113–139.

[60] F. Umehara, W. Matsumoto, M. Kuriyama, K. Sukegawa, S. Gasa, M. Osame, Mucolipidosis III (pseudo-Hurler polydystrophy); clinical studies in aged patients in one family, J. Neurol. Sci. 146 (1997) 167–172.

[61] Human Gene Mutation Database websiteHGMD, http://www.hgmd.org.

[62] K. Kollmann, S. Pohl, K. Marschner, M. Encarnação, I. Sakwa, S. Tiede, B.J. Poorthuis, T. Lübke, S. Müller-Loennies, S. Storch, T. Braulke, Mannose phosphorylation in health and disease, Eur. J. Cell Biol. 89 (1) (2010) 117–123.

[63] B. Tappino, S. Regis, F. Corsolini, M. Filocamo, An Alu insertion in compound heterozygosity with a microduplication in GNPTAB gene underlies Mucolipidosis II, Mol. Genet. Metab. 93 (2008) 129–133.

[64] T. Otomo, T. Muramatsu, T. Yorifuji, T. Okuyama, H. Nakabayashi, T. Fukao, T. Ohura, M. Yoshino, A. Tanaka, N. Okamoto, K. Inui, K. Ozono, N. Sakai, Mucolipidosis II and III alpha/beta: mutation analysis of 40 Japanese patients showed genotype-phenotype correlation, J. Hum. Genet. 54 (2009) 145–151.

[65] Coutinho F., Silva Santos L., Lacerda L., Quental S., Wibrand F., Lund A.M., Johansen K.B., Prata M.J., Alves S. Alu–Alu Recombination Underlying the First Large Genomic Deletion in GlcNAc-Phosphotransferase Alpha/Beta (GNPTAB) Gene in a MLII Alpha/Beta Patient. J. Inherit. Metab. Dis. (in press)

[66] R. Bargal, M. Zeigler, B. Abu-Libdeh, V. Zuri, H. Mandel, Z. Ben Neriah, F. Stewart, N. Elcioglu, T. Hindi, M. Le Merrer, G. Bach, A. Raas-Rothschild, When Mucolipidosis III meets Mucolipidosis II: GNPTA gene mutations in 24 patients, Mol. Genet. Metab. 88 (2006) 359–363.

[67] B. Tappino, N.A. Chuzhanova, S. Regis, A. Dardis, F. Corsolini, M. Stroppiano, E. Tonoli, T. Beccari, C. Rosano, J. Mucha, M. Blanco, M. Szlago, M. Di Rocco, D.N. Cooper, M. Filocamo, Molecular characterization of 22 novel UDP-N-acetylglucosamine-1-phosphate transferase alpha- and beta-subunit (GNPTAB) gene mutations causing mucolipidosis types IIalpha/beta and IIIalpha/beta in 46 patients, Hum. Mutat. 30 (2009) E956–E973.

[68] M. Encarnação, L. Lacerda, R. Costa, M.J. Prata, M.F. Coutinho, H. Ribeiro, L. Lopes, M. Pineda, J. Ignatius, H. Galvez, A. Mustonen, P. Vieira, M.R. Lima, S. Alves, Molecular analysis of the GNPTAB and GNPTG genes in 13 patients with mucolipidosis type II or type III — identification of eight novel mutations, Clin. Genet. 76 (2009) 76–84.

[69] M. Kudo, M.S. Brem, W.M. Canfield, Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha / beta-subunits precursor gene, Am. J. Hum. Genet. 78 (2006) 451–463.

[70] M. Plante, S. Claveau, P. Lepage, E.M. Lavoie, S. Brunet, D. Roquis, C. Morin, H. Vézina, C. Laprise, Mucolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTAB) in a French Canadian founder population, Clin. Genet. 73 (2008) 236–244.

[71] M. Coutinho, M. Encarnação, R. Gomes, L. Da Silva Santos, S. Martins, D. Sirois-Gagnon, R. Bargal, M. Filocamo, A. Raas-Rothschild, B. Tappino, C. Laprise, G. Cury, I. Schwartz, O. Artigalás, M. Prata, S. Alves, Origin and spread of a common deletion causing mucolipidosis type II: insights from patterns of haplotypic diversity, Clin. Genet. (2010) 273–280.

[72] Y.Y. Shugart, J. Mundorff, J. Kilshaw, K. Doheny, B. Doan, J. Wanyee, E.D. Green, D. Drayna, Results of a genome-wide linkage scan for stuttering, Am. J. Med. Genet. A 124A (2004) 133–135.

[73] N. Riaz, S. Steinberg, J. Ahmad, A. Pluzhnikov, S. Riazuddin, N.J. Cox, D. Drayna, Genomewide significant linkage to stuttering on chromosome 12, Am. J. Hum. Genet. 76 (2005) 647–651.

[74] R. Suresh, N. Ambrose, C. Roe, A. Pluzhnikov, J.K. Wittke-Thompson, M.C. Ng, X. Wu, E.H. Cook, C. Lundstrom, M. Garsten, R. Ezrati, E. Yairi, N.J. Cox, New complexities in the genetics of stuttering: significant sex-specific linkage signals, Am. J. Hum. Genet. 78 (2006) 554–563.

[75] J.K. Wittke-Thompson, N. Ambrose, E. Yairi, C. Roe, E.H. Cook, C. Ober, N.J. Cox, Genetic studies of stuttering in a founder population, J. Fluency Disord. 32 (2007) 33–50.

[76] C. Kang, S. Riazuddin, J. Mundorff, D. Krasnewich, P. Friedman, J.C. Mullikin, D. Drayna, Mutations in the lysosomal enzyme-targeting pathway and persistent stuttering, N. Engl. J. Med. 362 (2010) 677–685.

[77] W.S. Lee, C. Kang, D. Drayna, S. Kornfeld, Analysis of mannose 6-phosphate uncovering enzyme mutations associated with persistent stuttering, J. Biol. Chem. 286 (2011) 39786–39793.

[78] D. Klein, A. Yaghootfam, U. Matzner, B. Koch, T. Braulke, V. Gieselmann, Mannose 6-phosphate receptor-dependent endocytosis of lysosomal enzymes is increased in sulfatide-storing kidney cells, Biol. Chem. 390 (2009) 41–48.

[79] B. Sun, S. Li, A. Bird, H. Yi, A. Kemper, B.L. Thurberg, D.D. Koeberl, Antibody formation and mannose-6-phosphate receptor expression impact the efficacy of

muscle-specific transgene expression in murine Pompe disease, J. Gene Med. 12 (2010) 881–891.
[80] D. Reczek, M. Schwake, J. Schröder, H. Hughes, J. Blanz, X. Jin, W. Brondyk, S. Van Patten, T. Edmunds, P. Saftig, LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase, Cell 131 (2007) 770–783.
[81] A. Balreira, P. Gaspar, D. Caiola, J. Chaves, I. Beirão, J.L. Lima, J.E. Azevedo, M.C. Miranda, A nonsense mutation in the LIMP-2 gene associated with progressive myoclonic epilepsy and nephrotic syndrome, Hum. Mol. Genet. 17 (2008) 2238–2243.
[82] S.F. Berkovic, L.M. Dibbens, A. Oshlack, J.D. Silver, M. Katerelos, D.F. Vears, R. Lüllmann-Rauch, J. Blanz, K.W. Zhang, J. Stankovich, R.M. Kalnins, J.P. Dowling, E. Andermann, F. Andermann, E. Faldini, R. D'Hooge, L. Vadlamudi, R.A. Macdonell, B.L. Hodgson, M.A. Bayly, J. Savige, J.C. Mulley, G.K. Smyth, D.A. Power, P. Saftig, M. Bahlo, Array-based gene discovery with three unrelated subjects shows SCARB2/LIMP-2 deficiency causes myoclonus epilepsy and glomerulosclerosis, Am. J. Hum. Genet. 82 (2008) 673–684.
[83] C.M. Petersen, M.S. Nielsen, A. Nykjaer, L. Jacobsen, N. Tommerup, H.H. Rasmussen, H. Roigaard, J. Gliemann, P. Madsen, S.K. Moestrup, Molecular identification of a novel candidate sorting receptor purified from human brain by receptor-associated protein affinity chromatography, J. Biol. Chem. 272 (1997) 3599–3605.
[84] J. Mazella, N. Zsürger, V. Navarro, J. Chabry, M. Kaghad, D. Caput, P. Ferrara, N. Vita, D. Gully, J.P. Maffrand, J.P. Vincent, The 100-kDa neurotensin receptor is gp95/sortilin, a non-G-protein-coupled receptor, J. Biol. Chem. 273 (1998) 26273–26276.
[85] M.S. Nielsen, C. Jacobsen, G. Olivecrona, J. Gliemann, C.M. Petersen, Sortilin/neurotensin receptor-3 binds and mediates degradation of lipoprotein lipase, J. Biol. Chem. 274 (1999) 8832–8836.
[86] S. Lefrancois, J. Zeng, A.J. Hassan, M. Canuel, C.R. Morales, The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin, EMBO J. 22 (2003) 6430–6437.
[87] X. Ni, C.R. Morales, The lysosomal trafficking of acid sphingomyelinase is mediated by sortilin and mannose 6-phosphate receptor, Traffic 7 (2006) 889–902.
[88] M. Canuel, A. Korkidakis, K. Konnyu, et al., Sortilin mediates the lysosomal targeting of cathepsins D and H, Biochem. Biophys. Res. Commun. 373 (2008) 292–297.

Molecular Genetics and Metabolism 107 (2012) 257–266

Contents lists available at SciVerse ScienceDirect

Molecular Genetics and Metabolism

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

Minireview

A shortcut to the lysosome: The mannose-6-phosphate-independent pathway

Maria Francisca Coutinho [a,b,c,*], Maria João Prata [b,c], Sandra Alves [a]

[a] Research and Development Unit, Department of Genetics, CGMJM, INSA, Portugal
[b] IPATIMUP, Porto, Portugal
[c] Department of Biology, Faculty of Sciences, Porto, Portugal

ARTICLE INFO

Article history:
Received 12 July 2012
Accepted 12 July 2012
Available online 20 July 2012

Keywords:
Mannose-6-phosphate-dependent pathway
Sortilin/neurotensin receptor-3
Lysosomal integral membrane protein 2 (LIMP2)
Action myoclonus-renal failure syndrome (AMRF)

ABSTRACT

Lysosomal hydrolases have long been known to be responsible for the degradation of different substrates in the cell. These acid hydrolases are synthesized in the rough endoplasmic reticulum and transported through the Golgi apparatus to the *trans*-Golgi network (TGN). From there, they are delivered to endosomal/lysosomal compartments, where they finally become active due to the acidic pH characteristic of the lysosomal compartment. The majority of the enzymes leave the TGN after modification with mannose-6-phosphate (M6P) residues, which are specifically recognized by M6P receptors (MPRs), ensuring their transport to the endosomal/lysosomal system. Although M6P receptors play a major role in the intracellular transport of newly synthesized lysosomal enzymes in mammalian cells, several lines of evidence suggest the existence of alternative processes of lysosomal targeting. Among them, the two that are mediated by the M6P alternative receptors, lysosomal integral membrane protein (LIMP-2) and sortilin, have gained unequivocal support. LIMP-2 was shown to be implicated in the delivery of beta-glucocerebrosidase (GCase) to the lysosomes, whereas sortilin has been suggested to be a multifunctional receptor capable of binding several different ligands, including neurotensin and receptor-associated protein (RAP), and of targeting several proteins to the lysosome, including sphingolipid activator proteins (prosaposin and GM2 activator protein), acid sphingomyelinase and cathepsins D and H.
Here, we review the current knowledge on these two proteins: their discovery, study, structural features and cellular function, with special attention to their role as alternative receptors to lysosomal trafficking. Recent studies associating both LIMP2 and sortilin to disease are also extensively reviewed.

Contents

* Corresponding author at: Research and Development Unit, Department of Genetics, CGMJM, INSA, Praça Pedro Nunes, 88, 4099-028 Porto, Portugal. Fax: +351 226070399.
E-mail addresses: francisca_coutinho@yahoo.com, francisca.coutinho@insa.min-saude.pt (M.F. Coutinho).

1096-7192/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2012.07.012

1. Introduction

Lysosomes have long been known to be responsible for the degradation of different substrates in the cell, a crucial process only possible because of the existence of soluble enzymes inside the acidic organelle. Most of those acid hydrolases require previous modification with mannose-6-phosphate (M6P) residues, which allow their recognition by specific M6P receptors present in the *trans*-Golgi network (TGN), ensuring the transport to the endosomal/lysosomal system [1].

Although M6P receptors play a major role in the intracellular transport of newly synthesized lysosomal enzymes in mammalian cells, there are several lines of evidence suggesting the existence of an alternative mechanism of lysosomal targeting [2], the most relevant of which have come from studies with cells from patients with I-cell disease. Because these individuals harbor mutations in the gene encoding for the phosphotransferase that adds M6P residues to hydrolases, they were expected to present reduced or null levels of hydrolases inside the lysosome. Nevertheless, in some types of cells such as hepatocytes, Kupffer cells and lymphocytes, the cellular levels of lysosomal enzymes can be normal [3–5] despite the inability to add mannose-6-phosphate to newly synthesized hydrolases [5]. These findings pivotally suggested the existence of M6P-independent pathways for lysosomal sorting and trafficking.

Other experimental approaches have proven that there are soluble enzymes as well as non-enzymatic proteins that are transported to lysosomes in a M6P-independent manner, specifically by means of two alternative receptors: the lysosomal integral membrane protein (LIMP-2) and sortilin. LIMP-2 was implicated in the delivery of GCase (the defective enzyme in patients with Gaucher disease) to the lysosomes [6]. Sortilin is a multifunctional receptor capable of binding different ligands, including neurotensin and receptor-associated protein (RAP), and has been suggested to mediate Golgi-to-lysosome transport of the sphingolipid activator proteins (SAPs) prosaposin (PSAP) and GM2 activator protein (GM2AP), acid sphingomyelinase (ASM) and cathepsins D and H [2,7–11].

Herein, we review the current knowledge on these two proteins, their discovery, study, structural features and cellular function, with special attention to their role as alternative receptors to lysosomal trafficking. We will also review the recently discovered disease associations: the causal role of LIMP-2 mutations in a severe autosomal recessive lysosomal storage disorder (action myoclonus-renal failure syndrome, AMRF) and their roles as modifiers of Gaucher disease (GD) and sortilin expression levels as regulators of hepatic lipoprotein export, associated to increased risk of myocardial infarction (MI) in humans.

2. LIMP2

2.1. Historical perspective

The lysosomal integral membrane protein, LIMP-2, was recently shown to be implicated in the delivery of GCase (the enzyme defective in patients with Gaucher disease) to the lysosomes [6].

LIMP-2 was first described in 1995 [12], but its role as a sorting receptor was only unveiled several years later when Reczek and colleagues unexpectedly identified LIMP-2 as a specific receptor for GCase [6]. In the following year, Berkovic et al. [13], aiming to map a severe genetic disease of unknown etiology characterized by action myoclonus associated with renal failure (AMRF), found that *SCARB2*, the gene that codes for LIMP-2, was responsible for the condition. In this study, the disorder was first mapped in 4q13-q21, and using microarray expression analysis, it was possible to identify the gene encoding a lysosomal membrane protein as the most likely candidate found in the critical region of chromosome 4. Then, the screening of the *SCARB2* gene led to the identification of mutations associated with lack of the LIMP2 protein in the 3 families diagnosed with AMRF included in the study. Subsequent reanalysis of a pre-existing *Limp2* knockout mouse showed intracellular inclusions in the cerebral and cerebellar cortex, as well as subtle glomerular changes in the kidneys. Notably, in 2003, Gamp et al. [14] had already reported on Limp2-deficient mice with increased postnatal mortality associated with uni- or bilateral hydronephrosis caused by ureteropelvic junction obstruction. Additionally, those mice showed an accumulation of lysosomes in epithelial cells of the ureter as well as a disturbed apical expression of uroplakin, further suggesting an impairment of membrane transport processes. Limp2-deficient mice suffered from a massive decline of spiral ganglia in the cochlea concomitant with that of the inner and outer hair cells. However, these pathologic changes were thought to be secondary to degeneration of the stria vascularis. Limp2-deficient mice were also characterized by peripheral demyelinating neuropathy, associated with massive loss of peripheral myelin proteins and increased activity and expression of lysosomal proteins.

2.2. A specific lysosomal receptor

LIMP-2 was shown to be a specific lysosomal receptor involved in the transport of GCase (GBA) from the TGN to the endosomal/lysosomal compartment. GCase has no membrane-spanning domain [15]; it is glycosylated after synthesis, becoming membrane-associated [3]. Only in 2007 was the machinery required for its transport to the endosomal/lysosomal compartment unveiled. It was discovered that LIMP-2 binds GCase in a pH-dependent fashion, enabling its association in the ER and transport to the lysosome, where the complex likely dissociates due to the acidic pH [6]. In fact, during the passage through the various organelles of the secretory pathway, the receptor-ligand complex experiences a gradient of decreasing pH: the nearly neutral milieu of the ER allows GCase to associate with LIMP2, whereas the low pH of the endosomes/lysosomes leads to the dissociation of GCase from LIMP2 [16]. Additional studies from Blanz et al. [17] demonstrated that, through disruption of either the helical arrangement or the amphiphatic nature of the LIMP2 coiled-coil domain (residues 152 to 167), GCase binding was abolished. Furthermore, a synthetic peptide comprising the coiled-coil domain of LIMP2 was shown to display pH-selective multimerization properties. In contrast to the reduced binding properties of nonsense mutations affecting residues 145 to 288, the H363N mutation led to increased binding of GBA, indicating that the highly conserved His363 residue may modify the affinity of LIMP2 to its ligand. All these data led to the conclusion that both disruption of the coiled-coil structure and AMRF disease-causing mutations abolish GBA binding, indicating the importance of an intact coiled-coil structure for the interaction of LIMP2 and GCase. In addition, Blanz et al. [17] investigated the biochemical function of several *SCARB2* mutations in transfected cells, having observed that all three nonsense mutations under study led to the retention of the mutant protein in the endoplasmic reticulum (ER) but affected the binding to GCase in different ways: in the presence of the Q288X mutation, LIMP2 was still able to bind to GBA as efficiently as the wild-type protein, whereas the W146SfsX16 and W178X mutant proteins lost their GBA-binding capacity almost completely.

Further analysis of the pH dependency of LIMP-2 and GCase binding by perturbing the intra-lysosomal pH in cultured cells revealed that lumenal acidification mediated by the vacuolar (H^+)-ATPase triggers the dissociation of LIMP-2 and GC in late endosomal/lysosomal compartments [16]. Moreover, the same authors identified a single histidine residue in LIMP2 (H171) that likely serves as the critical pH sensor. Interestingly, that particular residue is in close proximity to the previously described coiled-coil motif in the luminal domain of LIMP2, which is necessary for GCase binding. Interestingly, that residue is not conserved in chicken or *Xenopus*; instead, an arginine residue is present at the corresponding position, which would favor the dissociation of the receptor and its ligand. In addition, the protein sequences of these species are also devoid of the proposed coiled-coil domain, which is also an important GCase binding factor. This led the authors to suggest that it is unlikely that non-mammalian LIMP2 proteins are able to bind GCase

orthologs in these species. Thus, such species may be of great value to decipher additional roles of LIMP2 other than transporting GCase to lysosomes [16].

2.3. LIMP2 and disease: action myoclonus-renal failure syndrome

As mentioned in Section 2.1, mutations in the gene coding for this receptor prevent GCase from reaching the lysosome and from assuming its function inside the lysosome. Not surprisingly, *SCARB2* mutations underlie a serious autosomal-recessive disorder presently known as action myoclonus-renal failure syndrome (AMRF; OMIM#254900) [13,18]. This disease combines progressive myoclonus epilepsy (PME), associated with storage material in the brain, and focal glomerulosclerosis, frequently with glomerular collapse.

At the biochemical level, this condition was shown to be characterized by pathological levels of GCase activity in fibroblasts, normal or slightly reduced levels in leukocytes, increased levels in plasma and the absence of other markers of Gaucher disease (GD), such as elevated chitotriosidase activity [19].

Recently, Hopfner and colleagues [20], by screening a German AMRF family with these features associated with renal failure, provided data indicating that demyelinating polineuropathy and dilated cardiomyopathy are part of the AMRF syndrome. The findings were confirmed almost simultaneously by Dibbens et al. [21], who reported on a patient with progressive myoclonus epilepsy and demyelinating peripheral neuropathy, also carrying a *SCARB2* mutation that could explain the phenotype. This particular case had previously been published in the *Archives of Neurology* and presented with PME, preserved intellect and a nonprogressive generalized demyelination neuropathy [22]. By that time, failure to establish the cause of the phenotype, despite the extensive evaluation, led the authors to describe a novel PME syndrome, but due to the absence of renal failure, they withdrew the initial AMRF diagnosis.

Currently, the presentation of PME without renal impairment [23] prompts screening of the *SCARB2* gene for causative mutations, and new findings exemplify the study of Rubboli et al. [24], which reports on *SCARB2* mutations in five non-related Italian patients who presented with PME of unknown origin of adolescent or adult onset.

So far, no *SCARB2* genotype–phenotype correlations have been established. From the inspection of the SCARB2 mutations found to date, no clear distinction emerges between the class (e.g., missense or nonsense) or position of the mutation and the patient phenotype. The phenotypic heterogeneity encompasses a wide range of tissues, but it is still unclear why some patients with *SCARB2* mutations develop PME with or without renal failure, or why others develop mild hearing impairment or peripheral neuropathy [24].

Other symptoms may also occur, as suggested by the clinical manifestations observed in LIMP2-deficient mice [14], which presented increased postnatal mortality and uni- or bilateral hydronephrosis caused by an obstruction of the ureteropelvic junction. The obstruction was associated with metaplasia of the uretetic urothelium into simple columnar epithelium and hypertrophy of the smooth muscle layer. In addition, LIMP2-deficient mice suffered from a peripheral demyelinating neuropathy. Demyelination was found to be associated with a loss of peripheral myelin proteins. Serious hearing impairments were also observed and were associated with massive spiral ganglion neuron losses, concomitant with loss of the inner and outer hair cells and a strongly impaired capacity to generate endocochlear potential [14]. In a subsequent study, those mice were also shown to display progressive high-frequency hearing loss and otoacoustic emissions. In fact, more in depth studies on that model revealed that the potassium channel KCNQ1 and its beta-subunit KCNE1 were almost completely lost in the luminal part of marginal cells in the stria vascularis, affecting first higher and then lower frequency processing cochlear turns. The study of this mouse model and of its deafness was particularly important because it suggests an important and previously unsuspected role for LIMP2 in controlling the localization and the level of apically expressed membrane proteins such as KCNQ1, KCNE2 and megalin in the stria vascularis [25]. Extrapolation of mouse studies to humans is still premature, but it is tempting to hypothesize that LIMP2 may assume additional roles in the endocytic pathway. Nevertheless, because for the moment only a few clinical and neurophysiologic descriptions of AMRF syndrome have been reported [13,18,26–30], there is still much to learn about the clinical presentation and molecular basis of this disease. Future research exploring the function of *SCARB2* gene and its role in cerebral and renal function is warranted [24]. Further insights into its function in the endocytic pathway and its unidentified interacting partners [25] are also needed.

2.4. LIMP2 and disease: SCARB2 mutations as modifiers of Gaucher disease

Recently, *SCARB2* mutations have also been demonstrated to act as modifiers in Gaucher disease. Considering that LIMP-2 is the only known receptor for the enzyme that is deficient in Gaucher disease, Velayati and co-workers hypothesized that *SCARB2* mutations could impact the Gaucher phenotype. In order to test that hypothesis, they re-evaluated a previously reported case of two siblings with GD with very disparate phenotypes [31,32]. One sibling had been followed closely for over a decade with progressive myoclonic epilepsy (PME) and dementia, while the second was subsequently diagnosed through family screenings, having only a few disease manifestations throughout his life, with no neurologic involvement. Both shared three GBA alterations: the maternally inherited c.535G>C (p.Asp140His) and c.10936G>A (p.Glu326Lys) and the paternally inherited c.586A>C (p.Lys157Gln) [31]. Extensive analyses, including studies of endoplasmic reticulum-associated degradation (ERAD), have previously been performed on fibroblasts from both siblings, demonstrating that the one with myoclonic epilepsy had increased ERAD of GCase and elevated intracellular cholesterol [32]. Velayati et al. screened this sib-pair for mutations in *SCARB2* and found one novel, heterozygous, maternally-inherited mutation, c.1412A>G (p.Glu471Gly), in the brother with GD and myoclonic epilepsy, which was absent in his sibling and controls. This finding demonstrated that LIMP-2 could serve as a modifier in GD. Although myoclonic epilepsy had previously been reported in heterozygotes with *SCARB2* mutations [6], only in this case was it possible to demonstrate that the *SCARB2* mutation alone was not responsible for the development of such a phenotype, as it was also present in the sib's mother, who had a normal clinical phenotype [33]. In the severely affected brother, however, the deficient GBA activity, accompanied by a mutation in the transporter, resulted in mistrafficking of the enzyme and contributed to the observed phenotype.

3. Sortilin

3.1. Historical perspective: from a neurotensin alternative receptor to an essential sorting receptor

Sortilin was first discovered in 1997, when Petersen and colleagues [7] performed receptor-associated protein affinity chromatography and isolated, identified and purified a previously unknown membrane glycoprotein of approximately 95 kDa. This protein was numbered gp95 and later coined as sortilin. After cloning of the encoding gene by cDNA library screening, it was possible to verify that it has a type I receptor structure and striking sequence homologies with yeast Vps10p and the CD-MRP and CI-MPR receptors [7]. Transfection analysis of this novel receptor demonstrated that it is able to bind receptor-associated binding protein (RAP), which had previously only been described for LDL receptors, and also showed that its cytoplasmic tail provides localization to the Golgi and CI-MRP vesicles. Yet, the cell functions of sortilin were far from being totally unveiled. In the next

years, three major discoveries provided support for the hypothesis that sortilin could be a candidate sorting receptor, targeted for transport by ligands in the synthetic pathway as well as on the surface membrane. First, as demonstrated by Mazella et al. [8] and Nielsen et al. [9], sortilin binds a variety of unrelated ligands in mammals, including lipoprotein lipase and neurotensin. These ligands have in common the ability to bind to other receptor(s) and candidates for regulated transport [34–36]. Therefore, even though sortilin appeared to contribute to endocytosis and perhaps even signaling as evidenced by different works, its main role seems plausibly related to functions not covered by other receptors (e.g., regulated transport). Secondly, sortilin has striking structural similarities to known receptors involved in intracellular sorting and transport. Indeed, the C-terminal segment in sortilin's cytoplasmic tail is closely related to the corresponding and functionally important segment in Vps10p, the sorting receptor for carboxypeptidase Y (CPY) in yeast [37–39]. Moreover, the sortilin cytoplasmic domain contains several potential signal sequences that conform to established consensus motifs known to be involved in adaptor protein binding, endocytosis, basolateral targeting and Golgi-endosome sorting [40]. Thirdly, sortilin predominates in certain intracellular compartments; only a minor fraction of the protein is expressed on the cell surface, whereas as much as 90% of the receptor pool is found in the Golgi and its derived vesicles, showing extensive co-localization with CI-MPR.

This prompted the next steps to examine the nature of sortilin as a sorting receptor. In 2001, when examining the cellular trafficking of chimeric receptors containing constructs of the sortilin tail, Nielsen et al. [34] demonstrated that CI-MPR-sortilin chimeras expressed in CI-MPR knockout cells were almost as efficient as the MPR itself in the transport of newly synthesized beta-hexosaminidase and beta-glucoronidase to lysosomes. It then became established that the sortilin tail contains potent signals for Golgi-endosome sorting, while the entire receptor had potential to convey transport from the synthetic pathway to lysosomes. Later, direct binding was shown between this luminal receptor and a member of the GGA family of cytosolic proteins (multidomain protein family involved in protein trafficking between Golgi and endosomes), suggesting that GGA2 participates in sortilin sorting. After this study, a number of works reinforced the ability of sortilin to mediate lysosomal trafficking of several enzymes and non-enzymatic proteins such as prosaposin, GM2AP [10], acid sphingomyelinase [2] and cathepsins D and H [11] (see Section 3.3).

3.2. Structural features: insights into sortilin function and evolution

Sortilin is one of five members of a Vps10p domain receptor family, often found in the trans-Golgi network and early endosomes: sortilin, SorLA, SorCS1, SorCS2 and SorCS3 [41]. The Vps10 domain is a 10-bladed beta-propeller consisting of three structural domains: the N-terminal domain (residues 45–576), which was the first example of a 10-bladed beta-propeller structure, followed by two smaller domains named 10CC-a (residues 577–633) and 10CC-b (residues 634–716). The 10CC domains have low secondary structure content and both interact extensively with the beta-propeller [41]. Being the protein with the largest number of blades at the Vps10 domain (10), sortilin has an increased potential to accommodate larger ligands when compared to similar proteins. With several highly conserved regions in the inner rim of the tunnel that correspond to different binding sites, sortilin has a unique structural organization that allows it to bind a series of different ligands. Furthermore, it is translated as an inactive propeptide, which is only activated after furine cleavage. In fact, it has been proven that newly synthesized sortilin (and SorLA) cannot bind ligands to their Vps10p domain until they have been cleaved by furin in the TGN [42]. This simple structural organization seems to be a key element of protein regulation because the sortilin propeptide blocks premature binding of any ligand, preventing simultaneous binding of ligands with conflicting functions [reviewed in 41]. Over the next years, evidence started gathering that both sortilin and SorLA shuttle between the TGN and late endosomes [34,43]. This type of trafficking was shown to involve interactions between the cytoplasmic domain of the receptors and various adaptors, including adaptor-protein-1 (AP1), Golgi-localized, γ ear-containing, ARF-binding proteins (GGA1, GGA2 and GGA3) and the retromer complex [43]. Other structural features exist, giving more insights into this protein's function and evolution. The cytosolic tail of sortilin closely resembles the one of the mannose-6-phosphate receptors (MPRs), containing motifs known to be involved in trafficking from the Golgi to the endosome and vice versa: FLV, YSVL and an acidic-cluster dileucine motif, all interacting with the same adaptor proteins (GGAs and AP-1) responsible for the forward transport of the MPRs. In addition, the retrograde recycling of sortilin also involves an interaction with the retromer complex through a YXXϕ site in the cytosolic tail [reviewed in 44]. There are also four particular amino acids in the cytoplasmic domain shared by both sortilin and SorLA (the MVIA motif), which are crucial for binding to the cytosolic adaptors GGA1, 2 and 3 [45].

There are some clues suggesting that the sortilin sorting pathway may represent an evolutionary older mechanism for lysosomal transport than the MPR pathway. Actually, even though the MPR is usually described as analogous to the yeast Vps10p receptor, sortilin shares more similarities with Vps10p than the MPR because both sortilin and Vps10p contain the Vps10 domain, which is not present in the MPR [46], and sortilin and Vps10p both interact with their ligands through proteinaceous interactions, while the MPR recognizes its ligands through an oligosaccharide side-chain, the M6P tag [37,47,48, reviewed in 44].

It appears, therefore, that the MPR pathway emerged as a very specific mechanism for sorting soluble lysosomal hydrolases. Nevertheless, it did not completely substitute the mechanism mediated by sortilin, which retained the ability to translocate different cargo proteins (see Section 3.3).

3.3. A multiligand lysosomal receptor?

In 2003, the accumulated evidence of a possible role of sortilin as an alternative sorting receptor in the M6P-independent transport of newly synthesized hydrolases lead Lefrancois and colleagues [10] to investigate whether sortilin could be involved in the sorting of SAPs. Their attention was drawn to these activator proteins because earlier reports firmly demonstrated that SAPs (prosaposin and GM2AP) could reach the lysosomes in an M6P-independent manner.

A few years earlier, a carbohydrate-independent pathway of native GM2AP had been characterized [49]. Formerly, immunocytochemical analysis of tunicamycin-treated cells had shown that non-glycosylated prosaposin could be targeted more efficiently to lysosomes, whereas biochemical analysis revealed that prosaposin was associated with Golgi membrane fractions in an association not disrupted by free mannose-6-phosphate [50].

When Lefrancois et al. performed the first study, they applied three approaches: a dominant-negative competition experiment, siRNA and Co-IP assays, which as a whole suggested that the SAPs (prosaposin and GM2AP) constituted a class of proteins whose lysosomal sorting and trafficking were mediated by sortilin. The dominant-negative construct was a truncated sortilin lacking the acidic cluster dileucine signal (Δcytosolic), which is implicated in the binding of GGAs [34]. That construct was retained in the Golgi apparatus and localized to the same compartment as the Golgi marker. Full-length sortilin, in contrast, was not restricted to the perinuclear region but extended into punctuate structures. A C-terminal myc-tag was added to the construct, to discriminate between truncate and endogenous sortilin. The overexpression of such a construct abolished the punctuate staining seen in non-transfected COS-7 cells with anti-GM2AP antibody, suggesting that the truncated sortilin competed for the binding of prosaposin and abolished its transport to the lysosomal compartment. Through siRNA

assays, the authors saw that ablation of sortilin increases the secretion of prosaposin into the culture media. Finally, co-immunoprecipitation assays showed that sortilin interacts with both SAPs [10].

Then in 2006, using the same dominant-negative sortilin construct (essential to recruit adaptor proteins and clathrin), Ni and colleagues [2] demonstrated that sortilin was also involved in the lysosomal targeting of acid sphingomyelinase (ASM). Confocal microscopy revealed that truncated sortilin partially inhibited the lysosomal trafficking of ASM in COS-7 cells and abolished the lysosomal targeting of ASM in I-cells. Pulse-chase experiments corroborated that sortilin was involved in normal sorting of newly synthesized ASM. Furthermore, over-expression of truncated sortilin accelerated and enhanced the secretion of ASM from COS-7 cells and I-cells. Co-immunoprecipitation assays confirmed the interaction between sortilin and ASM [2].

Two years later, sortilin was also implicated in the lysosomal targeting of cathepsins D and H [11] in a study that was also based on Co-IP assays but that used cathepsin B as a negative control and prosaposin as a positive control. Interestingly, when analyzing the cathepsins' interaction with the CI-MPR, the authors observed that, unlike cathepsin H, which bound exclusively to sortilin, cathepsin D was also able to bind to the CI-MPR. In the same study, COS7 cells were transfected with a myc-tagged truncated sortilin construct, and its effect on the transport of the lysosomal proteins was analyzed by confocal immuno-microscopy. Untransfected cells labeled with anti-cathepsin D or anti-cathepsin H antibodies revealed the presence of cathepsins D and H in punctate lysosomal structures that overlapped with LAMP-1. However, anti-cathepsin D and anti-cathepsin H antibodies did not yield granular staining in cells expressing truncated sortilin but did label the perinuclear region, suggesting that truncated sortilin binds and retains cathepsins D and H within the Golgi apparatus. This finding further showed that truncated sortilin-myc competed not only with endogenous sortilin, but also with the mannose-6-phosphate receptor to abrogate sorting of cathepsin D. Canuel et al. [11] also addressed the effect of sortilin inactivation on the transport of cathepsins D and H through sortilin siRNA transfection and confocal microscopy analysis after staining with anti-cathepsin D or H antibodies. The granular labeling and perinuclear labeling of prosaposin and cathepsin H in untransfected cells were absent in cells transfected with sortilin siRNA, but the granular cytoplasmic staining of both anti-cathepsin B and D antibodies was unaffected by the absence of sortilin. Thus the effect of sortilin ablation on cathepsin D differed from the dominant-negative experiment. To better understand the role of sortilin and the M6PR in the trafficking of cathepsins D and H, cells with both the sortilin and the M6P pathways blocked were examined, demonstrating that the transport of cathepsin D to the lysosomes occurs through both sorting pathways. In conclusion, sortilin appears to be the sole receptor necessary for lysosomal translocation of cathepsin H. Alternatively, the efficient transport of cathepsin D requires both the M6P receptor and sortilin.

Other soluble proteases that are known to be transported in a M6P-independent fashion exist, but this transport is not mediated by sortilin. This is the case with cathepsins K and L, where evidence of MPR-independent transport are well established. When the M6P tag was added to CHO cells expressing cathepsin K, its subcellular localization remained unaltered, showing no competition between cathepsin K and exogenous M6P [51]. For cathepsin L, low affinity levels for the MPR were registered [52], and mutagenic analysis of vacuolar sorting determinants suggested a proteinaceous nature for its vacuolar sorting signal, in contrast with the oligosaccharide nature of the M6P signal expected for MPR-dependent transport [43,53]. Nevertheless, when both cathepsins were studied using an approach similar to that applied to cathepsins D and H, it was observed that neither cathepsins K nor L interacted and immunoprecipitated with the MPR, nor did they interact with sortilin. Their subcellular localization remained unaltered in cells transfected with truncated sortilin when compared to untransfected cells, clearly implying that cathepsins K and L do not require sortilin for transport to the lysosomes [44].

Based on their assumption that sortilin was the sorting and trafficking receptor for at least five soluble lysosomal proteins, Carlos Morales' group tested the hypothesis that inactivation of the sortilin gene results in lysosomal storage disorders. To their surprise, none of the sortilin nullizygous mice exhibited clinical signs of lysosomal pathologies. Based on that observation, the authors decided to quantify the concentration of prosaposin in the lysosomes of nonciliated epithelial cells lining the efferent ducts of sortilin$^{-\backslash-}$ and sortilin$^{+\backslash+}$ mice. Nonciliated cells are able to target both exogenous and endogenous prosaposin to the lysosomes and are an excellent in vivo system with which to analyze the effect of sortilin gene inactivation in cells with or without luminal prosaposin. The authors observed that inactivation of sortilin resulted in a significant decrease in the level of prosaposin in endosomes and lysosomes. Nevertheless, that inhibition was only partial. In summary, the creation, characterization and analysis of a sortilin knockout mouse, while supporting the idea that sortilin transports prosaposin, also suggests the existence of an alternative mechanism of sorting and transport of the sortilin receptor [54]. On the frontline of the most probable alternative receptors stands SorLA, the only other known Vps10p family member containing motifs that have been shown to be involved in trafficking from the Golgi to the endosome and vice versa. Nevertheless, no studies to evaluate its function in the M6P-independent transport of lysosomal hydrolases have yet been conducted.

Some points concerning sortilin's function as an alternative receptor for lysosomal transport remain unclear, however. Additional experiments with non-tagged sortilin constructs and specific anti-sortilin antibodies instead of myc-tagged ones would help to clarify sortilin's subcellular location and interaction with its putative ligands by preventing any disruption of the normal activation process of the protein by furin cleavage.

In general, there is still much to learn about the possible role of Vps10p family receptors on the M6P-independent trafficking pathway. Further studies are needed on both sortilin and SorLA. A compelling experiment would be the generation of a double knockout for both proteins. In fact, the gene for SorLA has also been inactivated in a previous study, and the knockout mice did not exhibit abnormal phenotypes, which also suggests the existence of a complementary receptor [55]. Finally, it would also be desirable to seek out rare human patients with a large-effect mutation in *SORT1* (e.g., a loss of function mutation) and to carefully define the phenotype [56].

3.4. A multifunctional protein

Even though sortilin is an ancient lysosomal sorting receptor, during the evolutionary process it has developed additional and novel functions [Table 1 and references therein] that are gradually being discovered. Several lines of evidence exist showing that in various pathways sortilin plays key roles, which can range from important functions in normal biological processes, such as glucose metabolism [60,68] and neuronal cell proliferation and death [41,57], to influencing pathological processes like those underlying Alzheimer's disease [69]. Sortilin is also acknowledged to play a part in lipoprotein metabolism by interacting with RAP [7,46] and binding lipoprotein lipase (LPL) [9] and apolipoprotein (apo) A-V [63].

Since its early description, it has been known that sortilin is highly expressed in several tissues, especially in the brain [7], which may be easily explained by its role as a receptor of neuromediators [8] and growth factors and also by its involvement in the fronto-temporal lobar degeneration pathway. Indeed, sortilin may form part of a signaling complex that regulates cell survival as evidenced by Teng and colleagues after performing coexpression assays of sortilin and p75 neurotrophin receptor (p75NTR) in sympathetic neurons [70]. The authors found that the effects of proBDNF, an apoptotic ligand that induces

Table 1
Pleiotropic activities of sortilin in various biological pathways.
Adapted from Dubé et al. [56].

Pathway	Sortilin mediated role
Lysosomal transport	Mediates Golgi-to-lysosome transport of sphingolipid activator proteins (prosaposin and GM2 activator protein) [10]
	Mediates Golgi-to-lysosme transport of acid sphingomyelinase [2]
	Mediates Golgi-to-lysosme transport of cathepsins D and H [11]
Neurotrophic signaling	Binds proNGF forming an apoptotic signaling complex with p75NTR facilitating neuronal cell death [57]
	Transports TrkA, TrkB and TrkC to nerve synapse terminals enhancing their interaction with neurotrophins; facilitates neuronal differentiation, survival and plasticity [58]
Fronto-temporal lobar degeneration	Modulates extracellular PGRN levels via binding and endocytosis; pathogenesis via decreased extracellular PGRN levels through haploinsufficiency [59]
Carbohydrate metabolism	Binds and translocates Glut4 to the plasma membrane from Glut4 storage vesicles in response to insulin signaling [60]
Osteoblast differentiation	Limits LPL-mediated suppression of mineralization activity in differentiating hMSC [61]
	Identified as a marker of late osteogenesis [62]
Lipoprotein metabolism	Binds LDL RAP [46]
	Surface-expression mediates binding to LPL and subsequent degradation by endocytosis [9]
	In vitro binding and endocytosis of apo A-V; facilitates lysosomal degradation of apo A-V [63]
	Robust association between SORT1 variation (chromosome 1p13) and LDL-C concentrations [64]
	Expression assays show regulatory effects on plasma LDL-C concentrations [65–67]
	Implicated in apo B-100 metabolism and VLDL secretion [67]

death at subnanomolar concentrations, are dependent on the cellular coexpression of both p75NTR and sortilin because neurons deficient in p75NTR are resistant to proBDNF-induced apoptosis, and competitive antagonists of sortilin block sympathetic neuron death [70]. In addition, sortilin facilitates neuronal differentiation, survival and plasticity by transporting TrkA, TrkB and TrkC receptors to nerve synapse terminals, enhancing their interaction with neurotrophins, as shown in the study of Vaegter et al. [58], who demonstrated that sortilin is a positive modulator of neurotrophin-induced neuronal survival. Neuronal survival, differentiation and synaptic plasticity depend on the binding of target-derived neurotrophins to Trk receptors at nerve terminals. Nevertheless, the distance between the soma and the nerve terminal is large enough to make efficient anterograde Trk transport a critical step for Trk synaptic translocation and signaling. Vaegter and colleagues also observed that sortilin interacted with TrkA, TrkB and TrkC receptors to enable their anterograde axonal transport, consequently concluding that it enhanced neurotrophin signaling. Finding that sortilin was frequently coexpressed with Trk and p75NTR in tissues that are not destined for apoptosis (e.g., adult dorsal root ganglia (DRG)), the authors hypothesized that sortilin might affect trafficking and subcellular distribution of Trk receptors. In order to test that hypothesis, they cultured DRG neurons lacking sortilin having verified that those cells presented blunted MAP kinase signaling and reduced neurite outgrowth upon stimulation with nerve growth factor (NGF). These results were also supported by in vivo experiments in double knockout mice ($Sort1^{-/-}$; $Ngfr^{-/-}$-knockout for the p75NTR coding gene), which presented an aggravated phenotype when compared to the typical Trk phenotypes observed in $Ngfr^{-/-}$ mice. Moreover, when sortilin knockouts ($Sort1^{-/-}$) were crossed with mice heterozygous for a mutant TrkA allele ($Ntrk1^{+/-}$), an embryonic lethality and increased sympathetic neuropathy with smaller superior cervical ganglion (SCG) were observed [58].

Very recently, new evidence emerged indicating that sortilin also has a role during phagosome maturation. Wähe and colleagues [71], using a latex-bead phagosome (LBP) model, showed that in macrophages, sortilin is mainly localized in the Golgi and transported to phagosomes. Through live-cell imaging and electron microscopy, the authors observed that the delivery of sortilin to LBP was dependent on its cytoplasmic tail and that sortilin participates in the direct delivery of acid sphingomyelinase (ASM) and prosaposin (PS) to the phagosome, bypassing fusion with lysosomal compartments. Furthermore, an analysis of primary macrophages isolated from $Sort1^{-/-}$ mice indicated that the delivery of ASM and PS to LBP was severely impaired, but the transport of pro-cathepsin D remained normal. Considering these results, the authors proposed a sortilin-mediated pathway by which selected lysosomal proteins are transported to the phagosome along a Golgi-dependent route during phagosomal maturation.

In summary, additional roles of sortilin are still being unveiled, suggesting that its function is far from being fully understood.

3.5. Sortilin and disease: the cardiovascular risk association

Over the last few years, novel technological advances such as next-generation sequencing catapulted genome wide association studies (GWAS) of population-based cohorts into the limelight.

GWAS are reasonably successful in the identification of single nucleotide polymorphisms (SNPs) within or near genes associated with variations in plasma lipid and lipoprotein levels. More than 100 loci were reported to account for a significant part of the genetic variation in triglyceride, low-density lipoprotein (LDL) and high-density lipoprotein cholesterol levels. In coronary artery disease or MI, GWAS have identified a smaller number of genetic loci, some of which are also associated with changes in traditional lipoprotein risk factors. Notably, within a widely replicated chromosome 1p13 locus associated with both MI and LDL cholesterol levels [66,72–74], the relevant SNPs are located in a gene cluster containing four genes, one of them *SORT1*, the gene coding for human sortilin. It has a key role contributing to the effects of the chromosome 1p13q locus on LDL cholesterol and coronary artery disease, which was described by three major studies. The first was developed by Linsel–Nitschke and colleagues who, from their observations, proposed that the overexpression of sortilin increases the internalization of LDL, with a consequent decrease of its plasma levels [65]. Soon after, through studies in human cohorts, hepatocytes and mice, Musunuru et al. reported an inverse relationship between sortilin expression and circulating LDL-C levels and proposed an explanatory mechanism through transcriptional regulation (liver-specific) of the *SORT1* gene by the transcription factor C/EBPα [66]. Conversely, Kjolby and his team observed a direct relationship between the expression of *Sort1* and the concentration of circulating LDL, suggesting that it could result from increased VLDL secretion [67].

Several explanations have been presented to justify the discrepancy between these results and the answer seems to lie in sortilin itself, which appears to be a multifaceted protein that may assume different functions depending on the circumstances.

Overall however, the three studies presented strong evidence demonstrating that *SORT1* is a regulatory element of plasma LDL-C

levels, adding a significant role to the sortilin-coding gene, which was ignored until very recently.

3.6. Sortilin and disease: molecular basis for familial amyotrophic lateral sclerosis?

Recently, Belzil and co-workers [75] hypothesized that, by reducing progranulin levels and promoting neurodegeneration, *SORT1* mutations or *SORT1* aberrant splicing could cause amyotrophic lateral sclerosis (ALS).

That hypothesis was based on independent assumptions. First, causative mutations had just been reported in the *TARDBP* and *FUS* genes, encoding respectively for TAR DNA binding protein TDP-43 and RNA binding protein FUS (both involved in the regulation of transcription and RNA splicing) [76–79], in both amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) [80,81]. The latter is a condition commonly seen in comorbidity with ALS. It had also been found that depletion of TDP-43 in the mouse adult brain modifies the expression levels of *FUS* and *GRN* (gene encoding progranulin) and alters the splicing of *SORT1* [82]. Finally, sortilin had also recently been identified as a receptor that mediates progranulin uptake [59]. Thus, the authors speculated that *TARDBP*, *GRN*, and *SORT1* act in a common pathological pathway and that mutations in any of those genes would result in a similar phenotype.

To test this notion, they screened the *SORT1* coding sequence for mutations in a cohort of 112 unrelated individuals with familial ALS and tested for aberrant splicing by RT-PCR using RNA samples from cell lines expressing six different ALS-associated *TARDBP* mutations [75]. They identified one unique missense and two unique silent mutations; none of them were predicted to have functional effects. No aberrant *SORT1* splicing events were observed. SORT1 mutations are not a common cause of familial ALS, and the influence of TARDBP mutations on SORT1 splicing still requires clarification.

4. Conclusion

Although the M6P pathway is the primary sorting mechanism implicated in lysosomal targeting, it is becoming clear that alternative receptors are also involved in the lysosomal sorting of some soluble proteins. Initial evidence for the existence of additional receptors came from studies of immortalized cell lines of patients suffering from I-cell disease, a lysosomal storage disorder in which the M6P sorting pathway is disabled due to mutations in the genes that code for UDP-N-acetylglucosamine-1-phosphotransferase. As a result of the loss of enzyme function, lysosomal hydrolases are not modified with M6P residues, and consequently, their traffic to the endosomal–lysosomal compartment is impaired. Despite that, under this serious condition, some soluble hydrolases such as SAPs and/or soluble hydrolases can still be delivered to the lysosomes, implying that their transport is necessarily conducted in a M6P-independent fashion [3,83]. Presently, there are two potential alternative receptors: LIMP2 and possibly sortilin. LIMP2 transports GCase (the enzyme defective in patients with Gaucher disease) to the lysosomes. Recently, mutations in its encoding gene, *SCRAB2*, were shown to cause a severe autosomal recessive disorder characterized by PME and nephrotic syndrome — AMRF. Sortilin is a multifunctional protein that, apart from its role as a multiligand receptor, was proposed to be responsible for the lysosomal transport of SAPs (prosaposin and GM2AP), acid sphingomyelinase and cathepsins D and H [2,10,11]. Like MPRs, sortilin is thought to be an ancient receptor involved in a conserved trafficking mechanism, but evolutionarily, the sortilin sorting receptor has developed new functions essential in embryonic development, neural function and apoptosis [44,11,84,85]. Further understanding this network topology will permit better insights into why certain lysosomal enzymes travel to the lysosomes in a M6P-independent fashion while others do not, eventually leading to the understanding of the mechanism underlying lysosomal storage disorders of unknown origin [Table 2 and references therein].

The disruption of mechanisms involved in the trafficking of lysosomal enzymes might have dramatic pathological consequences. Until recently, attention was focused on lysosomal diseases resulting from mutations in the gene that codes GlcNAc-phosphotransferase, which plays a crucial role in lysosomal transport in a M6P-dependent manner, because the enzyme is responsible for adding M6P residues to enzymes transported to the endosomal/lysosomal system. Total or near-total deficiency of GlcNAc-phosphotransferase results in the nearly complete absence of lysosomal targeting in many cell types and tissues, together with the secretion of most lysosomal enzymes. This defect is clinically recognized as mucolipidosis II (ML II or I-cell disease; MIM# 252500), an autosomal recessive disease often resulting in death in the 1st decade of life [92]. Another type of mucolipidosis, mucolipidosis III (ML III or pseudo-Hurler polydystrophy; MIM# 252600), is a milder disease in which GlcNAc-phosphotransferase activity is reduced rather than absent [93]. Both these pathologies belong to a group of rare genetic diseases known as lysosomal storage diseases (LSD).

Now is the time to delve deeper to gain insights into the alternative trafficking pathways to the lysosome. Strong evidence is accumulating indicating that there is a series of proteins whose lysosomal transport is done neither by the MPRs nor by any of the known alternative receptors. Cathepsins K and L, for example, have long been known to reach the lysosomes through a M6P-independent pathway [43,51,53] and were recently proven not to be transported by any of the known alternative receptors [44]. Other hydrolases exist, though, that do not acquire M6P residues and consequently travel to the lysosome in a MPR-independent fashion. Such is the case of lysosomal acid phosphatase, for example. This lysosomal hydrolase is a type I transmembrane protein that follows the constitutive secretory pathway to the plasma membrane and reaches the lysosome via endocytosis. In lysosomes, the transmembrane precursor is processed into a soluble form [94]. It contains the tyrosine-based motif YRHV in its cytosolic domain, which is required for endocytosis but does not mediate direct Golgi-to-endosome targeting [95–97]. It is clear that there is still much to learn about Golgi-to-lysosome transport of lysosomal hydrolases.

A proper understanding of these trafficking pathways is essential to gain a better understanding of lysosomal transport mechanisms and may also unveil the molecular defects underlying LSD phenotypes

Table 2
Lysosomal storage disorders associated to deficiencies of alternative lysosomal receptors and/or of their cargos.

Lysosomal receptor	Transported proteins	Associated lysosomal storage disorder	Reference
Sortilin		No known associated LSD	
	GM2AP	AB-variant of GM2 gangliosidosis	[86]
	Saposin B	Variant of metachromatic leukodystrophy	[87]
	Saposin C	Variant of Gaucher's	[88]
	Cathepsin D	Neuronal ceroid lipofuscinoses (CNCL)	[89]
	Acid sphingomyelinase	Niemann–Pick types A and B	[90]
LIMP-2		Action myoclonus-renal failure	[13]
	Beta-glucocerebrosidase	Gaucher	[91]

presented by patients who remain unclassified because their biochemical profiles do not fit the characteristics of any known disorder.

Conflict of interest

None of the authors declares any conflict of interest.

Acknowledgments

This work was supported by FCT — project PTDC/SAU-GMG/102889/2008. Coutinho MF is a grantee from the FCT (SFRH/BD/48103/2008).

References

[1] T. Braukle, J.S. Bonifacino, Sorting of lysosomal proteins, Biochim. Biophys. Acta 1793 (2008) 605–614.
[2] X. Ni, M. Canuel, C.R. Morales, The sorting and trafficking of lysosomal proteins, Histol. Histopathol. 21 (2006) 899–913.
[3] S. Rijnboutt, H.M. Aerts, H.J. Geuze, J.M. Tager, G.J. Strous, Mannose 6-phosphate-independent membrane association of cathepsin D, glucocerebrosidase, and sphingolipid-activating protein in HepG2 cells, J. Biol. Chem. 266 (1991) 4862–4868.
[4] S. Rijnboutt, A.J. Kal, H.J. Geuze, H. Aerts, G.J. Strous, Mannose 6-phosphate-independent targeting of cathepsin D to lysosomes in HepG2 cells, J. Biol. Chem. 266 (1991) 23586–23592.
[5] A. Tsuji, K. Omura, Y. Suzuki, I-cell disease: evidence for a mannose 6-phosphate independent pathway for translocation of lysosomal enzymes in lymphoblastoid cells, Clin. Chim. Acta 176 (1988) 115–121.
[6] D. Reczek, M. Schwake, J. Schröder, H. Hughes, J. Blanz, X. Jin, W. Brondyk, S. Van Patten, T. Edmunds, P. Saftig, LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase, Cell 131 (2007) 770–783.
[7] C.M. Petersen, M.S. Nielsen, A. Nykjaer, L. Jacobsen, N. Tommerup, H.H. Rasmussen, H. Roigaard, J. Gliemann, P. Madsen, S.K. Moestrup, Molecular identification of a novel candidate sorting receptor purified from human brain by receptor-associated protein affinity chromatography, J. Biol. Chem. 272 (1997) 3599–3605.
[8] J. Mazella, N. Zsürger, V. Navarro, J. Chabry, M. Kaghad, D. Caput, P. Ferrara, N. Vita, D. Gully, J.P. Maffrand, J.P. Vincent, The 100-kDa neurotensin receptor is gp95/sortilin, a non-G-protein-coupled receptor, J. Biol. Chem. 273 (1998) 26273–26276.
[9] M.S. Nielsen, C. Jacobsen, G. Olivecrona, J. Gliemann, C.M. Petersen, Sortilin/neurotensin receptor-3 binds and mediates degradation of lipoprotein lipase, J. Biol. Chem. 274 (1999) 8832–8836.
[10] S. Lefrancois, J. Zeng, A.J. Hassan, M. Canuel, C.R. Morales, The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin, EMBO J. 22 (2003) 6430–6437.
[11] M. Canuel, A. Korkidakis, K. Konnyu, C.R. Morales, Sortilin mediates the lysosomal targeting of cathepsins D and H, Biochem. Biophys. Res. Commun. 373 (2008) 292–297.
[12] D. Calvo, J. Dopazo, M.A. Vega, The CD36, CLA-1 (CD36L1), and LIMPII (CD36L2) gene family: cellular distribution, chromosomal location, and genetic evolution, Genomics 25 (1995) 100–106.
[13] S.F. Berkovic, L.M. Dibbens, A. Oshlack, J.D. Silver, M. Katerelos, D.F. Vears, R. Lüllmann-Rauch, J. Blanz, K.W. Zhang, J. Stankovich, R.M. Kalnins, J.P. Dowling, E. Andermann, F. Andermann, E. Faldini, R. D'Hooge, L. Vadlamudi, R.A. Macdonell, B.L. Hodgson, M.A. Bayly, J. Savige, J.C. Mulley, G.K. Smyth, D.A. Power, P. Saftig, M. Bahlo, Array-based gene discovery with three unrelated subjects shows SCARB2/LIMP-2 deficiency causes myoclonus epilepsy and glomerulosclerosis, Am. J. Hum. Genet. 82 (2008) 673–684.
[14] A.C. Gamp, Y. Tanaka, R. Lüllmann-Rauch, D. Wittke, R. D'Hooge, P.P. De Deyn, T. Moser, H. Maier, D. Hartmann, K. Reiss, A.L. Illert, K. von Figura, P. Saftig, LIMP-2/LGP85 deficiency causes ureteric pelvic junction obstruction, deafness and peripheral neuropathy in mice, Hum. Mol. Genet. 12 (2003) 631–646.
[15] A.H. Erickson, E.I. Ginns, J.A. Barranger, Biosynthesis of the lysosomal enzyme glucocerebrosidase, J. Biol. Chem. 26 (1985) 319–324.
[16] C. Zachos, J. Blanz, P. Saftig, M. Schwake, A Critical Histidine Residue Within LIMP-2 Mediates pH Sensitive Binding to Its Ligand β-Glucocerebrosidase, Traffic 13 (2012) 1113–1123.
[17] J. Blanz, J. Groth, C. Zachos, C. Wehling, P. Saftig, M. Schwake, Disease-causing mutations within the lysosomal integral membrane protein type 2 (LIMP-2) reveal the nature of binding to its ligand beta-glucocerebrosidase, Hum. Mol. Genet. 19 (2010) 563–572.
[18] A. Balreira, P. Gaspar, D. Caiola, J. Chaves, I. Beirão, J.L. Lima, J.E. Azevedo, M.C. Miranda, A nonsense mutation in the LIMP-2 gene associated with progressive myoclonic epilepsy and nephrotic syndrome, Hum. Mol. Genet. 17 (2008) 2238–2243.
[19] A. Dardis, M. Filocamo, S. Grossi, G. Ciana, S. Franceschetti, S. Dominissini, G. Rubboli, M. Di Rocco, B. Bembi, Biochemical and molecular findings in a patient with myoclonic epilepsy due to a mistarget of the beta-glucosidase enzyme, Mol. Genet. Metab. 97 (2009) 309–311.
[20] F. Hopfner, B. Schormair, F. Knauf, A. Berthele, T.R. Tölle, R. Baron, C. Maier, R.D. Treede, A. Binder, C. Sommer, C. Maihöfner, W. Kunz, F. Zimprich, U. Heemann, A. Pfeufer, M. Näbauer, S. Kääb, B. Nowak, C. Gieger, P. Lichtner, C. Trenkwalder, K. Oexle, J. Winkelmann, Novel SCARB2 mutation in action myoclonus-renal failure syndrome and evaluation of SCARB2 mutations in isolated AMRF features, BMC Neurol. 27 (2011) 134.
[21] L.M. Dibbens, I. Karakis, M.A. Bayly, D.J. Costello, A.J. Cole, S.F. Berkovic, Mutation of SCARB2 in a patient with progressive myoclonus epilepsy and demyelinating peripheral neuropathy, Arch. Neurol. 68 (2011) 812–813.
[22] D.J. Costello, K.H. Chiappa, P. Siao, Progressive myoclonus epilepsy with demyelinating peripheral neuropathy and preserved intellect: a novel syndrome, Arch. Neurol. 66 (2009) 898–901.
[23] L.M. Dibbens, R. Michelucci, A. Gambardella, F. Andermann, G. Rubboli, M.A. Bayly, T. Joensuu, D.F. Vears, S. Franceschetti, L. Canafoglia, R. Wallace, A.G. Bassuk, D.A. Power, C.A. Tassinari, E. Andermann, A.E. Lehesjoki, S.F. Berkovic, SCARB2 mutations in progressive myoclonus epilepsy (PME) without renal failure, Ann. Neurol. 66 (2009) 532–536.
[24] G. Rubboli, S. Franceschetti, S.F. Berkovic, L. Canafoglia, A. Gambardella, L.M. Dibbens, P. Riguzzi, C. Campieri, A. Magaudda, C.A. Tassinari, R. Michelucci, Clinical and neurophysiologic features of progressive myoclonus epilepsy without renal failure caused by SCARB2 mutations, Epilepsia 52 (2011) 2356–2363.
[25] M. Knipper, C. Claussen, L. Rüttiger, U. Zimmermann, R. Lüllmann-Rauch, E.L. Eskelinen, J. Schröder, M. Schwake, P. Saftig, Deafness in LIMP2-deficient mice due to early loss of the potassium channel KCNQ1/KCNE1 in marginal cells of the stria vascularis, J. Physiol. 576 (2006) 73–86.
[26] D.S. Horoupian, R.T. Ross, Pigment variant of neuronal ceroid-lipofuscinosis (Kufs' disease), Can. J. Neurol. Sci. 4 (1977) 67–75.
[27] E. Andermann, F. Andermann, S. Carpenter, L.S. Wolfe, R. Nelson, G. Patry, J. Boileau, Action myoclonus-renal failure syndrome: a previously unrecognized neurological disorder unmasked by advances in nephrology, Adv. Neurol. 43 (1986) 87–103.
[28] A.J. Rothdach, T. Dietl, T. Kümpfel, M. Gottschalk, E.M. Schumann, C. Trenkwalder, Familial myoclonus-renal failure syndrome, Nervenarzt 72 (2001) 636–640.
[29] A. Badhwar, S.F. Berkovic, J.P. Dowling, M. Gonzales, S. Narayanan, A. Brodtmann, L. Berzen, J. Caviness, C. Trenkwalder, J. Winkelmann, J. Rivest, M. Lambert, O. Hernandez-Cossio, S. Carpenter, F. Andermann, E. Andermann, Action myoclonus-renal failure syndrome: characterization of a unique cerebro-renal disorder, Brain 127 (2004) 2173–2182.
[30] L. Vadlamudi, D.F. Vears, A. Hughes, E. Pedagogus, S.F. Berkovic, Action myoclonus-renal failure syndrome: a cause for worsening tremor in young adults, Neurology 67 (2006) 1310–1311.
[31] N. Eyal, N. Firon, S. Wilder, E.H. Kolodny, M. Horowitz, Three unique base pair changes in a family with Gaucher disease, Hum. Genet. 87 (1991) 328–332.
[32] I. Ron, M. Horowitz, Intracellular cholesterol modifies the ERAD of glucocerebrosidase in Gaucher disease patients, Mol. Genet. Metab. 93 (2008) 426–436.
[33] A. Velayati, J. DePaolo, N. Gupta, J.H. Choi, N. Moaven, W. Westbroek, O. Goker-Alpan, E. Goldin, B.K. Stubblefield, E. Kolodny, N. Tayebi, E. Sidransky, A mutation in SCARB2 is a modifier in Gaucher disease, Hum. Mutat. 32 (2011) 1232–1238.
[34] M.S. Nielsen, P. Madsen, E.I. Christensen, A. Nykjaer, J. Gliemann, D. Kasper, R. Pohlmann, C.M. Petersen, The sortilin cytoplasmic tail conveys Golgi-endosome transport and binds the VHS domain of the GGA2 sorting protein, EMBO J. 20 (2001) 2180–2190.
[35] J.E. Braun, D.L. Severson, Lipoprotein lipase release from cardiac myocytes is increased by decavanadate but not insulin, Am. J. Physiol. 262 (1992) E663–E670.
[36] P. Barbero, C. Rovère, I. De Bie, N. Seidah, A. Beaudet, P. Kitabgi, PC5-A-mediated processing of pro-neurotensin in early compartments of the regulated secretory pathway of PC5-transfected PC12 cells, J. Biol. Chem. 273 (1998) 25339–25346.
[37] E.G. Marcusson, B.F. Horazdovsky, J.L. Cereghino, E. Gharakhanian, S.D. Emr, The sorting receptor for yeast vacuolar carboxypeptidase Y is encoded by the VPS10 gene, Cell 77 (1994) 579–586.
[38] K.F. Johnson, S. Kornfeld, A His-Leu-Leu sequence near the carboxyl terminus of the cytoplasmic domain of the cation-dependent mannose 6-phosphate receptor is necessary for the lysosomal enzyme sorting function, J. Biol. Chem. 267 (1992) 17110–17115.
[39] J.F. Johnson, S. Kornfeld, The cytoplasmic tail of the mannose 6-phosphate/insulin-like growth factor-II receptor has two signals for lysosomal enzyme sorting in the Golgi, J. Cell Biol. 119 (1992) 249–257.
[40] N.T. Ktistakis, H.A. Brown, M.G. Waters, P.C. Sternweis, M.G. Roth, Evidence that phospholipase D mediates ADP ribosylation factor-dependent formation of Golgi coated vesicles, J. Cell Biol. 134 (1996) 295–306.
[41] T.E. Willnow, C.M. Petersen, A. Nykjaer, VPS10P-domain receptors — regulators of neuronal viability and function, Nat. Rev. Neurosci. 9 (2008) 899–909.
[42] C. Munck Petersen, M.S. Nielsen, C. Jacobsen, J. Tauris, L. Jacobsen, J. Gliemann, S.K. Moestrup, P. Madsen, Propeptide cleavage conditions sortilin/neurotensin receptor-3 for ligand binding, EMBO J. 18 (1999) 595–604.
[43] M.S. Nielsen, C. Gustafsen, P. Madsen, J.R. Nyengaard, G. Hermey, O. Bakke, M. Mari, P. Schu, R. Pohlmann, A. Dennes, C.M. Petersen, Sorting by the cytoplasmic domain of the amyloid precursor protein binding receptor SorLA, Mol. Cell. Biol. 27 (2007) 6842–6851.
[44] M. Canuel, Y. Libin, C.R. Morales, The interactomics of sortilin: an ancient lysosomal receptor evolving new functions, Histol. Histopathol. 24 (4) (2009 Apr) 481–492.
[45] S.C. Klinger, S. Glerup, M.K. Raarup, M.C. Mari, M. Nyegaard, G. Koster, T. Prabakaran, S.K. Nilsson, M.M. Kjaergaard, O. Bakke, A. Nykjær, G. Olivecrona, C.M. Petersen, M.S. Nielsen, SorLA regulates the activity of lipoprotein lipase by intracellular trafficking, J. Cell Sci. 124 (2011) 1095–1105.
[46] J. Tauris, L. Ellgaard, C. Jacobsen, M.S. Nielsen, P. Madsen, H.C. Thøgersen, J. Gliemann, C.M. Petersen, S.K. Moestrup, The carboxy-terminal domain of the

receptor-associated protein binds to the Vps10p domain of sortilin, FEBS Lett. 429 (1998) 27–30.
[47] R. Kornfeld, S. Kornfeld, Assembly of asparagine-linked oligosaccharides, Annu. Rev. Biochem. 54 (1985) 631–664.
[48] E.T. Baldwin, T.N. Bhat, S. Gulnik, M.V. Hosur, R.C. Sowder, R.E. Cachau, J. Collins, A.M. Silva, J.W. Erickson, Crystal structures of native and inhibited forms of human cathepsin D: implications for lysosomal targeting and drug design, Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 6796–6800.
[49] B. Rigat, W. Wang, A. Leung, D.J. Mahuran, Two mechanisms for the recapture of extracellular GM2 activator protein: evidence for a major secretory form of the protein, Biochemistry 36 (1997) 8325–8331.
[50] S.A. Igdoura, A. Rasky, C.R. Morales, Trafficking of sulfated glycoprotein-1 (prosaposin) to lysosomes or to the extracellular space in rat Sertoli cells, Cell Tissue Res. 283 (1996) 385–394.
[51] M.M. Gottesman, Transformation-dependent secretion of a low molecular weight protein by murine fibroblasts, Proc. Natl. Acad. Sci. U. S. A. 75 (1978) 2767–2771.
[52] L.M. Johnson, V.A. Bankaitis, S.D. Emr, Distinct sequence determinants direct intracellular sorting and modification of a yeast vacuolar protease, Cell 48 (1987) 875–885.
[53] L.A. Valls, C.P. Hunter, J.H. Rothman, T.H. Stevens, Protein sorting in yeast: the localization determinant of yeast vacuolar carboxypeptidase Y resides in the propeptide, Cell 48 (1987) 887–897.
[54] J. Zeng, J. Racicott, C.R. Morales, The inactivation of the sortilin gene leads to a partial disruption of prosaposin trafficking to the lysosomes, Exp. Cell Res. 315 (2009) 3112–3124.
[55] Y. Zhu, H. Bujo, H. Yamazaki, K. Ohwaki, M. Jiang, S. Hirayama, T. Kanaki, M. Shibasaki, K. Takahashi, W.J. Schneider, Y. Saito, LR11, an LDL receptor gene family member, is a novel regulator of smooth muscle cell migration, Circ. Res. 94 (2004) 752–758.
[56] J.B. Dubé, C.T. Johansen, R.A. Hegele, Sortilin: an unusual suspect in cholesterol metabolism: from GWAS identification to in vivo biochemical analyses, sortilin has been identified as a novel mediator of human lipoprotein metabolism, Bioessays 33 (2011) 430–437.
[57] A. Nykjaer, R. Lee, K.K. Teng, P. Jansen, P. Madsen, M.S. Nielsen, C. Jacobsen, M. Kliemannel, E. Schwarz, T.E. Willnow, B.L. Hempstead, C.M. Petersen, Sortilin is essential for proNGF-induced neuronal cell death, Nature 427 (2004) 843–848.
[58] C.B. Vaegter, P. Jansen, A.W. Fjorback, S. Glerup, S. Skeldal, M. Kjolby, M. Richner, B. Erdmann, J.R. Nyengaard, L. Tessarollo, G.R. Lewin, T.E. Willnow, M.V. Chao, A. Nykjaer, Sortilin associates with Trk receptors to enhance anterograde transport and neurotrophin signaling, Nat. Neurosci. 14 (2011) 54–61.
[59] F. Hu, T. Padukkavidana, C.B. Vægter, O.A. Brady, Y. Zheng, I.R. Mackenzie, H.H. Feldman, A. Nykjaer, S.M. Strittmatter, Sortilin-mediated endocytosis determines levels of the frontotemporal dementia protein, progranulin, Neuron 68 (2010) 654–667.
[60] J. Shi, K.V. Kandror, The luminal Vps10p domain of sortilin plays the predominant role in targeting to insulin-responsive Glut4-containing vesicles, J. Biol. Chem. 282 (2007) 9008–9016.
[61] S. Maeda, T. Nobukuni, K. Shimo-Onoda, K. Hayashi, K. Yone, S. Komiya, I. Inoue, Sortilin is upregulated during osteoblastic differentiation of mesenchymal stem cells and promotes extracellular matrix mineralization, J. Cell. Physiol. 193 (2002) 73–79.
[62] B. Kulterer, G. Friedl, A. Jandrositz, F. Sanchez-Cabo, A. Prokesch, C. Paar, M. Scheideler, R. Windhager, K.H. Preisegger, Z. Trajanoski, Gene expression profiling of human mesenchymal stem cells derived from bone marrow during expansion and osteoblast differentiation, BMC Genomics 8 (2007) 70.
[63] S.K. Nilsson, S. Christensen, M.K. Raarup, R.O. Ryan, M.S. Nielsen, G. Olivecrona, Endocytosis of apolipoprotein A-V by members of the low density lipoprotein receptor and the VPS10p domain receptor families, J. Biol. Chem. 283 (2008) 25920–25927.
[64] N.J. Samani, J. Erdmann, A.S. Hall, C. Hengstenberg, M. Mangino, B. Mayer, R.J. Dixon, T. Meitinger, P. Braund, H.E. Wichmann, J.H. Barrett, I.R. König, S.E. Stevens, S. Szymczak, D.A. Tregouet, M.M. Iles, F. Pahlke, H. Pollard, W. Lieb, F. Cambien, M. Fischer, W. Ouwehand, S. Blankenberg, A.J. Balmforth, A. Baessler, A.S.G. Ball, T.M. Strom, I. Braenne, C. Gieger, P. Deloukas, M.D. Tobin, A. Ziegler, J.R. Thompson, H. Schunkert, WTCCC and the Cardiogenics consortium, genomewide association analysis of coronary artery disease, N. Engl. J. Med. 357 (2007) 443–453.
[65] P. Linsel-Nitschke, N.J. Samani, H. Schunkert, Sorting out cholesterol and coronary artery disease, N. Engl. J. Med. 363 (2010) 2462–2463.
[66] K. Musunuru, A. Strong, M. Frank-Kamenetsky, N.E. Lee, T. Ahfeldt, K.V. Sachs, X. Li, H. Li, N. Kuperwasser, V.M. Ruda, J.P. Pirruccello, B. Muchmore, L. Prokunina-Olsson, J.L. Hall, E.E. Schadt, C.R. Morales, S. Lund-Katz, M.C. Phillips, J. Wong, W. Cantley, T. Racie, K.G. Ejebe, M. Orho-Melander, O. Melander, V. Kotelianskiy, K. Fitzgerald, R.M. Krauss, C.A. Cowan, S. Kathiresan, D.J. Rader, From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus, Nature 466 (2010) 714–719.
[67] M. Kjolby, O.M. Andersen, T. Breiderhoff, A.W. Fjorback, K.M. Pedersen, P. Madsen, P. Jansen, J. Heeren, T.E. Willnow, A. Nykjaer, Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export, Cell Metab. 12 (2010) 213–223.
[68] J. Shi, K.V. Kandror, Sortilin is essential and sufficient for the formation of Glut4 storage vesicles in 3T3-L1 adipocytes, Dev. Cell 9 (2005) 99–108.
[69] E.J. Mufson, J. Wuu, S.E. Counts, A. Nykjaer, Preservation of cortical sortilin protein levels in MCI and Alzheimer's disease, Neurosci. Lett. 471 (2010) 129–133.
[70] H.K. Teng, K.K. Teng, R. Lee, S. Wright, S. Tevar, R.D. Almeida, P. Kermani, R. Torkin, Z.Y. Chen, F.S. Lee, R.T. Kraemer, A. Nykjae, B.L. Hempstead, ProBDNF induces neuronal apoptosis via activation of a receptor complex of p75NTR and sortilin, J. Neurosci. 25 (2005) 5455–5463.
[71] A. Wähe, B. Kasmapour, C. Schmaderer, D. Liebl, K. Sandhoff, A. Nykjaer, G. Griffiths, M.G. Gutierrez, Golgi-to-phagosome transport of acid sphingomyelinase and prosaposin is mediated by sortilin, J. Cell Sci. 123 (2010) 2502–2511.
[72] S. Kathiresan, O. Melander, C. Guiducci, A. Surti, N.P. Burtt, M.J. Rieder, G.M. Cooper, C. Roos, B.F. Voight, A.S. Havulinna, B. Wahlstrand, T. Hedner, D. Corella, E.S. Tai, J.M. Ordovas, G. Berglund, E. Vartiainen, P. Jousilahti, B. Hedblad, M.R. Taskinen, C. Newton-Cheh, V. Salomaa, L. Peltonen, L. Groop, D.M. Altshuler, M. Orho-Melander, Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans, Nat. Genet. 40 (2008) 189–197.
[73] S. Kathiresan, C.J. Willer, G.M. Peloso, S. Demissie, K. Musunuru, E.E. Schadt, L. Kaplan, D. Bennett, Y. Li, T. Tanaka, B.F. Voight, L.L. Bonnycastle, A.U. Jackson, G. Crawford, A. Surti, C. Guiducci, N.P. Burtt, S. Parish, R. Clarke, D. Zelenika, K.A. Kubalanza, M.A. Morken, L.J. Scott, H.M. Stringham, P. Galan, A.J. Swift, J. Kuusisto, R.N. Bergman, J. Sundvall, M. Laakso, L. Ferrucci, P. Scheet, S. Sanna, M. Uda, Q. Yang, K.L. Lunetta, J. Dupuis, P.I. de Bakker, C.J. O'Donnell, J.C. Chambers, J.S. Kooner, S. Hercberg, P. Meneton, E.G. Lakatta, A. Scuteri, D. Schlessinger, J. Tuomilehto, F.S. Collins, L. Groop, D. Altshuler, R. Collins, G.M. Lathrop, O. Melander, V. Salomaa, L. Peltone, M. Orho-Melander, J.M. Ordovas, M. Boehnke, G.R. Abecasis, K.L. Mohlke, L.A. Cupples, Common variants at 30 loci contribute to polygenic dyslipidemia, Nat. Genet. 41 (2009) 56–65.
[74] A.R. Tall, D. Ai, Sorting out sortilin, Circ. Res. 108 (2011) 158–160.
[75] V.V. Belzil, C. André-Guimont, M.R. Atallah, H. Daoud, N. Dupré, J.P. Bouchard, W. Camu, P.A. Dion, G.A. Rouleau, Analysis of the SORT1 gene in familial amyotrophic lateral sclerosis, Neurobiol. Aging 33 (2012) 1845.e7-9.
[76] E. Kabashi, P.N. Valdmanis, P. Dion, D. Spiegelman, B.J. McConkey, C. Vande Velde, J.P. Bouchard, L. Lacomblez, K. Pochigaeva, F. Salachas, P.F. Pradat, W. Camu, V. Meininger, N. Dupre, G.A. Rouleau, TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis, Nat. Genet. 40 (2008) 572–574.
[77] J. Sreedharan, I.P. Blair, V.B. Tripathi, X. Hu, C. Vance, B. Rogelj, S. Ackerley, J.C. Durnall, K.L. Williams, E. Buratti, F. Baralle, J. de Belleroche, J.D. Mitchell, P.N. Leigh, A. Al-Chalabi, C.C. Miller, G. Nicholson, C.E. Shaw, TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis, Science 319 (2008) 1668–1672.
[78] T.J. Kwiatkowski, D.A. Bosco, A.L. Leclerc, E. Tamrazian, C.R. Vanderburg, C. Russ, A. Davis, J. Gilchrist, E.J. Kasarskis, T. Munsat, P. Valdmanis, G.A. Rouleau, B.A. Hosler, P. Cortelli, P.J. de Jong, Y. Yoshinaga, J.L. Haines, M.A. Pericak-Vance, J. Yan, N. Ticozzi, T. Siddique, D. McKenna-Yasek, P.C. Sapp, H.R. Horvitz, J.E. Landers, R.H. Brown, Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis, Science 323 (2009) 1205–1208.
[79] C. Vance, B. Rogelj, T. Hortobágyi, K.J. De Vos, A.L. Nishimura, J. Sreedharan, X. Hu, B. Smith, D. Ruddy, P. Wright, J. Ganesalingam, K.L. Williams, V. Tripathi, S. Al-Saraj, A. Al-Chalabi, P.N. Leigh, I.P. Blair, G. Nicholson, J. de Belleroche, J.M. Gallo, C.C. Miller, C.A. Shaw, Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6, Science 323 (2009) 1208–1211.
[80] O. Broustal, A. Camuzat, L. Guillot-Noel, N. Guy, S. Millecamps, D. Deffond, L. Lacomblez, V. Golfier, D. Hannequin, F. Salachas, W. Camu, M. Didic, B. Dubois, V. Meininger, I. Le Ber, A. Brice, French clinical and genetic research network on FTD/FTD-MND, FUS mutations in frontotemporal lobar degeneration with amyotrophic lateral sclerosis, J. Alzheimers Dis. 22 (2010) 765–769.
[81] E.D. Huey, R. Ferrari, J.H. Moreno, C. Jensen, C.M. Morris, F. Potocnik, R.N. Kalaria, M. Tierney, E.M. Wassermann, J. Hardy, J. Grafman, P. Momeni, FUS and TDP43 genetic variability in FTD and CBS, Neurobiol. Aging 33 (2012) 1016.e9-17.
[82] M. Polymenidou, C. Lagier-Tourenne, K.R. Hutt, S.C. Huelga, J. Moran, T.Y. Liang, S.C. Ling, E. Sun, E. Wancewicz, C. Mazur, H. Kordasiewicz, Y. Sedaghat, J.P. Donohue, L. Shiue, C.F. Bennett, G.W. Yeo, D.W. Cleveland, Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43, Nat. Neurosci. 14 (2011) 459–468.
[83] D.J. Mahuran, The GM2 activator protein, its roles as a co-factor in GM2 hydrolysis and as a general glycolipid transport protein, Biochim. Biophys. Acta 1393 (1998) 1–18.
[84] I. Hermans-Borgmeyer, G. Hermey, A. Nykjaer, C. Schaller, Expression of the 100-kDa neurotensin receptor sortilin during mouse embryonal development, Brain Res. Mol. Brain Res. 65 (1999) 216–219.
[85] J. Mazella, J.P. Vincent, Internalization and recycling properties of neurotensin receptors, Peptides 27 (2006) 2488–2492.
[86] B. Chen, B. Rigat, C. Curry, D.J. Mahuran, Structure of the GM2A gene: identification of an exon 2 nonsense mutation and a naturally occurring transcript with an in-frame deletion of exon 2, Am. J. Hum. Genet. 65 (1999) 77–87.
[87] R.L. Stevens, A.L. Fluharty, H. Kihara, M.M. Kaback, L.J. Shapiro, B. Marsh, K. Sandhoff, G. Fischer, Cerebroside sulfatase activator deficiency induced metachromatic leukodystrophy, Am. J. Hum. Genet. 33 (1981) 900–906.
[88] H. Christomanou, A. Aignesberger, R.P. Linke, Immunochemical characterization of two activator proteins stimulating enzymic sphingomyelin degradation in vitro. Absence of one of them in a human Gaucher disease variant, Biol. Chem. Hoppe Seyler 367 (1986) 879–890.
[89] D. Ramirez-Montealegre, P.G. Rothberg, D.A. Pearce, Another disorder finds its gene, Brain 129 (2006) 1353–1356.
[90] R. Dhami, M.A. Passini, E.H. Schuchman, Identification of novel biomarkers for Niemann–Pick disease using gene expression analysis of acid sphingomyelinase knockout mice, Mol. Ther. 13 (2006) 556–564.
[91] R.O. Brady, J.N. Kanfer, D. Shapiro, Metabolism of glucocerebrosides. II. Evidence of an enzymatic deficiency in Gaucher's disease, Biochem. Biophys. Res. Commun. 18 (1965) 221–225.

[92] M.L. Reitman, A. Varki, S. Kornfeld, Fibroblasts from patients with I-cell disease and pseudo-Hurler polydystrophy are deficient in uridine 5′-diphosphate-N-acetylglucosamine: glycoprotein N-acetylglucosaminylphosphotransferase activity, J. Clin. Invest. 67 (1981) 1574–1579.
[93] N.K. Honey, O.T. Mueller, L.E. Little, A.L. Miller, T.B. Shows, Mucolipidosis III is genetically heterogeneous, Proc. Natl. Acad. Sci. U. S. A. 79 (1982) 7420–7424.
[94] A. Waheed, S. Gottschalk, A. Hille, C. Krentler, R. Pohlmann, T. Braulke, H. Hauser, H. Geuze, K. von Figura, Human lysosomal acid phosphatase is transported as a transmembrane protein to lysosomes in transfected baby hamster kidney cells, EMBO J. 7 (1988) 2351–2358.
[95] S. Obermüller, C. Kiecke, K. von Figura, S. Höning, The tyrosine motifs of Lamp 1 and LAP determine their direct and indirect targetting to lysosomes, J. Cell Sci. 115 (2002) 185–194.
[96] C. Peters, M. Braun, B. Weber, M. Wendland, B. Schmidt, R. Pohlmann, A. Waheed, K. von Figura, Targeting of a lysosomal membrane protein: a tyrosine-containing endocytosis signal in the cytoplasmic tail of lysosomal acid phosphatase is necessary and sufficient for targeting to lysosomes, EMBO J. 9 (1990) 3497–3506.
[97] R. Pohlmann, C. Krentler, B. Schmidt, W. Schröder, G. Lorkowski, J. Culley, G. Mersmann, C. Geier, A. Waheed, S. Gottschalk, Human lysosomal acid phosphatase: cloning, expression and chromosomal assignment, EMBO J. 7 (1988) 2343–2350.

Sortilina e Risco de Doença Cardiovascular

Sortilin and the risk of Cardiovascular Disease

Maria Francisca Coutinho[1,2,3]; Mafalda Bourbon[4]; Maria João Prata[2,3]; Sandra Alves[1]

[1] Grupo de Investigação em Doenças Lisossomais de Sobrecarga, Unidade de I&D; Departamento de Genética Humana, INSA, Portugal

[2] IPATIMUP, Porto, Portugal

[3] Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal

[4] Grupo de Investigação Cardiovascular, Unidade de I&D; Departamento de Promoção da Saúde e Doenças Crónicas. INSA, Portugal

(Revista Portuguesa de Cardiologia, *in press*)

Pedido de separatas:

Maria Francisca Coutinho

Grupo de Investigação em Doenças Lisossomais de Sobrecarga, Unidade de I&D

Departamento de Genética Humana, INSA

Praça Pedro Nunes, 88, 4099-028 Porto; Portugal.

Tel.: +351 226070300; Fax: +351 226070399

E-mail: francisca_coutinho@yahoo.com; francisca.coutinho@insa.min-saude.pt

Palavras Chave

Genome wide association studies (GWAS); Doença das artérias coronárias (DAC); Colesterol LDL; Sortilina; Genómica funcional; Metabolismo das lipoproteínas.

Key Words

Genome wide association studies (GWAS); Coronary Artery Disease (CAD); LDL-cholesterol; Sortilin; Functional Genetics; Lipoprotein Metabolism.

Resumo

O nível plasmático de c-LDL constitui um determinante chave para o risco de doença cardiovascular, razão pela qual muitos estudos têm procurado elucidar as vias que regulam o seu metabolismo. As novas técnicas de sequenciação de última geração permitiram identificar um forte sinal de associação entre o *locus* 1p13 e o risco de doença cardiovascular causada por alteração dos níveis de LDL no plasma. Como seria de esperar para um fenótipo complexo, os efeitos da variação nesse *locus* são apenas moderados, mas ainda assim, o conhecimento da associação foi de grande importância uma vez conduziu à descoberta de uma nova via metabólica reguladora dos níveis de colesterol no plasma. Para tal, foram fundamentais os trabalhos efetuados por três equipas independentes, que ao procurarem esclarecer as bases biológicas da associação em causa, conseguiram provar que o gene *SORT1*, codificador da sortilina, era o gene do *locus* 1p13 implicado no metabolismo do LDL. *SORT1* foi o primeiro dos genes identificados como determinantes dos níveis plasmáticos de LDL a ser alvo de avaliação mecanística, e embora cada uma das equipas recorresse a metodologias experimentais diferentes, mas igualmente apropriadas face à questão em investigação, os resultados que obtiveram foram contraditórios em alguns aspetos. Neste trabalho, revemos o caminho percorrido até descoberta da nova via que relaciona a sortilina com os níveis plasmáticos de LDL e com o risco de enfarte do miocárdio. Ainda por esclarecer permanece o mecanismo regulador dessa ligação, mas a sua descoberta sugere novos alvos terapêuticos até há bem pouco tempo desconhecidos.

Abstract

Plasma LDL levels are a key determinant for the risk of cardiovascular disease. This is the major reason why many studies have attempted to elucidate the pathways that regulate its metabolism. The novel next-generation sequencing techniques allowed the identification of a strong association between the 1p13 locus and the risk of cardiovascular disease caused by changes in the levels of LDL in plasma. As expected for a complex phenotype, effects of variation in that locus are only moderate. Still, knowledge of that association was of major importance since it unveiled a new metabolic pathway regulating plasma cholesterol levels. Crucial for this discovery were the works performed by three independent teams who seek to clarify the biological basis of such association and managed to prove that *SORT1*, encoding sortilin, was the gene in the 1p13 locus involved in the metabolism of LDL. *SORT1* was the first gene identified as determinant of plasma LDL levels to be mechanistically evaluated and, although each of the three teams used different, even though valuable, experimental methodologies, their results were somehow contradictory. Here we review all experiments that led to the identification of the new pathway connecting sortilin with plasma LDL levels and risk of myocardial infarction. The regulatory mechanism underlying such association remains unclear but its discovery has paved the way for considering previously unsuspected therapeutic targets and/or approaches.

Abreviaturas

apo – apolipoproteína	apo – apolipoprotein
C/EBP - *CCAAT-enhancer binding protein*	
CT- colesterol total	TC – total cholesterol
DAC – doença das artérias coronárias	CAD – coronary artery disease
DALYs – *disability-adjusted life years*	
DCI – doença cardíaca isquémica	IHD – ischaemic heart disease
DCV – doença cerebrovascular	CVC – cerebrovascular disease
EAM – enfarte agudo do miocárdio	MI – myocardial infarction
GLUT4 – *glucose transport isoform 4*	
GM2AP – *GM2 activator protein*	
GWAS – *Genome-wide association studies*	
LDL – *Low-density lipoproteins*	
HDL – *High-density lipoproteins*	
LPL – lipoproteina lípase	LPL – lipoprotein lípase
RAP - *receptor associated protein*	
SAP – *sphingolipid activator protein*	
SNP – *single nucleotide polymorphism*	

Introdução

As doenças cardiovasculares constituem a maior causa de mortalidade nos países desenvolvidos [(1)], sendo responsáveis por 32% das mortes em Portugal, segundo dados do Instituto Nacional de Estatística [(2)]. A Doença Arterial Coronária (DAC), em particular, representa um enorme problema clínico, provocando uma em cada cerca de cinco mortes nos Estados Unidos [(3,4)]. Se para o desenvolvimento de DAC contribuem múltiplos fatores, está bem estabelecido que um dos determinantes chave de risco são os níveis de colesterol LDL no plasma. Segundo estimativa da OMS, cerca de 9 milhões de mortes/ano e mais de 75 milhões de anos de vida perdidos/ano podem ser atribuíveis a hipertensão ou Hipercolesterolemia [(5)]. Globalmente a hipercolesterolemia é responsável por 18% da doença cerebrovascular (DCV), maioritariamente eventos não fatais, e por 56% da doença cardíaca isquémica (DCI) [(5)]. Os dados disponíveis relativamente à Europa indicam que a hipercolesterolemia possa ser responsável por até 12 % dos anos de vida perdidos ajustados à incapacidade (DALYs – Disability-Adjusted Life Years) [(5)]. Face à ordem de grandeza destes números, muitas têm sido as tentativas no sentido de esclarecer as vias que regulam o metabolismo do colesterol associado às lipoproteínas de baixa densidade (c-LDL). Atualmente sabe-se que em subgrupos minoritários de indivíduos o colesterol elevado pode ser de origem genética simples, resultando em doenças mendelianas como a Hipercolesterolemia Familiar [(6)] A maioria dos doentes com esta patologia apresenta mutações patogénicas no gene que codifica o recetor

das LDL (*LDLR*), mas tem-se verificado que defeitos no gene que codifica a apolipoproteína B (*APOB)* ou, mais raramente, no que codifica a proproteína convertase subtilisina quexina tipo 9 (*PCSK9*) também estão associados ao mesmo fenótipo clínico [6, 7] As mutações encontradas em qualquer um destes genes provocam perda (no caso do *LDLR* e *APOB*) ou ganho de função (no caso do *PCSK9*) da proteína que codificam, alterando significativamente a sua função, sendo consideradas fator de elevado risco para o desenvolvimento de DCV.

Todavia, poucos são os casos em que se pode relacionar DCV com os efeitos drásticos de uma mutação num dado gene. Na patogénese da maioria das formas estão envolvidos fatores comportamentais, ambientais e genéticos, sabendo-se ainda que a componente genética é por si complexa podendo implicar a interação entre múltiplos determinantes genéticos [8]. Conhecem-se já uma série de polimorfismos nos genes referidos ou outros envolvidos no metabolismo lipídico que podem contribuir significativamente para o risco de DCV [revisto em 6], embora o efeito individual de cada variação seja pequeno.

Com o advento das novas tecnologias de sequenciação, ganhou-se novo fôlego o esforço para alcançar uma compreensão mais profunda dos fatores genéticos subjacentes ao desenvolvimento de doenças complexas. Assistiu-se assim a um aumento substancial dos estudos de epidemiologia molecular e a avanços importantes no entendimento da etiologia de muitas dessas doenças, o que em última análise é fundamental para desenhar novas estratégias de prevenção e tratamento [9].

Recentemente, através de estudos de associação envolvendo rastreios genómicos de grande escala (GWAS, *Genome Wide Association Studies*) foi possível identificar uma série nova de variantes de DNA que influenciam os níveis de LDL no plasma. Uma das associações mais consistentes foi inicialmente observada numa região do cromossoma 1p13, onde acabaria por se identificar uma variação genética comum fortemente associada ao c-LDL e enfarte de miocárdio. A importância clínica desta descoberta está patente na diferença de 40% no risco de enfarte do miocárdio verificada em indivíduos homozigóticos para os alelos maioritário (mais comum) e minoritário (menos comum) dessa variação. O efeito é comparável ao das variantes comuns dos já mencionados *LDLR* e *PCSK9* e maior do que descrito para as variantes mais comuns de *HMGCR* (gene que codifica a 3-hidroxi-3-methilglutaril-CoA reductase) [10], o gene que codifica o alvo terapêutico das estatinas, os fármacos mais usados no tratamento das hiperlipidemias. No *locus* 1p13, está localizado *SORT1,* o gene que codifica a sortilina, uma proteína multifuncional cuja relevância a nível biológico tem vindo a ser cada vez mais valorizada à medida que lhe vão sendo decifradas novas funcionalidades. Embora já fossem conhecidas as suas funções como recetor de vários ligandos, começaram a aparecer indícios claros de que a sortilina pode influenciar os níveis plasmáticos de c-LDL e, consequentemente, o risco de DAC. Se inicialmente, diferentes GWAS foram fundamentais para chegar a *SORT1* como provável gene de risco para a associação c-LDL/DAC, em 2010 surgiram quase em simultâneo 3 estudos desenvolvidos por equipas independentes [10-12] que procuraram explorar o mecanismo biológico que poderia relacionar a sortilina com os níveis

de c-LDL. Para tal, utilizaram abordagens mecanicistas diferentes tendo, curiosamente, chegado a conclusões também diferentes. Neste artigo, resumimos cada uma dessas abordagens e suas principais conclusões, avançando ainda com as razões que têm sido invocadas para conciliar os resultados aparentemente discrepantes.

Em busca de uma agulha num palheiro: os GWAS

Ao longo dos últimos anos, os avanços ao nível das técnicas da Biologia Molecular, em grande parte proporcionados pelo aparecimento e generalização do uso da terceira geração de sequenciadores, catapultaram os GWAS para a linha da frente dos estudos populacionais, com enfoque na identificação de moduladores genéticos de doenças, em especial de doenças complexas. Os GWAS baseiam-se na premissa de que em muitas dessas doenças as variantes hereditárias que concorrem para explicar o desenvolvimento do processo patogénico podem ser relativamente comuns, com frequência do alelo minoritário (*minor allele frequency*, MAF) superior a 5%, o que torna possível identificar associações entre doenças e variações genómicas através de grandes rastreios populacionais. Essas variações não são necessariamente codificantes, e de facto, através de GWAS têm sido identificadas associações com uma série de variantes genéticas localizadas em zonas não codificantes do genoma. A interpretação do significado destas associações depende, em grande parte, do conhecimento fino sobre as regiões em que se encontram as variantes. No fundo, questiona-se se uma variante para a qual se detetou um sinal positivo de associação não influenciará a expressão de um gene, ou se o sinal não será consequência da ocorrência de desequilíbrio de ligação com outras variações em genes localizados nessa mesma região genómica. Procura-se, então, identificar os genes com potencial para explicar a associação detetada, seguindo-se um outro desafio, que é compreender a base biológica dos sinais revelados pelos GWAS. Embora as expectativas nem sempre sejam alcançadas, este tipo de estudos tem desvendado importantes fatores genéticos que estão por detrás de uma série de doenças complexas. Um dos casos de maior sucesso, é ilustrado pela identificação de SNPs (SNP, *single nucleotide polimorphism*) relevantes em doenças cujos quadros patogénicos dependem dos níveis de lípidos e de lipoproteínas no plasma, como são exemplo a dislipidemia ou o enfarte agudo do miocárdio (EAM). Ao longo dos últimos anos, foram já descritos mais de 100 *loci* associados a variação dos níveis de triglicerídeos, c-LDL e colesterol HDL [13-20]. Relativamente a DAC ou EAM, os GWAS levaram à identificação de um menor número de *loci*, alguns dos quais igualmente associados a alterações nos fatores de risco tradicionais. Um balanço conjunto dos vários GWAS que resultaram na anotação de *loci* associados a DAC [14, 17, 21-23], entre os quais se incluíam os estudos do *"Welcome Trust Case Control Consortium"* e do *"German MI Family Study"*, saldou-se no registo de indícios de associação com 7 *loci* cromossómicos [14], nomeadamente localizados em 1p13.3 (compreendendo os genes *SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SORT1*, *PSMA5* e *SYPL2*), 1q41 (gene *MIA3*), 2q36.3 (região intergénica), 6q25.1 (gene *MTHFD1L*), 9p21.3 (genes *CDKN2A* e *CDKN2B*), 10q11.21 (região intergénica) e 15q22.33

(gene *SMAD3*). A questão imediata que se colocou foi a de saber se estes novos *loci* afetavam os fatores de risco cardiovascular já conhecidos. Para a esclarecer, Samani e colaboradores [15] investigaram a relação destes sete *loci* com uma série de parâmetros mensuráveis de reconhecida relevância em termos de doença cardiovascular, tendo demonstrado que apenas o *locus* de risco de DAC no cromossoma 1p13 estava significativamente associado a níveis elevados de c-LDL, sendo na região intergénica localizada entre os genes *PSRC1* e *CELSR2* que detetaram o sinal de associação mais forte. Os genes *PSRC1* e *CELSR2* codificam respetivamente as proteínas *proline/serine-rich coiled-coil 1* e caderina EGF, cuja função permanece desconhecida, e estão muito próximos do gene codificador da sortilina, *SORT1*. Nenhum destes 3 genes ou dos outros contidos no locus 1p13 tinha sido conotado com alguma das doenças mendelianas conhecidas que afetam os níveis de colesterol LDL [10, 13, 20].

Os suspeitos do costume e um suspeito inesperado: LDL e sortilina

Face à grande robustez estatística da associação detetada entre o *locus* 1p13 e os níveis plasmáticos de c-LDL, a busca do mecanismo biológico capaz de explicar a observação tornou-se o foco de pesquisa de diversas equipas.

Em primeiro lugar, importava conseguir uma melhor aproximação à causa genética da associação, impossível de obter através dos GWAS devido aos efeitos de *linkage disequilibrium* (relação não aleatória) entre vários SNPs do *locus* 1p13.

Para alcançar o fim em vista, seria necessário proceder à re-sequenciação fina da região contendo os genes *SORT1, PSRC1, CELSR2,* e efetuar diversos estudos *in silico*, *in vitro* e *in vivo*. À medida que os resultados foram surgindo, um gene começou a destacar-se de entre os candidatos que compõem este *locus* de risco de DAC: *SORT1*, o gene que codifica a sortilina [10-12].

A sortilina pertence à família de recetores de domínio Vps10p, constituída por cinco membros atualmente conhecidos. É sintetizada sob a forma de uma proproteína e clivada no Golgi por convertases de proproteínas, processo após o qual assume a forma matura que lhe permite a correta ligação aos ligandos. Em termos funcionais, é um recetor de múltiplos ligandos, incluindo a lipoproteína lípase (LPL) [24], as apolipoproteínas A-V (apo A-V) [25], a neurotensina [26] e a proteína RAP (*receptor-associated protein*) [27]. É ainda responsável por mediar o transporte do Golgi para o lisossoma de uma série de proteínas, enzimáticas ou não: proteínas ativadoras dos esfingolípidos (SAPs, *sphingolipid activator proteins*: prosaposina e, GM2AP, GM2 *activator protein*), esfingomielinase ácida, catepsina D e catepsina H [24, 28-30]. Demonstrou-se ainda que a sortilina está envolvida numa série de processos de grande relevância biológica de que é exemplo a formação de vesículas de armazenamento de GLUT-4 (*glucose transport isoform 4*) na resposta à insulina durante a diferenciação dos adipócitos [31]. No cérebro, faz parte de um complexo de sinalização que regula a sobrevivência da célula [32].

A relevância *in vivo* destas propriedades multifacetadas ainda permanece por esclarecer, mas cada vez mais se tem reforçado a noção de que se trata de uma proteína com importantes papéis a nível biológico e cuja desregulação parece passível de causar efeitos indesejáveis, podendo ir muito além do possível transtorno dos níveis de colesterol LDL no plasma.

Uma abordagem exemplar: análise mecanística

Em 2010, três equipas independentes publicaram resultados de trabalhos pioneiros na tentativa de esclarecer o mecanismo biológico que poderia explicar a associação do *locus* 1p13 aos níveis de colesterol LDL no plasma [10-12]. Recorrendo a sólidas abordagens experimentais mas distintas entre si, os 3 estudos apontaram para o gene *SORT1* como sendo o responsável pelo aumento do risco de DAC e/ou EAM. Pese embora esta unanimidade, curioso é o facto de os 3 estudos terem chegado não apenas a conclusões diferentes mas até opostas em alguns aspetos quanto ao papel da sortilina na secreção de colesterol das lipoproteínas de muito baixa densidade (VLDL) [revisto em 33, 34].

As primeiras evidências *in vitro* da interação entre a sortilina e partículas de c-LDL foram apresentadas por Linsel-Nitschke e colaboradores. Através do mapeamento fino do *locus* 1p13, os autores começaram por procurar a variação com sinais mais fortes de associação, chegando ao SNP rs599839 e verificando que o alelo G era o que se associava a redução dos níveis plasmáticos de LDL e diminuição do risco de doença cardiovascular. Demostraram depois que os indivíduos homozigóticos para o mesmo alelo apresentavam expressão aumentada dos genes *SORT1*, *CELSR2* e *PSRC1* em glóbulos brancos do sangue periférico, sendo, no entanto, nos níveis de mRNA *SORT1* que observaram o efeito mais acentuado. Estes resultados foram confirmados em células embrionárias de fígado humano (HEK293) induzidas a sobre-expressar *SORT1*, em que verificaram um aumento da internalização de partículas c-LDL, com consequente diminuição dos respetivos níveis plasmáticos [11].

No mesmo ano, Musunuru *et al*. apostaram numa abordagem multifacetada, vista como um autêntico "*tour de force*" exemplar para dar sentido aos resultados obtidos com os GWAS [35]. Baseando-se no conhecimento prévio de que os SNPs rs646776, rs599839, rs12740374, e rs629301 do locus 1p13 eram os mais altamente associados ao c-LDL no plasma, e partindo do princípio que variantes de DNA não codificantes, como o caso das 4 em que se centraram, podem alterar a expressão génica, Musunuru e seus colaboradores começaram por analisar os seus efeitos nos níveis de mRNA dos seis genes localizados no *locus* 1p13: *SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SORT1* e *SYP2*. Verificaram, então, que no fígado humano o alelo minoritário de rs646776 estava associado a um aumento da expressão de 3 genes, *SORT1*, *CELSR2* e *PSRC1* [36], tendo verificado maior alteração de expressão ao nível de *SORT1,* traduzida no aumento da quantidade do seu produto proteico, a sortilina. O mapeamento fino da região genómica em estudo, permitiu-lhes identificar os haplótipos definidos pelos SNPs presentes numa região de 6.1 kilobases localizada entre os genes *CELSR2* e *PSRC1*, e chegar ao SNP rs12740374 como sendo afinal o responsável pela associação evidenciada pelos GWAS. Os indícios começaram

por surgir na sequência da análise bioinformática, indicando que ao alterar a sequência G**G**TGCTCAAT para G**T**TGCTCAAT, o alelo minoritário deste SNP criava um novo local de ligação de fator de transcrição, especificamente para a proteína C/EBP (CCAAT-enhancer binding protein) α, capaz de aumentar a atividade do promotor e o nível de expressão de *SORT1*. O efeito previsto *in silico* foi depois confirmado *in vitro*. Refira-se que este resultado era consistente com o reportado pelo grupo de Linsel-Nitschke *et al*. [11] relativamente aos níveis de mRNA expressos no fígado. Finalmente, Musunuru *et al*. [10] efetuaram estudos em células de fígado de ratinhos mutantes, em que o gene que codifica a sortilina tinha sido inativado ou sobrexpresso, demonstrando que o nível de expressão da sortilina era capaz de modular a secreção hepática das VLDL. O ratinho transgénico escolhido por esta equipa, com o gene *Apobec1* (que efetua a correção das transições C>U nos mRNAs da apo B, entre outros) suprimido, *Apobec1*$^{-/-}$, era um ratinho "humanizado", apresentando um perfil lipídico diferente do que é normal naquela espécie e mais próximo do que se observa em humanos, em que o c-LDL é o transportador de colesterol predominante na circulação. Quando Musunuru *et al*. [10] sobre-expressaram o gene *Sort1* nas células hepáticas desses ratinhos, observaram uma redução 70% nas concentrações de colesterol total (CT) e de colesterol LDL. Pelo contrário, quando procederam à inativação do gene *Sort1* por siRNA, registaram um aumento de 46% dos níveis de CT e de 125% dos níveis de LDL.

De um modo geral, os dados apresentados por estas duas equipas vinham reforçar, e também dar sentido, às conclusões extraídas dos GWAS, na medida em que ao demostrarem a existência de uma correlação negativa entre os níveis de mRNA *SORT1* e a concentração de c-LDL no plasma, sustentavam um mecanismo biológico explicativo de como variações genéticas no locus 1p13 podiam influenciar os níveis plasmáticos de LDL e, portanto, alterar o risco de DAC/EAM.

No entanto, ainda no mesmo ano, foi publicado um terceiro estudo mecanístico abordando a mesma questão, cujos resultados nem se conciliavam tão facilmente com os dois anteriores nem com as evidências inferidas através dos GWAS.

Referimo-nos ao trabalho de Kjolby e sua equipa [12], que utilizaram como modelo um ratinho *knockout* duplo, *Sort1*$^{-/-}$,*Ldlr*$^{-/-}$, tendo observado que os seus hepatócitos apresentavam reduções de 30% nos níveis de CT, de ~50% nos níveis de proteínas contendo apo B100 (VLDL e LDL) e de ~60% na área das placas arterosclerótica, quando comparados com o ratinho *knockout* simples, *Ldlr*$^{-/-}$. De seguida, induziram sobre-expressão do gene *Sort1* ao nível do fígado dos ratinhos, verificando, resumidamente, que enquanto a deficiência de sortilina levava a uma redução de 50% na secreção de lipoproteínas, a sua sobre-expressão se traduzia num aumento de 50% na secreção das mesmas proteínas. No conjunto, os resultados apontavam para uma correlação positiva entre a expressão *Sort1* e a concentração de c-LDL, ou seja, de sentido oposto ao observado por Musunuru *et al* [10].

Descubra as diferenças: análise dos resultados

A discrepância entre os resultados destes três estudos tem vindo a ser discutida entre a comunidade científica, como ilustram as publicações surgidas em 2011 de Dubé ou de Tall e Ai [33, 34]. Nos seus comentários, estes autores chamam a atenção para as diferenças experimentais entre os trabalhos de Linsel-Nitschke *et al.* [11], Kjolby *et al.* [12] e Musunuru *et al.* [10], e para como tais diferenças podem ter condicionado algum desacordo nos resultados obtidos. Tendo em vista esclarecer o mecanismo através do qual o *locus* 1p13 afetava os níveis de LDL e risco de DAC, as 3 equipas optaram por modelos diferentes, e embora cada uma delas tivesse recorrido a abordagens experimentais bem desenhadas e delas extraído ilações apropriadas, as diferentes conclusões podem não ser estritamente comparáveis entre si.

Primeiro, devido à diferença de *background* metabólico dos modelos animais usados nas experiências efetuadas. Linsel-Nitschke *et al.* [11] conduziram as suas investigações apenas em humanos, ao contrário das equipas de Musunuru [10] e Kjolby [12] que apostaram na análise de modelos animais não-humanos, especificamente em modelos de ratinho mas, mesmo assim, os ratinhos que escolheram tinham diferentes perfis metabólicos: Musunuru *et al.* [10] trabalharam com células hepáticas de um ratinho "humanizado", *Apobec*$^{-/-}$; Kjolby *et al.* estudaram um *knockout* duplo para a sortilina e para o recetor das LDL (*Sort1*$^{-/-}$,*Ldlr*$^{-/-}$). Enquanto o modelo de Musunuru [10] sobre-produzia e secretava quantidades anormalmente elevadas de lipoproteínas de modo a mimetizar o perfil lipídico humano, o que pode ter alterado artificialmente as vias secretórias e a disponibilidade de sortilina, o de Kjolby [12] tinha um catabolismo deficiente em lipoproteínas, criado pela repressão da expressão no interior dos hepatócitos e acentuado pela dieta "*high fat western*". Outro ponto a ter em linha de conta, são as diferenças na regulação génica entre as duas espécies, homem e ratinho, como ficou demonstrado, por exemplo, pela inexistência do local C/EBP α no ratinho [10; 37]. No que se refere à sortilina, isto significa que as conclusões retiradas de observações no ratinho não podem ser linearmente extrapoladas para o homem, nem vice-versa [34].

Em conjunto, as observações *in vivo* dos estudos em ratinhos parecem demonstrar que a sortilina pode assumir funções hepáticas complementares dependentes do meio metabólico, que em última análise regulam a secreção das VLDL. É de admitir que a sortilina possa regular a secreção e o tráfego das VLDL para o lisossoma quando os níveis intracelulares de apo B-100 são extremamente elevados. Pelo contrário, em condições de menor expressão da apo B-100, a sortilina poderá regular a formação e secreção das VLDL [33, 34].

Porém, esta possível função da sortilina na formação e secreção das VLDL não se harmoniza facilmente com os resultados obtidos com os GWAS que demonstravam uma associação específica com o c-LDL e não com os triglicerídeos, que são o componente principal das partículas das VLDL [34]. Fica assim em aberto a possibilidade de a sortilina ter um papel noutras vias de regulação dos níveis lipídicos no plasma que não impliquem as VLDL.

Se os 3 estudos em apreço forneceram pistas que em parte são contraditórias, apresentaram também provas da existência de uma nova via regulatória para o metabolismo das lipoproteínas e da possibilidade da modulação desta via alterar o risco de DCV em humanos, não obstante haver ainda muito a percorrer para obter uma visão clara de como se processa e dos fatores que a modulam.

Conclusão

Como os GWAS envolvem a genotipagem de variantes comuns em termos populacionais, têm a limitação de só terem capacidade para identificar alelos cujo efeito num fenótipo associado a um quadro patogénico é mínimo, senão negligenciável [(38)]. Ou seja, permitem identificar alelos relativamente comuns mas que explicam apenas uma pequena porção da variação genética associada a um fenótipo [(39)]. Assim, mesmo depois de todos os esforços de investimento em GWAS, uma parte significativa da hereditariedade das doenças complexas, como é o caso das doenças cardiovasculares e das DAC em particular, permanece desconhecida. Esta porção da hereditariedade foi denominada "a hereditariedade perdida" ou "a matéria negra da hereditariedade" [(35)]. Os grandes defensores dos GWAS argumentam que aumentando o tamanho das amostras e reforçando substancialmente a densidade dos SNPs a estudar, será possível ir desvendando a porção da hereditariedade que se mantém desconhecida [(35)]. Porém, há quem defenda uma estratégia diferente que consiste na sequenciação direta de todo o genoma tendo em vista a identificação de alelos raros que exerçam grandes efeitos no fenótipo [(40)]. As portas para o recurso generalizado a esta estratégia começam agora a abrir-se com o aparecimento e aperfeiçoamento de novas plataformas de sequenciação de terceira geração que permitem a sequenciação do genoma completo a um custo cada vez mais acessível. Estima-se que o genoma de cada indivíduo contenha aproximadamente 10000 variantes não sinónimas entre cerca de 3.5 milhões de SNPs. Tendo em mente a ordem de grandeza destes valores é compreensível que este tipo de sequenciação venha a dominar os estudos genéticos nos próximos anos, tendência que já hoje se verifica [(41-45)].

Mas a descoberta de novas variantes genéticas associadas a doenças complexas pode não ser suficiente para descodificar a "matéria negra da hereditariedade", dado que esta pode ser o produto de interações complexas entre fatores genéticos, genómicos e epigenéticos [(39, 46)]. E a um outro nível, também o fenótipo pode ser consequência de interações não lineares entre diversos fatores de natureza genética e ambiental.

Apesar disso, é muito importante ir enriquecendo o catálogo das variantes genéticas associadas a um determinado fenótipo complexo, até porque a descoberta de novas variantes pode abrir caminho para chegar a via moleculares até então insuspeitas de influenciar a doença, como aconteceu com o exemplo que temos vindo a referir.

Os três *reports* aqui revistos representam abordagens exemplares e paradigmáticas de como se pode partir de uma associação estatística "cega" fornecida pelos GWAS, até encontrar uma

explicação mecanística, mais ou menos esclarecida, do modo como uma variação genética pode modular um certo fenótipo. Neste caso, os GWAS orientaram para um ponto de partida particular, o *locus* 1p13 [14, 21-23], que acabaria por despertar a atenção de três equipas independentes, levando-as a enveredar por diferentes abordagens experimentais para compreender em que se fundamentava a associação estatística, mas todas apostando no gene *SORT1* como o modulador chave dos níveis de c-LDL e risco de EAM. Se os resultados foram concordantes na indicação de que *SORT1* tinha um papel importante na regulação do metabolismo lipoproteíco, já foram menos consonantes as interpretações sobre a direção do efeito da expressão do gene nos níveis plasmáticos do c-LDL bem como sobre o mecanismo subjacente [10-12].

Face às suas observações, Linsel-Nitschke [11] e colaboradores propuseram que a sobre-expressão da sortilina aumentava a internalização de LDL, com consequente diminuição dos respectivos níveis plasmáticos [11]. Pouco depois, Musunuru *et al*. [10] reportaram uma relação inversa entre a expressão da sortilina e a concentração de LDL circulante, avançando com um mecanismo explicativo que passava pela regulação transcricional, específica ao nível do fígado, do gene *SORT1* por fatores de transcrição C/EBP α, em que a sobre-expressão da *SORT1* reduziria a produção/secreção de VLDL [10]. Mas pelo contrário, Kjolby e a sua equipa [12] observaram uma relação direta entre a expressão da *Sort1* e a concentração de LDL circulante, sugerindo que tal poderia resultar de secreção aumentada das VLDL [11].

Entre as várias explicações que foram já apresentadas para justificar a discrepância entre resultados, a resposta parece em parte residir na própria sortilina, que emerge cada vez mais como uma proteína multifacetada podendo assumir diferentes funções consoante as circunstâncias.

Em síntese, os estudos aqui revistos proporcionaram uma série de provas consistentes de que *SORT1* é um elemento regulador dos níveis de c-LDL no plasma, o que veio acrescentar ao gene que codifica a sortilina uma importante função de que não se desconfiava até há bem pouco tempo. De momento, é ainda controversa a via celular que relaciona a sortilina com o metabolismo lipídico, mas certamente que nos próximos anos a questão vai continuar a ser explorada, sendo fundamental para avaliar se a sortilina pode representar um alvo potencial para abordagens terapêuticas direcionadas para hipercolesterolemia ou DCV [revisto em 33-35].

Introduction

Cardiovascular disease is the leading cause of death in developed countries [1], being responsible for 32% of deaths registered in Portugal, according to our National Institute of Statistics [2]. Coronary artery disease (CAD), in particular, represents a major clinical problem, accounting for one in every five deaths in the U.S. [3, 4]. For the development of CAD multiple factors contribute but it is well established that one of its key determinants are the plasma LDL levels. According to the WHO estimate, about 9 million deaths/year and more than 75 million years of life lost/year are due to hypertension or hypercholesterolemia [5]. Overall, hypercholesterolemia is responsible for 18% of registered events of cerebrovascular disease (CVD), mostly non-fatal events, and 56% of ischaemic heart disease (IHD) [5]. The data available for Europe suggest that hypercholesterolemia may be responsible for up to 12% of disability-adjusted life years (DALYs) [5]. Given the magnitude of these numbers, many attempts have been made to elucidate the pathways that regulate LDL metabolism. Presently, it is known that, for minor groups of individuals, high cholesterol levels may be of genetic origin. There is even a Mendelian disease associated to it: familial hypercholesterolemia [6]. Most patients suffering from this pathology present pathogenic mutations in the gene that codes for the LDL receptor (*LDLR*), but it has been reported that also defects in the the apolipoprotein B gene (*APOB*), or less commonly, in the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene, may also be associated to that clinical phenotype [6, 7]. Mutations in any of these genes lead to either loss (*LDLR* and *APOB*) or gain (*PCSK9*) of function of its associated protein and high cardiovascular risk.

However, there are few cases in which it is possible to relate the presence of a specific gene mutation with CVD. The pathogenesis of the major forms of CVD involves behavioral, environmental and genetic factors and it is known that the genetic component is very complex resulting from the interaction of multiple genetic determinants [8]. There are, however, several polymorphisms in these or other genes involved in lipid metabolism that, even though presenting a smaller effect in the protein which they code, may play a significant part in CVD risk [reviewed in 6].

With the advent of new sequencing technologies, the search for a deeper understanding of these mechanisms, as well as the genetic basis of other risk factors had a new incentive. It became possible to screen large populations for the genetic basis for complex diseases. Ultimately, such epidemiological studies may lead to better and deeper understanding of etiological pathways and contribute to the development of new strategies for prevention and treatment [9].

Recently, through association studies involving large-scale genome screenings (genome-wide association studies, GWAS) it was possible to identify a novel set of DNA variants that influence plasma LDL-c levels. The most consistent of these associations was observed in a cluster of genes on chromosome 1p13. The clinical relevance of this novel pathway is defined by the 40% difference in risk of myocardial infarction (MI) verified between individuals

homozygous for the minor (less common) and major (more common) alleles of the p13 locus on chromosome 1. The effect is comparable to the one attributed to common variants of *LDLR* and *PCSK9* and greater than that described for the most common variants in *HMGCR* (gene that codes for 3-hydroxy-3-methylglutaryl-CoA reductase) [10] the gene that encodes the therapeutic target of statins, which is the class of drugs more commonly used in the treatment of hyperlipidemias. In the 1p13 cluster is located the *SORT1* gene. This gene encodes for sortilin, a multifunctional protein whose biological relevance is becoming greater as novel and unexpected functions are being unveiled for it. Even though its functions as a receptor for various ligands were already known, given the clear association reported by different GWAS, three independent teams [10-12] tried to understand the biological mechanism relating sortilin to LDL cholesterol (LDL-c) levels and, ultimately, to risk of CAD. To this end, they used different mechanistic approaches and, interestingly, they also came to different conclusions. Here we summarize each of those approaches and their main conclusions, trying to reconcile the apparently discrepant results.

Looking for a needle in a haystack: GWAS

Over the past few years, advances in sequencing and genotyping, with the emergence and spread of third-generation sequencers, catapulted GWAS to the forefront of population studies, with special focus on the relationship between genotype and common disease. These studies are based on the premise that, for a large number of such diseases, the underlying hereditary variations have a minor allele frequency (MAF) higher than 5%. It is, then, possible through the analysis of large populations samples, to identify associations of certain diseases to certain regions of the genome, both coding and non-coding. In fact, several genetic variants identified through GWAS are located in non-coding regions of the genome. The interpretation of the effects of the identified variants depends largely on the knowledge available for those regions. It is then estimated the potential of genes located in the target region being responsible for the detected association, without excluding the possibility that it may result from long-range genetic interactions or from other unknown reasons. The challenge is to understand the biological basis of the signs revealed in GWAS. And, although this is may be a difficult task, GWAS have already uncovered important genetic factors underlying a number of complex diseases. One of the most successful cases in terms of identification of SNPs (single nucleotide polymorphisms) which are relevant to the pathogenesis of a complex disease is precisely the annotation of genes correlated with plasma lipid and lipoprotein levels, factors which are long known to have great importance in pathological conditions such as dyslipidemia and/or myocardial infarction (MI). Over the last few years more than 100 *loci* have been described as associated with the genetic variation of triglycerides, LDL cholesterol and HDL cholesterol levels [13-20]. In cases of CAD or MI, GWAS identified a smaller number of genetic loci, some of which were also associated with changes in traditional risk factors. A comprehensive analysis of several GWAS identified and annotated

CAD-associated *loci* [14, 17, 21-23], by combining the data revealed in the studies of "Welcome Trust Case Control Consortium" and "German MI Family Study". It presented evidences of associations between 7 chromosomal *loci* and CAD risk [14]: 1p13 (*SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SORT1*, *PSMA5* and *SYPL2* genes), 1q41 (*MIA3* gene), 2q36 (intergenic region), 6q25.1 (*MTHFD1L* gene), 9p21 (*CDKN2A* and *CDKN2B* genes), 10q11 (intergenic region) and 15q22.33 (*SMAD3* gene). The immediate question that arose was whether these new loci affected already known cardiovascular risk factors. To clarify it, Samani *et al*. [15] investigated the association of these 7 *loci* with a number of quantitative traits of known relevance to cardiovascular disease, showing that only the risk locus on chromosome 1p13 was significantly associated with higher levels of LDL cholesterol. The strongest association was located in the intergenic region comprehending the *PSRC1* and *CELSR2* genes. *PSRC1* and *CELSR2* genes code for the protein proline/serine-rich coiled-coil 1 and for cadherin EGF, respectively. The function of these proteins remains unknown and its coding genes are very closely located to the one that codes for sortilin, *SORT1*. None of these 3 genes nor any of the others present in the 1p13 locus have ever been associated to a known Mendelian disease affecting LDL-c levels [10, 13, 20].

Usual and unusual suspects: LDL-cholesterol and sortilin

Given such a statistically significant association between the 1p13 locus and plasma LDL-c levels, the search for its explanatory mechanism became the research focus of various teams. Firstly, it was important to clarify which was the particular genomic variant that was causing this association. Due to the *linkage desequilibrium* (extensive and non-random relationship) between multiple SNPs at the 1p13 locus (comprising the genes *SORT1*, *PSRC1* and *CELSR2*), it was impossible to identify solely through GWAS, the causal variant. To clarify this point, as well as the mechanisms behind this association *in silico*, *in vitro* and *in vivo* studies would be required. And, as those studies have been carried out, a gene began to detach itself from all the others comprising this CAD risk *locus*: the *SORT1* gene that codes for sortilin [10-12].

Sortilin belongs to the Vps10p domain receptor family, which consists of five known members. It is synthesized as a propeptide, cleaved in the Golgi by proprotein convertases, after which the process takes the mature form, which allows the proper ligand binding. Functionally, sortilin is a receptor of multiple ligands, including lipoprotein lipase (LPL; 24), the A-V apolipoproteins (apo A-V) [25], neurotensin [26] and receptor associated protein (RAP; 27). It is also responsible for mediating Golgi-to-lysosome transport of a number of lysosomal proteins, enzymatic or not: sphingolipid activator proteins (SAPs: prosaposin and GM2AP), acid sphingomyelinase, cathepsin H and cathepsin D [24, 28-30]. In recent years, it has been demonstrated that sortilin is involved in a number of highly relevant biological processes such as the formation of GLUT-4 (glucose transport isoform 4) storage vesicles in response to insulin during adipocyte differentiation [31]. In the brain, is part of a signaling complex that regulates cell survival [32].

The *in vivo* relevance of these multifaceted properties still remains unclear but reinforces the notion that sortilin is a protein with an important biological role, and whose deregulation seems likely to cause severe side effects that may go beyond its effect on the levels of plasma LDL-c levels.

An exemplary approach: mechanistic analysis

In 2010, three independent teams published results from pioneer studies, trying to clarify the biological mechanism underlying the association between the 1p13 locus and plasma LDL levels [10-12]. Based on solid, even though different, experimental approaches, all three studies indicate the *SORT1* gene as the one responsible for the increased risk of CAD and/or MI. Curiously, different studies have reached conclusions that were not only different but, in some cases, even opposite about the role of sortilin on the secretion of very low density cholesterol (VLDL-c) [reviewed in 33, 34].

The first *in vitro* evidences of the interaction between sortilin and LDL particles were given by Linsel-Nitschke *et al.* [11] Through a fine mapping of the 1p13 locus, the authors started by searching the variant with stronger signs for that association and reached to the rs599839 SNP, verifying that the G allele was the one associated to reduced LDL-c plasma levels and lower cardiovascular disease risk. The authors demonstrated that homozygous individuals for the G allele showed increased expression of the *SORT1*, *CELSR2* and *PSRC1* genes in peripheral white blood cells. The strongest and most consistent association, however, was seen for *SORT1* mRNA levels. These results were confirmed in human embryonic kidney cells (HEK293) over-expressing *SORT1* that showed increased internalization of LDL-c particles, with consequent reduction of its plasma levels [11].

In the same year, Musunuru *et al.* (2010) presented a multifaceted approach, an exemplar *"tour de force"* for the follow-up of GWAS [35]. Based on the previous recognition that the rs646776, rs599839, rs12740374 and rs629301 SNPs from the 1p13 locus were the most strongly associated to LDL-c plasma levels and on the assumption that non-coding DNA variants may alter gene expression, Musunuru *et al.* [10] started by analysing the effects if these 4 variants on the mRNA levels of the 6 genes located in that locus: *SARS*, *CELSR2*, *PSRC1*, *MYBPHL*, *SYP2* and *SORT1*. The authors verified that in human liver the minor allele for the rs646776 SNP was associated to increased expression of the *SORT1, CELSR2* and *PSRC1* genes [36], with the strongest association observed for *SORT1* mRNA levels and its corresponding protein: sortilin. The fine mapping of the interest region led to the definition of the haplotypes defined by the SNPs present in 6.1 kilobases located between the *CELSR2* and the *PSRC1* genes, and to the identification of the SNP rs12740374 as the ultimate responsible for the association observed in GWAS. Bioinformatic analysis showed that, by altering the wild-type sequence from GGTGCTCAAT to GTTGCTCAAT, the minor allele of this variant created a binding site for the C/EBP (CCAAT-enhancer binding protein) α protein increasing promoter activity and *SORT1* expression level. This was later confirmed *in vitro*. It

should be noted that these results are in full agreement with the analyses published by Linsel-Nitsche's group [11], concerning the mRNA expression levels in the liver. Finally, through studies on liver cells from mutant mice in which the gene coding for sortilin was over-expressed or inactivated, Musunuru *et al*. [10] demonstrated that sortilin expression levels modulate the hepatic secretion of VLDL. The transgenic mouse chosen by this team was *Apobec1*$^{-/-}$ (having suppressed the gene that encodes the C->U-editing enzyme APOBEC-1) , a "humanized" mouse, with a lipid profile closer to that seen in humans, where LDL is the predominant cholesterol transporter in circulation, rather than typical of mice. When Musunuru *et al*. [10] over-expressed the *Sort1* gene in *Apobec1*$^{-/-}$ liver cells, a 70% reduction in plasma total cholesterol (TC) and LDL-c was observed. Similarly, *Sort1* inactivation by siRNA, led to increases of 46% in TC and 125% in LDL.

In general, the data presented by these two teams supports GWAS' findings and reinforces the idea of a negative correlation between *SORT1* mRNA levels and plasma LDL-c concentration. However, in the same year a third mechanistic study addressing this association was published and in this latter case, the results were not so easily reconciled neither with the results from the previous studies, nor with the previous assumptions inferred through GWAS.

We are referring to the results presented by Kjolby et al. [12] that were published almost simultaneously. These authors used as a model a double knockout mouse, *Sort1*$^{-/-}$,*Ldlr*$^{-/-}$ having observed that its' hepatocytes presented reductions of 30% of the TC levels, of ~50% of proteins containing apo B100 (VLDL and LDL) and of ~60% in atherosclerotic plaque area compared to *Ldlr*$^{-/-}$ single KO mice. Next, they performed liver-specific *Sort1* over-expression. Briefly, they found that sortilin deficiency led to a 50% reduction in the secretion of lipoproteins, whereas its over-expression resulted in a 50% increase in the secretion of those proteins. Together, these results indicate a positive correlation between *Sort1* expression and LDL-c levels, opposite to the one observed by Musunuru *et al* [10].

Discover the differences: results' analysis

The question of the discrepancy between the results of these three studies has been discussed by various experts, as illustrated by the publications of Dubé [34] and Tall and Ai [33], from 2011. In their comments, these authors drew attention to the differences between the experimental works by Linsel-Nitschke *et al*. [11], Musunuru *et al*. [10] and Kjolby *et al*. [12] and the consequences that those differences may have had in their results. The three teams that tried to clarify the mechanism by which the 1p13 *locus* affects the LDL levels and the risk of CAD, have opted for different experimental models which seem to have influenced the final results. This implies that, even though their conclusions are consistent and have resulted from well designed and consistent experiments, may not be comparable.

The first important point to focus is the metabolic background in which each of the experiments was carried out. Linsel-Nitschke, and his team [11] conduced their investigation only in humans, unlike Musunuru [10] and Kjolby [12] who, together with their collaborators

analysed non-human animal models. Nevertheless, these latter teams chose mice models with different metabolic profiles: Musunuru *et al*. [10] worked with liver cells from a "humanized" mouse, *Apobec*$^{-/-}$; Kjolby et al. studied a double knockout for sortilin and LDL receptor (*Sort1*$^{-/-}$, *Ldlr*$^{-/-}$). Musunuru's mouse produced and secreted abnormally high amounts of lipoproteins in order to mimic the human lipid profile. This may have artificially modified sortilin's secretory pathways and availability; Kjolby's [12] mouse had a deficient lipoprotein catabolism, created by the repression of *Sort1* expression within hepatocytes and marked by "high fat western diet". Finally, it is also known that there are differences in gene regulation between the two species, man and mouse, as demonstrated by the absence of the C/EBP α binding site in mice [10, 37]. This may hinder extrapolation of mouse studies to human when referring to sortilin [34].

It should be highlighted that, altogether, the *in vivo* observations from studies in mouse models showed that sortilin assumes complementary liver functions, depending on the metabolic milieu in which it operates and, ultimately, regulates VLDL secretion. The different results here reviewed seem to suggest that sortilin can regulate VLDL secretion and traffic to the lysosome when the intracellular levels of apo B-100 are extremely high. Conversely, at low apo B-100 expression levels, sortilin can regulate the formation and secretion of VLDL [33, 34].

Nevertheless, this putative role of sortilin in the formation and secretion of VLDL hardly reconciles with the GWAS' results that pointed to a specific association with LDL-c and not to triglycerides, which are the major components of VLDL particles [34].

Taken together, even though these studies provide evidences which are somehow contrary, they also give strong evidence of the existence of a novel regulatory pathway for lipoprotein metabolism and the possibility that its modulation alters the risk of cardiovascular disease in humans. Nevertheless, there's still a long way to walk in order to get a clear vision of the whole process and its modulator factors.

Conclusion

By genotyping alleles whose frequency is higher, GWAS are limited to the identification of alleles which exert minimal, if not negligible, effects on the phenotype [38]. Furthermore, these most common alleles seem to explain only a small portion of the phenotype [39]. Thus, a significant fraction of the inheritance of complex phenotypes such as cardiovascular diseases (CAD in particular) remains unknown, even after all efforts have been directed to this area through numerous GWAS. This portion of heredity is often called "the missing heritability" or "the dark matter of heritability" [35]. GWAS' supporters argue that, by increasing the size of the studied samples and the SNP density, it will be possible to detect alleles with very small effect sizes, revealing the portion of inheritance which remains unknown [35]. Nevertheless, some authors support an alternative strategy based on whole genome direct sequencing as a way of identifying rare alleles carrying large effects on the phenotype [40]. Crucial to this are

the recent advances in DNA sequencing technology, with the onset and development of new platforms for third-generation sequencing. These platforms allow low cost whole genome sequencing, with consequent identification of rare and/or new variants. It is believed that each genome has approximately 10000 non-synonymous variants, among approximately 3,5 millions of SNPs. Having in mind the magnitude of these values it is expectable that this kind of sequencing dominates genetic studies in the coming years, a trend that is already seen nowadays [41, 42, 43, 44, 45].

It seems, however, correct to state that the "dark matter of heritability" is the product of complex interactions between factors of different nature: genetic, genomic and epigenetic [39, 46]. Similarly, also the phenotype is the result of nonlinear and stochastic interactions between different genetic and non-genetic factors.

Nevertheless, the discovery and systematization of new genetic variants associated with a particular complex phenotype is very important, particularly for the case here reported, in which this variant has pioneered the discovery of a previously unknown molecular pathway.

The three reports here reviewed are exemplary and paradigmatic approaches of how to get from a "blind" statistical association given by GWAS to a mechanistic explanation of how a certain genetic variation may modulate a particular phenotype. In this case, GWAS' results pointed to a particular starting point, the 1p13 locus [14, 21-23], which would end up catching the attention of three independent teams who relied on different experimental approaches to unveil the basis of the statistical association, all of them highlighting the *SORT1* gene as the modulator of LDL-c levels and MI risk. But, if their results were in accordance concerning the relevance of SORT1's role on the lipoprotein metabolism regulation, their interpretations on the direction of the effect of its expression over plasma LDL-c levels and its underlying mechanism are far from concordant [10-12].

Linsel-Nitschke and colleagues [11] have proposed, due to their observations, that the overexpression of sortilin increases the internalization of LDL, with a consequent decrease of its plasma levels [11]. Soon after, through studies in human cohorts, hepatocytes and mice, Musunuru *et al*. [10] reported an inverse relationship between the sortilin expression and circulating LDL-c levels and proposed an explanatory mechanism through transcriptional regulation (liver-specific) of the *SORT1* gene by transcription factor C/EBPα [10]. On the contrary, Kjolby and his team [12] observed a direct relationship between the expression of *Sort1* and the concentration of circulating LDL, suggesting it could result from increased VLDL secretion [12].

Several explanations have been presented to justify the discrepancy between these results and the answer seems to depend on sortilin itself, which appears to be a multifaceted protein that may assume different functions depending on the circumstances.

In summary, the studies here revised, presented a number of strong evidence demonstrating that *SORT1* is a regulatory element of plasma LDL-c levels, adding a significant role to the sortilin-coding gene, until very recently ignored. Presently, the cellular pathway relating

sortilin to lipid metabolism is still controversial but it is surely an issue that will be further explored. A full understanding of this pathway will be crucial to evaluate if sortilin may represent a potential target for therapeutic interventions for hypercolesteronemia or CAD (reviewed in 33-35).

Referências/ References

1. Deaton C, Froelicher E, Wuc L et al. The global burden of cardiovascular disease. Eur J Cardiovasc. Nurs, 2011; 10:S5–S13.

2. Anuário Estatístico de Portugal (2010). *Instituto Nacional de Estatística*. (http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=133813349&PUBLICACOESmodo=2, acedido a 18 de agosto de 2012).

3. Rosamond W, Flegal K, Friday G, et al. Heart disease and stroke statistics—2007 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation, 2007; 115:e69–e171.

4. Prospective Studies Collaboration. Blood cholesterol and vascular mortality by age, sex and blood pressure: meta-analysis of individual data from 61 prospective studies with 55 000 vascular deaths. Lancet, 2007; 370:1829–1839.

5. Costa J, Borges M, Oliveira E, et al. Incidence and prevalence of hypercholesterolemia in Portugal: a systematic review. Part I. Rev Port Cardiol, 2003; 22:569-77.

6. Bourbon M. Genetic factors and cardiovascular disease. Rev Port Cardiol, 2008; 27:1559-63.

7. Medeiros AM, Alves AC, Francisco V, Bourbon M; investigators of the Portuguese FH Study. Update of the Portuguese Familial Hypercholesterolaemia Study. Atherosclerosis, 2010; 212:553-8.

8. Watkins H, Farrall M. Genetic susceptibility to coronary artery disease: from promise to progress. Nat Rev Genet, 2006; 7:163–173.

9. Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. Lancet; 2005, 366:941–951.

10. Musunuru K, Strong A, Frank-Kamenetsky M, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature, 2010; 466:714-9.

11. Linsel-Nitschke P, Samani NJ, Schunkert H. Sorting out cholesterol and coronary artery disease. N Engl J Med, 2010; 363:2462-3.

12. Kjolby M, Andersen OM, Breiderhoff T, et al. Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. Cell Metab, 2010; 12:213-23.

13. Kathiresan S, Melander O, Guiducci C, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat Genet, 2008a; 40:189–197.

14. Samani NJ, Erdmann J, Hall AS, et al. Genomewide association analysis of coronary artery disease. N Engl J Med, 2007; 357:443–453.

15. Samani NJ, Braund PS, Erdmann J, et al. The novel genetic variant predisposing to coronary artery disease in the region of the PSRC1 and CELSR2 genes on chromosome 1 associates with serum cholesterol. J Mol Med (Berl), 2008; 86:1233-41.

16. Willer CJ, Sanna S, Jackson AU, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet, 2008; 40:161–169.

17. Sandhu MS, Waterworth DM, Debenham SL, et al. LDL-cholesterol concentrations: a genome-wide association study. Lancet, 2008; 371:483–491.

18. Wallace C, Newhouse SJ, Braund P, et al. Genomewide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. Am J Hum Genet, 2008; 82:139–49.

19. Samani NJ, Deloukas P, Erdmann J, et al. Large scale association analysis of novel genetic loci for coronary artery disease. Arterioscler Thromb Vasc Biol, 2009; 29:774-80.

20. Kathiresan S, Willer CJ, Peloso GM, et al. Cupples, Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet, 2009; 41:56-65.
21. McPherson R, Pertsemlidis A, Kavaslar N, et al. A common allele on chromosome 9 associated with coronary heart disease. Science, 2007; 316:1488–1491.

22. Helgadottir A, Thorleifsson G, Manolescu A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. Science, 2007; 316:1491–1493.

23. The Wellcome Trust Case Control Consortium. Genomewide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature, 2007; 447:661–678.

24. Nielsen MS, Jacobsen C, Olivecrona G, et al. Sortilin/neurotensin receptor-3 binds and mediates degradation of lipoprotein lipase. J Biol Chem, 1999; 274:8832-6.

25. Nilsson SK, Christensen S, Raarup MK, et al. Endocytosis of apolipoprotein A-V by members of the low density lipoprotein receptor and the VPS10p domain receptor families. J Biol Chem, 2008; 283:25920-7.

26. Mazella J, Zsürger N, Navarro V, et al. The 100-kDa neurotensin receptor is gp95/sortilin, a non-G-protein-coupled receptor. J Biol Chem, 1998; 273:26273-6.

27. Petersen CM, Nielsen MS, Nykjaer A et al. Molecular identification of a novel candidate sorting receptor purified from human brain by receptor-associated protein affinity chromatography. J Biol Chem, 1997; 272: 3599-605.

28. Lefrancois S, Zeng J, Hassan AJ, et al. The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin. EMBO J, 2003; 22: 6430-7.

29. Ni X, Canuel M, Morales CR. The sorting and trafficking of lysosomal proteins. Histol Histopathol, 2006; 21: 899-913.

30. Canuel M, Korkidakis A, Konnyu K. Sortilin mediates the lysosomal targeting of cathepsins D and H. Biochem Biophys Res Commun, 2008; 373:292-7.

31. Shi J, Kandror KV. Sortilin is essential and sufficient for the formation of Glut4 storage vesicles in 3T3-L1 adipocytes. Dev Cell, 2005; 9: 99-108.

32. Vaegter CB, Jansen P, Fjorback AW, Glerup S, Skeldal S, Kjolby M, Richner M, Erdmann B, Nyengaard JR, Tessarollo L, Lewin GR, Willnow TE, Chao MV, Nykjaer A. Sortilin associates with Trk receptors to enhance anterograde transport and neurotrophin signaling. Nat Neurosci. 2011; 1:54-61.

33. Tall AR, Ai D. Sorting out sortilin. Circ Res, 2011; 108:158-60.

34. Dubé JB, Johansen CT, Hegele RA. Sortilin: an unusual suspect in cholesterol metabolism: from GWAS identification to in vivo biochemical analyses, sortilin has been identified as a novel mediator of human lipoprotein metabolism. Bioessays, 2011; 33:430.

35. Marian AJ. Genome-wide association studies complemented with mechanistic biological studies identify sortilin 1 as a novel regulator of cholesterol trafficking. Curr Atheroscler Rep, 2011; 13:190-2.

36. Schadt EE, Molony C, Chudin E, et al. Mapping the genetic architecture of gene expression in human liver. PLoS Biol, 2008; 6:e107.

37. Russell JC, Proctor SD. Small animal models of cardiovascular disease: tools for the study of the roles of metabolic syndrome, dyslipidemia, and atherosclerosis. Cardiovasc Pathol. 2006; 15:318-30.

38. Marian AJ. Nature's genetic gradients and the clinical phenotype. Circ Cardiovasc Genet, 2009; 2:537-9.

39. Eichler EE, Flint J, Gibson G, et al. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet, 2010; 11:446-50.

40. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet, 2010; 11:415-25.

41. Wang J, Wang W, Li R, et al.. The diploid genome sequence of an Asian individual. Nature, 2008; 456:60-5.

42. Levy S, Sutton G, Ng PC, et al. The diploid genome sequence of an individual human. PLoS Biol, 2007; 5:e254.

43. Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. Nature, 2008; 452:872-6.

44. Ng PC, Levy S, Huang J, et al. Genetic variation in an individual human exome. PLoS Genet, 2008; 15:e1000160.

45. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature, 2009; 461:272-6.

46. Kapranov P, Laurent GS. Genomic "dark matter": implications for understanding human disease mechanisms, diagnostics, and cures. Front Genet. 2012; 3:95.

Appendix 2

Book Chapter

N-acetylglucosamine-1-phosphate transferase, gamma subunit; N-acetylglucosamine-1- (GNPTG, GNPTAB). *In*: Taniguchi N., Honke K., Fukuda M., Narimatsu H., Yamaguchi Y., Angata T. (Ed.) *Handbook of Glycosyltransferases and Related Genes: SpringerReference* (www.springerreference.com). Springer-Verlag Berlin Heidelberg, 2013 [in press].

N-acetylglucosamine-1-phosphate transferase, gamma subunit; N-acetylglucosamine-1- (GNPTG, GNPTAB)

Introduction

GlcNAc-1-phosphotransferase catalyses the transfer of a GlcNAc-1-phosphate residue from UDP-GlcNAc to C6 positions of selected mannoses in high-mannose type oligosaccharides of the hydrolases (Goldberg and Kornfeld, 1981; Natowicz et al., 1982; Varki and Kornfeld, 1983). At a biological level this reaction is followed by the removal of the terminal GlcNAc by an N-acetylglucosamine-1-phosphodiester α-N-acetyl-glucosaminidase, usually referred to as 'uncovering enzyme' (UCE; see Chapter 78; Article ID: 332135). Sequential action of these two enzymes results in the formation of the mannose-6-phosphate (Man-6-P) marker, a specific tag acquired by lysosomal hydrolases that ensures recognition by M6P receptors and delivery to the endosomal/lysosomal system (Braulke and Bonifacino, 2009).
GlcNAc-1-phosphotransferase was purified from bovine mammary glands and proposed to exist as an hexameric complex with a molecular mass of 540 kDa composed of two α-, two β- and two γ-subunits (Bao et al., 1996).
Two independent genes code for GlcNAc-1-phosphotransferase: GNPTAB, which codes for the α/β-subunit precursor (Raas-Rothschild et al., 2000) and GNPTG that codes for the γ-subunit (Kudo et al., 2005; Tiede et al., 2005).
Absence or deficient activity of GlcNAc-1-phosphotransferase leads to lysosomal storage disorder, underlying Mucolipidosis type II or III, respectively.

Databanks

IUBMB enzyme nomenclature: EC 2.7.8.17

Species	Gene Symbol	GeneBank accession number	Uniprot ID	PDB accession number
Homo sapiens				
α/β-subunit precursor	GNPTAB (alternatively GNPTA; KIAA1208)	NP_077288	Q3T906	N/A
γ-subunit	GNPTG (alternatively GNPTAG; C16orf27)	NP_115909	Q9UJJ9	N/A
Mus musculus				
α/β-subunit precursor	Gnptab (alternatively Gnpta; Kiaa1208)	NM_001004164.2	Q69ZN6	N/A
γ-subunit	Gnptg (alternatively Gnptag; C16orf27)	BC055872.1	Q6S5C2	N/A

Name and History

GlcNAc-1-phosphotransferase was first demonstrated in the membrane fraction of rat liver, skin fibroblasts and CHO cells (Reitman and Kornfeld, 1981a; Hasilik et al., 1981).
Early reports on GlcNAc-1-phosphotransferase activity have shown that it phosphorylates acid hydrolases at least 100-fold more than nonlysosomal glycoproteins containing identical oligosaccharide units (Reitman and Kornfeld, 1981b; Lang et al., 1984; Ketcham and Kornfeld, 1992). Additionally, heat-denaturated acid hydrolases as well as high-mannose-type oligosaccharides and glycopeptides were shown to be extremely poor substrates, while deglycosylated lysosomal enzymes acted as potent inhibitors of the phosphorylation of intact lysosomal enzymes (Lang et al., 1984). Altogether, these data suggested that GlcNAc-1-phosphotransferase recognizes a conformation-dependent protein domain that is common to all acid hydrolases but absent in non-lysosomal glycoproteins (Lang et al., 1984).
It is this specific ability to recognize lysosomal hydrolases that allows for the critical role GlcNAc-1-phosphotransferase

plays in the cell: to selectively target those enzymes to their final destination, the endosomal/lysosomal compartment, by catalyzing the first of two steps that result in the biosynthesis of the Man-6-P recognition signal (Fig. 1). The modified proteins are then recognized by two independent receptors that bind the Man-6-P residue of the newly synthesized lysosomal hydrolases in the trans-Golgi network (TGN) and finally, the ligand-receptor complex is packaged into clathrin-coated transport vesicles for delivery to endosomes and lysosomes (Braulke and Bonifacino, 2009).

Numerous researchers tried to purify and clone this enzyme. In 1981, rat GlcNAc-1-phosphotransferase was partially purified (Reitman and Kornfeld, 1981b) but the complete enzyme was only accessed in 1996, when the group of Canfield succeeded in purifying it from bovine mammary glands and demonstrated that GlcNAc-1-phosphotransferase is, indeed, a 540-kDa complex composed of three different subunits: two α, two β and two γ (Bao et al., 1996a). Latter, three independent teams have identified and/or cloned the genes that code for those subunits: GNPTAB, which codes for the α/β-subunits (Raas-Rothschild et al., 2000) and GNPTG, which codes for the γ-subunit (Kudo et al., 2005; Tiede et al ., 2005). The GNPTAB gene contains 21 exons and spans 85 kb on chromosome 12q23.3 (Tiede et al., 2005), whilst the GNPTG gene spans 11 exons on chromosome 16p13.3 (Raas-Rothschild et al., 2000).

UDP-N-acetylglucosamine (UDP-GlcNAc): lysosomal enzyme N-acetylglucosamine-1-phosphotransferase may also be referred to as stealth protein GNPTAB or UDP-N-acetylglucosamine-1-phosphotransferase, and is usually abbreviated as GlcNAc-1-phosphotransferase.

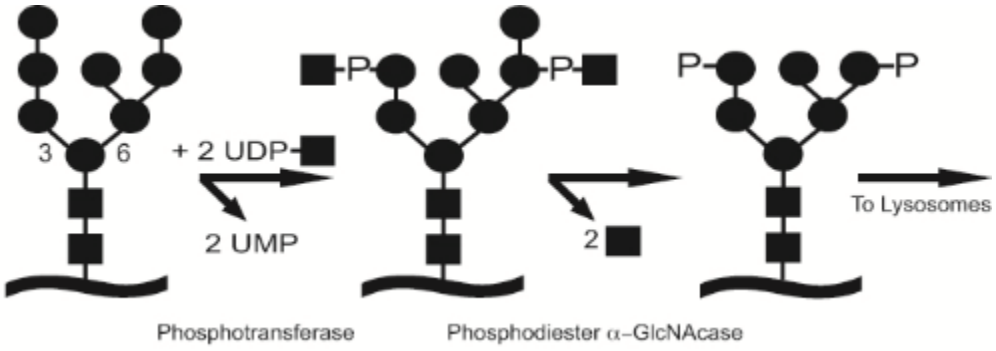


Figure 1: Phosphorylation of mannose residues of lysosomal enzyme oligosaccharides. Circles,mannose; Squares, GlcNAc.

Structure

GlcNAc-1-phosphotransferase is a Golgi-resident 540-kDa hexameric transmembrane enzyme composed by three subunits, $\alpha_2\beta_2\gamma_2$.

The α/β-subunit precursor is synthesized as an N-glycosylated 190-kDa type III membrane protein. Proteolytic cleavage of the α/β-precursor between Lys928 and Asp929 by site-1 protease (S1P; also known as subtilisin kexin isoenzyme–1, SKI-1) generates the individual α- and β- subunits with molecular masses of 145 kDa and 45 kDa, respectively (Tiede et al., 2005; Marschner et al., 2011). The mature human α-subunit is a type II membrane protein with an N-terminal tail of 19 amino acids, a transmembrane domain consisting of 22 amino acids and a subsequent luminal domain comprising 886 residues (Kudo et al., 2005; Tiede et al., 2005). The luminal domain of the α-subunit contains 17 potential N-glycosylation sites and exhibits a conserved modular structure composed of at least six domains with homologies to bacterial synthesis proteins plus two repeats of the Notch receptor and a binding domain for the transcriptional co-repressor DMAP1 (Fig. 2). Additionally, a potential conserved coiled-coil domain comprising amino acids 80 to 120 is also present (Kudo et al., 2005; Tiede et al., 2005; reviewed in Kollmann et al., 2010).

The mature β- subunit is a type I membrane protein with three potential N-glycosylation sites, and contains cytoplasmic, transmembrane and luminal domains of 21, 23 and 284 amino acids, respectively.

The human γ-subunit, on the other hand, is not a transmembranar but a soluble glycoprotein of 305 amino acids with a predicted signal sequence of 24 amino acids and can form disulfide-linked dimmers (Raas-Rothschild et al., 2000). It contains two in vivo used N-glycosylation sites at positions 88 and 115, equipped with high mannose-type oligosaccharides (Encarnação et al., 2010).

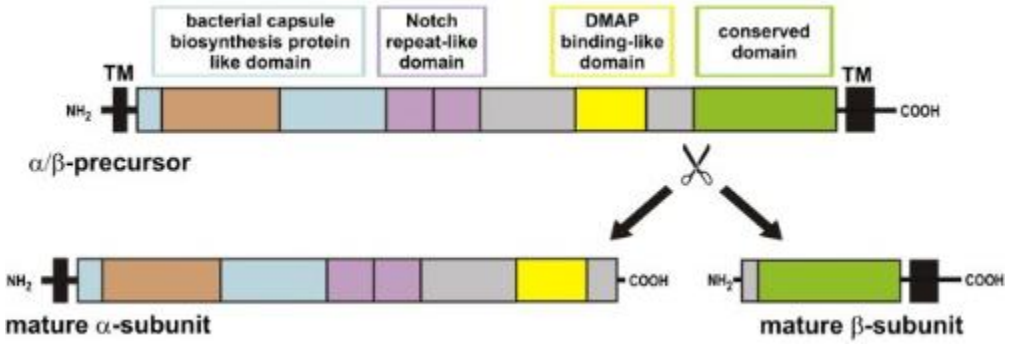


Figure 2: Schematic presentation of the modular structure of the type III membrane protein α/β-subunit precursor of GlcNAc-1-phosphotransferase. The α/β-subunit precursor is cleaved by site-1-protease into the mature type II α-subunit and the mature type I β-subunit membrane protein (adapted from Kollmann et al., 2010).

Enzyme Activity Assay and Substrate Specificity

GlcNac-1-phosphotransferase catalyzes the first step in the two-step reaction sequence shown in Fig. 1 by selectively recognizing and phosphorylating lysosomal enzymes (Reitman and Kornfeld, 1981; Waheed et al., 1982):
UDP-N-acetyl-D-glucosamine + lysosomal-enzyme D-mannose =
UMP + lysosomal-enzyme N-acetyl-D-glucosaminyl-phospho-D-mannose

GlcNac-1-phosphotransferase activity assays are carried out with a radioactively-labeled sugar donor: [b-^{32}P] UDP-GlcNAc. Even though not being commercially available, [β-^{32}P] UDP-GlcNAc is the most suitable substrate for determining this enzyme's activity and may be easily prepared from [γ-^{32}P] ATP using commercially available enzymes (Lang and Kornfeld, 1984) and α-methylmannoside.The transfer of ^{32}P-labeled GlcNAc-1-phosphate from [β-^{32}P] UDP-GlcNAc to various acceptor substrates is quantified by one of two methods. In the first method, α-methylmannoside is the acceptor. Ion-exchange chromatography is used to separate the phosphodiester product, which has only one negative charge at neutral pH, from free phosphate, phosphomonoesters, and nucleotide sugars. The second method is used to measure GlcNAc-1-phosphate transfer to glycoprotein and glycopeptide acceptors. After incubation, the acceptor glycoproteins are digested with Pronase, producing glycopeptides that are then subjected to affinity chromatography on concanavalin A (Con A)-Sepharose. The ^{32}P radioactivity bound to the Con A-Sepharose is the product (Reitman et al., 1984).

The enzyme activity is dependent on a divalent cation, Mg^{2+} or Mn^{2+} (Reitman and Kornfeld, 1981a). The Km of bovine GlcNAc-1-phosphotransferase for UDP-GlcNAc is 30 μM and that for the lysosomal enzyme cathepsin D is 18 μM. On the other hand, the Km for the nonlysosomal enzyme RNase B is 1,2 mM and that for a-methylmannoside is 64 mM (Bao et al., 1996b).

The recognition site of GlcNAc-1-phosphotransferase is separate from its catalytic domain (Reitman and Kornfeld, 1981b; Waheed et al., 1982). The recognition site binds to a signal patch that is present in lysosomal hydrolases but absent in most secretory glycoproteins. Early reports by the Kornfeld's lab demonstrated that lysine 203 and amino acids 265-292 of cathepsin D were the minimum number of amino acids on the signal patch for recognition by GlcNAc-1-phosphotransferase (Baranski et al., 1990; Baranski et al., 1991; Cantor et al., 1992). Latter, antibody inhibition experiments with arylsulfatase A were also used to show that GlcNAc-1-phosphotransferase recognizes common conformation-dependent protein structures of lysosomal enzymes in which lysine residues are the major determinants (Sommerlade et al. 1994). In addition, the interaction of both subunits with the protein determinant of acid hydrolases stimulates the catalytic function of GlcNAc-1-phosphotransferase (Qian et al. 2010).

GlcNAc-1-phosphate is transferred to selected mannose residues on high mannose type N-glycans at the catalytic site of GlcNAc-1-phosphotransferase. The first mannose to be phosphorylated is generally the terminal mannose on the α1,6

branch linked to the core mannose, while phosphorylated oligosaccharides formed subsequently contain one or two Man-6-P residues, located at different positions on the oligosaccharide (Goldberg and Kornfeld, 1981; Natowicz et al., 1982; Varki and Kornfeld, 1983).

Preparation

GlcNAc-1-phosphotransferase activity has been demonstrated in membranes of a multitude of tissues and cells, including rat liver, skin fibroblasts, placenta, bovine pancreas, CHO cells and COS1-cells (Waheed et al., 1981; Hasilik et al., 1981; Reitman and Kornfeld, 1981a,b; Reitman et al., 1984; Zhao et al., 1992; Nishikawaet al., 1997; Kudo et al., 2006).
Rat GlcNAc-1-phosphotransferase was partially purified through diethylaminoethanol-cellulose chromatography (Reitman and Kornfeld, 1981b). The bovine form of the same enzyme was purified at approximately 490.000-fold from mammary glands using monoclonal antibody immunoaffinity chromatography (Bao et al., 1996a).

Biological Aspects

The $\alpha_2\beta_2\gamma_2$ heterohexamericcomplex that constitutes the active GlcNAc-1-phosphotransferase is encoded by two different genes, GNPTAB and GNPTG, which code for the synthesized α/β- and the γ-subunits, respectively. The α/β-subunit precursor is synthesized as an inactive 190-kD precursor protein in a hairpin orientation that contains cytosolic N and C termini and a complex modular structure. Subsequent cleavage of the α/β-precursor by the S1P gives origin to the mature α- and β-subunits, which are catalytically active (Tiede et al., 2005; Kudo et al., 2005; Marschner et al., 2011).
The γ-subunit, on the other hand, is synthesized as a soluble glycoprotein of 305 amino acids that forms disulfide-linked homodimers (Raas-Rothschild et al., 2000). Its function is still unclear even though recent studies have suggested that the γ-subunit is important to facilitate the proper folding of the subunits that constitute GlcNAc-1-phosphotransferase and to maintain them in a conformation competent for substrate recognition and binding or to regulate the activity/expression of the α/β-subunits (Lee et al., 1981; Pohl et al., 2009; Pohl et al., 2010). Furthermore, the γ-subunit was proposed to facilitate the addition of the second GlcNAc-P to high mannose oligosaccharides of protein substrates (Qian et al. 2010).
It is now known that the assembly of the GlcNAc-1-phosphotransferase complex takes place in the ER and depends on a proper dimerization of the γ-subunits (Encarnação et al., 2010). Then, thanks to a combinatorial sorting motif that is present in the α/β-subunit precursor protein, the oligomeric type III GlcNAc-1-phosphotransferase is exported from the ER to the Golgi in a COPII-dependent fashion (Franke et al., 2013). Once in the Golgi, the α/β-subunit precursors of the complex may finally be cleaved by the S1P, thus becoming active (Marschner et al., 2011).
GlcNAc-1-phosphotransferase activity is essential to a proper delivery of acid hydrolases to the endosomal/lysosomal compartment. The interaction of GlcNAc-1-phosphotransferase with those hydrolases results in the transfer of GlcNAc-1-phosphate to mannose residues on the asparagines-linked high-mannose oligosaccharides of the acid hydrolases. GlcNAc residues are subsequently removed by the sequential action of a second enzyme, N -acetylglucosamine-1-phosphodiester α-N-acetyl-glucosaminidase (see Chapter 78) that exposes the Man-6-P recognition signal which is then recognized by specific receptors at the TGN: the cation-independent M6P receptor and/or the cation-dependent M6P receptor. Ligand-receptor complexes are packed intoclathrin-coated that bud from the TGN and, ultimately, fuse with late endosomes. At the low pH of the late endosome, the hydrolases dissociate from the Man-6-P receptors and the empty receptors are recycled to the Golgi apparatus for further rounds of transport (Braulke and Bonifacino, 2009). Ultimately, the biological relevance of GlcNAc-1-phosphotransferase and of the targeting pathway in which it is involved, may be demonstrated by the severity of pathology triggered by GlcNAc-1-phosphotransferase impaired function or, in extreme, by its complete absence.

Knockout Mouse and Transgenic mice

The first mouse model for GNPTAB deficiencies was generated in 2007 by Gelfman and colleagues,through microinjection of embryonic stem (ES) cell clones GNPTAB$^{-/-}$ into host blastocysts. Homozygous mice lacking α/β-subunits of GlcNAc-1-phosphotransferase presented with growth retardation, retinal degeneration and secretory cell lesions that do not exactly mimic the human disease (Gelfman et al., 2007). Latter, a second mouse model was generated, to prevent compensatory mechanisms in the complete absence of GlcNAc-1-phosphotransferase. This 'knock-in' mouse model, generated by Kollmann and collaborators, has a single cytosine inserted (c.3082insC), which disrupts the open reading frame causing a premature translational termination in the C-terminal conserved region (p.G1049RfsX16) (Kollmann et al., 2012). This insertion mimics a human pathogenic mutation (p.G1049RfsX16) already reported and extensively analysed (Tiede et al., 2005). The 'knock-in' mice are severely affected and show all of the clinical and biological symptoms and features of the human ML II disease thus being the ideal model to further investigate pathology (Kollmann et al., 2012).

Finally, the only mouse model for ML III gamma reported so far was generated by direct Gnptg targeting, leading to a truncated γ-subunit with only 59 amino acids (Lee et al., 2007). Gnptg-deficient mice presented with normal growth and no cartilage defects or retinal degeneration, even though the serum levels of numerous enzymes were elevated (Vogel et al., 2009).

With the exception of the Gnptab 'knock-in' mouse generated in the Braulke lab (Kollmann et al., 2012), mice lacking either α/β- or γ-subunits displayed clinical and pathologic features that differed substantially from those reported in humans having mutations in orthologous genes (Vogel et al., 2009).

Human Disease

Defective GlcNAc-1-phosphotransferase causes two distinct human lysosomal storage diseases, Mucolipidosis II (ML II) and Mucolipidosis III (ML III), which are among the few lysosomal storage disorders related to defects in non-lysosomal proteins. ML II, often referred to as I-cell disease, is characterized by a total loss of GlcNAc-1-phosphotransferase activity whether ML III, also known as pseudo-Hurler polydystrophy, manifests when enzymatic activity is reduced (Leroy and Spranger, 1970; Spranger and Wiedemann, 1970). In both ML II and ML III patients, newly synthesized lysosomal enzymes fail to be correctly sorted to the endosome/lysosome compartment due to the absence or weak equipment in M6P residues. As a consequence, lysosomal dysfunction develops leading to accumulation of non-degraded material, the hallmark of this group of diseases. Unlike the majority of lysosomal storage disorders, which involve single enzymes acting in a catabolic pathway, MLII and III results from impaired sorting of multiple enzymes to lysosomes that instead are over-secreted from cells. The excessive accumulation of non-degraded substrates results in the subsequent formation of large inclusion bodies.

While ML II and ML III share similar clinical features, including skeletal abnormalities, ML II is the more severe in terms of phenotype (Leroy and Spranger, 1970). In this pathology the skeletal system is severely affected, with abnormalities in both cartilage and bone, cardiomegaly and developmental delay. Linear growth decelerates during the first year of life almost stopping during the second year and death usually occurs between 5 and 8 years of age (Kornfeld and Sly in Scriver, 2001).

ML III is a much milder disorder, being characterized by latter onset of clinical symptoms and slower, progressive course, which may allow the survival into the eighth decade. Usual clinical findings in ML III patients include restricted joint mobility, short stature and mild Hurler-like dysmorphism, among other less severe features (Spranger and Wiedemann, 1970; Umehara et al., 1997). Only 50% of the patients present with mental retardation (Umehara et al., 1997).

Once GlcNAc-1-phosphotransferase is an hexameric complex whose protein subunits are encoded by two genes, depending on which of them harbors the causal mutation(s) and simultaneously on the severity/clinical course of the disease, the associated pathologies are classified as ML II alpha/beta (OMIM: 252500) and ML III alpha/beta (OMIM: 252600) if mutations are present in the GNPTAB gene, or ML III gamma (OMIM: 252605) if mutations occur in the GNPTG gene (Cathey et al., 2008).

To date, more than 100 different GNPTAB mutations have been reported, causing either ML II alpha/beta or ML III alpha/beta, including missense, nonsense, small deletions, small insertions, small indels and splice site mutations (HGMD and references therein). Large genomic rearrangements appear to be rare although two gross insertions (Otomo et al.,

2009; Tapino et al., 2008) and one large deletion have already been detected (Coutinho et al., 2011a). Most of these mutations are private or rare. One exception does exist, however: the microdeletion c.3503_3504delTC. Being the most common ML II alpha/beta-causing allele, c.3503_3504delTC presents a remarkably wide geographical distribution, having been detected among Israeli and Palestinian Arab-Muslim, Turkish, Irish traveler (Bargal et al., 2006), Portuguese (Encarnação et al., 2009), Italian (Tapino et al., 2009) and U.S. patients (Kudo et al., 2006). This pathogenic mutation was due to an ancient and unique founder molecular lesion that arose, most probably, in a peri-Mediterranean region (Coutinho et al., 2011b).

Concerning the GNPTG gene, the mutations until now reported associated to Mucolipidosis type III gamma include 4 missense, 2 nonsense, 5 small deletions, 4 small insertions, 4 splice site mutations and 2 gross deletions [HGMDand references therein].

Recently, genomewide scans have unveiled a curious and unexpected relation between mutations in the Man-6-P dependent lysosomal enzyme-targeting pathway and persistent stuttering. While studying individuals with persistent stuttering, Kang and colleagues (2010) found mutations in all the three genes which encode the two enzymes that generate the Man-6-P recognition marker: GNPTAB, GNPTG and NAGPA. Ten different mutations were found in total, including four mutations in the GNPTAB gene and three mutations in the GNPTG. Yet, the affected subjects studied by Kang et al. (2010) stuttered but were otherwise normal, without presenting any of the typical symptoms of lysosomal malfunction, even when being homozygous for some of the identified mutations on the GNPTAB and GNPTG genes. Eventually, this can be explained admitting that the efficiency of lysosomal targeting was only partially reduced in the presence of the detected mutations.

Future Perspectives

Over the last two decades many important progresses were made with the purification of the bovine GlcNAc-1-phosphotransferase and the subsequent identification of the genes encoding its α/β-precursor and γ-subunit. All these achievements allowed improved pre- and postnatal diagnosis of ML II and III patients with immediate consequences for affected families.

Nevertheless, several questions remain unanswered and important issues remain to be solved concerning GlcNAc-1-phosphotransferase itself. Detailed analysis of the modular structure of the α/β-precursor for the binding of lysosomal enzymes, high-mannose-type oligosaccharides, and UDP-GlcNAc is still missing. Knowledge of the GlcNAc-1-phosphotransferase crystallographic structure and subunit oligomerization is also missing (Kollmann et al., 2010).

Only a deeper understanding of this enzyme, the travelling pathway in which it is involved and the alternative mechanisms of M6P-independent lysosomal enzyme transport may allow establishment and improvement of therapeutic approaches not only for ML II and III but also for other lysosomal storage disorders.

Finally, also the remarkable finding that mutations in the GNPTAB and GNPTG may be associated with persistent stuttering in individuals who are otherwise normal, deserves further attention. Additional studies in this area may contribute to a better understanding of the neural structures and functions within the brain that generate human speech, which are still poorly elucidated.

Cross Reference

N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase (NAGPA)

Further Reading

Bao et al., 1996 a and b: purification of bovine GlcNAc-1-phosphotransferase and insights into its heterohexameric structure, enzymatic activity and catalytic subunit.

Raas-Rothschild et al., 2000; Kudo et al., 2005 and Tiede et al., 2005: cloning of the two genes that code for GlcNAc-1-phosphotransferase and their involvement in disease.

Braulke and Bonifacino, 2008 and Coutinho et al., 2012: comprehensive revisions on the newly synthesized lysosomal enzymes' travelling routes to the lysosome, particularly of the Man-6-P-dependent pathway in which

GlcNAc-1-phosphotransferase plays a key role.
Kollmann et al., 2010: complete, easy-to-follow review on the involvement of mannose-phosphorylation in human disease.

References

Bao M, Booth JL, Elmendorf BJ, Canfield WM (1996a) Bovine UDP-N-acetylglucosamine: lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure. J Biol Chem 271:31437-45.
Bao M, Elmendorf BJ, Booth JL, Drake RR, Canfield WM (1996b) Bovine UDP-N-acetylglucosamine:lysosomal-enzyme N-acetylglucosamine-1-phosphotransferase. II. Enzymatic characterization and identification of the catalytic subunit. J Biol Chem 271:31446-51.
Baranski TJ, Faust PL, Kornfeld S (1990) Generation of a lysosomal enzyme targeting signal in the secretory protein pepsinogen. Cell 63:281-91.
Baranski TJ, Koelsch G, Hartsuck JA, Kornfeld S (1991) Mapping and molecular modeling of a recognition domain for lysosomal enzyme targeting. J Biol Chem 266:23365-72.
Bargal R, Zeigler M, Abu-Libdeh B, Zuri V, Mandel H, Ben Neriah Z, Stewart F, Elcioglu N, Hindi T, Le Merrer M, Bach G, Raas-Rothschild A (2006) When Mucolipidosis III meets Mucolipidosis II: GNPTA gene mutations in 24 patients. Mol Genet Metab 88:359-63. Erratum in: Mol Genet Metab. 2007; 91:299.
Braulke T, Bonifacino JS (2009) Sorting of lysosomal proteins. Biochim Biophys Acta 1793:605-14. doi: 10.1016/j.bbamcr.2008.10.016.
Cantor AB, Baranski TJ, Kornfeld S (1992) Lysosomal enzyme phosphorylation. II. Protein recognition determinants in either lobe of procathepsin D are sufficient for phosphorylation of both the amino and carboxyl lobe oligosaccharides. J Biol Chem 267:23349-56.
Cathey SS, Kudo M, Tiede S, Raas-Rothschild A, Braulke T, Beck M, Taylor HA, Canfield WM, Leroy JG, Neufeld EF, McKusick VA (2008) Molecular order in mucolipidosis II and III nomenclature. Am J Med Genet A 146A:512-3. doi: 10.1002/ajmg.a.32193.
Coutinho MF, da Silva Santos L, Lacerda L, Quental S, Wibrand F, Lund AM, Johansen KB, Prata MJ, Alves S (2011ª) Alu-Alu Recombination Underlying the First Large Genomic Deletion in GlcNAc-Phosphotransferase Alpha/Beta (GNPTAB) Gene in a MLII Alpha/Beta Patient. JIMD Rep 4:117-24. doi: 10.1007/8904_2011_83.
Coutinho MF, Encarnação M, Gomes R, da Silva Santos L, Martins S, Sirois-Gagnon D, Bargal R, Filocamo M, Raas-Rothschild A, Tappino B, Laprise C, Cury GK, Schwartz IV, Artigalás O, Prata MJ, Alves S (2011b) Origin and spread of a common deletion causing mucolipidosis type II: insights from patterns of haplotypic diversity. Clin Genet 80:273-80. doi: 10.1111/j.1399-0004.2010.01539.x.
Coutinho MF, Prata MJ, Alves S (2012) Mannose-6-phosphate pathway: a review on its role in lysosomal function and dysfunction. Mol Genet Metab 105:542-50. doi: 10.1016/j.ymgme.2011.12.012.
Encarnação M, Lacerda L, Costa R, Prata MJ, Coutinho MF, Ribeiro H, Lopes L, Pineda M, Ignatius J, Galvez H, Mustonen A, Vieira P, Lima MR, Alves S (2009) Molecular analysis of the GNPTAB and GNPTG genes in 13 patients with mucolipidosis type II or type III - identification of eight novel mutations. Clin Genet 76:76-84. doi: 10.1111/j.1399-0004.2009.01185.x.
Encarnação M, Kollmann K, Trusch M, Braulke T, Pohl S (2010) Post-translational modifications of the gamma-subunit affect intracellular trafficking and complex assembly of GlcNAc-1-phosphotransferase. J Biol Chem 286:5311-8. doi: 10.1074/jbc.M110.202382.
Franke M, Braulke T, Storch S (2013) Transport of the GlcNAc-1-phosphotransferase α/β-subunit precursor protein to the Golgi apparatus requires a combinatorial sorting motif. J Biol Chem 288:1238-49. doi: 10.1074/jbc.M112.407676.
Gelfman CM, Vogel P, Issa TM, Turner CA, Lee WS, Kornfeld S, Rice DS (2007) Mice lacking alpha/beta subunits of GlcNAc-1-phosphotransferase exhibit growth retardation, retinal degeneration, and secretory cell lesions. Invest Ophthalmol Vis Sci 48:5221-8.
Goldberg DE, Kornfeld S (1981) The phosphorylation of beta-glucuronidase oligosaccharides in mouse P388D1 cells. J Biol Chem 256:13060-7.
Hasilik A, Waheed A, von Figura K (1981) Enzymatic phosphorylation of lysosomal enzymes in the presence of UDP-N-acetylglucosamine. Absence of the activity in I-cell fibroblasts. Biochem Biophys Res Commun 98:761-7.
Kang C, Riazuddin S, Mundorff J, Krasnewich D, Friedman P, Mullikin JC, Drayna D (2010) Mutations in the lysosomal enzyme-targeting pathway and persistent stuttering. N Engl J Med 362:677-85. doi: 10.1056/NEJMoa0902630.

Ketcham CM, Kornfeld S (1992) Characterization of UDP-N-acetylglucosamine:glycoprotein N-acetylglucosamine-1-phosphotransferase from Acanthamoeba castellanii. J Biol Chem 267:11654-9.

Kollmann K, Damme M, Markmann S, Morelle W, Schweizer M, Hermans-Borgmeyer I, Röchert AK, Pohl S, Lübke T, Michalski JC, Käkelä R, Walkley SU, Braulke T (2012) Lysosomal dysfunction causes neurodegeneration in mucolipidosis II 'knock-in' mice. Brain 135:2661-75. doi: 10.1093/brain/aws209.

Kollmann K, Pohl S, Marschner K, Encarnação M, Sakwa I, Tiede S, Poorthuis BJ, Lübke T, Müller-Loennies S, Storch S, Braulke T (2010) Mannose phosphorylation in health and disease. Eur J Cell Biol 89:117-23. doi: 10.1016/j.ejcb.2009.10.008.

Kornfeld S, Sly WS (2001) I-cell disease and pseudo-Hurler polydystrophy: disorders of lysosomal enzyme phosphorylation and localization. In: Scriver CR, Beaudet AL, Sly WS, Valle D (ed) The metabolic and molecular bases of inherited disease, 8th edn. MacGraw-Hill Inc, New York, pp 3421-3452.

Kudo M, Bao M, D'Souza A, Ying F, Pan H, Roe BA, Canfield WM (2005) The alpha- and beta-subunits of the human UDP-N-acetylglucosamine:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase [corrected] are encoded by a single cDNA. J Biol Chem 280:36141-9.

Kudo M, Brem MS, Canfield WM (2006) Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha/beta -subunits precursor gene. Am J Hum Genet 78:451-63.

Lang L, Kornfeld S (1984) A simplified procedure for synthesizing large quantities of highly purified uridine [beta-32P]diphospho-N-acetylglucosamine. Anal Biochem 140:264-9.

Lang L, Reitman M, Tang J, Roberts RM, Kornfeld S (1984) Lysosomal enzyme phosphorylation. Recognition of a protein-dependent determinant allows specific phosphorylation of oligosaccharides present on lysosomal enzymes. J Biol Chem 259:14663-71.

Lee WS, Payne BJ, Gelfman CM, Vogel P, Kornfeld S (2007) Murine UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase lacking the gamma-subunit retains substantial activity toward acid hydrolases. J Biol Chem 282:27198-203.

Leroy JG, Spranger JW (1970) I-cell disease. N Engl J Med 283:598-9.

Marschner K, Kollmann K, Schweizer M, Braulke T, Pohl S (2011) A key enzyme in the biogenesis of lysosomes is a protease that regulates cholesterol metabolism. Science 333:87-90. doi: 10.1126/science.1205677.

Natowicz M, Baenziger JU, Sly WS (1982) Structural studies of the phosphorylated high mannose-type oligosaccharides on human beta-glucuronidase. J Biol Chem 257:4412-20.

Nishikawa A, Gregory W, Frenz J, Cacia J, Kornfeld S (1997) The phosphorylation of bovine DNase I Asn-linked oligosaccharides is dependent on specific lysine and arginine residues. J Biol Chem 272:19408-12.

Nishikawa A, Nanda A, Gregory W, Frenz J, Kornfeld S (1999) Identification of amino acids that modulate mannose phosphorylation of mouse DNase I, a secretory glycoprotein. J Biol Chem 274:19309-15.

Otomo T, Muramatsu T, Yorifuji T, Okuyama T, Nakabayashi H, Fukao T, Ohura T, Yoshino M, Tanaka A, Okamoto N, Inui K, Ozono K, Sakai N (2009) Mucolipidosis II and III alpha/beta: mutation analysis of 40 Japanese patients showed genotype-phenotype correlation. J Hum Genet 54:145-51. doi: 10.1038/jhg.2009.3.

Pohl S, Encarnacão M, Castrichini M, Müller-Loennies S, Muschol N, Braulke T (2010) Loss of N-acetylglucosamine-1-phosphotransferase gamma subunit due to intronic mutation in GNPTG causes mucolipidosis type III gamma: Implications for molecular and cellular diagnostics. Am J Med Genet A 152A:124-32. doi: 10.1002/ajmg.a.33170.

Pohl S, Tiede S, Castrichini M, Cantz M, Gieselmann V, Braulke T (2009) Compensatory expression of human N-acetylglucosaminyl-1-phosphotransferase subunits in mucolipidosis type III gamma. Biochim Biophys Acta 1792:221-5. doi: 10.1016/j.bbadis.2009.01.009.

Qian Y, Lee I, Lee WS, Qian M, Kudo M, Canfield WM, Lobel P, Kornfeld S (2010) Functions of the alpha, beta, and gamma subunits of UDP-GlcNAc:lysosomal enzyme N-acetylglucosamine-1-phosphotransferase. J Biol Chem 285:3360-70. doi: 10.1074/jbc.M109.068650.

Raas-Rothschild A, Cormier-Daire V, Bao M, Genin E, Salomon R, Brewer K, Zeigler M, Mandel H, Toth S, Roe B, Munnich A, Canfield WM (2000) Molecular basis of variant pseudo-hurler polydystrophy (mucolipidosis IIIC). J Clin Invest 105:673-81.

Reitman ML, Kornfeld S (1981a) UDP-N-acetylglucosamine:glycoprotein N-acetylglucosamine-1-phosphotransferase. Proposed enzyme for the phosphorylation of the high mannose oligosaccharide units of lysosomal enzymes. J Biol Chem

256:4275-81.
Reitman ML, Kornfeld S (1981b) Lysosomal enzyme targeting. N-Acetylglucosaminylphosphotransferase selectively phosphorylates native lysosomal enzymes. J Biol Chem 256:11977-80.
Reitman ML, Lang L, Kornfeld S (1984) UDP-N-acetylglucosamine: lysosomal enzyme N-acetylglucosamine-1-phosphotransferase. Methods Enzymol 107:163-72.
Sommerlade HJ, Selmer T, Ingendoh A, Gieselmann V, von Figura K, Neifer K, Schmidt B (1994) Glycosylation and phosphorylation of arylsulfatase A. J Biol Chem 269:20977-81.
Spranger JW, Wiedemann HR (1970) The genetic mucolipidoses. Diagnosis and differential diagnosis. Humangenetik 9:113-39.
Tappino B, Regis S, Corsolini F, Filocamo M (2008) An Alu insertion in compound heterozygosity with a microduplication in GNPTAB gene underlies Mucolipidosis II. Mol Genet Metab 93:129-33.
Tappino B, Chuzhanova NA, Regis S, Dardis A, Corsolini F, Stroppiano M, Tonoli E, Beccari T, Rosano C, Mucha J, Blanco M, Szlago M, Di Rocco M, Cooper DN, Filocamo M (2009) Molecular characterization of 22 novel UDP-N-acetylglucosamine-1-phosphate transferase alpha- and beta-subunit (GNPTAB) gene mutations causing mucolipidosis types IIalpha/beta and IIIalpha/beta in 46 patients. Hum Mutat 30:E956-73. doi: 10.1002/humu.21099.
Tiede S, Storch S, Lübke T, Henrissat B, Bargal R, Raas-Rothschild A, Braulke T (2005) Mucolipidosis II is caused by mutations in GNPTA encoding the alpha/beta GlcNAc-1-phosphotransferase. Nat Med 11:1109-12.
Umehara F, Matsumoto W, Kuriyama M, Sukegawa K, Gasa S, Osame M (1997) Mucolipidosis III (pseudo-Hurler polydystrophy); clinical studies in aged patients in one family. J Neurol Sci 146:167-72.
Varki A, Kornfeld S (1983) The spectrum of anionic oligosaccharides released by endo-beta-N-acetylglucosaminidase H from glycoproteins. Structural studies and interactions with the phosphomannosyl receptor. J Biol Chem 258:2808-18.
Vogel P, Payne BJ, Read R, Lee WS, Gelfman CM, Kornfeld S (2009) Comparative pathology of murine mucolipidosis types II and IIIC. Vet Pathol 46:313-24. doi: 10.1354/vp.46-2-313.
Waheed A, Hasilik A, von Figura K (1982) UDP-N-acetylglucosamine:lysosomal enzyme precursor N-acetylglucosamine-1-phosphotransferase. Partial purification and characterization of the rat liver Golgi enzyme. J Biol Chem 257:12322-31.
Zhao KW, Yeh R, Miller AL (1992) Purification and characterization of human lymphoblast N-acetylglucosamine-1-phosphotransferase. Glycobiology 2:119-25.

Keywords

GlcNAc-1-phosphotransferase
Uncovering enzyme
Site-1-protease
Lysosomal enzymes
Mannose 6-phosphate
GNPTAB
GNPTG
Mucolipidosis type II alpha/beta
Mucolipidosis type III alpha/beta
Mucolipidosis type III gamma
Lysosomal storage disorders

N-acetylglucosamine-1-phosphate transferase, gamma subunit; N-acetylglucosamine-1- (GNPTG, GNPTAB)

Dr. Maria Francisca Coutinho — Department of Genetics, Medical Genetics Center Jacinto Magalhaes, INSA, Porto, Portugal

DOI: 10.1007/SpringerReference_332158

URL: http://www.springerreference.com/index/chapterdbid/332158

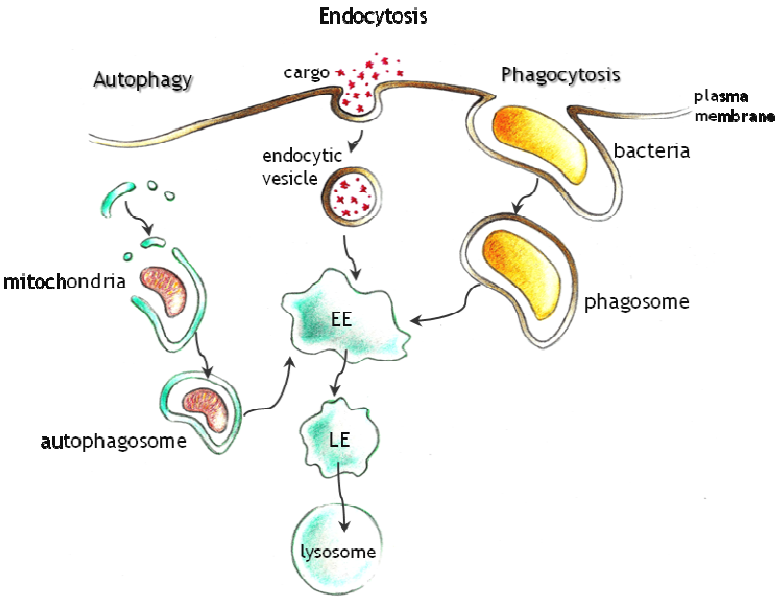
Part of: Handbook of Glycosyltransferases and Related Genes

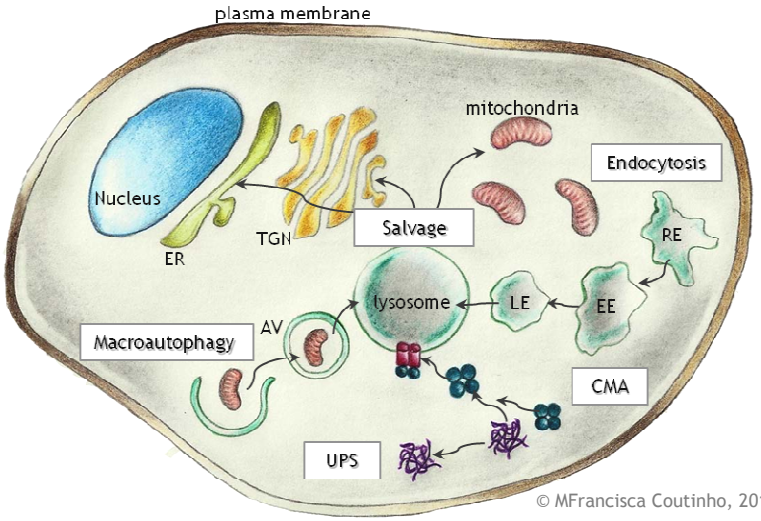
Editors: Professor Naoyuki Taniguchi, Professor Koichi Honke, Professor Minoru Fukuda, Professor Hisashi Narimatsu, Dr. Yoshiki Yamaguchi and Dr. Takashi Angata

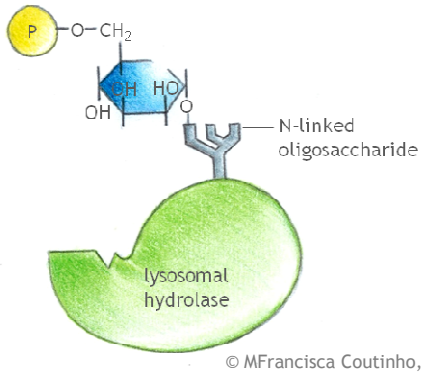
PDF created on: May, 26, 2013 16:20

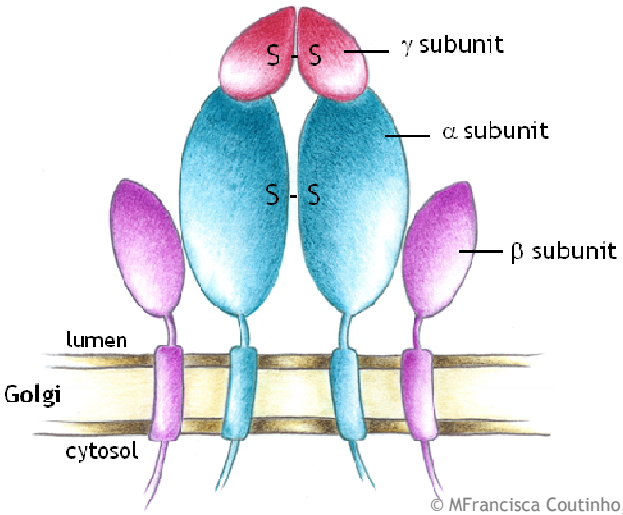
Appendix 3

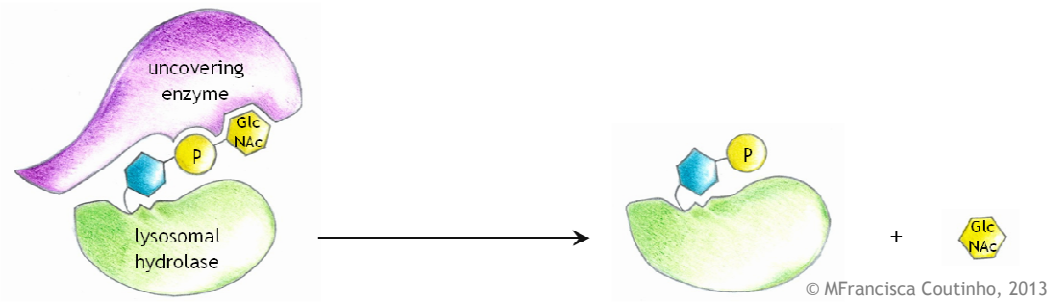
Scientific Illustration



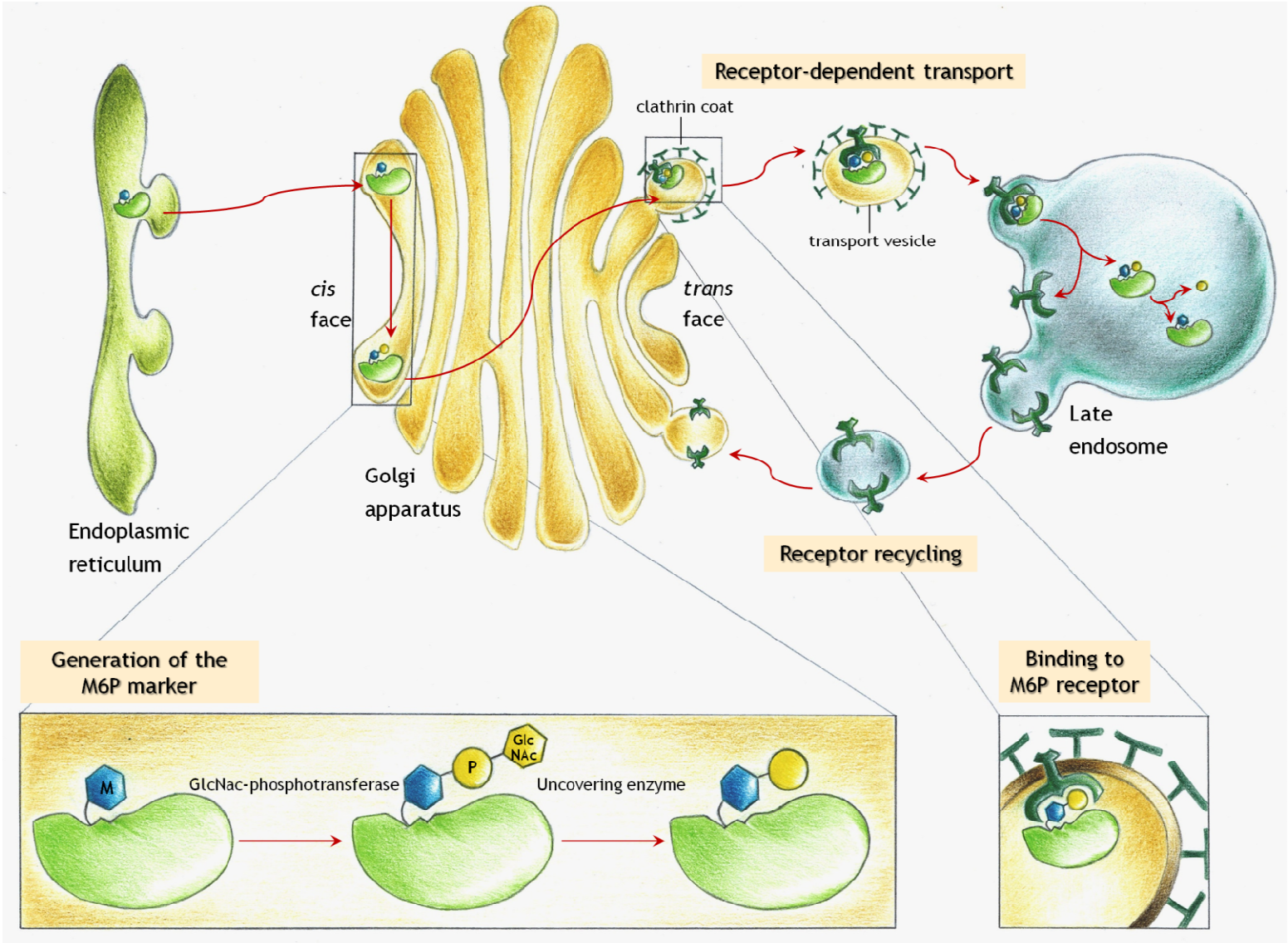


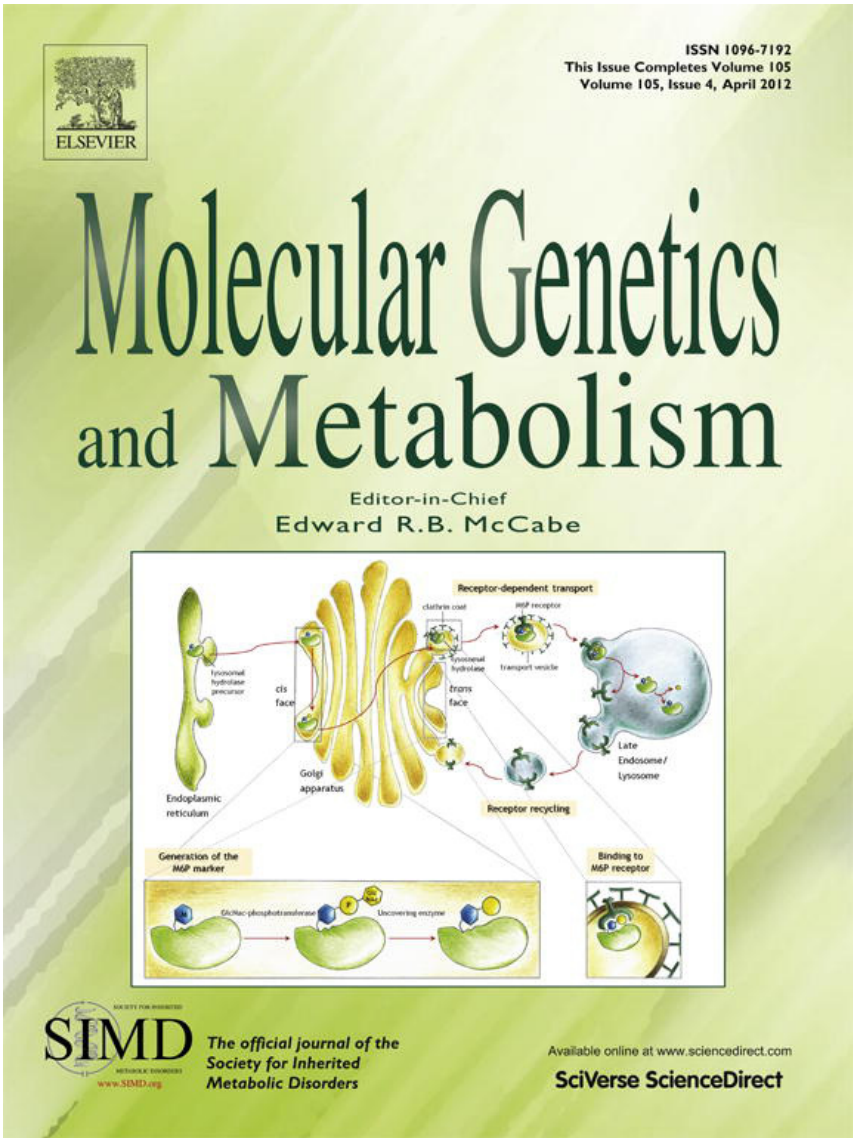






Panel 1





Panel 1 was chosen as the cover art for the April 2012 issue of *Molecular Genetics and Metabolism*, volume 105, issue 4.

Panel 2