

**CARBOHYDRATE-ACTIVE ENZYMES:  
FUNCTIONAL AND APPLICATIVE ASPECTS  
OF GLYCOSIDE HYDROLASES**

**Maria Carmina Ferrara**

**DOCTORAL THESIS**



# **Carbohydrate-Active enZymes: functional and applicative aspects of glycoside hydrolases**



**Dottorato di Ricerca in Biologia Applicata-XXV ciclo  
Indirizzo Microbiologia ed Ecologia  
Università di Napoli Federico II**

<b>Dottoranda:</b>	<b>Maria Carmina Ferrara</b> <i>Maria Carmina Ferrara</i>
<b>Supervisore:</b>	<b>Dr. Marco Moracci</b> <i>Marco Moracci</i>
<b>Coordinatore:</b>	<b>Prof. Ezio Ricca</b> <i>Ezio Ricca</i>

*The important thing in science is not so much to obtain new facts  
as to discover new ways of thinking about them.*

[Sir William Henry Bragg]

## TABLE OF CONTENTS

Keywords.....	1
List of abbreviations .....	2
<b>Chapter 1</b>	
<b>General Introduction.....</b>	<b>3</b>
1.1 Glycoside hydrolases: general features and classification .....	5
1.2 CAZy database.....	7
1.3 Glycoside hydrolases: catalytic mechanisms .....	10
1.4 Glycoside hydrolases: biological roles and applications .....	13
1.5 Post-genomic era: what's about function and structure? .....	16
1.6 Purpose of the thesis .....	19
1.7 References .....	21
<b>Chapter 2</b>	
<b>Identification and molecular characterization of a</b>	
<b><i>N</i>-acetyl-<math>\beta</math>-glucosaminidase from the crenarchaeon <i>S. solfataricus</i> .....</b>	<b>24</b>
2.1 General overview.....	25
2.2 <i>N</i> -acetyl- $\beta$ -hexosaminidases: structural and functional features .....	26
2.3 <i>N</i> -acetyl- $\beta$ -hexosaminidases: biological roles and applications.....	35
2.4 <i>Sulfolobus solfataricus</i> : a hyperthermophilic archaeon .....	39
2.5 Purpose of the work .....	46
2.6 Experimental procedures.....	48
2.7 Results and Discussion .....	59
2.8 References .....	87
<b>Chapter 3</b>	
<b>Pharmacological enhancement of acid alpha-glucosidase</b>	
<b>by the allosteric chaperone <i>N</i>-acetylcysteine .....</b>	<b>99</b>
3.1 General overview .....	100
3.2 acid alpha-glucosidase: structure and function.....	101
3.3 Lysosomal storage diseases .....	108
3.4 Pompe disease.....	113
3.5 Therapeutic approaches .....	117
3.5.1 Enzyme Replacement Therapy in Pompe disease.....	119
3.5.2 Pharmacological Chaperone Therapy.....	122

3.5.3 Other therapeutic approaches .....	128
3.6 Purpose of the work .....	131
3.7 Experimental procedures .....	133
3.8 Results and Discussion .....	137
3.9 References .....	161
List of Publications .....	172
Conferences .....	173
Acknowledgments .....	176

## KEYWORDS

Glycoside hydrolases

Archaea

*Sulfolobus solfataricus*

N-acetyl-glucosaminidase

Alpha-glucosidase

Lysosomal storage disease

Enzyme replacement therapy

Pharmacological chaperones

## LIST OF ABBREVIATIONS

2- or 4-NP	2- or 4-nitrophenyl
2-ME	2-mercaptoethanol
aa	amino acids
CAZy	Carbohydrate-Active enZYme
CBE	conduritol b-epoxide
Cys	cysteine
DNJ	1-deoxynojirimycin
ER	endoplasmic reticulum
ERT	enzyme replacement therapy
FCE	free cell extract
GAA	acid $\alpha$ -glucosidase
Gal	galactopyranoside
GalNAc	<i>N</i> -acetylgalactosaminide
GH	glycoside hydrolase
Glc	glucopyranoside
GlcNAc	<i>N</i> -acetylglucosaminide
Gly	glycine
His	histidine
LSD	lysosomal storage disease
M	molar
mM	millimolar
MPS	mucopolisaccharidosis
MU	4-methylumbelliferyl
NAC	<i>N</i> -acetylcysteine
NAG	<i>N</i> -acetyl glycine
NAS	<i>N</i> -acetylserine
NB-DNJ	<i>N</i> -Butyldeoxynojirimycin
ORF	open reading frame
PCR	polymerase chain reaction
PCs	pharmacological chaperones
PCT	pharmacological chaperone therapy
PD	Pompe disease
Ser	serine
v/v	volume/volume

# **Chapter 1**

## **GENERAL INTRODUCTION**

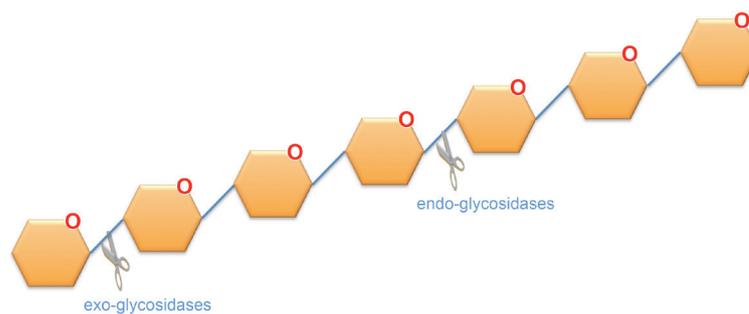
Carbohydrates constitute one of the four fundamental classes of biopolymers, along with nucleic acids, proteins, and lipids. In contrast to DNA, RNA, and proteins that are linear polymers containing only one basic type of linkage between monomers, carbohydrates can form branching structures and the multiple monosaccharide building blocks can be linked to various regio- and stereochemistries. Accordingly, a relatively simple set of sugars can form a huge number of complex structures. Moreover, the resulting oligosaccharides can be assembled on various scaffolds such as protein, nucleic acids, lipids, antibiotics, etc. generating the so-called glycoconjugates. The carbohydrates alone or linked to other biomolecules constitute an “information-rich” system capable of participating in a wide range of biological functions. These molecules can serve as intermediates in generating energy and as signalling effectors, recognition markers, and structural components. They constitute the main structural element of the plant cell walls such as cellulose, one of the most abundant organic compounds in the biosphere, and are the major constituents of the outer covering of insects and crustaceans, such as chitin, a polymer of *N*-acetylglucosamine (Khor, 2001). In addition, carbohydrates can constitute a nutritional reservoir, such as glycogen in animal cells and starch in plants. Glycans are also important in infections (Hooper and Gordon, 2001): bacteria, viruses and parasites may use lectins on their cell surface as a ‘key’ to open a sugar ‘lock’ on a host cell surface. For instance, influenza virus begins its infection cycle through the binding to sialic acid on host cell surface (Sauter et al., 1992).

Although glycans are found in all organisms, considerable diversity of their structure and expression exists in nature, both within and between evolutionary lineages (Varki, 2006). Sufficient data are available, however, to indicate that there is no universal “glycan code” akin to the genetic code. Unlike protein sequences, which are primary gene products, glycan chain structures are not encoded directly in the genome and are secondary gene products. The glycan chains represent numerous combinatorial possibilities, generated by a variety of competing and sequentially acting enzymes, glycosidases and glycosyltransferase.

# **1.1 GLYCOSIDE HYDROLASES: GENERAL FEATURES AND CLASSIFICATION**

## Glycoside hydrolases: general features and classification 1.1

Glycoside hydrolases (GHs), also known as glycosidases or glycosyl hydrolases, are enzymes that catalyze the hydrolysis of glycosidic bonds between two or more carbohydrate residues or a carbohydrate unit linked to an aglycon group. This class of enzymes is ubiquitous in nature, about 1% of the genome of any organism encodes for GHs (Vocadlo and Davies 2008), and is involved in several biological processes. Selective hydrolysis of glycosidic bonds is crucial for energy uptake, cell wall expansion and degradation, and turnover of signalling molecules. As a consequence of saccharides diversity, there is a great variety amongst GHs. The IUBMB (International Union of Biochemistry and Molecular Biology) nomenclature of GHs is based on their substrate specificity and molecular mechanism: the Enzyme Commission (EC) number corresponding to GHs is EC 3.2.1.x, where x represents substrate specificity. According to the two catalytic mechanisms proposed by Koshland in 1953, GHs are classified in *inverting* or *retaining* (see section 1.3 for details). They can be further classified on the basis of their action mode on long sugar chains: *exo*-glycosidases catalyze the hydrolysis of the glycosidic bond from the reducing or the non-reducing termini of carbohydrate polymers, while the *endo*-glycosidases act in the middle of a poly- or oligosaccharide (Fig. 1.1.1). Another classification of GHs is based on the amino acid sequence similarities among the various enzymes and is reported in CAZy database, which will be described in more details in the next section.



**Figure 1.1.1:** Mechanism of action of *exo*- and *endo*-glycosidases

## **1.2 CAZY DATABASE**

IUBMB enzyme classification does not reflect the structural and mechanistic features of GHs. To meet these needs, in the 1991 Henrissat proposed the classification of GHs based on protein sequence and structure similarities that correlates with enzyme mechanisms more than enzyme specificity (Bernard Henrissat 1991). CAZy database is founded on this type of classification and is available on the web since September 1998 at URL <http://www.cazy.org> (Cantarel et al., 2009). This resource includes not only GHs, but also other enzymes involved in the synthesis, degradation and modification of carbohydrates and glycoconjugates, collectively known as Carbohydrate-Active enZymes (CAZymes).

Each family is annotated with information regarding known enzyme activities and includes catalytic and structural features. At present, CAZy covers approximately 300 protein families in four classes of enzyme activities (Tab. 1.2.1).

Enzymatic Class	No of families
Glycosyde hydrolases (GHs)	132
Glycosyltransferases (GTs)	94
Polysaccharide lyases (PLs)	22
Carbohydrate esterases (CEs)	16

**Table 1.2.1:** CAZy families, updated to March 2013

Members of the same family have the same protein fold and closely related catalytic mechanism. Furthermore, the GH families with similar tertiary structure are grouped into clans which are thought to have evolved from a common ancestral gene; to date there are 14 clans (Henrissat and Bairoch, 1996). In addition to enzymes, CAZy database also includes Carbohydrate-binding modules (CBMs) that have no enzymatic activity *per se*, but potentiate the activity of many CAZymes by targeting to and promoting a prolonged interaction with the substrate.

The CAZy database is a convenient tool to derive mechanistic informations. It is regularly update through addition of new sequences, new families and further information related to the families. New sequences lacking biochemical and enzymatic information that cannot be included in an existing family or grouped in a new family are included in the GH0 family.

## 1.2 CAZy database

---

The principal structural and functional features of each CAZyme family can be retrieved by CAZypedia (<http://www.cazypedia.org>), a comprehensive encyclopedia of the CAZymes. In particular, each page of this resource cites primary literature which reports on the first identification of enzyme active site residues, catalytic mechanism, three dimensional structure, or other important works which define catalysis in a GH family or subfamily. This project is inspired by, and closely connected with, the CAZy database. All contributors to CAZypedia, from the Authors to the Board of Curators, are selected experts in the field.

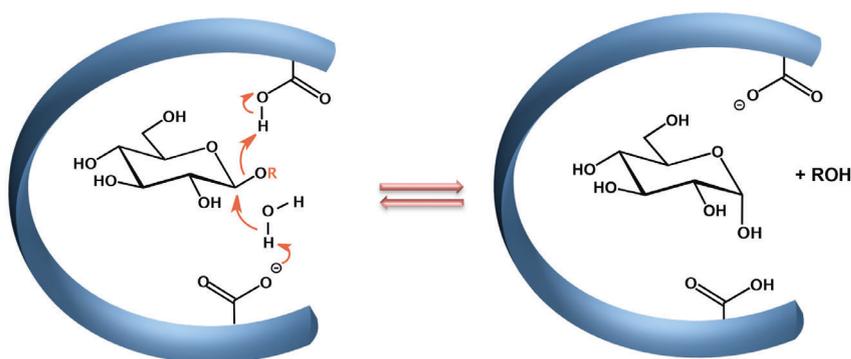
In definitive, glycoscience is a field in continuous development and ongoing efforts are directed towards structural and biochemical characterization of novel GHs in order to expand our understanding in this field.

## **1.3 GLYCOSIDE HYDROLASES: CATALYTIC MECHANISMS**

### 1.3 Glycoside hydrolases: catalytic mechanisms

Glycoside hydrolases carry out cleavage of glycosidic linkages in a stereospecific manner. According to the two catalytic mechanisms proposed by Koshland in 1953, GHs can be classified in *inverting* or *retaining*, depending on the stereochemistry at the anomeric center of the product generated.

The *inverting* enzymes hydrolyze the glycosidic bonds generating a product with anomeric carbon inverted with respect to that of substrate. The catalysis proceeds through a single-displacement mechanism (Fig.1.3.1).

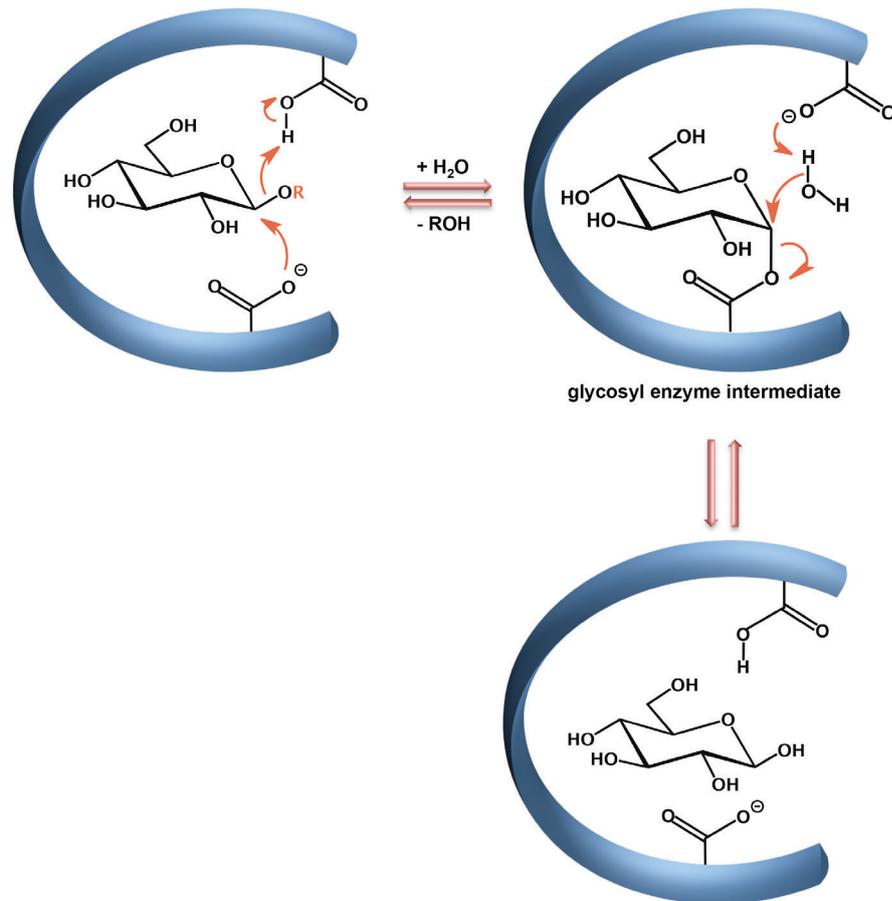


**Figure 1.3.1:** Catalytic mechanism of an *inverting*  $\beta$ -glycoside hydrolase

The catalytic machinery of these enzymes involves two catalytic carboxylates normally glutamic or aspartic acids: one functions as general acid, and the other as a general base that assists the nucleophilic attack by a water molecule.

The retaining mechanism leads to a net retention of the configuration at the anomeric carbon of the substrate after cleavage. This reaction proceeds via a double displacement mechanism with the assistance of two catalytic carboxylate residues and involves the formation of a covalent glycosyl-enzyme intermediate (Fig.1.3.2). In the first step (glycosylation), one of the two carboxylic groups functions as a general acid catalyst by protonating the glycosidic oxygen, simultaneously the second carboxylate acts as a nucleophile by attacking the anomeric carbon of the glycosidic bond. This step leads to the formation of glycosyl-enzyme intermediate. In the second step (deglycosylation), the first carboxylate residue now functions as a general base and activates an incoming water molecule by deprotonation. The water molecule acts as nucleophile and hydrolyses the glycosyl-enzyme intermediate resulting in the release of the sugar molecule.

## Glycoside hydrolases: catalytic mechanisms 1.3



**Figure 1.3.2:** Catalytic mechanism of a *retaining*  $\beta$ -glycoside hydrolase

## **1.4 GLYCOSIDE HYDROLASES: BIOLOGICAL ROLES AND APPLICATIONS**

## Glycoside hydrolases: biological roles and applications 1.4

Glycosidases play an important role in a wide range of biological process, from primary metabolism through glycoprotein glycan maturation. Therefore, it is not surprising that their function or dysfunction is implicated in a number of different diseases, leading to an interest in glycosidases as potential therapeutics *per se* or as drug targets. For instance, intravenously administration of the recombinant GHs is at the basis of the Enzyme Replacement Therapy (ERT) (for details see paragraph 3.5.1), a therapeutic approach aimed to restore the defective metabolic function in Lysosomal Storage Disorders (LSDs) (see paragraph 3.3 for a detailed description). In addition, several pharmaceutical treatments exploit the GHs as therapeutic targets modulating their synthesis and/or activity. For example, neuraminidase has been the target for the development of two influenza drugs (Relenza™, GlaxoSmithKline, and Tamiflu™, Roche). These compounds are synthetic structural analogues of neuraminic acid that bind specifically to the neuraminidase active site, and thus inhibit the transmission of the virus. The therapy for diabetes mellitus type 2 takes advantage of competitive inhibitors of  $\alpha$ -glucosidase (Glucobay™, Bayer, and Glyset™, Pfizer) to delay the digestion and absorption of carbohydrates in the small intestine and hence reduces the increase in blood-glucose concentrations after a carbohydrate load. Iminocyclitols (also known as azasugars or iminosugars) are particularly well-studied glycosidase inhibitors that have shown to be promising in the treatment of cancer (Goss et al., 1994; Sun et al., 2009), LSDs (Valenzano et al. 2011) and viral infections such as HIV (Gruters et al., 1987), and hepatitis B and C (Mehta et al., 1998).

The great deal of attention for GHs can be also attributed to their role in a variety of industrial and biotechnological applications. The current efforts are directed to optimize the role of GHs in the emerging field of renewable energies. The last 15 years have seen a massive growth of so-called first-generation processes that are based upon hydrolysis by amylases, pullulanases and glucoamylases of polysaccharides extracted from biomasses of crops like corn, wheat and sugar cane. However, these plants are also important components of the human food web, so using them for ethanol production has the potential to affect the price and availability of these basic commodities (Naik et al., 2010). By contrast, the production of second generation bioethanol exploits

## 1.4 Glycoside hydrolases: biological roles and applications

lignocellulosic materials, deriving from abundant and underutilized biological resources such as forestry crops (poplar, willow and eucalyptus), and perennial grasses (miscanthus, switch grass and reed canary grass). Cellulases and hemicellulases, such as xylanases,  $\beta$ -glucosidases,  $\beta$ -xylosidases, *endo*- and *exo*-glucanases, have been investigated as means to hydrolyze cellulose and hemicellulose in the lignocellulosic materials to fermentable reducing sugars (Sun and Cheng, 2002).

GHs are “friendly” catalysts also in other industrial process. For instance, xylanases constitute a major group of industrial enzymes with significant application in paper industry: the hydrolysis of xylan facilitates release of lignin from paper pulp and reduces the level of usage of chlorine as the bleaching agent (Shoham et al., 1992). The hydrolytic activity of the amylases is harnessed in the fermentation processes at the bases of baking and brewing industries or for production of high-fructose corn syrup. In the dairy industry,  $\beta$ -galactosidases are used to hydrolyze lactose from milk products to minimize effects on lactose-intolerant individuals (Perotti et al., 2012). Finally, GHs such as  $\beta$ -*N*-acetylgalactosidase and  $\beta$ -galactosidase are able to hydrolyze the A and B blood antigens respectively and so they are exploited for the production of the common H structure found in the universal O group (Liu et al., 2007).

These are only some examples that delineate the role of GHs in a variety of biological and industrial processes and unveil the importance to study the structure and function of this class of enzymes. Understand how these enzymes function and what is their functional structure is essential to increase our knowledge in this field and to expand their biological and industrial implications.

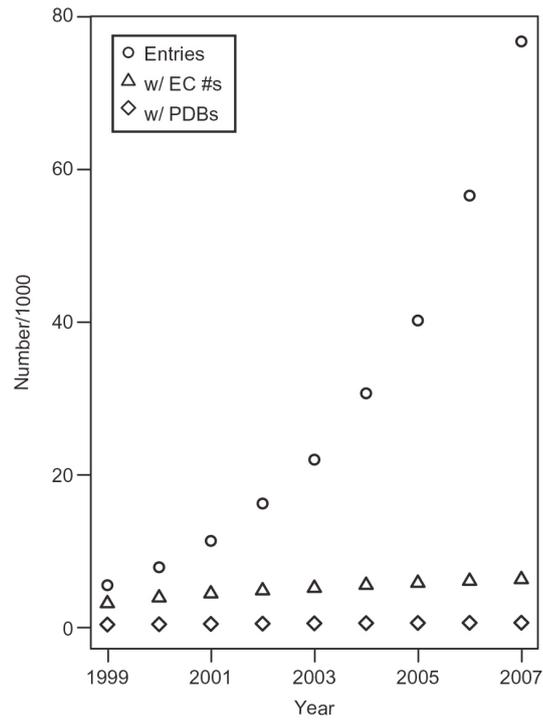
**1.5 POST-GENOMIC ERA:  
WHAT'S ABOUT  
FUNCTION AND STRUCTURE?**

## **1.5 Post-genomic era: what's about function and structure?**

Over the last decade, genome sequencing projects have provided a huge amount of data that need to be interpreted and integrated with functions. Many of the presently available annotations are inferred by bioinformatic analysis that predict the functions of newly sequenced genes on the basis of their similarity to genes or gene products of known or predicted function. This method displays at least two issues: 1) the small size of the genes set with experimentally established functions and 2) difficulties to define which level of identity or similarity is required to establish that two proteins have the same function or otherwise related functions. To date, a significant fraction of every genome sequenced, from bacterial to human, is composed of uncharacterized proteins. Additionally, one gene can code for different proteins as a result of post-transcriptional and post-translational processing and modification, so the number of proteins requiring characterization greatly exceeds the number of genes in the genome (Walsh 2006). Thus, genome sequences alone offer only limited biochemical and structural insights.

The post-genomic era shows great challenges and opportunities in the field of glycobiology. The content of genomes from a glycobiological perspective can be analyzed (e.g., by listing candidate GHs or GTs in a genome) and compared across genomes to see whether families have expanded or disappeared during evolution. The genomic sequences have unveiled a plethora of open reading frames (ORFs) likely to be involved in the synthesis, modification and degradation of sugars. The CAZy classification provides an efficient tool for the competent annotation of ORFs found during genome sequencing (e.g. prediction of the general function, fold and mechanism). However, progress on the biochemical, physiological and structural characterization of unknown ORFs are much more slower than sequencing. In CAZy database, the number of sequences has increased 14-fold in eight years, while the number of proteins with enzymatic and structural characterization has barely doubled (Fig. 1.5.1). Therefore, it is essential to understand how the sequences are related to structure, mechanism and specificity of the encoded enzyme. Future efforts should be directed to translate the data acquired by the genomic sequencing projects into the detailed structural and biochemical informations so as reducing the current gap between the genome sequencing and the proteomic characterization.

## Post-genomic era: what's about function and structure? 1.5



**Figure 1.5.1:** The number of CAZymes noted in December of the years 1999-2007 (Cantarel et al., 2009)

## **1.6 PURPOSE OF THE THESIS**

Carbohydrates play an important role in a variety of biological and industrial process. Elucidate the structure/function relationship of enzymes involved in their modification, CAZymes, can help us to understand the implication of this class of enzymes in various biological phenomena and how we can exploit their catalytic features, for example in pharmacological therapies or to optimize industrial process.

Starting from these observations my thesis work is dedicated to the study of GHs structure/function relationship as common denominator and can be subdivided in two principal sections, chapter 2 and 3, respectively. In chapter 2, the research plan is focused on the identification and biochemical characterization of novel glycosidases activities in the hyperthermophilic crenarchaeon *Sulfolobus solfataricus*, in particular *N*-acetylglucosaminidase activities. This study allows to increase our knowledge on the phylogeny of this class of enzymes and to achieve additional data concerning their structural and functional features. In chapter 3, the study of structure/function of lysosomal  $\alpha$ -glucosidase is aimed to understand what are the factors responsible of the enzyme stabilization and the gained informations are exploited to identify novel pharmacological chaperones for the therapy of the lysosomal storage disorder known as Pompe disease.

## **1.7 REFERENCES**

- Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic acids research* **37**:D233–238.
- Goss PE, Baptiste J, Fernandes B, Baker M, Dennis JW. 1994. A phase I study of swainsonine in patients with advanced malignancies. *Cancer research* **54**:1450–1457.
- Gruters RA, Neefjes JJ, Tersmette M, de Goede REY, Tulp A, Huisman HG, Miedema F, Ploegh HL. 1987. Interference with HIV-induced syncytium formation and viral infectivity by inhibitors of trimming glucosidase. *Nature* **330**:74–77.
- Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* **280**:309–316.
- Henrissat B, Bairoch A. 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal Letters* **316**:695–696.
- Hooper LV, Gordon JI. 2001. Glycans as legislators of host-microbial interactions: spanning the spectrum from symbiosis to pathogenicity. *Glycobiology* **11**:1R–10R.
- Khor E. 2001. Chitin: Fulfilling a Biomaterials Promise. Ed. Elsevier.
- Koshland DE. 1953. Stereochemistry and the mechanism of enzymatic reactions. *Biological Reviews* **28**:416–436.
- Liu QP, Sulzenbacher G, Yuan H, Bennett EP, Pietz G, Saunders K, Spence J, Nudelman E, Levery SB, White T, Neveu JM, Lane WS, Bourne Y, Olsson ML, Henrissat B, Clausen H. 2007. Bacterial glycosidases for the production of universal red blood cells. *Nature biotechnology* **25**:454–464.
- Mehta A, Zitzmann N, Rudd PM, Block TM, Dwek RA. 1998. Alpha-glucosidase inhibitors as potential broad based anti-viral agents. *FEBS letters* **430**:17–22.
- Naik SN, Goud VV., Rout PK, Dalai AK. 2010. Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews* **14**:578–597.
- Perotti MC, Veronica WI, Venica CI, Bergamini CV. 2012. Dairy Products Modified in their Lactose Content. *Current Nutrition & Food Science* **8**:8–18.
- Sauter NK, Hanson JE, Glick GD, Brown JH, Crowther RL, Park SJ, Skehel JJ, Wiley DC. 1992. Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and x-ray crystallography. *Biochemistry* **31**:9609–9621.

## 1.7 References

---

- Shoham Y, Schwartz Z, Khasin A, Gat O, Zosim Z, Rosenberg E. 1992. Delignification of wood pulp by a thermostable xylanase from *Bacillus stearothermophilus* strain T-6. *Biodegradation* **3**:207–218.
- Sun Y, Cheng J. 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology* **83**:1–11.
- Sun JY, Yang H, Miao S, Li JP, Wang SW, Zhu MZ, Xie YH, Wang JB, Liu Z, Yang Q. 2009. Suppressive effects of swainsonine on C6 glioma cell in vitro and in vivo. *Phytomedicine: international journal of phytotherapy and phytopharmacology* **16**:1070–1074.
- Valenzano KJ, Khanna R, Powe AC, Boyd R, Lee G, Flanagan JJ, Benjamin ER. 2011. Identification and characterization of pharmacological chaperones to correct enzyme deficiencies in lysosomal storage disorders. *Assay and drug development technologies* **9**:213–235.
- Varki A. 2006. Nothing in glycobiology makes sense, except in the light of evolution. *Cell* **126**:841–845.
- Vocadlo DJ, Davies GJ. 2008. Mechanistic insights into glycosidase chemistry. *Current opinion in chemical biology* **12**:539–555.
- Walsh C. 2006. Posttranslational Modification of Proteins. Ed. Roberts and Company Publishers.

## Chapter 2

Identification and molecular  
characterization of a  
*N*-acetyl- $\beta$ -glucosaminidase from  
the crenarchaeon *S. solfataricus*

## 2.1 General overview

---

Chapter 2 of this thesis work describes the identification and molecular characterization of a *N*-acetyl- $\beta$ -glucosaminidase from the crenarchaeon *S. solfataricus*. This section begins with a brief introduction on the structural/functional features and on the biological relevance of this class of enzymes (paragraphs 2.2 and 2.3). The chapter keeps on with the description of the model organism adopted in my thesis work, *S. solfataricus*, in which no such activity has been described so far (paragraph 2.4). The purpose of this work is shown in paragraph 2.5, while the experimental part is described in the last two paragraphs; in particular paragraph 2.6 reports on experimental procedures and paragraph 2.7 shows results and discussion. At last, the paragraph 2.8 reports the literature references included in this chapter.

## **2.2 *N*-ACETYL- $\beta$ -HEXOSAMINIDASES: STRUCTURAL AND FUNCTIONAL FEATURES**

## 2.2 *N*-acetyl- $\beta$ -hexosaminidases: structural and functional features

---

*N*-Acetyl- $\beta$ -hexosaminidases ( $\beta$ -HexNAcase, systematic name 2-acetamido-2-deoxy- $\beta$ -D-hexopyranoside acetamidodeoxyhexohydrolase, EC 3.2.1.52) are *exo*-glycosidases catalysing the removal of  $\beta$ -D-*N*-acetylglucosamine (GlcNAc) or  $\beta$ -D-*N*-acetylgalactosamine (GalNAc) from the non-reducing ends of a variety of *N*-acetyl- $\beta$ -D-hexosaminides. The activity on both gluco- and galacto-configurations is a typical feature of this class of enzymes (Muramatsu, 1968; Slámová et al., 2010). Generally, D-gluco-structures are preferred and the GlcNAcase/GalNAcase activity ratio is commonly in the range of 1.5-4.0 (Horsch et al., 1997). Nevertheless, in some limited cases, the GlcNAcase/GalNAcase ratio is much higher or inverted. For instance, the *Xenopus laevis* enzyme showed a GlcNAcase activity 167-fold higher than GalNAcase (Greve et al., 1985) or *N*-acetyl- $\beta$ -hexosaminidase from *Penicillium oxalicum* clearly preferred the axial hydroxyl at C-4 position, with a GalNAcase activity 2.5-fold higher than GlcNAcase (Weignerová et al., 2003). According to the CAZy database (<http://www.cazy.org>), *N*-acetyl- $\beta$ -hexosaminidases, as defined above, are classified into three families: GH3, GH20 and GH84. GH3 and GH84  $\beta$ -HexNAcases recognize only GlcNAc derivatives and thus are usually named  $\beta$ -*N*-acetyl-D-glucosaminidases. Since most of the GH20 enzymes recognize both GlcNAc and GalNAc, they are usually named  $\beta$ -*N*-acetyl-D-hexosaminidases (Liu et al., 2012). Although the enzymes grouped in these three distinct families possess no sequence homology, they share some basic structural and catalytic features.

### GH3 *N*-acetyl- $\beta$ -hexosaminidases

The CAZy family GH3 includes *N*-acetyl- $\beta$ -hexosaminidases (EC 3.2.1.52),  $\beta$ -glucosidases (EC 3.2.1.21), and, among other enzymatic activities, xylan 1,4- $\beta$ -xylosidase, EC 3.2.1.37; glucan 1,3- $\beta$ -glucosidase, EC 3.2.1.58; glucan 1,4- $\beta$ -glucosidase, EC 3.2.1.74; *exo*-1,3-1,4-glucanase, EC 3.2.1.-; and  $\alpha$ -L-arabinofuranosidase, EC 3.2.1.55 (Fig.2.2.1). GH3  $\beta$ -GlcNAcases constitute a small group of bacterial enzymes, also known as NagZ; no eukaryotic  $\beta$ -HexNAcases have been described in this family so far.

## N-acetyl- $\beta$ -hexosaminidases: structural and functional features 2.2

<b>Known Activities</b>	b-glucosidase (EC 3.2.1.21); xylan 1,4-b-xylosidase (EC 3.2.1.37); b-N-acetylhexosaminidase (EC 3.2.1.52); glucan 1,3-b-glucosidase (EC 3.2.1.58); glucan 1,4-b-glucosidase (EC 3.2.1.74); exo-1,3-1,4-glucanase (EC 3.2.1.-); $\alpha$ -L-arabinofuranosidase (EC 3.2.1.55).
<b>Mechanism</b>	Retaining
<b>Clan</b>	none
<b>Catalytic Nucleophile/Base</b>	Asp (experimental)
<b>Catalytic Proton Donor</b>	Glu (experimental)
<b>External resources</b>	<a href="#">CAZypedia</a> ; <a href="#">HOMSTRAD</a> ; <a href="#">PRINTS</a> ; <a href="#">PROSITE</a> ;
<b>Commercial Enzyme Provider(s)</b>	<a href="#">MEGAZYME</a> ; <a href="#">PROZOMIX</a> ;
<b>Statistics</b>	GenBank accession (4986); Uniprot accession (2556); PDB accession (44); 3D entries (16); cryst (5)
<b>Summary</b>	<a href="#">All</a> (4365)   <a href="#">Archaea</a> (54)   <a href="#">Bacteria</a> (3696)   <a href="#">Eukaryota</a> (576)   <a href="#">unclassified</a> (39)   <a href="#">Structure</a> (16 - 5 cryst)   <a href="#">Characterized</a> (219)

**Figure 2.2.1:** Principal structural and functional features of CAZy family GH3

Three dimensional structures have been solved for NagZ from *Vibrio cholerae* (Stubbs et al., 2007), *Bacillus subtilis* (Litzinger et al., 2010), and *Salmonella typhimurium* (Bacik et al., 2012). These enzymes are single domain proteins that adopt a TIM barrel fold with the strands of the barrel forming several loops at the C-terminal end delimiting the active centre (Fig. 2.2.2).

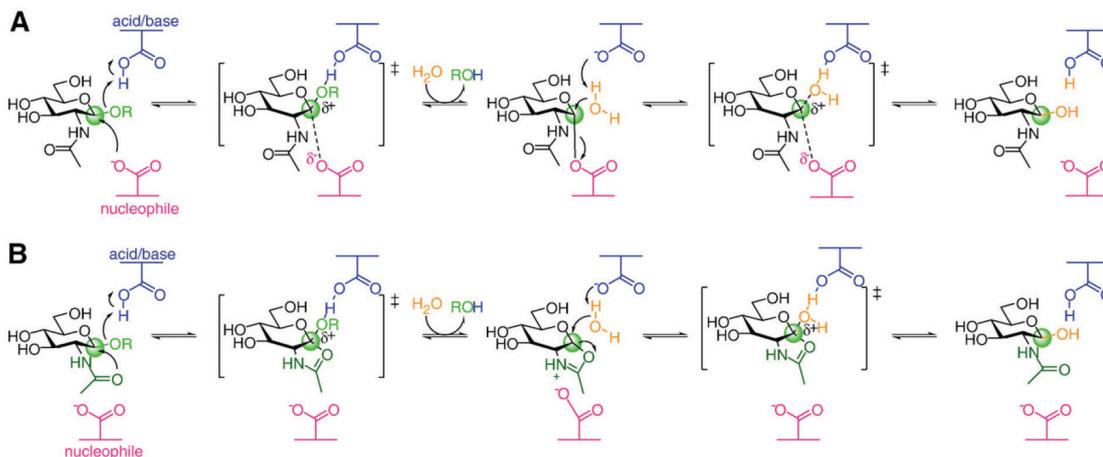


**Figure 2.2.2:** 3D structure of GH3 NagZ from *Vibrio cholerae* in complex with the inhibitor PUGNAc (PDB ID 20XN)

The peculiar architecture of the GH3 active site promotes catalysis via the classical double-displacement mechanism of *retaining* glycosidases which involves a covalent  $\alpha$ -glycosyl-enzyme intermediate (Fig. 2.2.3 A) (see paragraph 1.3 for details) (Vocadlo et al., 2000; Vocadlo and Withers, 2005). The substrate binding pocket is larger in the area accommodating N-acetyl moiety of the substrate and this extra space is occupied by a molecule of water. The catalytic nucleophile for NagZ *Vibrio furnisii* has been identified experimentally as Asp242, which is conserved throughout family GH3 (Vocadlo

## 2.2 *N*-acetyl- $\beta$ -hexosaminidases: structural and functional features

et al., 2000). A general acid residue has not been identified and a conserved histidine has been proposed as the general acid/base (Litzinger et al., 2010).



**Figure 2.2.3:** Double-displacement retaining mechanism for  $\beta$ -*N*-acetylhexosaminidases; **A)** classical mechanism utilised by GH3; **B)** substrate-assisted mechanism of GH 20 and 84

An unusual *N*-acetyl- $\beta$ -hexosaminidase was also reported, Nag3 from *Cellulomonas fimi*. This enzyme showed a broad substrate specificity, hydrolysing both  $\beta$ -*N*-acetylglucosaminides and  $\beta$ -glucosides (Mayer et al., 2006).

### **GH20 *N*-acetyl- $\beta$ -hexosaminidases**

Family 20 contains the highest number of *N*-acetyl- $\beta$ -hexosaminidases in CAZy classification and is also the most studied, mainly for the presence of the human lysosomal  $\beta$ -HexNAcases, HeXA and HexB, whose deficiency leads to GM2 gangliosidosis (see the paragraph 2.3 for more details) (Fig. 2.2.4).

## N-acetyl- $\beta$ -hexosaminidases: structural and functional features 2.2

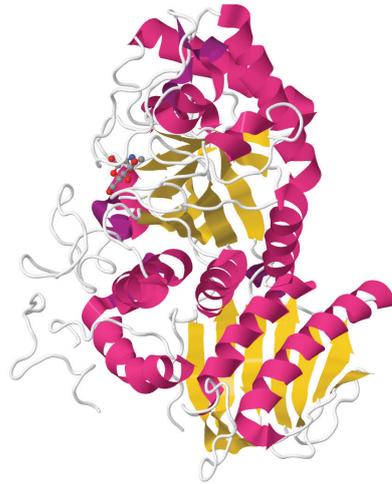
<b>Known Activities</b>	b-hexosaminidase (EC 3.2.1.52); lacto-N-biosidase (EC 3.2.1.140); b-1,6-N-acetylglucosaminidase (EC 3.2.1.-); b-6-SO3-N-acetylglucosaminidase (EC 3.2.1.-)
<b>Mechanism</b>	Retaining
<b>Clan</b>	GH-K
<b>3D Structure Status</b>	(b/a)8
<b>Catalytic Nucleophile/Base</b>	carbonyl oxygen of C-2 acetamido group of substrate
<b>Catalytic Proton Donor</b>	Glu
<b>External resources</b>	<a href="#">CAZypedia</a> ; <a href="#">HOMSTRAD</a> ; <a href="#">PRINTS</a> ; <a href="#">Tay-Sachs Disease</a> ;
<b>Statistics</b>	GenBank accession (1522); Uniprot accession (776); PDB accession (48); 3D entries (14); cryst (1)
<b>Summary</b>	<a href="#">All (1257)</a>   <a href="#">Archaea (4)</a>   <a href="#">Bacteria (1039)</a>   <a href="#">Eukaryota (214)</a>   <a href="#">Structure (14 - 1 cryst)</a>   <a href="#">Characterized (112)</a>

**Figure 2.2.4:** Principal structural and functional features of CAZy family GH20

The catalytic mechanism of GH20  $\beta$ -HexNAcases is a modification of the classical double-displacement process used by *retaining* glycosidases (Mark and James 2002) and is known as substrate-assisted mechanism (Fig. 2.2.3 B). In particular, such a catalysis, also called anchimeric assistance, involves, rather than an enzyme residue as a nucleophile, the 2-acetamido substrate moiety to form an oxazolinium ion intermediate (Mark et al., 2001; Tews et al., 1997). The N-acetyl group of the substrate is bound in a hydrophobic pocket defined by three conserved tryptophan residues. A detailed bioinformatic study of some GH20 hexosaminidases revealed a conserved sequence motif His/Asn-Xaa-Gly-Ala/Cys/Gly/Met-Asp-Glu-Ala/Ile/Leu/Val surrounding the catalytic residues (Gutternigg et al., 2007). The active site contains a highly conserved pair of catalytic residues Asp-Glu. The glutamate acts as a proton donor, while the aspartate, not functioning as nucleophile, could stabilize the oxazoline transition state and assist the correct orientation of the 2-acetamido moiety during catalysis (Prag et al., 2000; Williams et al., 2002). The crystal structures of several bacterial N-acetyl- $\beta$ -hexosaminidases have been solved, such as Hex1T from *Paenibacillus* sp. strain TS12 (Sumida et al., 2009), chitobiase from *Serratia marcescens* (Tews et al., 1996), StrH from *Streptococcus pneumoniae* R6 (Jiang et al., 2011) (Fig. 2.2.4). The eukaryotic kingdom is represented by the structures of human lysosomal hexosaminidases, HexA and HexB (Lemieux et al., 2006; Maier et al., 2003; Mark et al., 2003).

## 2.2 *N*-acetyl- $\beta$ -hexosaminidases: structural and functional features

---



**Figure 2.2.4:** 3D structure of GH20 *N*-acetyl- $\beta$ -hexosaminidase from *Paenibacillus* sp. TS12 in complex with GlcNAc (PDB ID 3GH5)

The common features of all known structures are the  $(\beta/\alpha)_8$ -barrel architecture of the catalytic domain and the location of the active site at the C-terminal end of the third of the four protein domains. Bacterial enzymes are monomeric, while the human  $\beta$ -HexNAcases form active dimers with two catalytic sites. Moreover, the subunits of human HexA and HexB are post-translationally cleaved, yielding the mature enzyme.

### **GH84 *N*-acetyl- $\beta$ -glucosaminidases**

The *N*-acetyl- $\beta$ -glucosaminidases belonging to family GH84 have recently gained great attention, primarily for the presence of human O-GlcNAcase (or HexC) which is involved in Alzheimer's disease (Yuzwa et al., 2008) (Fig. 2.2.5).

## N-acetyl- $\beta$ -hexosaminidases: structural and functional features 2.2

<b>Known Activities</b>	N-acetyl b-glucosaminidase (EC 3.2.1.52); hyaluronidase (EC 3.2.1.35); [protein]-3-O-(GlcNAc)-L-Ser/Thr b-N-acetylglucosaminidase (EC 3.2.1.169)
<b>Mechanism</b>	Retaining
<b>3D Structure Status</b>	(b/a) <sub>8</sub>
<b>Catalytic Nucleophile/Base</b>	carbonyl oxygen of C-2 acetamido group of substrate
<b>Catalytic Proton Donor</b>	Asp
<b>Note</b>	Mechanism shown in Macauley et al. (2005) JBC 280:25313-25322
<b>External resources</b>	<a href="#">CAZypedia</a> ;
<b>Commercial Enzyme Provider(s)</b>	<a href="#">PROZOMIX</a> ;
<b>Statistics</b>	GenBank accession (211); Uniprot accession (105); PDB accession (33); 3D entries (4); cryst (0)
<b>Summary</b>	<a href="#">All (161)</a>   <a href="#">Archaea (2)</a>   <a href="#">Bacteria (136)</a>   <a href="#">Eukaryota (22)</a>   <a href="#">unclassified (1)</a>   <a href="#">Structure (4)</a>   <a href="#">Characterized (10)</a>

**Figure 2.2.5:** Principal structural and functional features of CAZy family GH84

Although GH20 and GH84 enzymes differ in their amino acid sequence, they share the substrate-assisted catalytic mechanism (Fig. 2.2.3 B) and the  $(\beta/\alpha)_8$ -barrel architecture of the catalytic domain (Macauley et al., 2005). The catalytic residues are two aspartates (vs. Asp and Glu in GH20), actively involved in catalysis, while a tyrosine stabilizes the transition state (Dennis et al., 2006; Rao et al., 2006). Two bacterial enzymes of this family have been successfully crystallized and their structures resolved: the  $\beta$ -HexNAcases from the pathogen *Clostridium perfringens* (Fig. 2.2.6) and from the human gut symbiont *Bacteroides thetaiotaomicron* (Dennis et al., 2006). These structures serve as good models of both the active site and catalytic domain of human O-GlcNAcase.



**Figure 2.2.6:** 3D structure of GH84 N-acetyl- $\beta$ -hexosaminidase from *Clostridium perfringens* in complex with the inhibitor PUGNAc (PDB ID 2CBJ) (Rao et al., 2006)

## 2.2 *N*-acetyl- $\beta$ -hexosaminidases: structural and functional features

---

### *N*-acetyl- $\beta$ -glucosaminidases inhibitors

The inhibition of  $\beta$ -HexNAcases is useful both for elucidating the role of these enzymes in biological processes as well as for developing therapeutic tools. The principal issue in the generation of  $\beta$ -HexNAcases inhibitors relies on the design of highly selective molecules for the various forms of human enzymes. Promiscuous inhibitors acting on both lysosomal (HexA and HexB) and nuclear (O-GlcNAcase)  $\beta$ -HexNAcases generate complex phenotypes and are poorly useful in studying the physiological roles of individual enzymes.

In the past few years, two main inhibitor platforms have been studied: NAG-thiazoline (2'-methyl- $\alpha$ -D-glucopyrano-[2,1- $\sigma$ ]- $\Delta$ 2'-thiazoline) and PUGNAc (O-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino *N*-phenylcarbamate). The structural scaffolds of these compounds mimic the transition state of the substrate-assisted reaction mechanism shared by GH20 and 84 enzymes.

NAG-thiazolines are highly potent inhibitors of human HexA and HexB (GH20) as well as O-GlcNAcase (GH84). The active sites of GH84 enzymes tolerate moderately bulky substituents attached to the substrate acetamido group compared to those of GH20 enzymes. Exploiting these structural differences, new inhibitors highly selective for human O-GlcNAcase were developed. Indeed, the inhibition of HexA and HexB rapidly decreases with increasing *N*-acyl chain length of the inhibitor, while O-GlcNAcase tolerates longer chains. On these bases, functionalized NAG-thiazolines have been prepared, such as aliphatic, azide and fluoride derivatives, which showed outstanding selectivity for O-GlcNAcase. For instance, an aliphatic chain linked to the thiazoline ring allowed to increase 3100-fold the selectivity for O-GlcNAcase over the lysosomal forms (Macauley et al., 2005).

PUGNAc, a molecule that is thought to mimic the geometry of the dissociative transition states of most glycosidases, is another poorly selective inhibitor of  $\beta$ -GlcNAcases (Macauley et al., 2005). As in the case of NAG-thiazoline, the problem of the selectivity was addressed by the design and preparation of PUGNAc derivatives. Exploiting the ability of GH20 enzymes to accommodate substrates with the D-galacto-configuration, it has been possible to develop highly selective inhibitors of lysosomal *N*-acetyl- $\beta$ -hexosaminidases, such as Gal-PUGNAc, which is inactive on O-GlcNAcase (Stubbs et al., 2009).

## ***N*-acetyl- $\beta$ -hexosaminidases: structural and functional features 2.2**

---

However, the overall inhibition and selectivity of PUGNAc-derived compounds are somewhat poorer than NAG-thiazoline derivatives (Stubbs et al., 2006). Moreover, some PUGNAc derivatives, such as *N*-valeryl-PUGNAc and *N*-butyryl-PUGNAc, act as inhibitors also of NagZ (GH3) (Balcewich et al., 2009; Stubbs et al., 2007).

## **2.3 *N*-ACETYL- $\beta$ -HEXOSAMINIDASES: BIOLOGICAL ROLES AND APPLICATIONS**

## ***N*-acetyl- $\beta$ -hexosaminidases: biological roles and applications 2.3**

---

*N*-acetyl- $\beta$ -hexosaminidases are commonly involved in the degradation of polysaccharides and glycoconjugates containing *N*-acetylhexosamine residues. They are universally distributed among most types of organisms, both prokaryotes and eukaryotes, but not archaea, in which no such activity has been described so far. These enzymes show a very broad range of functions depending on the organism source and on the cellular localization. For instance, bacterial  $\beta$ -HexNAcases play an important role in recycling the cell wall. *N*-acetyl- $\beta$ -glucosaminidases from Gram-negative bacteria (NagZ, GH3), participate in the peptidoglycan recycling pathway hydrolysing the  $\beta$ -1,4-glycosidic bond between *N*-acetylglucosamine and *N*-acetylmuramic acid (Cheng et al., 2000). The resulting 1,6-anhydro-*N*-acetylmuramic acid peptides can induce the transcription of AmpC  $\beta$ -lactamase, and consequentially extend the resistance of Gram-negative bacteria to a broad spectrum of  $\beta$ -lactam antibiotics. The selective inhibition of NagZ represents an approach intensively investigated to attenuate/suppress the resistance to many  $\beta$ -lactam antibiotics (Stubbs et al., 2007).

Insect cells usually express several *N*-acetyl- $\beta$ -hexosaminidases, which are important for the turnover of the chitin exoskeleton (Hogenkamp et al., 2008), but also in N-glycan trimming producing paucimannosidic-type N-glycans that are different from the glycans produced by higher eukaryotes (Altmann et al., 1995; Gutternigg et al., 2007). It has been demonstrated that the inactivation of this class of enzymes in insect expression cell lines (Sf9, *Drosophila* S2 cells) could represent a useful tool to improve the production of eukaryotic glycoproteins in recombinant form via baculovirus expression systems (Geisler et al., 2008; Kim et al., 2009; Léonard et al., 2006).

The physiological role of  $\beta$ -HexNAcases is mainly studied in mammals, especially in humans, where a deficiency of these enzymes causes severe disorders. Mammalian cells express three distinct types of *N*-acetyl- $\beta$ -hexosaminidases:

1. HexA, HexB and HexS (GH20)
2. HexC or O-GlcNAcase (GH84)
3. HexD (GH20)

### 2.3 *N*-acetyl- $\beta$ -hexosaminidases: biological roles and applications

---

These enzymes differ in their physiological roles and structures.

HexD is a nucleocytoplasmic *N*-acetyl- $\beta$ -hexosaminidase discovered recently, which prefers *N*-acetylgalactosaminides as substrates, but its biological significance remains unclear (Gutternigg et al., 2009).

O-GlcNAcase (OGA) is another nucleocytoplasmic *N*-acetyl- $\beta$ -hexosaminidase, with a strict specificity to *N*-acetylglucosamine (GlcNAc) (Comtesse et al., 2001). It catalyzes the removal of GlcNAc moieties from post-translationally modified serine and threonine residues of nucleocytoplasmic proteins (Gao et al., 2001). This post-translational modification is abundant in mammalian cells (Torres and Hart, 1984) and is found on many cellular proteins involved in a wide range of vital functions, including transcription (Lamarre-Vincent and Hsieh-Wilson, 2003), proteasomal degradation (Zhang et al., 2003), and cellular signaling (Vosseller et al., 2002), and is also found on many structural proteins (Cole and Hart, 1999). This form of glycosylation has been implicated in various pathological states including cancer, type II diabetes, and Alzheimer's disease (Hart et al., 2007).

Lysosomal  $\beta$ -HexNAcases belong to the GH20 family and are composed of two subunits,  $\alpha$  and  $\beta$ , which are approximately 60% identical in their amino acid sequence. The two subunits are synthesized as precursor proteins and subsequently processed and assembled in a dimer form in the endoplasmic reticulum. The dimerization is essential for the enzymatic activity and yields three isoforms: HexA ( $\alpha\beta$ ), HexB ( $\beta\beta$ ) and HexS ( $\alpha\alpha$ ). Each subunit possesses an active site characterized by its own substrate specificity (Kytzia and Sandhoff, 1985). The  $\beta$ -subunit predominantly cleaves uncharged substrates like glycolipid GA2 and oligosaccharides with terminal *N*-acetylhexosamine residues, whereas the  $\alpha$ -subunit cleaves also negatively charged substrates (e.g., 6-sulphated GlcNAc) (Hou et al., 1996). Only the  $\alpha\beta$ -heterodimer HexA is able to degrade ganglioside GM2 at significant rates in the presence of the GM2 activator protein (GM2AP), which solubilizes the gangliosides for presentation to HexA (Hou et al., 1996). Mutations in the *HEXA* and *HEXB* genes lead to Tay-Sachs and Sandhoff disease, respectively. These LSDs are typically accompanied by increased concentration of the GM2 ganglioside in

## ***N*-acetyl- $\beta$ -hexosaminidases: biological roles and applications 2.3**

---

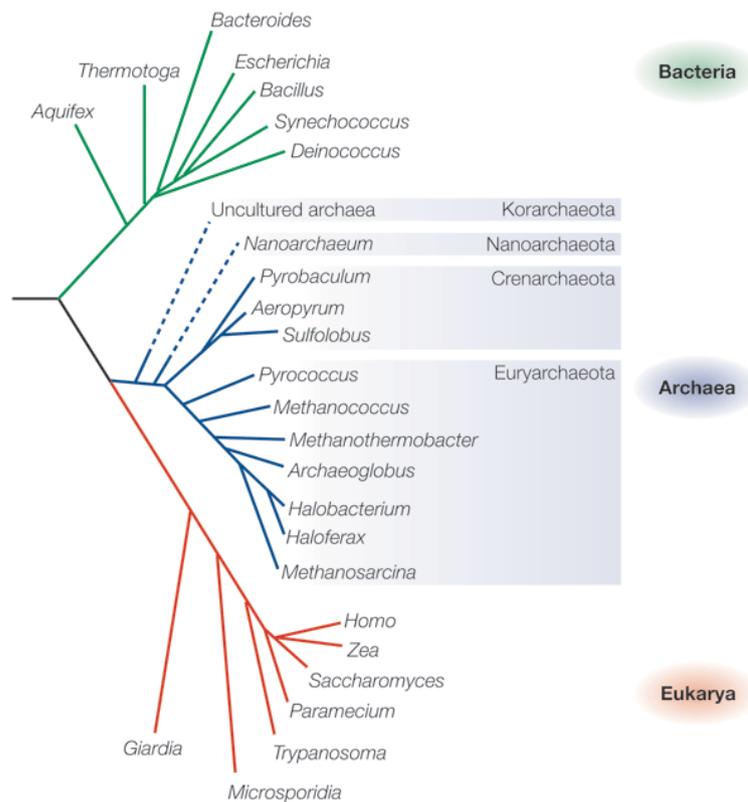
neuronal lysosomes, which gradually results in serious, or even deadly, neurodegeneration (Mahuran 1999). Several therapeutic strategies are currently being developed, such as direct gene transfer via brain capsules (Martino et al., 2005), pharmacological chaperone therapy based on the use of small molecules, primarily inhibitors, that enhance HexA activity (Tropak et al., 2007; Tropak and Mahuran, 2007) or enzyme replacement therapy with recombinant human HexA produced in yeast cells (Akeboshi et al., 2009).

## **2.4 *SULFOLOBUS SOLFATARICUS*: A HYPERTHERMOPHILIC ARCHAEON**

## ***Sulfolobus solfataricus*: a hyperthermophilic archaeon 2.4**

### **Archaea**

In the 1970s the pioneering work of Carl Woese, classifying all life forms upon sequence data from the small subunit ribosomal RNA (SSU rRNA), led to the surprising discovery of the Archaea (Woese and Fox, 1977). This domain represents a third line of evolutionary descent, distinct from Eukarya, and a second prokaryotic lineage distinct from the Bacteria (Fig. 2.4.1).



**Figure 2.4.1:** Universal phylogenetic tree based on SSU rRNA sequences (Allers and Mevarech, 2005)

Early studies on molecular characteristics and mechanisms of Archaea revealed their distinctive nature. They harbour unique traits radically distinguishing them from the other two domains, including the incorporation of ether-linked isoprenoid lipids in the cytoplasmic membrane, and methanogenesis to produce biological methane. In addition, they display bacterial-like traits, for example metabolism, and genomic complexity and organization, and eukaryal-like traits, such as DNA replication, transcription and translation (Tab. 2.4.1).

## 2.4 *Sulfolobus solfataricus*: a hyperthermophilic archaeon

Trait	Archaea	Bacteria	Eukarya
Carbon linkage of lipids	Ether	Ester	Ester
Phosphate backbone of lipids	Glycerol-1-phosphate	Glycerol-3-phosphate	Glycerol-3-phosphate
Nucleus and organelles	No	No	Yes
Methanogenesis	Yes	No	No
DNA topology	Relaxed or positively supercoiled	Negatively supercoiled	Negatively supercoiled
DNA replication, transcription and translation	Eukarial-like	Bacterial	Eukarial

**Table 2.4.1:** Comparison of Archaea, Bacteria and Eukarya, adapted from Cavicchioli 2011

Initial studies seemed to limit Archaea to various extreme environments, in terms of temperature, pH, salinity, and anaerobiosis. Members of Archaea domain inhabit a variety of the most hostile environments known to support life on earth: highest temperature, e.g. strain 121 grows at 121°C (Kashefi and Lovley, 2003), lowest pH, e.g. *Picrophilus* thrives optimally at pH ~0 (Fütterer et al., 2004; Schleper et al., 1995), and highest NaCl concentration, e.g. many Halobacteriales tolerate Na<sup>+</sup> saturation (~5.2 M NaCl) (Bowers and Wiegel, 2011). Typical habitats from which pure cultures of archaeal species have been isolated include hot springs, hydrothermal vents, solfataras, salt lakes, soda lakes, sewage digesters, and the rumen. However, within the past two decades, the use of molecular techniques, including PCR-based amplification of 16S rRNA genes, has unveiled a wide distribution of mostly uncultured archaea in normal habitats, such as ocean waters, lake waters, and soil.

Based upon SSU rRNA gene sequence analysis, the archaeal domain can be subdivided in five distinct phyla:

- ◆ Korarchaeota. A phylum proposed nearly 20 years ago by 16S rRNA gene analysis of various environmental samples (Barns et al., 1996) and still including exclusively uncultivated species.

## ***Sulfolobus solfataricus*: a hyperthermophilic archaeon 2.4**

- ◆ Nanoarchaeota. The first and sole representative of this phylum is *Nanoarchaeum equitans*, a nanosized hyperthermophilic anaerobe that lives in obligate symbiosis on the surface of *Ignicoccus* (Huber et al., 2002). From initial 16S rRNA gene sequence analysis, it was proposed that Nanoarchaeota form a deep branch in the archaeal phylogenetic tree and represents a novel archaeal phylum (Huber et al., 2002). However, recent phylogenetic analysis of the completed *Nanoarchaeum equitans* genome, which is the smallest cellular genome to date (480 kb), has called into question this placement. Extensive phylogenetic analysis comparing *Nanoarchaeum* sequences with Crenarchaeota and Euryarchaeota sequences have indicated that *N. equitans* branched more consistently with the euryarchaeal lineage, with notable ties to the *Thermococcales* (Brochier et al., 2005).
- ◆ Thaumarchaeota. Initially classified as mesophilic Crenarchaeota, comparative genomics has recently revealed that they form a separate and deep-branching phylum within the Archaea (Brochier-Armanet et al., 2008). This phylum has rapidly gained much attention after the discovery that some members are able to oxidize ammonia aerobically, providing the first example of nitrification in the Archaea (Könneke et al., 2005).
- ◆ Euryarchaeota. The name derives from the greek term εϋρυς, that means “broad” and reflects the phenotypic heterogeneity of this phylum. Included among the Euryarchaeota are various orders of methanogens (*Methanobacteriales*, *Methanococcales*, *Methanomicrobiales*, *Methanosarcinales*, and *Methanopyrales*), as well as extreme halophiles (*Halobacteriales*), sulfate reducers (*Archaeoglobales*), and various thermophiles (*Thermococcales* and *Thermoplasmatales*).
- ◆ Crenarchaeota. This group of microorganisms shows a limited phenotypic heterogeneity with respect to Euryarchaeota. All pure cultures of this phylum are thermophiles (optimum growth temperature >45°C) and hyperthermophiles (optimum growth temperature >80°C). The Bergey’s Manual of Systematic Bacteriology recognizes the cultured Crenarchaeota to be composed of four orders: *Caldisphaerales*, *Desulfurococcales*, *Sulfolobales*, and *Thermoproteales* (Chaban et al., 2006).

## **2.4 *Sulfolobus solfataricus*: a hyperthermophilic archaeon**

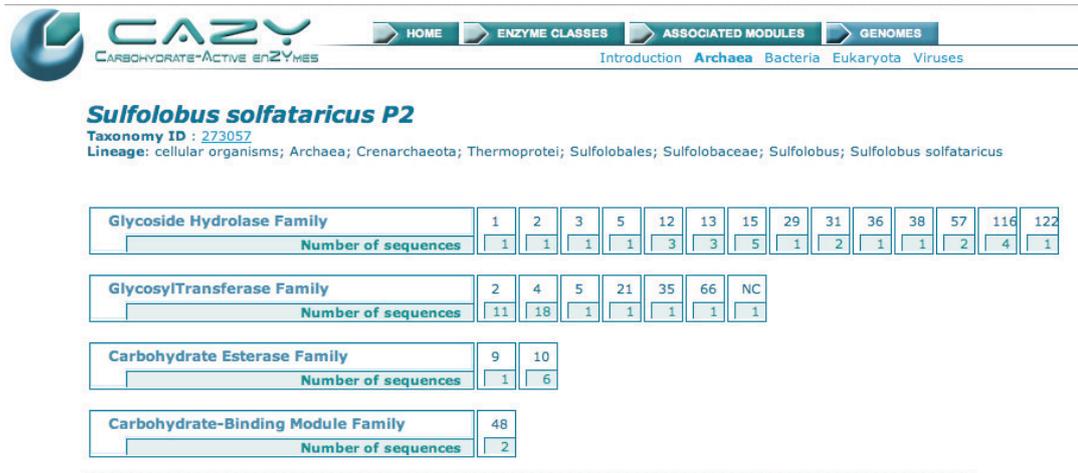
The study of archaeal microorganisms is important not only for unravelling questions about origin of life, but also in applicative perspective. The extremophilic properties of many archaea have made them a favourite starting point for theories concerning how life may have evolved in the hostile conditions of early Earth (Stetter, 2006). Moreover, the enzymes from extremophilic archaea, showing an increased stability to extreme conditions of temperature and pH, are excellent catalysts for industrial process. Several thermophilic archaeal enzymes can be potentially used in industrial biocatalysis, such as amidase/ $\gamma$ -lactamase from *Sulfolobus solfataricus* in the production of optically pure  $\gamma$ -lactam, the building block for antiviral carbocyclic nucleotides and alcohol dehydrogenase from *Aeropyrum pernix* in the production of optically pure alcohols (Littlechild, 2011).

### ***Sulfolobus solfataricus***

*Sulfolobus solfataricus* is a hyperthermoacidophilic archaeon and represents one of the best-studied organisms within the phylum Crenarchaeota. It colonizes terrestrial volcanic hot springs and was originally isolated from the Pisciarelli solfataric field near Naples, Italy. This microorganism grows optimally at 80°C and pH 2–4 (Brock et al., 1972; Zillig et al., 1980), but maintains its cytoplasmic pH at about 6.5 by generating a large pH gradient across the cytoplasmic membrane (Van de Vossenberg et al., 1995). It is a strict aerobe and grows heterotrophically on a variety of organic compounds as carbon and energy source such as sugars (e.g. glucose, galactose, arabinose, sucrose), amino acids, or peptides. Thus, *S. solfataricus* can be easily maintained in laboratory with relatively little special equipment (Grogan 1989). Its genome has been completely sequenced and published in 2001 (She et al., 2001) and a proteomic map of this microorganism was also developed leading to the identification of 1399 proteins, including 44 insoluble and 32 secreted proteins (Barry et al., 2006; Chong and Wright, 2005). Comparative biochemical studies on central carbohydrate metabolism revealed that *S. solfataricus* utilizes a modified Entner-Doudoroff (ED) pathway for glycolysis (De Rosa et al., 1984). Moreover, *Sulfolobus* species are well known for their saccharolytic capacity. This feature is particularly well reflected by the number of CAZymes that have

## Sulfolobus solfataricus: a hyperthermophilic archaeon 2.4

been found by either purification of their activity or by homology search methods (Fig. 2.4.2).



**Figure 2.4.2:** CAZymes annotated in the genome of *S. solfataricus* P2

Enzymes involved in the breakdown of starch have been described, such as extracellular surface-layer associated  $\alpha$ -amylase (SSO1172, GH57) (Worthington et al., 2003) and an intracellular maltase that releases glucose from maltodextrins (SSO3051, GH31) (Ernst et al., 2006; Rolfsmeier and Blum, 1995). The dextrins are also substrate for the GH13 enzymes TreY (SSO2095) and TreZ (SSO2093) that convert them to trehalose (Maruta et al., 1996), one of the main compatible solute in *Sulfolobus* spp. (Santos and Costa, 2001). A gene cluster involved in glycogen synthesis and breakdown can also be found in the *S. solfataricus* genome. Apart from a glycogen synthase (SSO0987, GT5), this cluster contains a glycogen associated  $\alpha$ -amylase (SSO0988, GH57) (Cardona et al., 2001), a glucoamylase (SSO0990, GH15) (Kim et al., 2004) and a putative glycogen debranching enzyme (SSO0991). The ability of *Sulfolobus* to use  $\alpha$ -linked galactosides, such as melibiose and raffinose, can be ascribed to  $\alpha$ -galactosidase activity of GalS (SSO3127, GH36) (Brouns et al., 2006). Moreover, the genome of *Sulfolobus* has also revealed an  $\alpha$ -fucosidase (SSO11867/SSO3060, GH29) encoded by an interrupted gene that was correctly translated by the ribosome programmed -1 frameshifting (Cobucci-Ponzano et al., 2003) and an  $\alpha$ -mannosidase (SSO3066, GH38), that seems involved in the process of protein glycosylation (Cobucci-Ponzano et al., 2010b). A  $\beta$ -galactosidase (SSO3019, GH1) involved in the degradation of lactose and cellobiose has been described (Pisani et al., 1990), which not only

## **2.4 *Sulfolobus solfataricus*: a hyperthermophilic archaeon**

has served as a thermostable model protein in engineering studies (Kaper et al., 2002), but also has allowed the development of genetic manipulation systems for *S. solfataricus* (Jonuscheit et al., 2003; Worthington et al., 2003). In the genomic vicinity of the  $\beta$ -galactosidase, there is an  $\alpha$ -xylosidase (SSO3022, GH31), and this may suggest a cooperative function of the two enzymes in the degradation of xyloglucan oligosaccharides (Moracci et al., 2000). The same gene cluster also harbours a bifunctional  $\beta$ -D-xylosidase/ $\alpha$ -L-arabinosidase (SSO3032, GH3) that is specifically induced in the presence of xylan polysaccharides (Morana et al., 2007). Finally, *Sulfolobus* genome harbours cellulases belonging to GH12 family (SSO1949, SSO1354 and SSO2534) that could operate in the extracellular environment characterized by extreme acidity and high temperature (Huang et al., 2005; Limauro et al., 2001; Maurelli et al., 2008).

In conclusion, *S. solfataricus* is of high interest for industry and biotechnology for various reasons: broad physiological versatility, unique properties of its thermostable proteins, and a model organism for physiology, biochemistry, and comparative and functional genomics.

## **2.5 PURPOSE OF THE WORK**

## 2.5 Purpose of the work

---

Glycosyl hydrolases plays a key role in a wide range of biological processes as well as in many industrial applications. Accordingly, unveil the structural and functional features of this class of enzymes is useful to increase our knowledge in this field and to expand the applicability of these catalysts. In particular, the enzymes isolated from hyperthermophilic microorganisms have elicited great interest in the scientific community for their unique features, such as thermostability and optimal activity at high temperatures. The efforts are focused on the study of the molecular basis of the resistance of these enzymes to extreme conditions of pH and temperature. In addition, their increased stability with respect to mesophilic enzymes makes them more suitable for harsh industrial processes.

On the basis of these observations, one part of my PhD project was aimed to identify novel extremophilic glycosidases from the hyperthermophilic crenarchaeon *S. solfataricus*. To date no enzymatic activities involved in the hydrolysis of *N*-acetyl-hexosaminides have been described in the archaeal organisms, although these sugars have been identified in archaeal glycoproteins (Zähringer et al., 2000) and exopolysaccharides (Zolghadr et al., 2010). These premises prompted us to seek *N*-acetyl-hexosaminidase activities in the extracts of *S. solfataricus*, strain P2. The first part of my work was focused on the isolation and identification of *N*-acetyl- $\beta$ -glucosaminidase from *S. solfataricus* P2. In the second part, the biochemical characterization of the recombinant form of this enzyme has allowed to unveil its catalytic features. The biochemical data associated to the phylogenetic analysis have provided major clues about the CAZy family GH116, in which *N*-acetyl- $\beta$ -glucosaminidase is included.

## **2.6 EXPERIMENTAL PROCEDURES**

## 2.6 Experimental procedures

### Bacterial Strains

*E. coli* strains used in this work are listed in the Table 2.6.1.

Strain	Genotype	Reference or Source
DH5 $\alpha$	F <sup>-</sup> <i>lacZ</i> $\Delta$ M15 <i>recA1 endA1 hsdR17 phoA supE44</i> $\Delta$ ( <i>lacZYA argF</i> ) <i>glnV44 thi-1 relA1 gyrA96 deoR nupG</i>	Hanahan 1983
TOP10	F <sup>-</sup> <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74 recA1 araD139 <math>\Delta</math>(<i>ara-leu</i>)7697 <i>galU galK rpsL</i> (Str<sup>R</sup>) <i>endA1 nupG</i></i>	Invitrogen
BL21 star (DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ) <i>gal dcm rne131</i> (DE3)	Invitrogen
BL21 Ril (DE3)	F <sup>-</sup> <i>ompT hsdS</i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ) <i>dcm+ Tetr gal</i> $\lambda$ (DE3) <i>endA Hte [argU ileY leuW Camr]</i>	Invitrogen
Rosetta (DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ) <i>gal dcm</i> (DE3) pRARE (Cam <sup>R</sup> )	Novagen

Table 2.6.1: *E. coli* strains used in this work

### Archaeal Strains

*Sulfolobus solfataricus* P2 (She et al., 2001)

*Sulfolobus solfataricus* 98/2 (Hochstein and Stan-Lotter, 1992)

*Sulfolobus solfataricus* PBL2025 (Schelert et al., 2004)

### Culture media

LB (Luria-Bertani Broth) (1 liter):

10 g NaCl

5 g yeast extract

10 g tryptone

*S. solfataricus* media

Brock medium (Brock et al., 1972) adjusted to pH 3.5 with sulfuric acid and supplemented with yeast extract, sucrose and casaminoacids (0.1 % each) as carbon source.

### Reagents

All commercially available substrates were purchased from Sigma-Aldrich and Carbosynth unless otherwise designated.

### Isolation of a *N*-acetyl- $\beta$ -glucosaminidase from *S. solfataricus* P2

*N*-acetylglucosaminidases activities were searched in free cell extract (FCE) of *S. solfataricus* P2. The enzymatic activities were followed through assays on the chromogenic substrates 4-nitrophenyl-*N*-acetyl- $\beta$ -glucosaminide (4Np- $\beta$ -GlcNAc) and 4-nitrophenyl-*N*-acetyl- $\alpha$ -glucosaminide (4Np- $\alpha$ -GlcNAc). In addition, *N*-acetyl- $\beta$ -glucosaminidase ( $\beta$ -GlcNAcase) activity was monitored by activity gel staining with 5-Bromo-4-Chloro-3-Indolyl 2-acetamido-2-deoxy- $\beta$ -D-glucopyranoside (X- $\beta$ -GlcNAc) (see below for details).

*S. solfataricus* cells, strain P2, were grown at 80°C and the growth was monitored spectrophotometrically at 600 nm. When the culture reached an  $A_{600\text{nm}}$  of 0.7-1.0 optical densities, cells were harvested by centrifugation at 5,000  $\times g$  for 15 min at 4°C. The resulting cell pellet was resuspended in 2 mL  $g^{-1}$  cells of PBS buffer (20 mM sodium phosphate buffer, pH 7.4, 150 mM NaCl) supplemented with 0.1% (v/v) Triton X-100. The cells were lysated with three cycles of freeze thawing (5 min at -70°C; 5 min at 37°C), and centrifuged for 30 min at 10,000  $\times g$ .

FCE was applied on a High Load 16/10 Q-Sepharose High Performance column (GE-Healthcare) equilibrated with Buffer A (20 mM phosphate buffer, pH 7.5) at a flow rate of 3 mL  $\text{min}^{-1}$ . The run was performed with an initial step of extensive wash with Buffer A (3-column volumes) followed by a linear ionic strength gradient from 0 to 1 M NaCl in Buffer A (3-column volumes) and a final step with 1 M NaCl in Buffer A (2-column volumes).

At these conditions,  $\beta$ -GlcNAcase activity was found primarily in the fractions eluted at about 150 mM NaCl. Active fractions were pooled, equilibrated with 1 M ammonium sulphate, and applied on a HiLoad 26/10 Phenyl Sepharose High performance (GE-Healthcare), which was equilibrated with Buffer A supplemented with 1 M ammonium sulphate at a flow rate of 3 mL  $\text{min}^{-1}$ . After 1-column volume of loading buffer, the protein was eluted with a two-step gradient of water (0-80%, 2-column volumes; 80-100%, 3-column volumes) followed by a final step at 100% of water (2-column volumes); the protein eluted in about 85% water.

Active fractions were pooled, dialyzed against PBS buffer and concentrated by ultrafiltration on an Amicon YM30 membrane (cut off 30,000 Da). The resulting

## **2.6 Experimental procedures**

---

sample was loaded on a Superdex 200 HR 10/300 gel filtration column (GE-Healthcare), run at a flow rate of 0.5 mL min<sup>-1</sup> in PBS buffer. Active fractions were pooled and concentrated. The protein concentration was determined with the method of Bradford (Bradford, 1976). The native enzyme was about 85% pure by SDS-PAGE and was stored at 4°C.

### **Activity assays of native *N*-acetyl- $\beta$ -glucosaminidase on X- $\beta$ -GlcNAc and on 4Np- $\beta$ -GlcNAc**

The activity on X- $\beta$ -GlcNAc was analysed by loading the non heat-denatured samples on 8% SDS-PAGE gel. After the run, the gel was washed with water and then equilibrated in 50 mM phosphate pH 6.5. The activity assay was carried out by incubating the gel at 65°C for 16 hours in 50 mM phosphate pH 6.5 with 2 mg ml<sup>-1</sup> of X- $\beta$ -GlcNAc and 10% methanol.

The hydrolysis of 4Np- $\beta$ -GlcNAc was analysed by incubating FCE or FPLC fractions in 50 mM phosphate pH 6.5 at 60°C for a suitable time (2-5 minutes). The reactions, performed on 12 mM substrate, were blocked in ice by adding 1M sodium carbonate pH 10.2. The release of nitrophenol group was monitored spectrophotometrically at 420 nm and the extinction coefficient to calculate enzymatic units was 17.2 mM<sup>-1</sup> cm<sup>-1</sup>. In all assays, spontaneous hydrolysis of the substrate was subtracted by using appropriate blank mixtures without enzyme.

### **Cloning of *sso3039* gene in *E. coli***

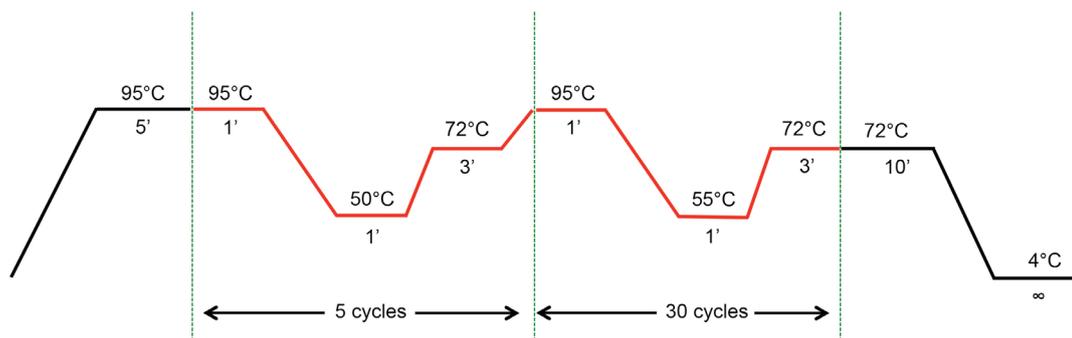
Cloning of *sso3039* gene was performed by using Champion™ pET directional TOPO® expression kit (Invitrogen).

Genomic DNA strain was extracted from *S. solfataricus* P2 cells by Nucleospin Tissue kit (Macherey Nagel). The *sso3039* gene was amplified by PCR from the genome of *S. solfataricus* P2 using the synthetic oligonucleotides 3039TOPO5' and 3039TOPO3'His (Tab. 2.6.2).

Primer	5'-sequence-3'	N° of bases	Tm (°C)
3039TOPO5'	CACCATGAAATACAAATACTCTTA	25	57.60
3039TOPO3'His	AAATTCAATCTCTTCATTTTCCC	23	55.60

**Table 2.6.2:** Sequence of oligonucleotides used for amplification of *ss03039* gene. Sequence indicated in red is introduced to pair with the overhang sequence, GTGG, in pET 101/D-TOPO vector

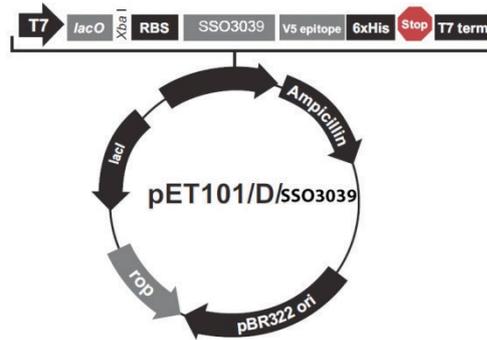
The amplification reaction was performed with the PfuUltra II Fusion HS DNA Polymerase (Stratagene) by using the program reported in Fig. 2.6.1.



**Figure 2.6.1:** Schematic representation of PCR protocol used to amplify *ss03039* gene

The PCR product obtained was verified by electrophoresis on 1% agarose gel, purified by PCR Kleen spin columns (Biorad) and then cloned in the expression vector pET101/D-TOPO (Invitrogen), according to the manufacturer. Briefly 50 ng of DNA fragment were incubated in salt solution with 1  $\mu$ l vector (15-20 ng) at room temperature for 30 minutes. The TOPO cloning reaction was used to transform One Shot TOP10 chemically competent cells according to the protocol of the manufacturer (see below for more details). Positive clones were selected through PCR colony and the absence of mutations was verified by sequencing. This procedure allowed to obtain the recombinant plasmid pET101/D-TOPO-SSO3039, in which the SSO3039 ORF is under the control of the isopropyl-1-thio- $\beta$ -D-galactopyranoside (IPTG) inducible T7 RNA polymerase promoter and the C-terminal of the protein was fused to V5 epitope and 6xHis tag (Fig. 2.6.2).

## 2.6 Experimental procedures



**Figure 2.6.2:** pET101/D-TOPO/SSO3039 features

### Transformation of One Shot TOP10 chemically competent cells

The recombinant plasmid pET101/D-TOPO/SSO3039 obtained by TOPO cloning reaction described above was transformed into competent One Shot® TOP10 chemically competent *E. coli* cells provided by the kit. The entire TOPO cloning reaction mixture has been added into a vial of cells. After incubation on ice for 30 minutes, the cells were subjected to a heat-shock at 42°C for 30 seconds and then immediately transferred to ice and supplemented with 250  $\mu\text{L}$  SOC medium. The cells were incubated at 37°C for 1 hour; then 200  $\mu\text{L}$  of culture were spread on LB-agar plates containing 50  $\mu\text{g mL}^{-1}$  ampicillin and incubated at 37°C for 16 hours.

### Expression trials of rSSO3039

To determine which strain allows to obtain the better yield of the recombinant SSO3039 (rSSO3039), expression trials were performed in three different *E. coli* strains: BL21 star (DE3), BL21 Ril (DE3) and Rosetta (DE3). For each strain, *E. coli* cells transformed with the construct pET101/D-TOPO-SSO3039 were used to prepare 100 mL cultures. The cells were incubated at 37°C and when reached an  $A_{600\text{nm}}$  of 0.6, each culture was divided in two aliquots: one was treated with 0.6 mM IPTG (I), while the other one was not treated (NI). The growth was allowed to proceed for 16 h and the cells were harvested by centrifugation at 5,000  $\text{xg}$  for 30 minutes. The cell pellet was resuspended in 50 mM sodium phosphate buffer, pH 8.0, 300 mM NaCl (Buffer B) with a ratio of 5 ml  $\text{g}^{-1}$  cells and then was incubated at 37°C for 1 h with lysozyme (Fluka) and 25 U  $\text{g}^{-1}$  cell of Benzonase (Novagen). The cell lysis was performed by three cycles of freeze/thawing and the cell debris were removed by centrifugation at 14000  $\text{xg}$  for 30 minutes. The rSSO3039 was partially

purified by one step of thermal precipitation: incubation at 80°C for 20 minutes followed by centrifugation at 16000 xg for 30 minutes. The resulting samples was used to evaluate rSSO3039 expression by two different approaches:

- SDS-PAGE
- Activity on 4Np-  $\beta$ -GlcNAc

### **Purification of rSSO3039**

The SSO3039 ORF was expressed in *E. coli* cells, strain BL21 star (DE3) (Invitrogen) transformed with pET101/D-TOPO-SSO3039 according to the manufacturer. Cells were grown at 37°C in 4.5 liters of Luria-Bertani (LB) broth supplemented with ampicillin (50  $\mu\text{g mL}^{-1}$ ). Growth was allowed to proceed for 16 h without addition of IPTG, and cells were harvested by centrifugation at 5,000 xg. The resulting cell pellet was resuspended in Buffer B with a ratio of 5 ml  $\text{g}^{-1}$  cells and then was incubated at 37°C for 1 h with 20 mg of lysozyme (Fluka) and 25 U  $\text{g}^{-1}$  cell of Benzonase (Novagen). The cells were lysed by French cell pressure treatment and cell debris were removed by centrifugation at 10,000 xg for 30 min. FCE was loaded on a His Trap FF crude column (GE-Healthcare) equilibrated with Buffer B at a flow rate of 1 mL  $\text{min}^{-1}$ . After an initial wash-step (20-column volumes) with buffer B, the protein was eluted with a two-step gradient of imidazole in Buffer B (250 mM imidazole, 20-column volumes; 500 mM imidazole, 15-column volumes).

The protein eluted at 250 mM imidazole. Active fractions were pooled, dialyzed against PBS buffer and then heat-fractionated for 20 min at 80°C. The resulting supernatant was concentrated to about 3 mg  $\text{mL}^{-1}$  and applied on a HiLoad 16/60 Superdex 200 gel filtration column (GE-Healthcare) equilibrated with PBS buffer at a flow rate of 0.5 mL  $\text{min}^{-1}$ . Active fractions were pooled, concentrated and stored at 4°C. After this procedure rSSO3039 was more than 95% pure by SDS-PAGE.

### **Molecular mass determination**

The molecular mass of native and recombinant SSO3039 was determined by gel filtration on a Superdex 200 HR 10/300 FPLC column (GE-Healthcare) performed as described above. Molecular weight markers were

## 2.6 Experimental procedures

apoferritin (443 kDa), amylase (200 kDa), alcohol dehydrogenase (150 kDa), albumin (66 kDa) and ribonuclease A (13.7 kDa).

### Temperature and pH influence on rSSO3039 activity

The temperature profile of rSSO3039 activity was determined in the range of 40-70°C in the standard assay conditions.

Thermal stability was evaluated by incubating pure rSSO3039 (0.33 mg mL<sup>-1</sup>) in PBS buffer, pH 7.3, at the indicated temperatures. At intervals, aliquots of 6 µl (2 µg enzyme) were withdrawn, transferred in ice and assayed at standard conditions. The residual activity was expressed as a percentage of the maximal enzymatic activity measured before the incubation at indicated temperatures.

Optimal pH and buffered saline were determined by assaying rSSO3039 in 50 mM of the indicated buffers and pHs in standard conditions.

### Enzymatic assays of rSSO3039

The standard assay for rSSO3039 activity was performed in 50 mM citrate phosphate buffer, pH 4.0 on 4Np-β-GlcNAc at the final concentration of 12 mM. Typically, in each assay, we used about 2–5 µg of rSSO3039 in the final volume of 0.2 mL. After 2 min of incubation at 60°C, the reaction was blocked in ice by adding 0.8 mL of 1 M sodium carbonate pH 10.2. The absorbance was measured at 420 nm at room temperature and the millimolar extinction coefficients of 4-nitrophenol and 2-nitrophenol were 17.2 and 4.7 mM<sup>-1</sup> cm<sup>-1</sup>, respectively. In all the assays, spontaneous hydrolysis of the substrate was subtracted by using appropriate blank mixtures without the enzyme. The specific activity, indicated as U mg<sup>-1</sup>, was determined with the formula reported in Fig. 2.6.3.

$$A = \frac{\Delta \text{Abs}/\text{min} \times 1000}{\epsilon \times V_e \times C_e}$$

**A**: Specific activity (U/mg)  
**ΔAbs/min**: absorbance per minute (420 nm)  
**ε**: molar extinction coefficient (mM<sup>-1</sup> × cm<sup>-1</sup>)  
**V<sub>e</sub>**: volume of enzyme (µL)  
**C<sub>e</sub>**: enzyme concentration (µg/µL)

**Figure 2.6.3:** Formula applied to calculate specific activity

One enzymatic unit is defined as the amount of enzyme catalysing the conversion of 1 µmole of substrate into product in 1 min, at

the indicated conditions. All kinetic data were calculated as the average of at least two experiments and were plotted and refined with the program GraphPad Prism.

The enzymatic activity on 4-Methylumbelliferyl- $\beta$ -D-glucoside (MU- $\beta$ -Glc) and 4-methylumbelliferyl N-acetyl- $\beta$ -D-glucosaminide (MU- $\beta$ -GlcNAc) was assayed in standard conditions with slight modifications; the reaction was performed in 0.25 mL with 0.3  $\mu$ g rSSO3039 and was blocked by adding 0.5 mL of 0.1M Glycine-NaOH rather than sodium carbonate. The release of 4-methylumbelliferone has been followed by fluorescence spectroscopy (Tab. 2.6.3). A standard curve of the free fluorophore, 4-methylumbelliferone, in the same conditions of assay was generated for the calculation of enzyme product.

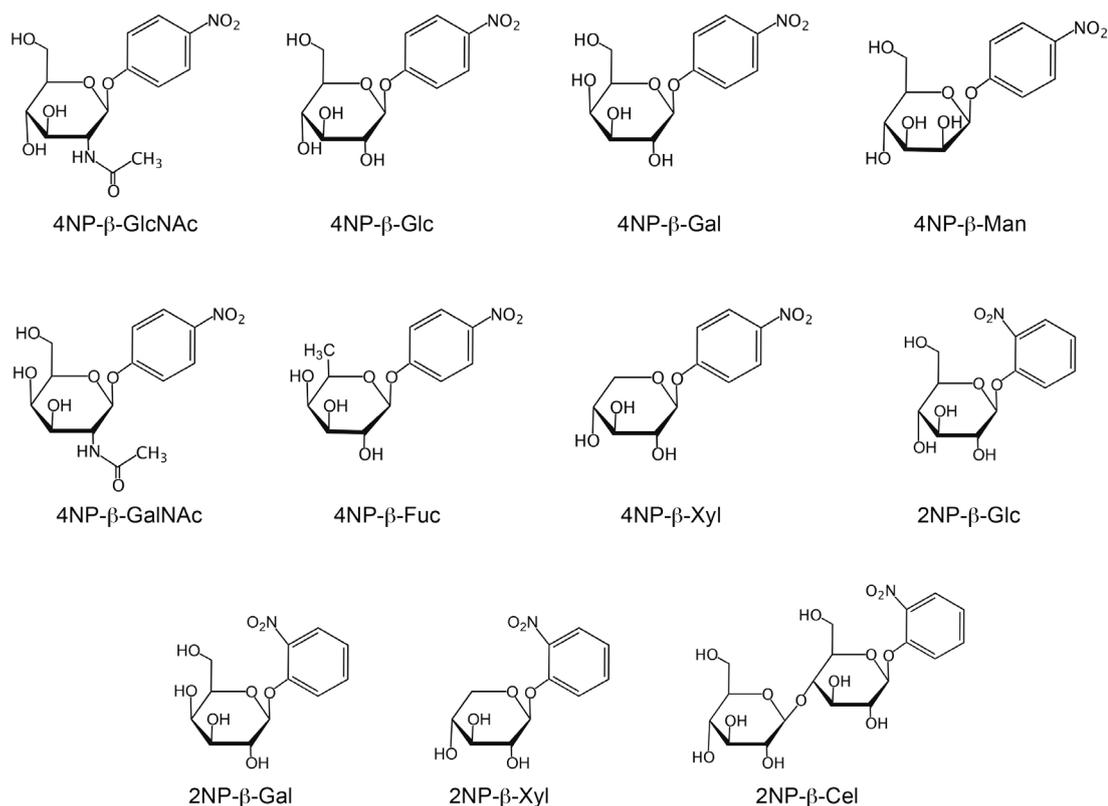
Kinetic constants of rSSO3039 were measured at standard conditions by using concentrations of substrate ranging between 0.5 and 20 mM for 4Np- $\beta$ -GlcNAc, 0.1 and 15 mM for 4Np- $\beta$ -Glc, 0.05 and 4 mM for MU- $\beta$ -Glc, 0.25 and 15 mM for MU- $\beta$ -GlcNAc.

Parameter	Set value
$\lambda_{\text{ex}}$	384 nm
$\lambda_{\text{em}}$	450 nm
Band width <sub>ex</sub>	2.5 nm
Band width <sub>em</sub>	5 nm

**Table 2.6.3:** Setting of parameters for the measurements on the spectrofluorometer JASCO FP-8200

The substrate specificity of rSS3039 was determined by analysing the activity on different aryl-glycosides (final concentration 12 mM) in standard conditions (Fig. 2.6.4).

## 2.6 Experimental procedures



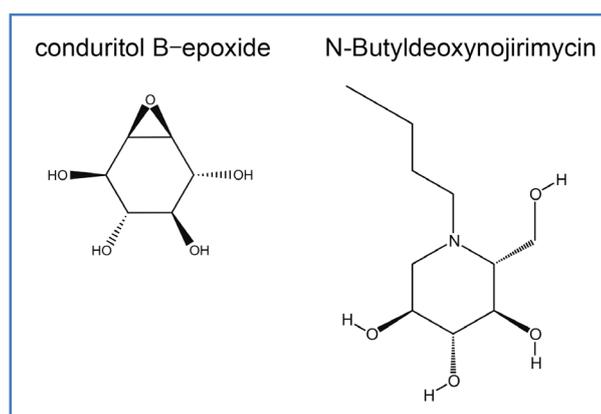
**Figure 2.6.4:** Structure of the  $\beta$ -aryl-glycosides used in the assays

The activity of rSSO3039 on tetra- and pentasaccharides of glucose and *N*-acetylglucosamine (Glc4, Glc5, GlcNAc4 and GlcNAc5), 2Np- $\beta$ -cellobioside (2Np- $\beta$ -Cel) and 4Np-*N,N'*-diacetyl- $\beta$ -chitobioside (4Np- $\beta$ -Chit) was measured by assaying the enzyme (5  $\mu$ g) at standard conditions. Aliquots were withdrawn at indicated times and stored at -20°C. The analysis of the reaction products was performed on a silica gel 60 F254 TLC or by High-Performance Anion-Exchange chromatography with Pulsed Amperometric Detection (HPAEC-PAD) equipped with a PA200 column (Dionex, USA). TLC analysis of the gluco-oligosaccharides was carried out by using acetone/isopropanol/ water (60:30:15) as the eluent and 5% sulfuric acid in methanol for the detection. HPAEC-PAD analyses of the reactions on 2Np- $\beta$ -Cel and gluco-oligosaccharides were performed at a flow-rate of 0.5 mL min<sup>-1</sup> in 100 mM NaOH/10 mM sodium acetate for 6 min followed by a two-step gradient of sodium acetate in 100 mM NaOH (10-100 mM in 10 min; 100-150 mM in 4 min); instead, an isocratic protocol (10 mM NaOH for 40

min) was applied for the analysis of the enzymatic reactions on *N*-acetylchitooligosaccharides and on 4Np- $\beta$ -Chit.

### Inhibition of rSSO3039

To evaluate the inhibition of rSSO3039 by *N*-butyldeoxyojirimycin (NB-DNJ) and conduritol  $\beta$ -epoxide (CBE) (Fig. 2.6.5), the enzyme (0.08 mg mL<sup>-1</sup>) was incubated in the presence of 0.1–50  $\mu$ M NB-DNJ or 2–15 mM CBE in 50 mM phosphate buffer, pH 7.0, in a final volume of 60  $\mu$ l. After incubation at 45°C for 30 min, aliquots of 25  $\mu$ l were withdrawn and the residual activity was measured in standard conditions. Residual activity was expressed as a percentage of the maximum activity measured with enzyme incubated in the same conditions described above with the exception of the inhibitor.



**Figure 2.6.5:** Structure of the inhibitors CBE and NB-DNJ

## **2.7 RESULTS AND DISCUSSION**

### Isolation of a *N*-acetyl- $\beta$ -glucosaminidase from *S. solfataricus* P2

According to the CAZy classification, *N*-acetyl- $\beta$ -glucosaminidase activities (EC 3.2.1.52) are classified in the GH families 3, 20 and 84. Although these families include archaeal GHs, no enzymatic activity involved in the hydrolysis of *N*-acetyl- $\beta$ -hexosaminides has been described in the archaeal organisms to date. In Archaea, only chitinase (EC 3.2.1.14) belonging to GH18 have been characterized, in particular the chitinolytic enzymes from *Pyrococcus furiosus* (Nakamura et al., 2008; Oku and Ishikawa, 2006; Tsuji et al., 2010), *Thermococcus kodakarensis* (Tanaka et al., 1999, 2001) and *Thermococcus chitinophagus* (Andronopoulou and Vorgias 2003, 2004). *N*-acetyl- $\beta$ -hexosaminides have been identified among the monosaccharides constituents of archaeal glycoproteins (Abu-Qarn et al., 2008; Namboori and Graham, 2008) and exopolysaccharides (Zolghadr et al., 2010). In addition chitobiose is the glycan linker to the *N*-glycosylation site, as observed in cytochrome  $b_{558/566}$  and in the S-layer proteins of the thermoacidophilic archaea *S. acidocaldarius* and *S. solfataricus* (Peyfoon et al., 2010; Zähringer et al., 2000). These premises prompted us to seek *N*-acetyl- $\beta$ -hexosaminidase activities in the extracts of *S. solfataricus*, strain P2.

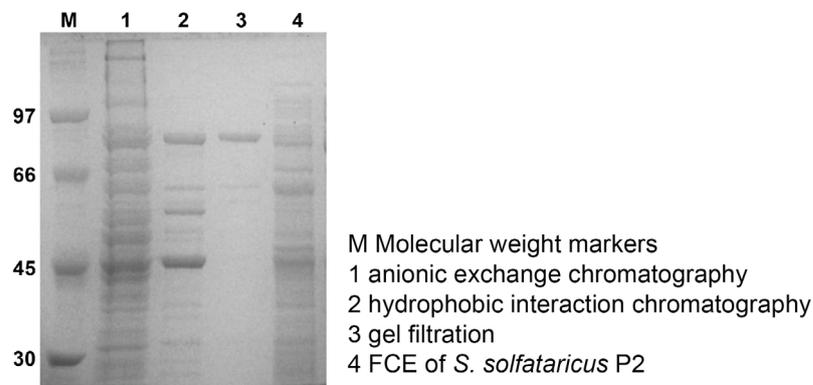
Free cell extracts (FCE) were prepared from a cell culture of *S. solfataricus* P2 stopped in the early stationary phase and assays on the chromogenic substrates, 4Np- $\beta$ -GlcNAc or 4Np- $\alpha$ -GlcNAc, allowed to reveal low, but detectable *N*-acetylglucosaminidase activities:  $4.6 \times 10^{-3}$  U  $\text{mg}^{-1}$  and  $5.5 \times 10^{-3}$  U  $\text{mg}^{-1}$  on 4Np- $\alpha$ -GlcNAc and 4Np- $\beta$ -GlcNAc, respectively. A purification procedure was set up to isolate the native enzymes involved in these catalytic processes. After three chromatography steps (anionic exchange, hydrophobic and gel filtration chromatographies), it was not possible to isolate the *N*-acetyl- $\alpha$ -glucosaminidase activity. Consequently our efforts were focused on the isolation and identification of *N*-acetyl- $\beta$ -glucosaminidase. To this aim we adopted the purification procedure summarized in Tab 2.7.1, that allowed to isolate the enzyme of interest through three chromatographic steps. Although the protein preparation was not pure to homogeneity (Fig. 2.7.1), SDS-PAGE activity stained with X- $\beta$ -GlcNAc revealed a clear band of high molecular weight

## 2.7 Results and discussion

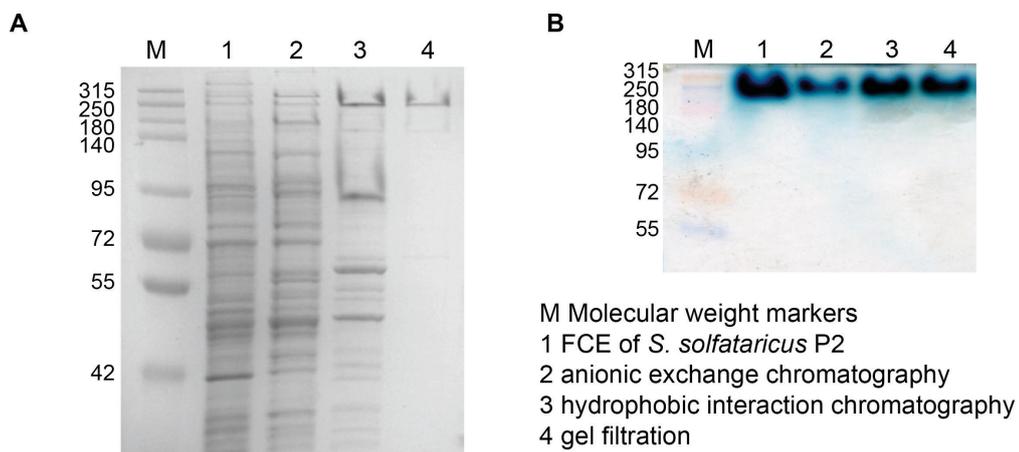
(Fig. 2.7.2A and 2.7.2B), indicating that a single *N*-acetyl- $\beta$ -glucosaminidase activity was present in *S. solfataricus* extracts.

Fraction	Volume (mL)	Protein (mg/mL)	Total Protein (mg)	Activity (U/mL)	Total Activity (U)	Specific Activity (U/mg)	Yield (%)	Fold Purification
Free cell extract	4	53	212	0.3	1.2	$5.5 \cdot 10^{-3}$	100	1
Anionic exchange	13	1.4	18.2	0.05	0.6	0.03	50	5.5
Hydrophobic chromatography	2.5	0.32	0.8	0.06	0.16	0.2	13	36
Gel filtration	0.5	0.04	0.02	0.05	0.03	1.3	2.5	236

**Table 2.7.1:** Purification table of native  $\beta$ -GlcNAcase from *S. solfataricus* P2 extracts. Enzymatic assays were performed in 50 mM phosphate pH 6.5 at 60°C on 12 mM 4Np- $\beta$ -GlcNAc



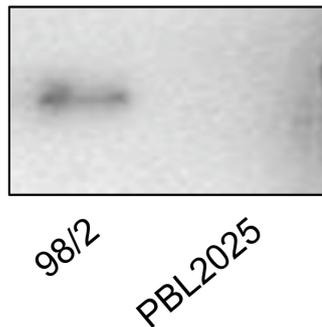
**Figure 2.7.1:** SDS-PAGE of native  $\beta$ -GlcNAcase: Coomassie staining of denatured native samples



**Figure 2.7.2:** SDS-PAGE of native  $\beta$ -GlcNAcase. **(A)** Coomassie staining of non-denatured native samples. **(B)** activity staining with X- $\beta$ -GlcNAc of non-denatured native samples

To identify the ORF encoding for this enzymatic activity, we have employed two complementary approaches: the use of a natural mutant and mass spectrometry analysis.

First, we assayed for *N*-acetyl- $\beta$ -glucosaminidase activity in extracts of a deletion mutant of *S. solfataricus* strain 98/2, named PBL2025, which lacks a chromosomal fragment of about 50 Kb, including the ORFs from SSO3004 to -3050 (Schelert et al., 2004). Interestingly, we could not detect any activity in PBL2025, while FCE of 98/2 strain, used as parental control, showed specific activity comparable to that observed in the P2 strain ( $3 \times 10^{-3}$  U mg<sup>-1</sup>). SDS-PAGE activity stained with X- $\beta$ -GlcNAc confirmed the absence of *N*-acetyl- $\beta$ -glucosaminidase activity in FCE of PBL2025 strain and its presence in the parental strain 98/2 (Fig. 2.7.3).

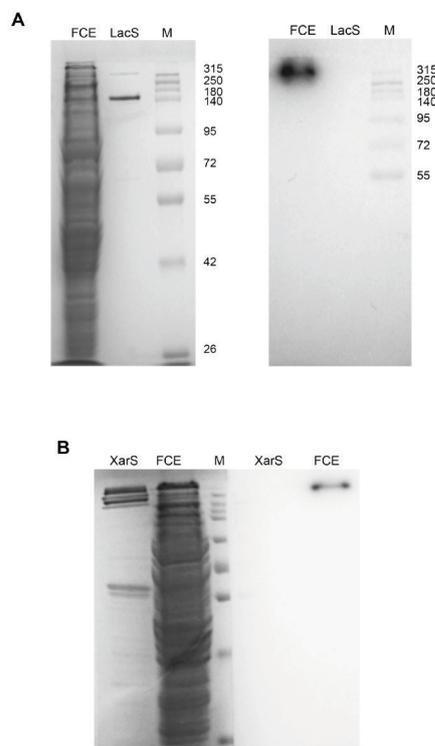


**Figure 2.7.3:** SDS-PAGE activity of FCE from *S. solfataricus* 98/2 (596  $\mu$ g) and PBL2025 (434  $\mu$ g) stained with X- $\beta$ -GlcNAc

These data strongly indicate that the *N*-acetyl- $\beta$ -glucosaminidase activity is encoded by an ORF deleted in PBL2025. The chromosomal region deleted in this strain includes a diverse set of genes possibly involved in sugar degradation and metabolism (Zolghadr et al., 2010), such as GHs, carbohydrate transporters and glucose dehydrogenases. Among GHs deleted in PBL2025, there are either already characterized enzymes, such as  $\alpha$ -mannosidase (Ss $\alpha$ -man, SSO3006, GH38) (Cobucci-Ponzano et al., 2010b),  $\beta$ -glycosidase (LacS, SSO3019, GH1) (Moracci et al., 1994),  $\alpha$ -xylosidase (XylIS, SSO3022, GH31) (Moracci et al. 2000),  $\beta$ -xylosidase/ $\alpha$ -L-arabinosidase (XarS, SSO3032, GH3) (Morana et al., 2007), or other still hypothetical, like putative endo-glucanase (SSO3007, GH5),  $\beta$ -glucuronidase (SSO3036, GH2),  $\alpha$ -

## 2.7 Results and discussion

glucosidase (SSO3037-3038, GH122), and bile-acid  $\beta$ -glucosidase (SSO3039, GH116). Taking into account only the characterized enzymes, GHs potentially involved in the hydrolysis of 4NP- $\beta$ -GlcNAc might be LacS and XarS, which are specific for  $\beta$ -O-anomeric substrates. Thus, the two purified enzymes and FCE of 98/2 strain (positive control) were assayed on 4Np- $\beta$ -GlcNAc and X- $\beta$ -GlcNAc. Assays on 4Np- $\beta$ -GlcNAc revealed that neither LacS nor XarS were able to hydrolyze this chromogenic substrate and SDS-PAGE activity stained with X- $\beta$ -GlcNAc confirmed this results for LacS (Fig. 2.7.4A) and XarS (Fig. 2.7.4B). Accordingly, these data fostered our hypothesis that *S. solfataricus* expresses another GH with *N*-acetyl- $\beta$ -glucosaminidase activity.



**Figure 2.7.4:** SDS-PAGE of the GHs LacS and XarS. **(A)** Coomassie and activity staining with X- $\beta$ -GlcNAc (left and right panels, respectively) of 98/2 FCE (434  $\mu$ g) and purified LacS (2  $\mu$ g). **(B)** Coomassie and activity staining with X- $\beta$ -GlcNAc (left and right panels, respectively) of 98/2 FCE (434  $\mu$ g) and purified XarS (4.74  $\mu$ g)

### Mass spectrometry analysis of *N*-acetyl- $\beta$ -glucosaminidase partially purified from *S. solfataricus* P2

To identify the ORF encoding for *N*-acetyl- $\beta$ -glucosaminidase activity we used a mass spectrometry approach. This study was performed by Dr. A. Amoresano and Dr. A. Carpentieri of the University of Naples “Federico II”. Mass spectrometry analysis was performed on the sample deriving from the final step of the purification procedure described above (gel filtration chromatography, Table 2.7.1). A gel free proteomic approach was carried out to identify the proteins in the sample showing *N*-acetyl- $\beta$ -glucosaminidase activity. The protein sample was reduced and alkylated and then hydrolyzed in solution using trypsin as a proteolytic enzyme. The peptide mixture obtained by this procedure was fractionated by nano LC and analyzed by tandem mass spectrometry (MS/MS). The MS/MS spectra submitted to MASCOT search led to the identification of at least 19 proteins occurring in the same fraction (Table 2.7.2). Among the identified proteins, only three ORFs (SSO3006, SSO3032 and SSO3039) resided in the chromosomal region deleted in PBL2025 strain, in which no *N*-acetyl- $\beta$ -glucosaminidase activity was revealed. SSO3006 encodes for an  $\alpha$ -mannosidase extensively characterized by Cobucci-Ponzano and colleagues (Cobucci-Ponzano et al., 2010b). SSO3032 encodes for the  $\beta$ -xylosidase/  $\alpha$ -L-arabinosidase XarS, which, as shown above, has no  $\beta$ -GlcNacase activity (Fig. 2.7.4B). SSO3039, with a sequence coverage of 55% and a MASCOT score of 474, has never been characterized so far and it is annotated in the genome of *S. solfataricus* P2 (<http://www-archbac.u-psud.fr/projects/sulfolobus/>) as a predicted bile acid  $\beta$ -glucosidase. It is worth to mention that SSO3039 is classified as GH116 in the CAZy database (<http://www.cazy.org/GH116.html>). This CAZy family, previously created in our laboratory by Cobucci-Ponzano and co-workers (Cobucci-Ponzano et al., 2010a), includes only acid  $\beta$ -glucosidase (EC 3.2.1.45),  $\beta$ -glucosidase (EC 3.2.1.21) and  $\beta$ -xylosidase (EC 3.2.1.37) and not  $\beta$ -GlcNacase (EC 3.2.1.52). These premises prompted us to explore the functional features of SSO3039 in more detail.

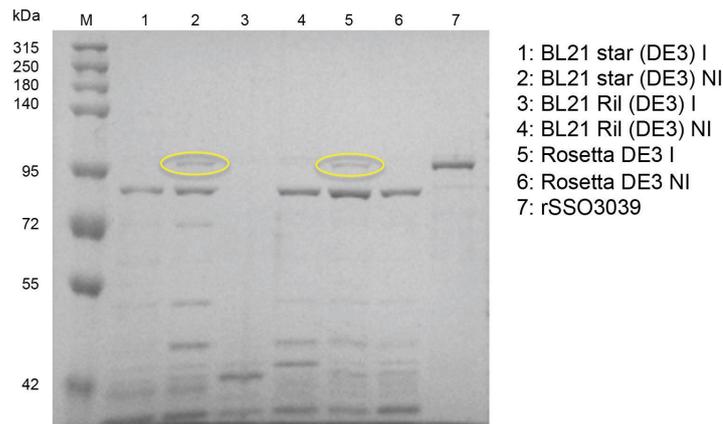
## 2.7 Results and discussion

	ORF	Functional annotation
1	SSO3006	$\alpha$ -mannosidase
2	SSO2810	Hydrolase, acetamidase/formamidase related, putative
3	SSO3032	$\beta$ -xylosidase
4	SSO3039	Predicted bile acid $\beta$ -glucosidase
5	SSO0341	alanyl-tRNA synthetase
6	SSO2742	Glucoamylase
7	SSO2766	Chromosome segregation ATPases
8	SSO2021	Acetyl-CoA synthetase, carboxy-end fragment
9	SSO1272	hypothetical protein
10	SSO0559	ATP synthase subunit I
11	SSO2434	Carbon monoxide dehydrogenase
12	SSO1370	pyruvate dehydrogenase, beta subunit
13	SSO2423	Uncharacterized proteins of the AP superfamily
14	SSO0633	Amidophosphoribosyltransferase
15	SSO0528	glyceraldehyde-3-phosphate dehydrogenase
16	SSO0428	hypothetical protein
17	SSO0561	ATP synthase subunit E
18	SSO2760	Carbon monoxide dehydrogenase
19	SSO1095	aconitate hydratase

**Table 2.7.2:** ORFs identified by mass spectrometry analysis in the protein sample showing  $\beta$ -GlcNacase activity

### Cloning, expression, and characterization of SSO3039

To unequivocally confirm that the ORF SSO3039 encodes for the *N*-acetyl- $\beta$ -glucosaminidase revealed in FCE of *S. solfataricus* P2, we produced and characterized the recombinant form of this enzyme. To this aim, *sso3039* gene was cloned in the expression vector pET101/D-TOPO obtaining the recombinant construct pET101/D-TOPO-SSO3039, which allowed to express a recombinant SSO3039 (rSSO3039) fused to the V5 epitope and to a 6xHis tag at the C-terminal end. To determine which strain allowed to obtain the better yield of rSSO3039, expression trials were performed in three different *E. coli* strains: BL21 star (DE3), BL21 Ril (DE3) and Rosetta (DE3) (see section 2.7 for details about the experimental procedures). SDS-PAGE analysis and activity assays on 4Np- $\beta$ -GlcNAc showed that the best yield of rSSO3039 can be obtained in *E. coli* BL21 (Star) DE3 strain without induction with IPTG or in Rosetta DE3 in which the protein expression was induced with 0.6 mM IPTG (Fig. 2.7.5).



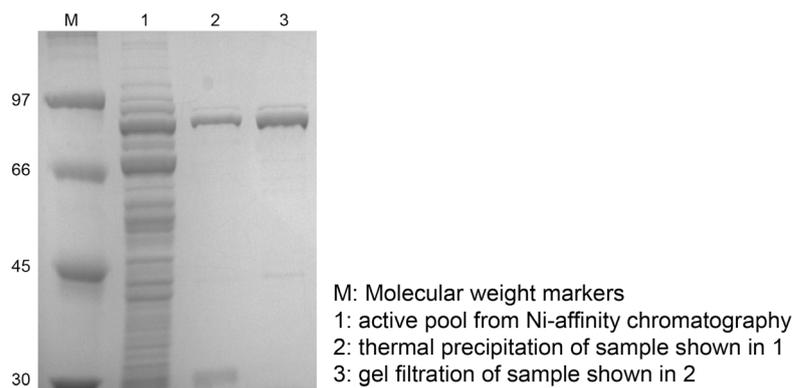
**Figure 2.7.5:** SDS-PAGE of the partially purified rSSO3039 expressed in three different *E. coli* strains with (I) or without (NI) the induction with IPTG. Yellow circles highlight rSSO3039 protein band

The levels of rSSO3039 expression in *E. coli* BL21 star (DE3) and Rosetta DE3 were comparable, and we decide to use the former strain for the expression and purification of rSSO3039, that is also the strain recommended by the manufacturer of the expression kit, Invitrogen. The purification procedure consists of three steps (Ni-affinity chromatography, thermal precipitation and gel filtration) (Tab. 2.7.3) and led to an enzyme preparation >95% pure (Fig. 2.7.6). The low final yield (~0.4 mg per liter of *E. coli*) demonstrated that the levels of rSSO3039 expression in *E. coli* were not high and were much lower than those obtained with another GH116 enzyme from *S. solfataricus*, the  $\beta$ -xylosidase/glucosidase SSO1353, whose yield was 1.5 mg per liter of *E. coli* (Cobucci-Ponzano et al., 2010a). Assays on 4Np- $\beta$ -GlcNAc revealed that the purified rSSO3039 showed a thermophilic *N*-acetyl- $\beta$ -glucosaminidase activity, confirming that SSO3039 ORF encodes for the enzyme involved in hydrolysis of *N*-acetyl- $\beta$ -glucosaminides in *S. solfataricus* P2.

## 2.7 Results and discussion

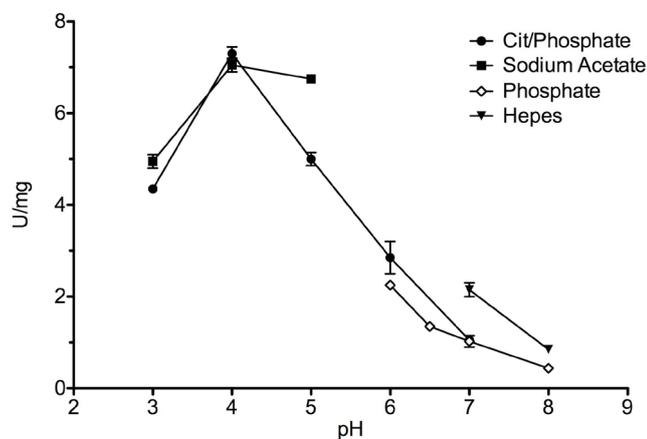
	Volume (mL)	[Protein] (mg/mL)	Total Proteins (mg)	Acitivity (U mL <sup>-1</sup> )	Total Activity (U)	Specific Activity (U mg <sup>-1</sup> )	Yield (%)	Purification folds
FCE	116	20.6	2390	0.37	43	0.018	100	1
His Trap	15	4.6	69	2.7	40	0.58	93	32
Heat-fractionation	14	1.07	15	2.5	34.5	2.3	80	128
Gel filtration	5	0.33	1.65	2.5	12.4	7.5	29	417

**Table 2.7.2:** Purification of rSSO3039 from *E. coli* BL21 star (DE3) strain transformed with pET101/D-TOPO-SSO3039. Assays were performed on 12 mM 4Np- $\beta$ -GlcNAc, in 50 mM citrate phosphate buffer, pH 4.0, at 60°C for 2 minutes



**Figure 2.7.6:** SDS-PAGE of rSSO3039 purification steps

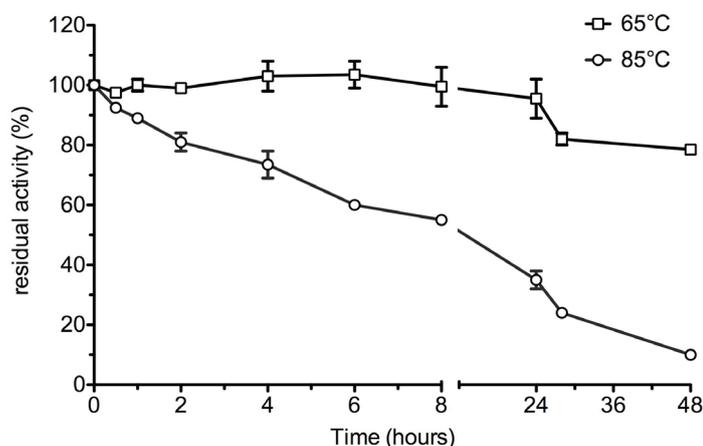
The initial biochemical analyses of rSSO3039 have been aimed to unveil the optimal conditions for the catalysis, such as pH and temperature optima, and thermal stability. To determine the pH optimum of rSSO3039 activity, the enzyme was assayed in 50 mM of different buffers in the range of pH 3.0-8.0. As illustrated in Fig. 2.7.7, rSSO3039 was optimally active in citrate phosphate or sodium acetate buffer at pH 4.0.



**Figure 2.7.7:** pH dependence of rSSO3039 activity on 12 mM 4Np- $\beta$ -GlcNAc

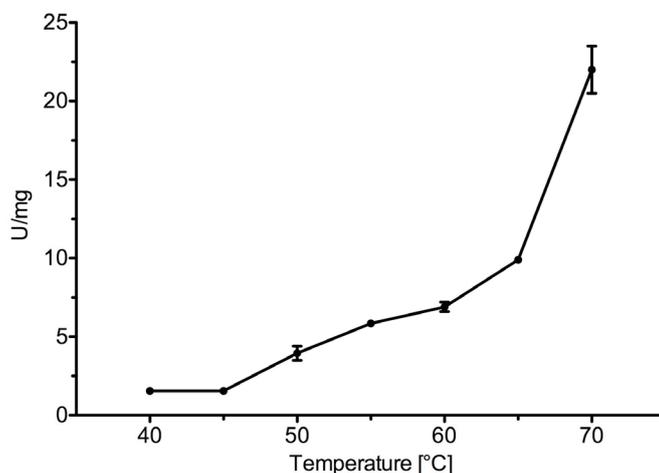
## Results and discussion 2.7

As expected for an enzyme from thermophilic source, the temperature profile assessed in the range of 40-70°C indicated the maximal activity ( $\sim 25 \text{ U mg}^{-1}$ ) at 70°C (Fig. 2.7.8).



**Figure 2.7.8:** rSSO3039 activity monitored in the range of 40-70°C

Moreover, rSSO3039 showed a noticeable stability at 65 and 85°C (Fig. 2.7.9). On the basis of this initial characterization, the standard assay used in all the following study was performed on 12 mM 4Np- $\beta$ -GlcNAc, in 50 mM citrate/phosphate buffer, pH 4.0 at 60°C.

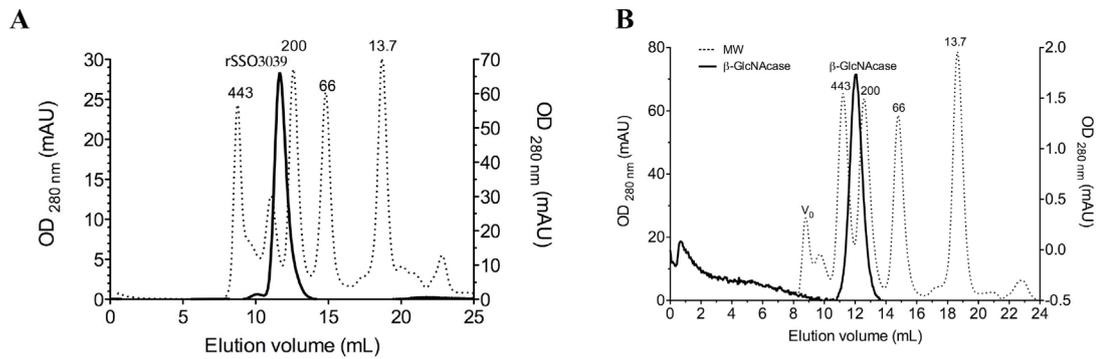


**Figure 2.7.9:** Thermal stability of rSSO3039 at 65°C and 85°C

The molecular weight of the recombinant enzyme in solution was of  $313 \text{ kDa} \pm 10 \text{ kDa}$  (Fig. 2.7.10A) comparable to that observed for the native enzyme in the same conditions ( $271 \pm 18.5 \text{ kDa}$ ) (Fig. 2.7.10B). Since the denatured monomer is 93.4 kDa for the native form and 96.4 kDa for the recombinant one (the

## 2.7 Results and discussion

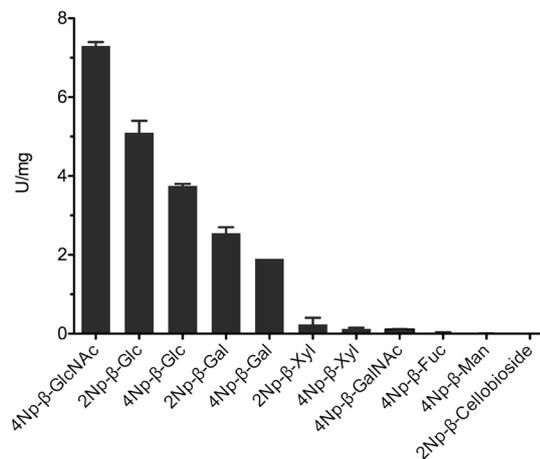
additional 3kDa derive from the fusion to V5 epitope and 6xHis tag), these data indicated that SSO3039 was a trimer in solution.



**Figure 2.7.10:** Molecular weight of rSSO3039 (A) and native β-GlcNAcase (B) determined by gel filtration

### Substrate specificity of rSSO3039

To determine the substrate specificity of rSSO3039, the enzyme was assayed on various aryl-glycosides as described in the Experimental procedures, section 2.7. Enzymatic assays on these substrates indicated the highest specificity for aryl-*N*-acetyl-glucosaminides, followed by aryl-glucosides and -galactosides, while the activity on aryl-xylosides, and -fucosides was only barely detectable, and no activity was observed on aryl-mannosides (Fig. 2.7.11) and glucocerebrosides (data not shown). The specific activity measured on 4Np-β-GlcNAc corresponds to 7.5 U mg<sup>-1</sup>, a value much higher than that observed with SSO1353 on 4Np-β-xylopyranoside (0.8 U mg<sup>-1</sup>), another GH116 enzyme (Cobucci-Ponzano et al., 2010a).



**Figure 2.7.11:** Substrate specificity of rSSO3039

## Results and discussion 2.7

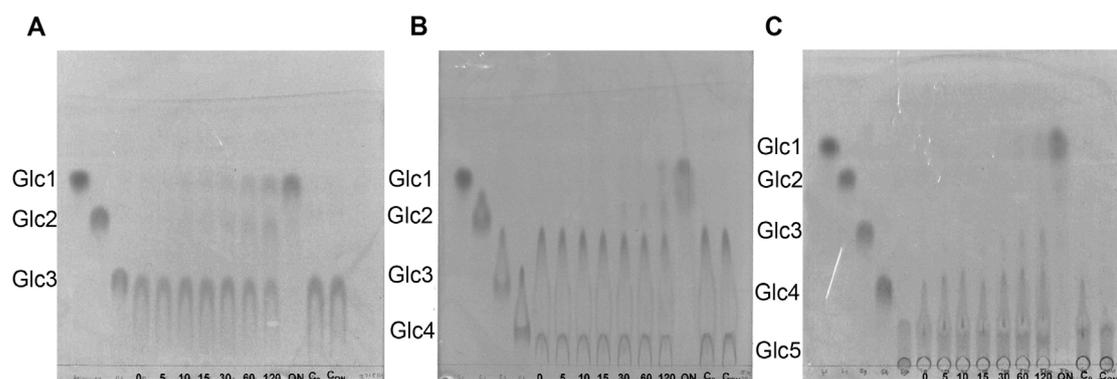
The steady state kinetic constant of rSSO3039 on 4Np- $\beta$ -GlcNAc, 4Np- $\beta$ -Glc, MU- $\beta$ -Glc and Mu- $\beta$ -GlcNAc were measured in standard conditions and listed in Tab. 2.7.4. This kinetic analysis confirmed that rSSO3039 is a *N*-acetyl- $\beta$ -glucosaminidase/ $\beta$ -glucosidase as the specificity constant for 4Np- $\beta$ -Glc is about 10-fold higher than that for 4Np- $\beta$ -GlcNAc (11.5 and 1.65 s<sup>-1</sup> mM<sup>-1</sup>, respectively) and for MU- $\beta$ -Glc is about 2-fold higher than that for MU- $\beta$ -GlcNAc (9.9 and 4.9 s<sup>-1</sup> mM<sup>-1</sup>, respectively), because of the lower  $K_M$ .

Substrate	$k_{cat}$ (s <sup>-1</sup> )	$K_M$ (mM)	$k_{cat}/K_M$ (s <sup>-1</sup> mM <sup>-1</sup> )
4NP- $\beta$ -GlcNAc	19.2±0.46	4.15±0.27	1.65
4NP- $\beta$ -Glc	9.15±0.19	0.8±0.08	11.4
MU- $\beta$ -GlcNAc	5.4±0.14	1.1±0.13	4.9
MU- $\beta$ -Glc	8.8±0.5	0.9 ±0.14	9.9

**Table 2.7.4:** Steady state kinetic constants of rSSO3039

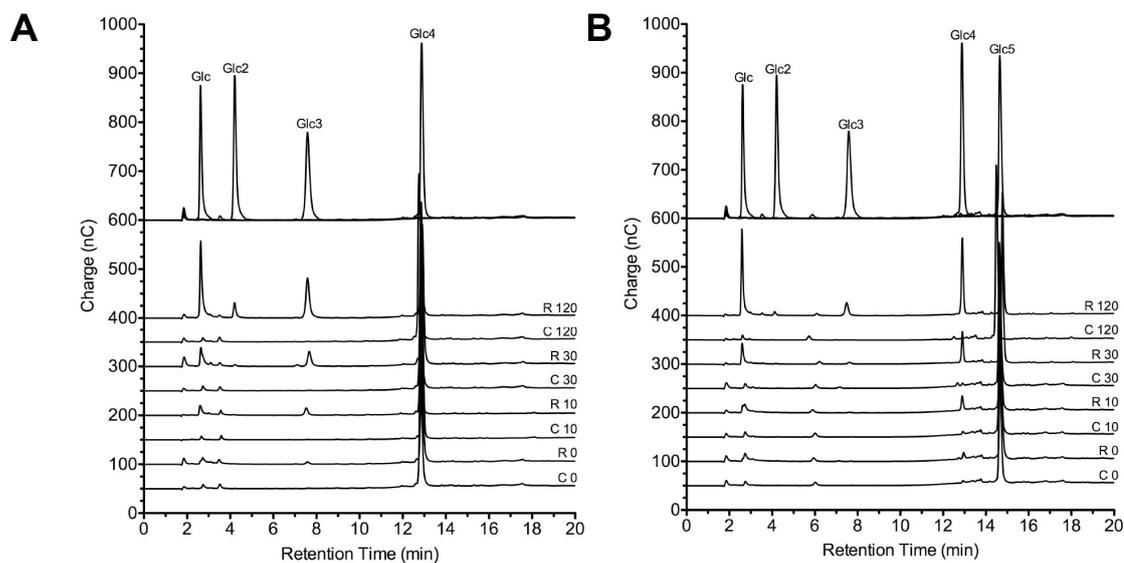
The hydrolytic activity of rSSO3039 was evaluated also on longer substrates, in particular on  $\beta$ -1,4 oligosaccharides of glucose (Glc) and *N*-acetylglucosamine (GlcNAc), namely celotriose (Glc3), cellotetraose (Glc4), cellopentaose (Glc5), tetra-*N*-acetylchitotetraose (GlcNAc4) and penta-*N*-acetylchitopentaose (GlcNAc5). The enzyme was active on all these substrates. The reaction mixtures with Glc3, Glc4 and Glc5 were first analysed by thin layer chromatography (TLC) showing in all the cases the progressive hydrolysis of the initial substrates with the accumulation of glucose after an overnight incubation (Fig. 2.7.12).

## 2.7 Results and discussion



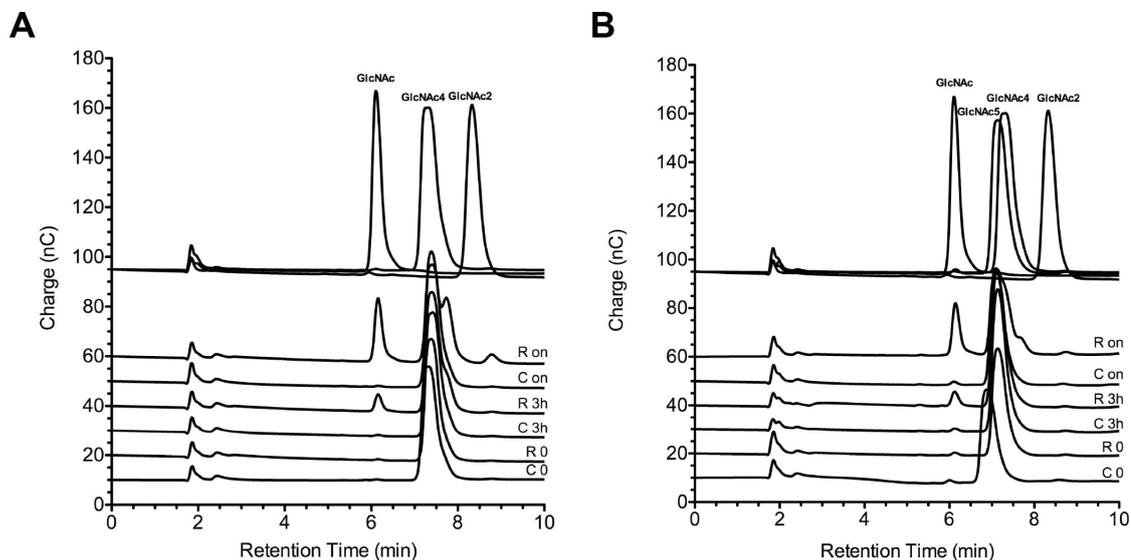
**Figure 2.7.12:** TLC analysis of the enzymatic assays on Glc3 (A), Glc4 (B) and Glc5 (C). The reaction mixtures were stopped in ice at the indicated times (minutes). C is the negative control. Glucose (Glc), cellobiose (Glc2), cellotriose (Glc3), cellotetraose (Glc4) and cellopentaose (Glc5) were used as markers

HPAEC-PAD analysis of the enzymatic reactions on Glc4 and Glc5 showed the release of glucose and of the oligosaccharide product shortened of one unit (cellotriose and cellotetraose from Glc4 and Glc5, respectively) after 10 min of incubation at 60°C (Fig. 2.7.13).



**Figure 2.7.13:** HPAEC-PAD analysis of the enzymatic assays on Glc4 (A) and Glc5 (B). The reaction mixtures were stopped in ice at the indicated times (minutes) and in all runs, C is the negative control, while R is the reaction mixture with rSSO3039. Glucose (Glc), cellobiose (Glc2), cellotriose (Glc3), cellotetraose (Glc4) and cellopentaose (Glc5) were used as markers

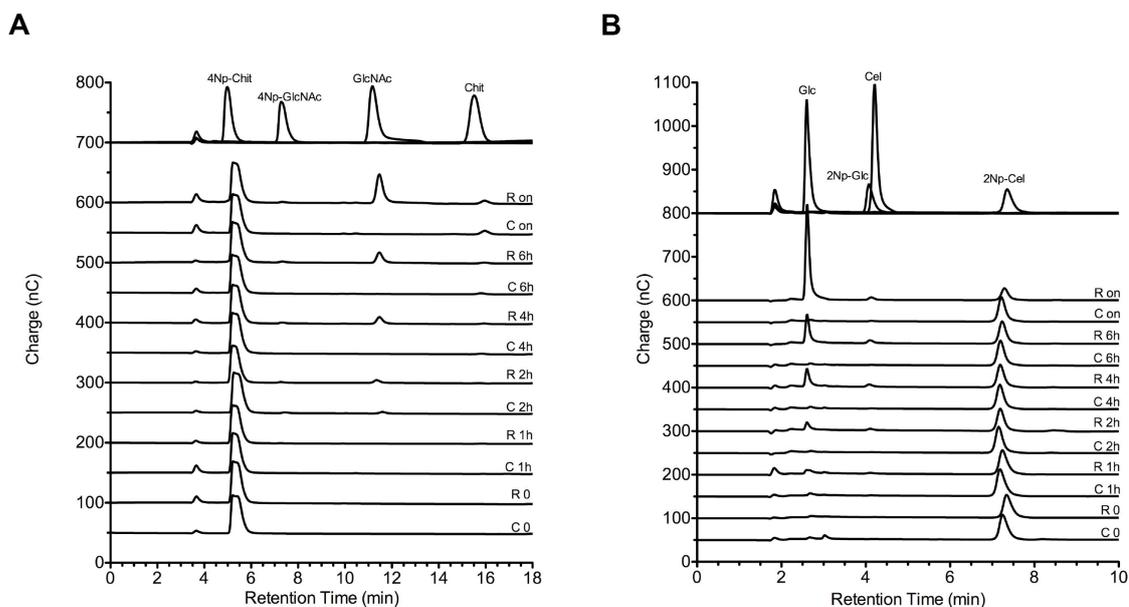
Similarly, rSSO3039 hydrolyzed GlcNAc4 and GlcNAc5 producing GlcNAc (Fig. 2.7.14) and the production of monosaccharides observed in the early times of the reaction strongly suggested that rSSO3039 was an *exo*-glycosidase acting from one of the ends of the substrates.



**Figure 2.7.14:** HPAEC-PAD analysis of the enzymatic assays on GlcNAc4 (**A**) and GlcNAc5 (**B**). The reaction mixtures were stopped in ice at the indicated times (hours) and in all runs, C is the negative control, while R is the reaction mixture with rSSO3039. *N*-acetylglucosamine (GlcNAc), chitobiose (GlcNAc2), tetra-*N*-acetylchitotetraose (GlcNAc4) and penta-*N*-acetylchitopentaose (GlcNAc5) were used as markers.

This hypothesis was confirmed also by the analysis of the catalytic activity of rSSO3039 on 2Np- $\beta$ -cellobioside and 4Np-*N,N'*-diacetyl- $\beta$ -chitobioside (2Np- $\beta$ -Cel and 4Np- $\beta$ -Chit, respectively). HPAEC-PAD analysis of the enzymatic reactions on these substrates indicated that rSSO3039 released the corresponding monosaccharide (either Glc or GlcNAc) after 2h incubation (Fig. 2.7.15), unequivocally demonstrating that rSSO3039 is an *exo*-glucosidase acting from the non-reducing end of the substrate.

## 2.7 Results and discussion

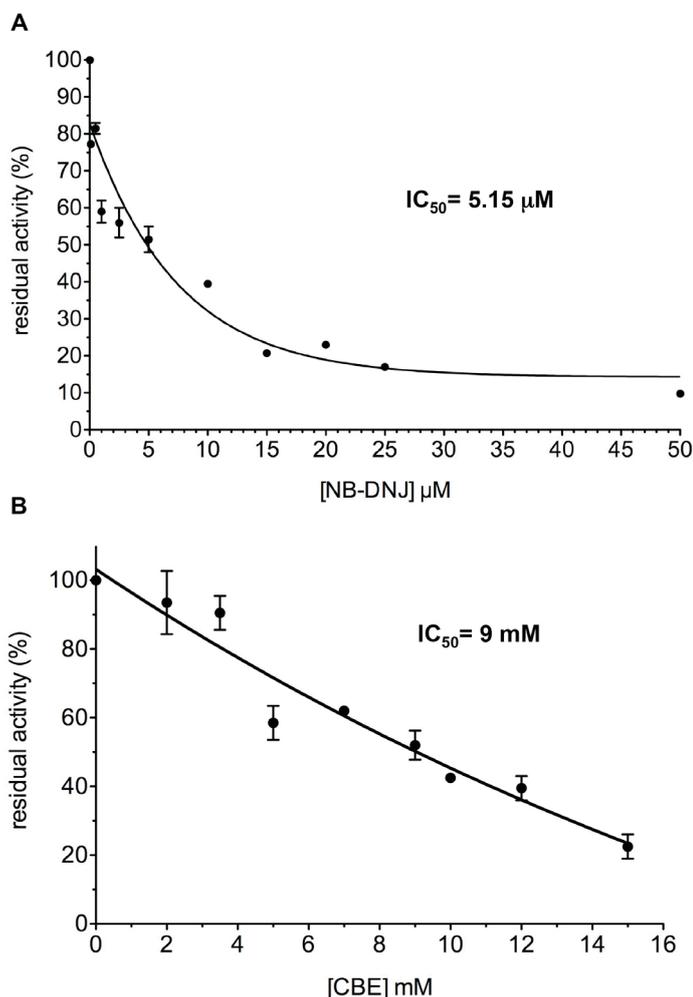


**Figure 2.7.15:** HPAEC-PAD analysis of the enzymatic assays on 4Np-Chit (**A**) and 2Np-Cel (**B**). The reaction mixtures were stopped in ice at the indicated times (hours) and in all runs, C is the negative control, while R is the reaction mixture with rSSO3039. 4Np-*N,N'*-diacetyl- $\beta$ -chitobioside (4Np-Chit), 4Np- $\beta$ -*N*-acetylglucosaminide (4Np-GlcNAc), *N*-acetylglucosamine (GlcNAc), chitobiose (Chit), glucose (Glc), 2Np- $\beta$ -glucoside (2Np-Glc), cellobiose (Cel) and 2Np- $\beta$ -cellobioside (2Np-Cel) were used as markers

### NB-DNJ and CBE inhibition profile

One of the peculiarities of GH116 enzymes is their sensitivity to known competitive inhibitors, such as conduritol  $\beta$ -epoxide (CBE) and *N*-butyldeoxynojirimycin (NB-DNJ) (Cobucci-Ponzano et al., 2010a). The inhibition profile of rSSO3039 revealed that the enzyme is inhibited by both NB-DNJ and CBE, but with different sensitivity. The IC<sub>50</sub> values calculated for both inhibitors differ of almost one order of magnitude: 5.15  $\mu$ M for NB-DNJ and 9 mM for CBE (Fig. 2.8.16). This sensitivity spectrum makes rSSO3039 different from other two characterized GH116 enzymes, the archaeal  $\beta$ -glycosidase SSO1353 and the human non-lysosomal glucosylceramidase (NLGase or GBA2). SSO1353 was inhibited by both NB-DNJ and CBE, like SSO3039, but it is worth noting the differences in term of sensitivity. In particular, both enzymes showed comparable mM sensitivity to CBE, but SSO3039 resulted more sensitive to NB-DNJ than SSO1353 (IC<sub>50</sub>= 5.15  $\mu$ M for SSO3039 and 1.5 mM for SSO1353) (Cobucci-Ponzano et al., 2010a). The comparison with GBA2 delineated a different picture as, conversely to SSO3039, the human enzyme is

insensitive to CBE and showed nM sensitivity to NB-DNJ (Boot et al., 2007). The different spectrum of sensitivity to NB-DNJ and CBE could represent an useful tool to distinguish and classify these enzyme within GH116.



**Figure 2.7.16:** Inhibition of rSSO3039 by NB-DNJ (**A**) and CBE (**B**)

### Phylogenetic analysis of GH116 enzymes

The CAZy family GH116 was created after the paper by Cobucci-Ponzano and co-workers reporting on the identification and characterization of a new  $\beta$ -glycosidase encoded by SSO1353 ORF of the hyperthermophilic crenarchaeon *S. solfataricus* P2 (Cobucci-Ponzano et al., 2010a). Presently this family includes  $\beta$ -glucosidases (EC 3.2.1.21),  $\beta$ -xylosidases (EC 3.2.1.37) and glucocerebrosidases (EC 3.2.1.45) from the three domains of life. The catalytic features were experimentally demonstrated for SSO1353 and indicate that the enzymes belonging to this family are *retaining* GHs following the classical Koshland double-displacement mechanism (see section 1.3 for a detailed description). The phylogenetic analysis (Fig. 2.7.17) has indicated that

## 2.7 Results and discussion

---

sequences included in the GH116 family can be subdivided into two major groups, one containing archaeal sequences and another one comprising mostly sequences from Cyanobacteria and Eukaryotes. Based on homology extent with SSO1353, the authors identified two other subgroups in the archaeal group: one includes SSO1353 homologs with identity >80% and the other one contains SSO1353 homologs with identities in the range of 21-33%, such as SSO2674 and SSO3039 from *S. solfataricus* P2. This family can be further subdivided into several subgroups according the organism source of sequences (Fig. 2.7.17). It is worth to note the presence of human non-lysosomal glucosylceramidase or  $\beta$ -glucosidase (GBA2) in the animal subgroup.

1.0

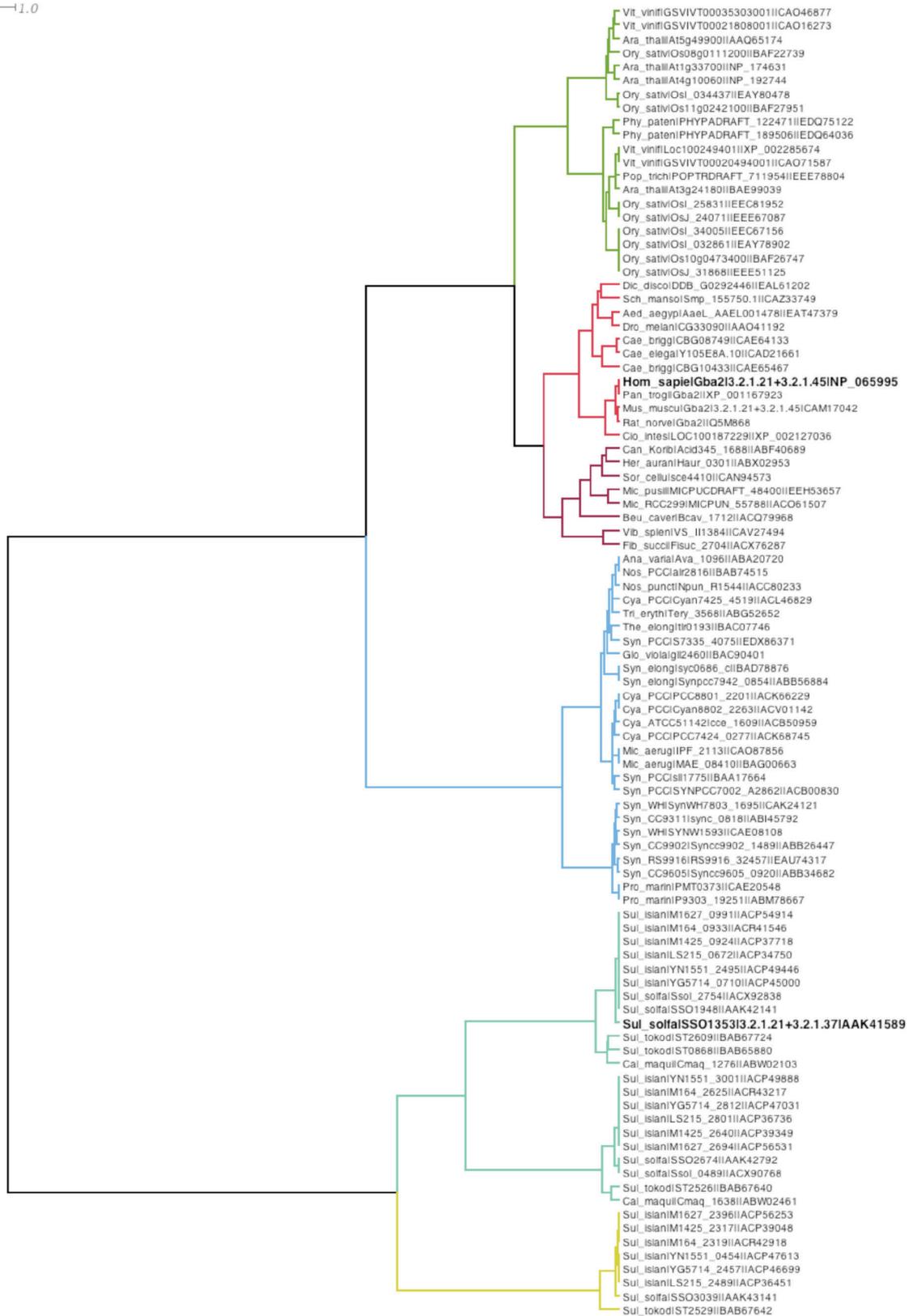


Figure 2.7.17: Phylogenetic tree of family GH116 proposed by Cobucci-Ponzano and co-workers (Cobucci-Ponzano et al., 2010a)

## 2.7 Results and discussion

---

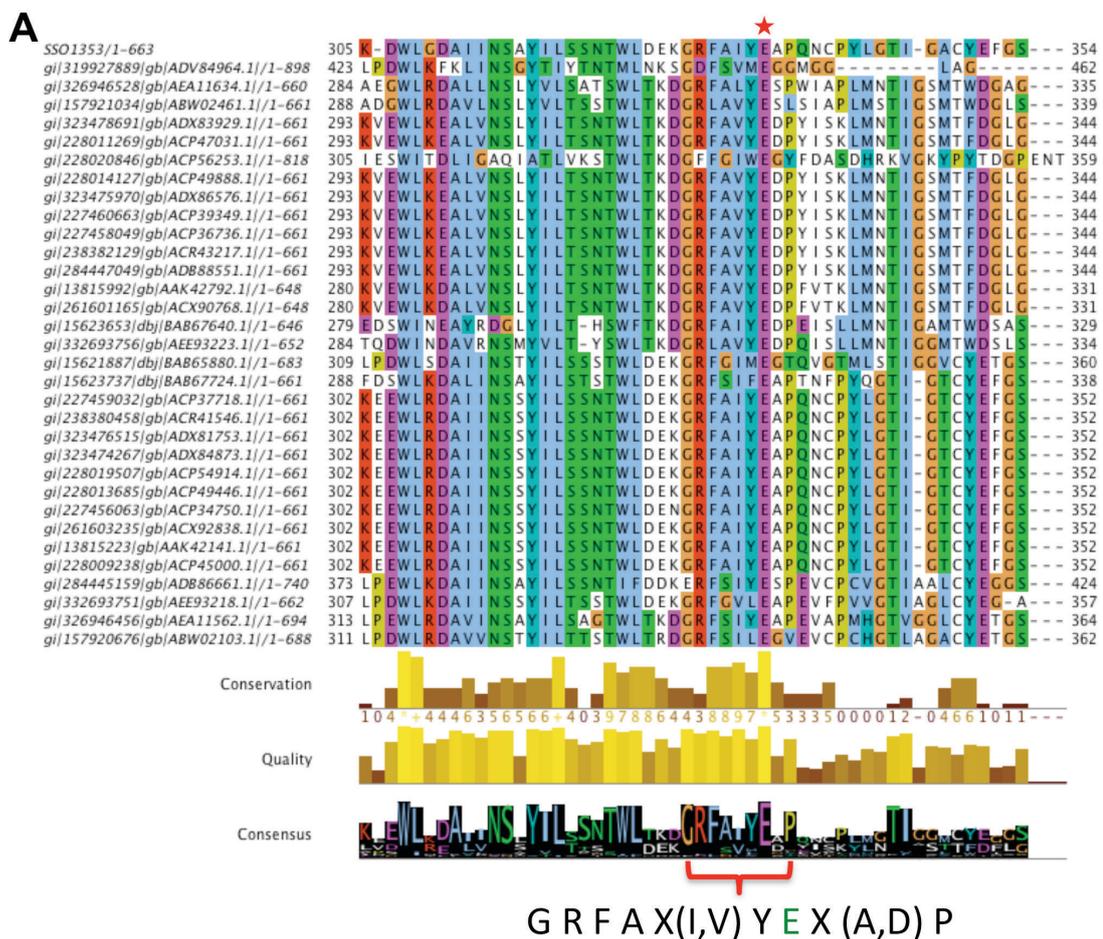
Following the biochemical characterization of SSO3039, a new phylogenetic analysis of GH116 enzymes was performed in collaboration with Dr. B. Henrissat, the curator of CAZy database. From this study it emerged that the CAZy family GH116 can be organized in three subfamilies (Fig. 2.7.18). The subfamily 1, indicated in blue, contains the human GBA2 (Boot et al., 2007), the subfamily 2, in green, includes SSO3039 described in this work and the archaeal  $\beta$ -glucosidase/xylosidase SSO1353 (Cobucci-Ponzano et al., 2010a) resides in the subfamily 3 depicted in yellow. Remarkably, the molecular characterization of the novel  $\beta$ -glucosidase/*N*-acetyl-glucosaminidase SSO3039 described here, gives experimental support to this phylogenetic analysis, demonstrating that the three subfamilies are functionally different. GH116 enzymes share the specificity for  $\beta$ -O-glycosides and the same reaction mechanism which is *retaining*-like (Cobucci-Ponzano et al., 2010a). However, we can recognize a different substrate specificity in each of the three subfamilies: enzymes included in subfamily 1 are specific for glucosylceramides, such as GBA2, those in subfamily 2 for *N*-acetyl-glucosaminides, such as SSO3039, and, lastly, those in subfamily 3 for xylosides, such as SSO1353. Additionally, as anticipated above, GH116 enzymes showed a different sensitivity to the competitive inhibitors NB-DNJ and CBE, and this feature further support their subdivision in three subfamilies. In particular, the subfamily 1 is insensitive to CBE and is inhibited by nM concentrations of NB-DNJ, while the subfamily 2 is inhibited by both compounds even if with a different sensitivity (IC<sub>50</sub> corresponds to 5.15  $\mu$ M for NB-DNJ and 9 mM for CBE). Instead, the subfamily 3 is sensitive to CBE conversely to subfamily 1 and differs from subfamily 2 for its lower sensitivity (mM) to NB-DNJ. The characterization of additional GH116 enzymes might be interesting to validate the effectiveness of this classification criterion.



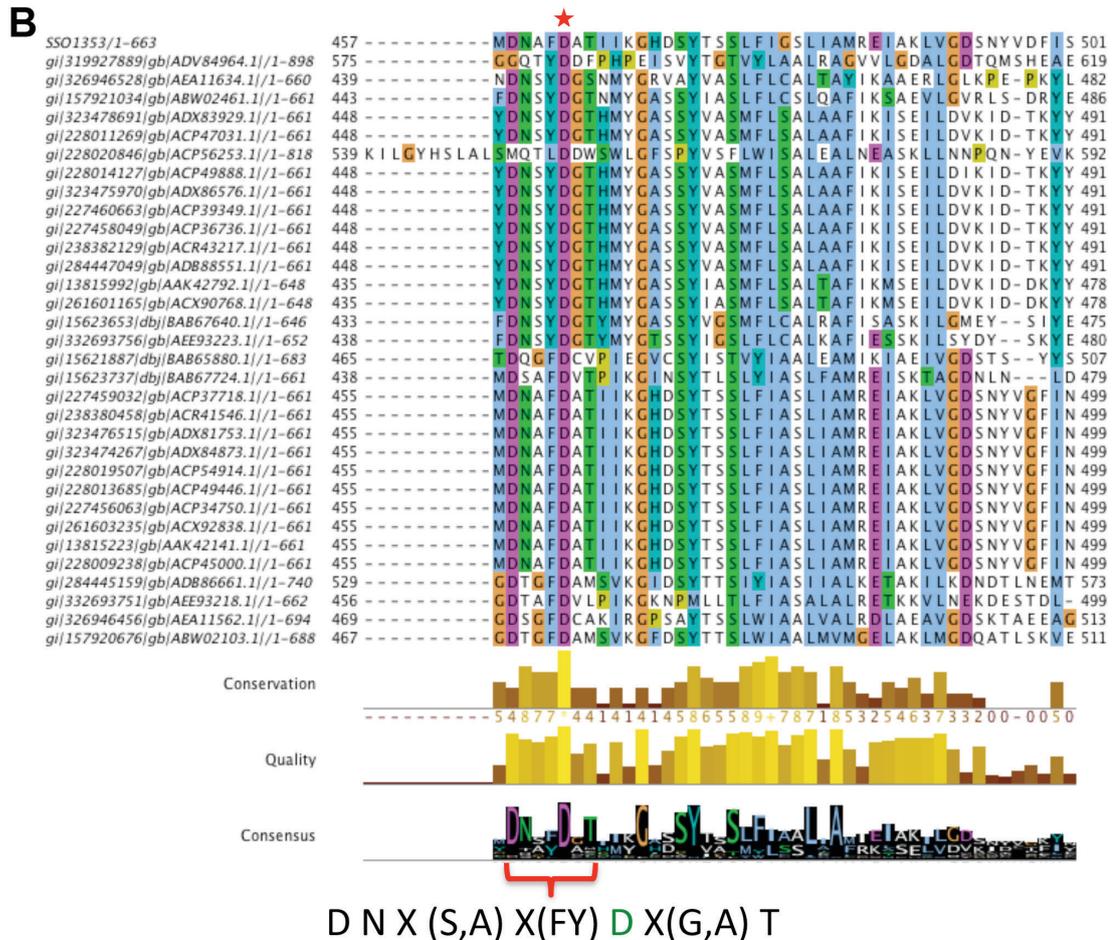
## 2.7 Results and discussion

Further bioinformatic studies were performed to predict potential active site signatures of the sequences grouped in the three subfamilies, namely amino acids surrounding the catalytic residues that may determine functional specificity.

The catalytic nucleophile and the general acid/base of SSO1353, belonging to subfamily 3, were identified experimentally as Glu335 and Asp462, respectively (Cobucci-Ponzano et al., 2010a). Multi-alignment of the sequences included in the subfamily 3 revealed the conserved sequence motifs GRFAX(I,V)YEX(A,D)P (Fig. 2.7.19A) and DNX(S,A)X(F,Y)D(X(G,A)T (Fig. 2.7.19B), surrounding the nucleophile Glu335 and the acid/base Asp462, respectively.



**Figure 2.7.19A:** Multi-alignment of sequences included in the subfamily 3 showing the conserved motif surrounding the nucleophile. Red star indicates Glu335

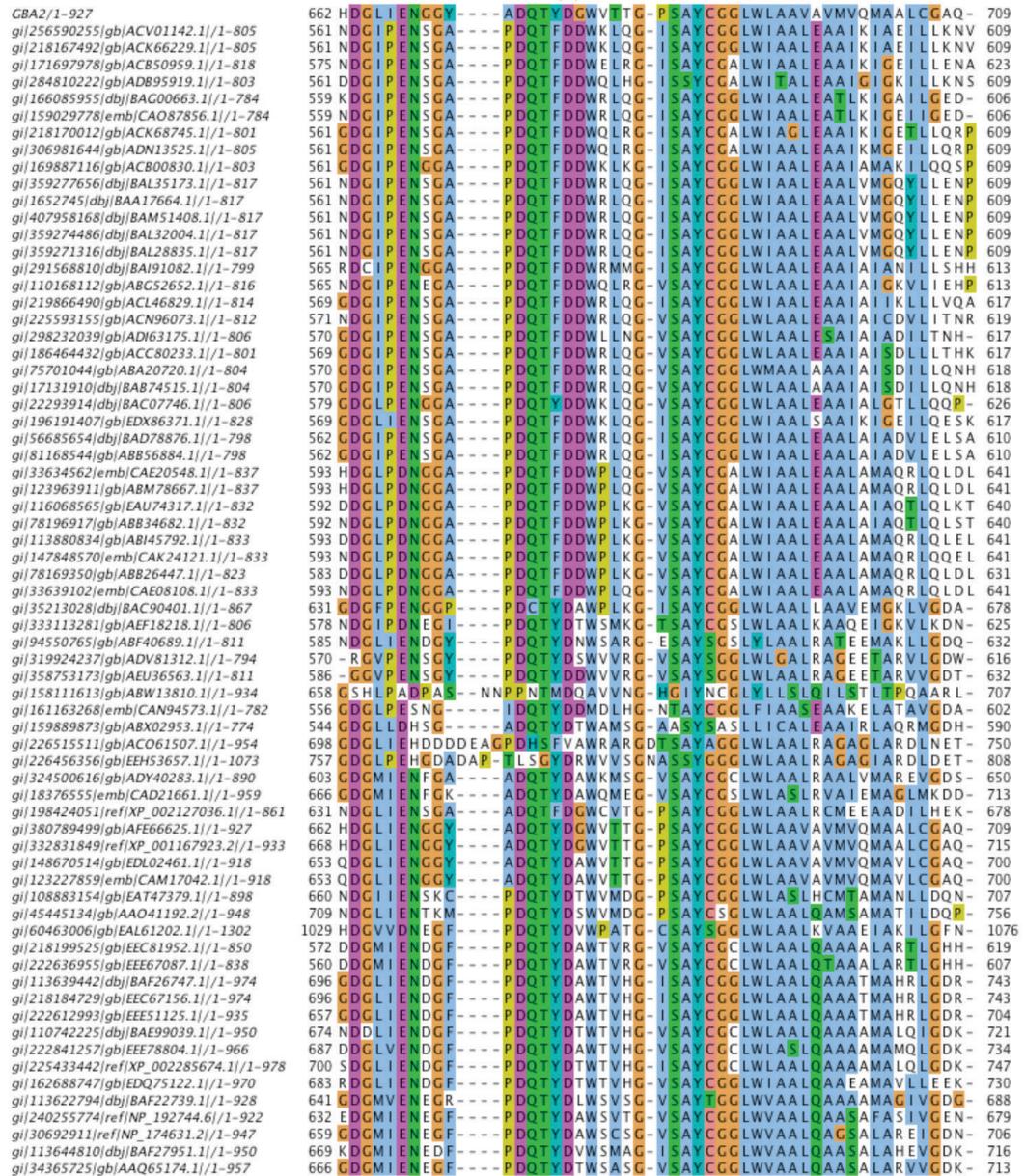


**Figure 2.7.19B:** Multi-alignment of sequences included in the subfamily 3 showing the conserved motif surrounding acid/base residue. Red star indicates Asp462

Multi-alignment of putative glucocerebrosidases from mammals, plants, and tunicates belonging to subfamily 1 of GH116 family allowed to predict GBA2 catalytic residues, Glu527 and Asp677 that act as nucleophile and acid/base, respectively (Cobucci-Ponzano et al., 2010a). Sequence analysis of subfamily 1 revealed the motifs GX(Q,R)FX(G,A,L)X(V,Y)LEX(C,G)X(L,I,V)X(D,E)Y (Fig. 2.7.20A) and PDQTX(F,Y)DX(D,A)W (Fig. 2.7.20B), surrounding the nucleophile and the acid/base, respectively.



B

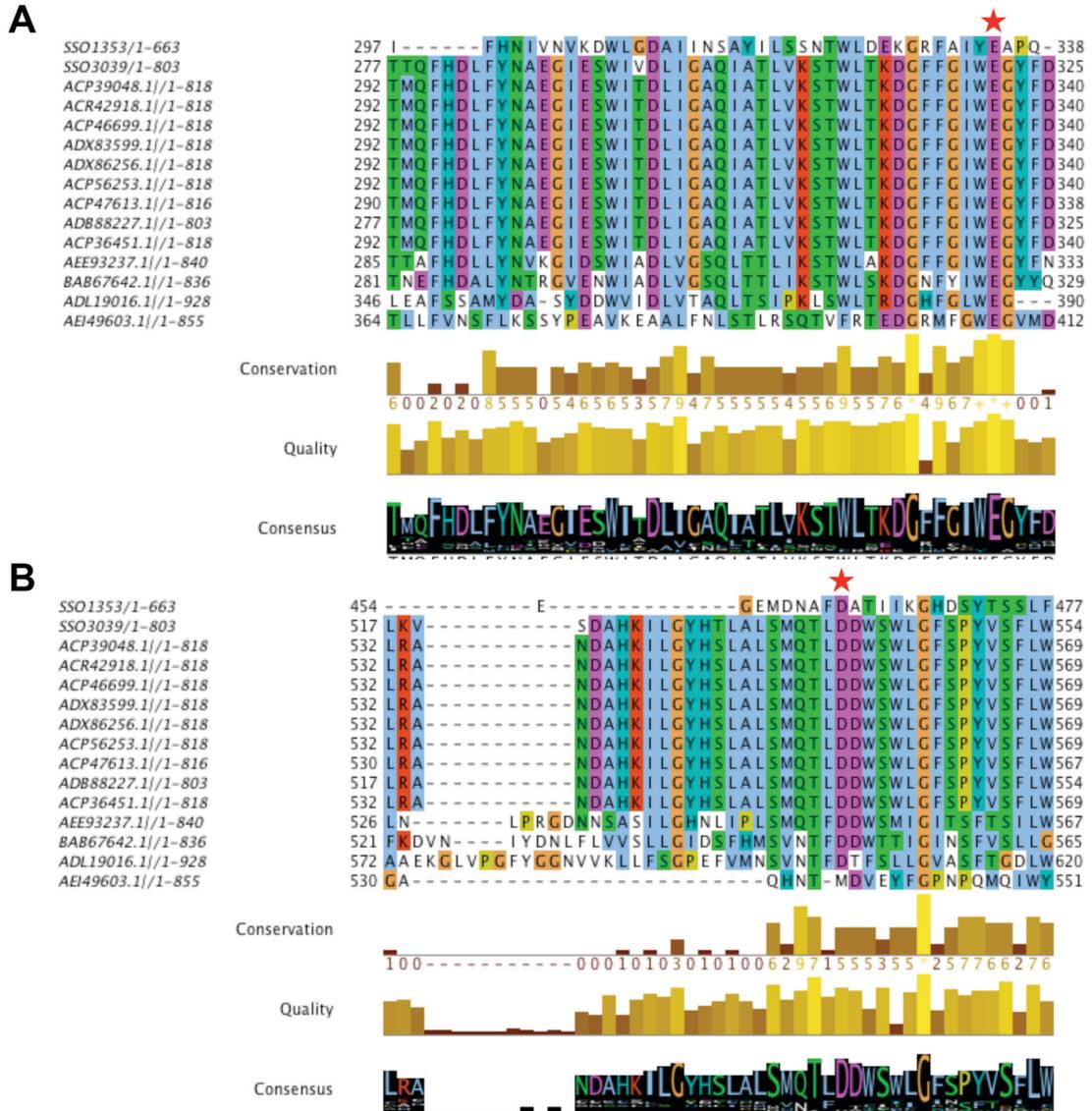


P D Q T X (F,Y) D X (D,A) W

Figure 2.7.20B: Multi-alignment of sequences included in the subfamily 1 showing the motif surrounding acid/base residue. Red star indicates Asp677

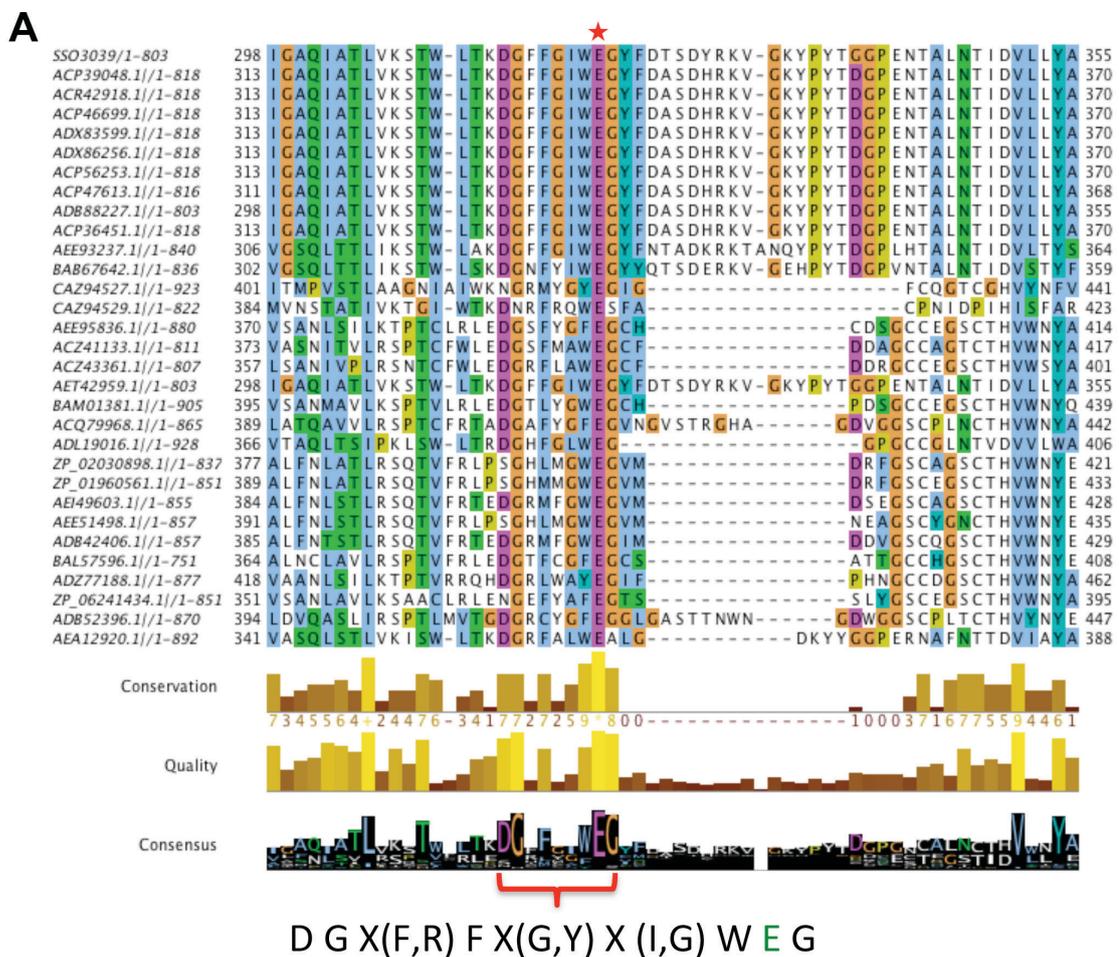
## 2.7 Results and discussion

In SSO3039 the nucleophile and the acid/base of the reaction were predicted to be Glu321 and Asp539, respectively, as observed in a multi-alignment of all GH116 enzymes (data not shown) and of SSO1353 with enzymes belonging to subfamily 2 (Fig. 2.7.21).



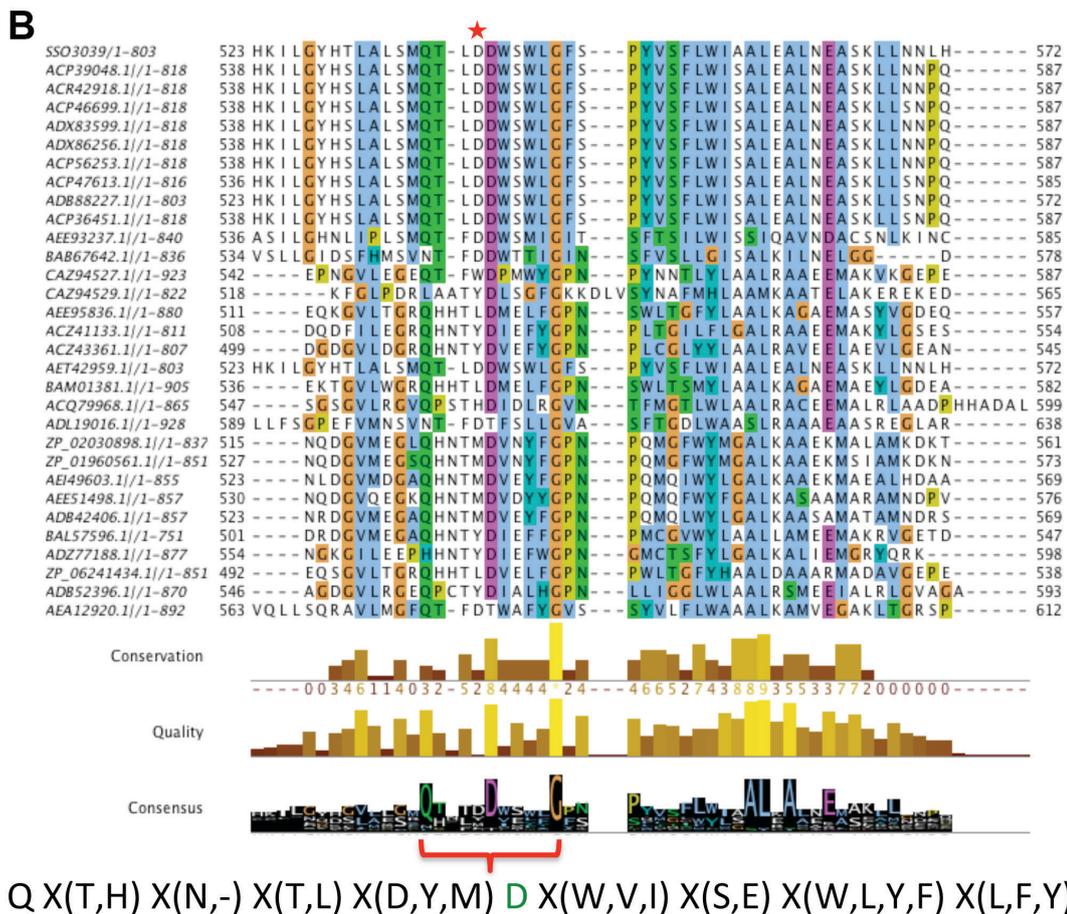
**Figure 2.7.21:** Multi-alignment of SSO1353 with enzymes belonging to subfamily 2. The residues corresponding to the nucleophile Glu335 (A) and acid/base Asp462 (B) are indicated with a red star

Multi-alignment of the sequences included in the subfamily 2 revealed the almost conserved motif DGX(F,R)FX(G,Y)X(I,G)WEG surrounding the putative nucleophile Glu321 (Fig. 2.7.22A), while the putative acid/base residue is located in a region showing a more variable sequence pattern (Fig. 2.7.22B). Moreover, it is worth to mention that the putative acid/base Asp539 is followed by another aspartate, Asp540, conserved among the enzymes belonging to the subfamily 2. Detailed biochemical studies are required to determine unequivocally the catalytic residues of these enzymes and to elucidate a possible role of Asp540.



**Figure 2.7.22A:** Multi-alignment of sequences included in the subfamily 2 showing the motifs surrounding the putative nucleophile Glu321, which is indicated with a red star

## 2.7 Results and discussion



**Figure 2.7.22B:** Multi-alignment of sequences included in the subfamily 2 showing the motif surrounding the putative acid/base residue Asp539, which is indicated with a red star

Comparing the motifs identified in the three subfamilies, it is possible to disclose an almost conserved sequence pattern surrounding the nucleophile (Fig. 2.7.23A). In addition, highlighting the residue preceding the nucleophilic glutamate, it is possible to note that its nature varies among the different subfamilies (Fig. 2.7.23A). In particular, the enzymes in subfamily 1 carry a leucine, an hydrophobic amino acid, while those in subfamily 2 a tryptophan, an aromatic amino acid and, lastly, a hydrophilic tyrosine precedes glutamate of subfamily 3 enzymes. Nevertheless, detailed biochemical analyses are required to establish an eventual role of this residue. Conversely, the sequence pattern surrounding the acid/base residues of the three subfamilies is more heterogeneous (Fig. 2.7.23B). Our findings can now allow the planning of more detailed site-directed mutagenesis studies to better understand the molecular bases of the substrate recognition of enzymes belonging to the three different subfamilies.

### A

Subfamily 1 **G** X(Q,R) **F** X(G,A,L) X(V,Y) L **E** X(C,G) X(L,I,V) X(D,E) Y  
 Subfamily 2 **D** **G** X(F,R) **F** X(G,Y) X(I,G) W **E** **G**  
 Subfamily 3 **G** R **F** A X(I,V) Y **E** X(A,D) P

### B

Subfamily 1 P **D** Q T X(F,Y) **D** X(D,A) W  
 Subfamily 2 Q X(T,H) X(N,-) X(T,L) X(D,Y,M) **D** X(W,V,I) X(S,E) X(W,L,Y,F) X(L,F,Y) **G**  
 Subfamily 3 **D** N X(S,A) X(FY) **D** X(G,A) T

**Figure 2.7.23:** Comparison of the sequence patterns surrounding the nucleophile (**A**) and the acid/base (**B**) residues of the three GH116 subfamilies. The catalytic residues are indicated in green, while the invariant amino acids are indicated in red

In conclusion, our study, which for the first time identified an *N*-acetylglucosaminidase from archaea, demonstrated that the detailed biochemical characterization of enzymes from natural sources is of utmost importance to functionally support genomic data, to increase our knowledge on enzyme classes, and to complete the information provided by bioinformatic data banks. In order to elucidate the function in vivo of these enzymes, it is needed a more detailed functional and structural characterization of this family. For instance, understanding the role of GBA2 and its interplay with the lysosomal counterpart, if there is, could be useful to set up new therapeutic strategies for Gaucher disease. But the production of GBA2 in abundant and homogenous form is difficult to achieve. These limits might be overcome by using GBA2 homologs from hyperthermophilic archaea that are easily produced in recombinant form and are stable in various conditions. The structural and functional studies on hyperthermophilic enzymes can provide useful informations that might be easily extended to the human counterpart.

## **2.8 REFERENCES**

- Abu-Qarn M, Eichler J, Sharon N. 2008. Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Current Opinion in Structural Biology* **18**:544–550.
- Akeboshi H, Kasahara Y, Tsuji D, Itoh K, Sakuraba H, Chiba Y, Jigami Y. 2009. Production of human beta-hexosaminidase A with highly phosphorylated N-glycans by the overexpression of the *Ogataea minuta* MNN4 gene. *Glycobiology* **19**:1002–9.
- Allers T, Mevarech M. 2005. Archaeal genetics - the third way. *Nature reviews. Genetics* **6**:58–73.
- Altmann F, Schwihla H, Staudacher E, Glössl J, März L. 1995. Insect cells contain an unusual, membrane-bound beta-N-acetylglucosaminidase probably involved in the processing of protein N-glycans. *The Journal of biological chemistry* **270**:17344–9.
- Andronopoulou E, Vorgias CE. 2003. Purification and characterization of a new hyperthermostable, allosamidin-insensitive and denaturation-resistant chitinase from the hyperthermophilic archaeon *Thermococcus chitonophagus*. *Extremophiles : life under extreme conditions* **7**:43–53.
- Andronopoulou E, Vorgias CE. 2004. Isolation, cloning, and overexpression of a chitinase gene fragment from the hyperthermophilic archaeon *Thermococcus chitonophagus*: semi-denaturing purification of the recombinant peptide and investigation of its relation with other chitinases. *Protein expression and purification* **35**:264–71.
- Bacik J-P, Whitworth GE, Stubbs KA, Vocadlo DJ, Mark BL. 2012. Active site plasticity within the glycoside hydrolase NagZ underlies a dynamic mechanism of substrate distortion. *Chemistry & biology* **19**:1471–82.
- Balcewich MD, Stubbs KA, He Y, James TW, Davies GJ, Vocadlo DJ, Mark BL. 2009. Insight into a strategy for attenuating AmpC-mediated beta-lactam resistance: structural basis for selective inhibition of the glycoside hydrolase NagZ. *Protein science : a publication of the Protein Society* **18**:1541–51.
- Barns SM, Delwiche CF, Palmer JD, Pace NR. 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proceedings of the National Academy of Sciences of the United States of America* **93**:9188–93.
- Barry RC, Young MJ, Stedman KM, Dratz E a. 2006. Proteomic mapping of the hyperthermophilic and acidophilic archaeon *Sulfolobus solfataricus* P2. *Electrophoresis* **27**:2970–83.
- Boot RG, Verhoek M, Donker-Koopman W, Strijland A, Van Marle J, Overkleeft HS, Wennekes T, Aerts JMFG. 2007. Identification of the non-lysosomal

## 2.8 References

---

- glucosylceramidase as beta-glucosidase 2. *The Journal of biological chemistry* **282**:1305–12.
- Bowers KJ, Wiegel J. 2011. Temperature and pH optima of extremely halophilic archaea: a mini-review. *Extremophiles: life under extreme conditions* **15**:119–28.
- Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* **72**:248–54.
- Brochier-Armanet C, Gribaldo S, Zivanovic Y, Confalonieri F, Forterre P. 2005. Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome biology* **6**:R42.
- Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P. 2008. Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature reviews. Microbiology* **6**:245–52.
- Brock TD, Brock KM, Belly RT, Weiss RL. 1972. Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Archiv für Mikrobiologie* **84**:54–68.
- Brouns SJJ, Smits N, Wu H, Snijders APL, Wright PC, De Vos WM, Van der Oost J. 2006. Identification of a novel alpha-galactosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Journal of bacteriology* **188**:2392–9.
- Cardona S, Remonsellez F, Guilliani N, Jerez CA. 2001. The glycogen-bound polyphosphate kinase from *Sulfolobus acidocaldarius* is actually a glycogen synthase. *Applied and environmental microbiology* **67**:4773–80.
- Cavicchioli R. 2011. Archaea-timeline of the third domain. *Nature reviews. Microbiology* **9**:51–61.
- Chaban B, Ng SYM, Jarrell KF. 2006. Archaeal habitats - from the extreme to the ordinary **116**:73–116.
- Cheng Q, Li H, Merdek K, Park JT. 2000. Molecular characterization of the beta-N-acetylglucosaminidase of *Escherichia coli* and its role in cell wall recycling. *Journal of bacteriology* **182**:4836–40.
- Chong PK, Wright PC. 2005. Identification and characterization of the *Sulfolobus solfataricus* P2 proteome. *Journal of proteome research* **4**:1789–1798.
- Cobucci-Ponzano B, Trincone A, Giordano A, Rossi M, Moracci M. 2003. Identification of an archaeal alpha-L-fucosidase encoded by an interrupted gene. Production of a functional enzyme by mutations mimicking

- programmed -1 frameshifting. *The Journal of biological chemistry* **278**:14622–31.
- Cobucci-Ponzano B, Aurilia V, Riccio G, Henrissat B, Coutinho PM, Strazzulli A, Padula A, Corsaro MM, Pieretti G, Pocsfalvi G, Fiume I, Cannio R, Rossi M, Moracci M. 2010a. A new archaeal beta-glycosidase from *Sulfolobus solfataricus*: seeding a novel retaining beta-glycan-specific glycoside hydrolase family along with the human non-lysosomal glucosylceramidase GBA2. *The Journal of biological chemistry* **285**:20691–703.
- Cobucci-Ponzano B, Conte F, Strazzulli A, Capasso C, Fiume I, Pocsfalvi G, Rossi M, Moracci M. 2010b. The molecular characterization of a novel GH38  $\alpha$ -mannosidase from the crenarchaeon *Sulfolobus solfataricus* revealed its ability in de-mannosylating glycoproteins. *Biochimie* **92**:1895–907.
- Cole RN, Hart GW. 1999. Glycosylation sites flank phosphorylation sites on synapsin I: O-linked *N*-acetylglucosamine residues are localized within domains mediating synapsin I interactions. *Journal of neurochemistry* **73**:418–28.
- Comtesse N, Maldener E, Meese E. 2001. Identification of a nuclear variant of MGEA5, a cytoplasmic hyaluronidase and a beta-*N*-acetylglucosaminidase. *Biochemical and biophysical research communications* **283**:634–40.
- Dennis RJ, Taylor EJ, Macauley MS, Stubbs KA, Turkenburg JP, Hart SJ, Black GN, Vocadlo DJ, Davies GJ. 2006. Structure and mechanism of a bacterial beta-glucosaminidase having O-GlcNAcase activity. *Nature structural & molecular biology* **13**:365–71.
- De Rosa M, Gambacorta a, Nicolaus B, Giardina P, Poerio E, Buonocore V. 1984. Glucose metabolism in the extreme thermoacidophilic archaeobacterium *Sulfolobus solfataricus*. *The Biochemical journal* **224**:407–14.
- Ernst HA, Lo Leggio L, Willemoës M, Leonard G, Blum P, Larsen S. 2006. Structure of the *Sulfolobus solfataricus* alpha-glucosidase: implications for domain conservation and substrate recognition in GH31. *Journal of molecular biology* **358**:1106–24.
- Fütterer O, Angelov a, Liesegang H, Gottschalk G, Schleper C, Schepers B, Dock C, Antranikian G, Liebl W. 2004. Genome sequence of *Picrophilus torridus* and its implications for life around pH 0. *Proceedings of the National Academy of Sciences of the United States of America* **101**:9091–6.
- Gao Y, Wells L, Comer FI, Parker GJ, Hart GW. 2001. Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-*N*-acetylglucosaminidase from human brain. *The Journal of biological chemistry* **276**:9838–45.

## 2.8 References

---

- Geisler C, Aumiller JJ, Jarvis DL. 2008. A fused lobes gene encodes the processing beta-*N*-acetylglucosaminidase in Sf9 cells. *The Journal of biological chemistry* **283**:11330–9.
- Greve LC, Prody GA, Hedrick JL. 1985. *N*-acetyl-beta-D-glucosaminidase activity in the cortical granules of *Xenopus laevis* eggs. *Gamete Research* **12**:305–312.
- Grogan DW. 1989. Phenotypic characterization of the archaeobacterial genus *Sulfolobus*: comparison of five wild-type strains. *Journal of bacteriology* **171**:6710–9.
- Gutternigg M, Kretschmer-Lubich D, Paschinger K, Rendić D, Hader J, Geier P, Ranftl R, Jantsch V, Lochnit G, Wilson IBH. 2007. Biosynthesis of truncated N-linked oligosaccharides results from non-orthologous hexosaminidase-mediated mechanisms in nematodes, plants, and insects. *The Journal of biological chemistry* **282**:27825–40.
- Gutternigg M, Rendić D, Voglauer R, Iskratsch T, Wilson IBH. 2009. Mammalian cells contain a second nucleocytoplasmic hexosaminidase. *The Biochemical journal* **419**:83–90.
- Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of molecular biology* **166**:557–80.
- Hart GW, Housley MP, Slawson C. 2007. Cycling of O-linked beta-*N*-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **446**:1017–22.
- Hochstein L, Stan-Lotter H. 1992. Purification and properties from *Sulfolobus solfataricus* of an ATPase. *Archives of biochemistry and biophysics* **295**:153–160.
- Hogenkamp DG, Arakane Y, Kramer KJ, Muthukrishnan S, Beeman RW. 2008. Characterization and expression of the beta-*N*-acetylhexosaminidase gene family of *Tribolium castaneum*. *Insect biochemistry and molecular biology* **38**:478–89.
- Horsch M, Mayer C, Sennhauser U, Rast DM. 1997. Beta-*N*-acetylhexosaminidase: a target for the design of antifungal agents. *Pharmacology & therapeutics* **76**:187–218.
- Hou Y, Tse R, Mahuran DJ. 1996. Direct determination of the substrate specificity of the alpha-active site in heterodimeric beta-hexosaminidase A. *Biochemistry* **35**:3963–9.
- Huang Y, Krauss G, Cottaz S, Driguez H, Lipps G. 2005. A highly acid-stable and thermostable endo-beta-glucanase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *The Biochemical journal* **385**:581–8.

- Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO. 2002. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**:63–7.
- Jiang Y-L, Yu W-L, Zhang J-W, Frolet C, Di Guilmi A-M, Zhou C-Z, Vernet T, Chen Y. 2011. Structural basis for the substrate specificity of a novel  $\beta$ -N-acetylhexosaminidase StrH protein from *Streptococcus pneumoniae* R6. *The Journal of biological chemistry* **286**:43004–12.
- Jonuscheit M, Martusewitsch E, Stedman KM, Schleper C. 2003. A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. *Molecular microbiology* **48**:1241–52.
- Kaper T, Brouns SJJ, Geerling ACM, De Vos WM, Van der Oost J. 2002. DNA family shuffling of hyperthermostable beta-glycosidases. *The Biochemical journal* **368**:461–70.
- Kashefi K, Lovley DR. 2003. Extending the upper temperature limit for life. *Science (New York, N.Y.)* **301**:934.
- Kim MS, Park JT, Kim YW, Lee HS, Nyawira R, Shin HS, Park CS, Yoo SH, Kim YR, Moon TW, Park KH. 2004. Properties of a novel thermostable glucoamylase from the hyperthermophilic archaeon *Sulfolobus solfataricus* in relation to starch processing. *Applied and environmental microbiology* **70**:3933–40.
- Kim YK, Kim KR, Kang DG, Jang SY, Kim YH, Cha HJ. 2009. Suppression of beta-N-acetylglucosaminidase in the N-glycosylation pathway for complex glycoprotein formation in *Drosophila* S2 cells. *Glycobiology* **19**:301–8.
- Könneke M, Bernhard AE, De la Torre JR, Walker CB, Waterbury JB, Stahl DA. 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**:543–6.
- Kytzia H, Sandhoff K. 1985. Evidence for Two Different Active Sites on Human. *The Journal of biological chemistry* **260**:7568–7572.
- Lamarre-Vincent N, Hsieh-Wilson LC. 2003. Dynamic glycosylation of the transcription factor CREB: a potential role in gene regulation. *Journal of the American Chemical Society* **125**:6612–3.
- Lemieux MJ, Mark BL, Cherney MM, Withers SG, Mahuran DJ, James MNG. 2006. Crystallographic structure of human beta-hexosaminidase A: interpretation of Tay-Sachs mutations and loss of GM2 ganglioside hydrolysis. *Journal of molecular biology* **359**:913–29.
- Léonard R, Rendic D, Rabouille C, Wilson IBH, Prétat T, Altmann F. 2006. The *Drosophila* fused lobes gene encodes an N-acetylglucosaminidase

## 2.8 References

---

- involved in *N*-glycan processing. *The Journal of biological chemistry* **281**:4867–75.
- Limauro D, Cannio R, Fiorentino G, Rossi M, Bartolucci S. 2001. Identification and molecular characterization of an endoglucanase gene, *celS*, from the extremely thermophilic archaeon *Sulfolobus solfataricus*. *Extremophiles* **5**:213–219.
- Littlechild JA 2011. Thermophilic archaeal enzymes and applications in biocatalysis. *Biochemical Society transactions* **39**:155–8.
- Litzinger S, Fischer S, Polzer P, Diederichs K, Welte W, Mayer C. 2010. Structural and kinetic analysis of *Bacillus subtilis* *N*-acetylglucosaminidase reveals a unique Asp-His dyad mechanism. *The Journal of biological chemistry* **285**:35675–84.
- Liu T, Yan J, Yang Q. 2012. Comparative Biochemistry of GH3, GH20 and GH84  $\beta$ -*N*-acetyl-D-hexosaminidases and Recent Progress in Selective Inhibitor Discovery. *Current Drug Targets*:512–525.
- Macauley MS, Stubbs KA, Vocadlo DJ. 2005. O-GlcNAcase catalyzes cleavage of thioglycosides without general acid catalysis. *Journal of the American Chemical Society* **127**:17202–3.
- Mahuran DJ. 1999. Biochemical consequences of mutations causing the GM2 gangliosidosis. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1455**:105–138.
- Maier T, Strater N, Schuette CG, Klingenstein R, Sandhoff K, Saenger W. 2003. The X-ray crystal structure of human beta-hexosaminidase B provides new insights into Sandhoff disease. *Journal of molecular biology* **328**:669–81.
- Mark BL, Vocadlo DJ, Knapp S, Triggs-Raine BL, Withers SG, James MN. 2001. Crystallographic evidence for substrate-assisted catalysis in a bacterial beta-hexosaminidase. *The Journal of biological chemistry* **276**:10330–7.
- Mark BL, James MN. 2002. Anchimeric assistance in hexosaminidases. *Canadian Journal of Chemistry* **80**:1064–1074.
- Mark BL, Mahuran DJ, Cherney MM, Zhao D, Knapp S, James MNG. 2003. Crystal structure of human beta-hexosaminidase B: understanding the molecular basis of Sandhoff and Tay-Sachs disease. *Journal of molecular biology* **327**:1093–109.
- Martino S, Marconi P, Tancini B, Dolcetta D, De Angelis MGC, Montanucci P, Bregola G, Sandhoff K, Bordignon C, Emiliani C, Manservigi R, Orlacchio A. 2005. A direct gene transfer strategy via brain internal capsule reverses the biochemical defect in Tay-Sachs disease. *Human molecular genetics* **14**:2113–23.

- Maruta K, Mitsuzumi H, Nakada T, Kubota M, Chaen H, Fukuda S, Sugimoto T, Kurimoto M. 1996. Cloning and sequencing of a cluster of genes encoding novel enzymes of trehalose biosynthesis from thermophilic archaeobacterium *Sulfolobus acidocaldarius*. *Biochimica et biophysica acta* **1291**:177–81.
- Maurelli L, Giovane A, Esposito A, Moracci M, Fiume I, Rossi M, Morana A. 2008. Evidence that the xylanase activity from *Sulfolobus solfataricus* Oalpha is encoded by the endoglucanase precursor gene (sso1354) and characterization of the associated cellulase activity. *Extremophiles: life under extreme conditions* **12**:689–700.
- Mayer C, Vocadlo DJ, Mah M, Rupitz K, Stoll D, Warren RAJ, Withers SG. 2006. Characterization of a beta-N-acetylhexosaminidase and a beta-N-acetylglucosaminidase/beta-glucosidase from *Cellulomonas fimi*. *Febs Journal* **273**:2929–2941.
- Moracci M, Ciaramella M, Nucci R, Pearl LH, Sanderson I, Trincone A, Rossi M. 1994. Thermostable beta-glycosidase from *Sulfolobus-Solfataricus*. *Biocatalysis* **11**:89–103.
- Moracci M, Cobucci Ponzano B, Trincone A, Fusco S, De Rosa M, Van Der Oost J, Sensen CW, Charlebois RL, Rossi M. 2000. Identification and molecular characterization of the first alpha-xylosidase from an archaeon. *The Journal of biological chemistry* **275**:22082–9.
- Morana A, Paris O, Maurelli L, Rossi M, Cannio R. 2007. Gene cloning and expression in *Escherichia coli* of a bi-functional beta-D-xylosidase/alpha-L-arabinosidase from *Sulfolobus solfataricus* involved in xylan degradation. *Extremophiles: life under extreme conditions* **11**:123–32.
- Muramatsu T. 1968. N-acetylhexosaminidases from a gastropod, *Turbo cornutus*. *Journal of biochemistry* **64**:521–31.
- Nakamura T, Mine S, Hagihara Y, Ishikawa K, Ikegami T, Uegaki K. 2008. Tertiary structure and carbohydrate recognition by the chitin-binding domain of a hyperthermophilic chitinase from *Pyrococcus furiosus*. *Journal of molecular biology* **381**:670–80.
- Namboori SC, Graham DE. 2008. Acetamido sugar biosynthesis in the Euryarchaea. *Journal of bacteriology* **190**:2987–96.
- Oku T, Ishikawa K. 2006. Analysis of the Hyperthermophilic Chitinase from *Pyrococcus furiosus*: Activity toward Crystalline Chitin. *Bioscience, Biotechnology, and Biochemistry* **70**:1696–1701.
- Peyfoon E, Meyer B, Hitchen PG, Panico M, Morris HR, Haslam SM, Albers SV, Dell A. 2010. The S-layer glycoprotein of the crenarchaeote *Sulfolobus acidocaldarius* is glycosylated at multiple sites with chitobiose-linked N-glycans. *Archaea (Vancouver, B.C.)* **2010**.

## 2.8 References

---

- Pisani FM, Rella R, Raia CA, Rozzo C, Nucci R, Gambacorta A, De Rosa M, Rossi M. 1990. Thermostable beta-galactosidase from the archaeobacterium *Sulfolobus solfataricus*. Purification and properties. *European journal of biochemistry / FEBS* **187**:321–8.
- Prag G, Papanikolau Y, Tavlas G, Vorgias CE, Petratos K, Oppenheim a B. 2000. Structures of chitobiase mutants complexed with the substrate Di-N-acetyl-D-glucosamine: the catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *Journal of molecular biology* **300**:611–7.
- Rao FV, Dorfmueller HC, Villa F, Allwood M, Eggleston IM, Van Aalten DMF. 2006. Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis. *The EMBO journal* **25**:1569–78.
- Rolfsmeier M, Blum P. 1995. Purification and characterization of a maltase from the extremely thermophilic crenarchaeote *Sulfolobus solfataricus*. *Journal of bacteriology* **177**:482–5.
- Santos H, da Costa MS. 2001. Organic solutes from thermophiles and hyperthermophiles. *Methods in Enzymology* **334**:302–315.
- Schelert J, Dixit V, Hoang V, Simbahan J, Drozda M, Blum P. 2004. Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon *Sulfolobus solfataricus* by use of gene disruption. *Journal of bacteriology* **186**:427–37.
- Schleper C, Puehler G, Holz I, Gambacorta A, Janekovic D, Santarius U, Klenk HP, Zillig W. 1995. Picrophilus gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0. *Journal of bacteriology* **177**:7050–9.
- She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, Awayez MJ, Chan-Weiher CC, Clausen IG, Curtis BA, De Moors A, Erauso G, Fletcher C, Gordon PM, Heikamp-de Jong I, Jeffries AC, Kozera CJ, Medina N, Peng X, Thi-Ngoc HP, Redder P, Schenk ME, Theriault C, Tolstrup N, Charlebois RL, Doolittle WF, Duguet M, Gaasterland T, Garrett RA, Ragan MA, Sensen CW, Van der Oost J. 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proceedings of the National Academy of Sciences of the United States of America* **98**:7835–40.
- Slámová K, Bojarová P, Petrásková L, Kren V. 2010.  $\beta$ -N-acetylhexosaminidase: what's in a name...? *Biotechnology advances* **28**:682–93.
- Stetter KO. 2006. Hyperthermophiles in the history of life. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **361**:1837–42; discussion 1842–3.

- Stubbs KA, Zhang N, Vocadlo DJ. 2006. A divergent synthesis of 2-acyl derivatives of PUGNAc yields selective inhibitors of O-GlcNAcase. *Organic & biomolecular chemistry* **4**:839–45.
- Stubbs KA, Balcewich M, Mark BL, Vocadlo DJ. 2007. Small molecule inhibitors of a glycoside hydrolase attenuate inducible AmpC-mediated beta-lactam resistance. *The Journal of biological chemistry* **282**:21382–91.
- Stubbs KA, Macauley MS, Vocadlo DJ. 2009. A selective inhibitor Gal-PUGNAc of human lysosomal beta-hexosaminidases modulates levels of the ganglioside GM2 in neuroblastoma cells. *Angewandte Chemie (International ed. in English)* **48**:1300–3.
- Sumida T, Ishii R, Yanagisawa T, Yokoyama S, Ito M. 2009. Molecular cloning and crystal structural analysis of a novel beta-N-acetylhexosaminidase from *Paenibacillus* sp. TS12 capable of degrading glycosphingolipids. *Journal of molecular biology* **392**:87–99.
- Tanaka T, Fujiwara S, Nishikori S, Fukui T, Takagi M, Imanaka T. 1999. A unique chitinase with dual active sites and triple substrate binding sites from the hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1. *Applied and environmental microbiology* **65**:5338–44.
- Tanaka T, Fukui T, Imanaka T. 2001. Different cleavage specificities of the dual catalytic domains in chitinase from the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *The Journal of biological chemistry* **276**:35629–35.
- Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE. 1996. Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. *Nature structural biology* **3**:638–48.
- Tews I, Scheltinga ACT Van, Perrakis A, Wilson KS, Dijkstra BW. 1997. Substrate-Assisted Catalysis Unifies Two Families of Chitinolytic Enzymes **7863**:7954–7959.
- Torres CR, Hart GW. 1984. Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *The Journal of biological chemistry* **259**:3308–17.
- Tropak MB, Blanchard JE, Withers SG, Brown ED, Mahuran D. 2007. High-throughput screening for human lysosomal beta-N-acetyl hexosaminidase inhibitors acting as pharmacological chaperones. *Chemistry & Biology* **14**:153–164.
- Tropak MB, Mahuran D. 2007. Lending a helping hand, screening chemical libraries for compounds that enhance beta-hexosaminidase A activity in GM2 gangliosidosis cells. *The FEBS journal* **274**:4951–61.

## 2.8 References

---

- Tsuji H, Nishimura S, Inui T, Kado Y, Ishikawa K, Nakamura T, Uegaki K. 2010. Kinetic and crystallographic analyses of the catalytic domain of chitinase from *Pyrococcus furiosus* - the role of conserved residues in the active site. *The FEBS journal* **277**:2683–95.
- Vocadlo DJ, Mayer C, He S, Withers SG. 2000. Mechanism of action and identification of Asp242 as the catalytic nucleophile of *Vibrio furnisii* N-acetyl-beta-D-glucosaminidase using 2-acetamido-2-deoxy-5-fluoro-alpha-L-idopyranosyl fluoride. *Biochemistry* **39**:117–26.
- Vocadlo DJ, Withers SG. 2005. Detailed comparative analysis of the catalytic mechanisms of beta-N-acetylglucosaminidases from families 3 and 20 of glycoside hydrolases. *Biochemistry* **44**:12809–12818.
- Vosseller K, Wells L, Lane MD, Hart GW. 2002. Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3-L1 adipocytes. *Proceedings of the National Academy of Sciences of the United States of America* **99**:5313–8.
- Van de Vossenberg JL, Ubbink-Kok T, Elferink MG, Driessen AJ, Konings WN. 1995. Ion permeability of the cytoplasmic membrane limits the maximum growth temperature of bacteria and archaea. *Molecular microbiology* **18**:925–32.
- Weignerová L, Vavrusková P, Pisvejcová A, Thiem J, Kren V. 2003. Fungal beta-N-acetylhexosaminidases with high beta-N-acetylgalactosaminidase activity and their use for synthesis of beta-GalNAc-containing oligosaccharides. *Carbohydrate research* **338**:1003–8.
- Williams SJ, Mark BL, Vocadlo DJ, James MNG, Withers SG. 2002. Aspartate 313 in the *Streptomyces plicatus* hexosaminidase plays a critical role in substrate-assisted catalysis by orienting the 2-acetamido group and stabilizing the transition state. *Journal of Biological Chemistry* **277**:40055–40065.
- Woese CR, Fox GE. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America* **74**:5088–90.
- Worthington P, Hoang V, Perez-pomares F, Blum P. 2003. Targeted Disruption of the alpha-Amylase Gene in the Hyperthermophilic Archaeon *Sulfolobus solfataricus* **185**:482–488.
- Yuzwa SA, Macauley MS, Heinonen JE, Shan X, Dennis RJ, He Y, Whitworth GE, Stubbs KA, McEachern EJ, Davies GJ, Vocadlo DJ. 2008. A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau in vivo. *Nature chemical biology* **4**:483–90.
- Zähringer U, Moll H, Hettmann T, Knirel YA, Schäfer G. 2000. Cytochrome b558/566 from the archaeon *Sulfolobus acidocaldarius* has a unique Asn-

- linked highly branched hexasaccharide chain containing 6-sulfoquinovose. *European journal of biochemistry / FEBS* **267**:4144–9.
- Zhang F, Su K, Yang X, Bowe DB, Paterson AJ, Kudlow JE. 2003. O-GlcNAc modification is an endogenous inhibitor of the proteasome. *Cell* **115**:715–25.
- Zillig W, Stetter KO, Wunderl S, Schulz W, Priess H, Scholz I. 1980. The Sulfolobus-"Caldariella" group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. *Archives of Microbiology* **125**:259–269.
- Zolghadr B, Klingl A, Koerdt A, Driessen AJM, Rachel R, Albers SV. 2010. Appendage-mediated surface adherence of *Sulfolobus solfataricus*. *J. Bacteriol.* **192**:104–110.

## **Chapter 3**

Pharmacological enhancement  
of acid alpha-glucosidase  
by the allosteric chaperone  
*N*-acetylcysteine

### 3.1 General overview

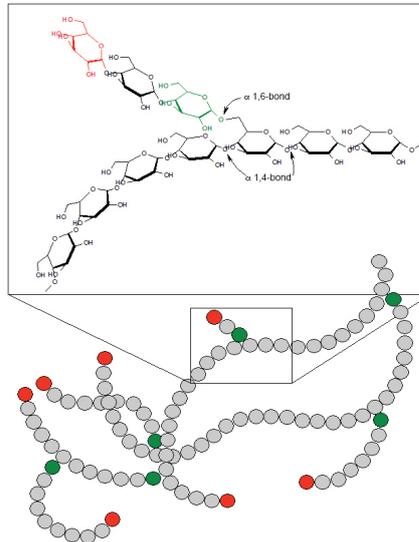
---

Chapter 3 focuses on pharmacological enhancement of acid  $\alpha$ -glucosidase (GAA), whose deficiency leads to Pompe disease (PD), belonging to lysosomal storage diseases (LSDs), by the allosteric chaperone *N*-acetylcysteine. This chapter starts with four introductory paragraphs (from 3.2 to 3.5) reporting the literature on the topic. In particular, in section 3.2 the functional and structural features of GAA will be exposed while the molecular aspects at the basis of LSDs will be discussed in the paragraph 3.3, followed by a deep description of PD in the section 3.4. The section 3.5 is further subdivided in three sub-paragraphs, each one highlighting a type of therapeutic approach adopted or under investigation for PD treatment. The purpose of this study is illustrated in the paragraph 3.6; while the experimental procedures and the results achieved are presented in sections 3.7 and 3.8, respectively. The references, instead, are reported in paragraph 3.9.

## **3.2 ACID ALPHA-GLUCOSIDASE: STRUCTURE AND FUNCTION**

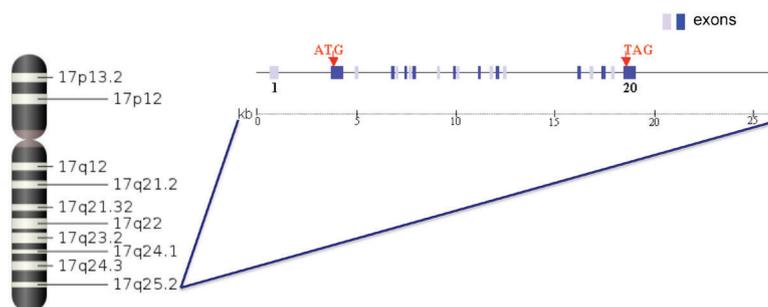
## acid alpha-glucosidase: structure and function 3.2

Acid  $\alpha$ -glucosidase (lysosomal  $\alpha$ -glucosidase, acid maltase, glucoamylase or GAA; EC 3.2.1.20) is an exo-glycosidase that catalyzes the hydrolysis at acidic pH (pH 4.0-5.0) of both  $\alpha$ -1,4 and  $\alpha$ -1,6 linkages in the natural substrate glycogen (Fig. 3.2.1) (Jeffrey et al., 1970a, 1970b).



**Figure 3.2.1:** Breakdown of glycogen catalyzed by GAA

Information on the primary structure of GAA has been obtained via molecular cloning and analysis of cDNA and genomic sequences (Hoefsloot et al., 1990; Martiniuk et al., 1991). The gene encoding for human GAA is located on the long arm of chromosome 17 and contains 20 exons and 19 introns spread over a distance of about 28 kb (Fig 3.2.2) (Martiniuk et al., 1991).

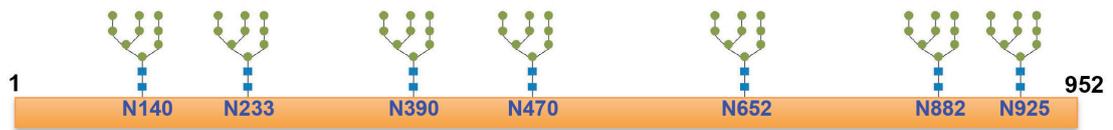


**Figure 3.2.2:** Organization of GAA gene

The cDNA encodes for a protein of 952 aa with an apparent molecular weight of 105 kDa (Hoefsloot et al., 1988). The newly synthesized precursor has an amino-terminal signal peptide for co-translational transport into the lumen of the endoplasmic reticulum (ER), where it is N-glycosylated at seven glycosylation

## 3.2 acid alpha-glucosidase: structure and function

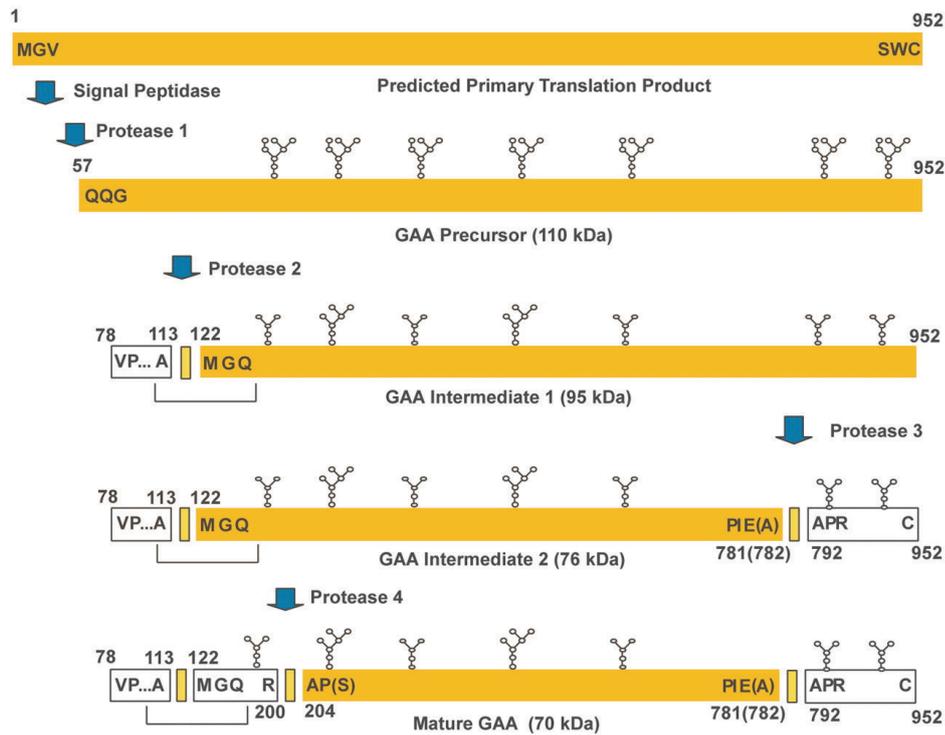
sites (Fig. 3.2.3), resulting in a glycosylated precursor with an apparent molecular mass of 110 kDa.



**Figure 3.2.3:** Glycosylation sites of GAA

During trafficking to the lysosome, the 110 kDa form is subject to several post-translational modifications involving both the N-glycan chains and the protein backbone (Fig. 3.2.4) (Moreland et al., 2005). The first processing step is the cleavage between aa 28 and 29 by a signal peptidase, followed by a proteolytic cleavage between aa 56 and 57. After transport through the Golgi complex and targeting to the endosome/lysosome, it is further proteolytically processed at the amino terminus, resulting in a 95 kDa intermediate with a sequence beginning at aa 122 (intermediate 1). Interestingly, almost all of the reduction in molecular mass from 110 to 95 kDa is attributed to the loss of part of the carbohydrate moiety due to extensive glycan trimming. This suggests that the processing yielding intermediate 1 takes place in a late endosome or lysosome, containing multiple glycosidases. In the next step, the 95 kDa intermediate is proteolytically cleaved at the carboxyl terminus, between aa 816 and 881, by an unknown protease generating the 76 kDa form that represents the major enzymatic species. The 76 kDa form is then proteolytically processed at the amino terminus, at aa 204, to give the 70 kDa mature form (Wisselaar et al., 1993). The nomenclature used for the processed forms of GAA is based on apparent molecular mass as determined by SDS-PAGE.

## acid alpha-glucosidase: structure and function 3.2



**Figure 3.2.4:** Model for the maturation of GAA

The identities of the proteases involved in the maturation of GAA are unknown. GAA has been purified from many different tissues such as bovine testis (Reuser et al. 1984), rat liver (Scheibe et al., 1985), pig liver (Tashiro et al., 1986), human liver (Murray et al., 1978), rabbit muscle (Matsui et al., 1984), human heart (Chambers and Williams, 1983), human urine (Oude Elferink et al., 1984) and human placenta (Hasilik and Neufeldg, 1980): the predominant species observed are the 76- and 70-kDa mature forms.

The peptides released during proteolytic processing remained tightly associated with the major species (Moreland et al., 2005). The first 95 kDa intermediate is covalently linked via a disulfide bond to a 3.9 kDa polypeptide (aa 78–113). The 76 kDa form is associated with peptides of 3.9 kDa (aa 78–113) and 19.4 kDa (aa 792–952), while the 70 kDa form contains the 3.9- and 19.4 kDa peptide species as well as a 10.3-kDa species (aa 122–199). The association of the proteolytically cleaved fragments of GAA is not unique for lysosomal enzymes. The lysosomal  $\alpha$ -mannosidase enzyme is synthesized as a single chain precursor and then is processed into three glycopeptides of 70, 42, and 15 kDa. The 70-kDa glycopeptide is further proteolyzed into three peptides that are joined by disulfide bridges (Nilssen et al., 1997).

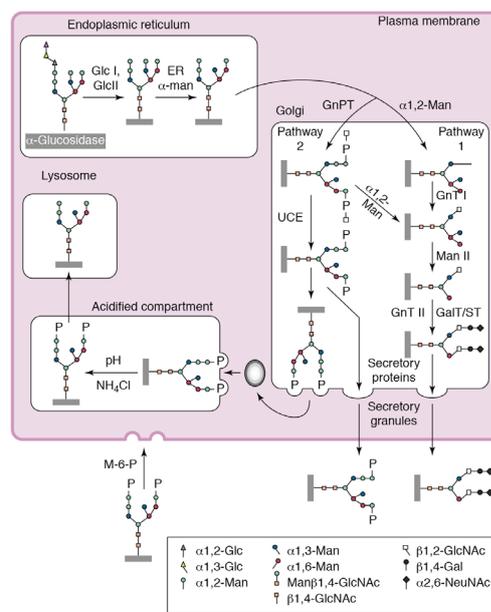
## 3.2 acid alpha-glucosidase: structure and function

Proteolytic processing appears to be required for optimal activity toward the natural substrate glycogen. The 76/70-kDa species display an affinity for glycogen 7–10-fold higher than the 110 kDa precursor (Tab. 3.2.1) (Bijvoet et al., 1998; Wisselaar et al., 1993).

Source	CHO medium	CHO medium	Mouse milk	Mouse milk	Placenta
Isoform (kDa)	110	76	110	76	76/70
Km (mg/ml)	425.3	43	484.6	54.8	42.2

**Table 3.2.1:** Affinity for glycogen of the recombinant forms of human GAA

In addition to maturation of the peptide backbone, GAA is subjected to an extensive remodelling of the carbohydrate chains (Fig. 3.2.5) (Chen and Amalfitano, 2000).



**Figure 3.2.5:** Remodelling of GAA N-glycans during trafficking to the lysosomes

The newly synthesized GAA harbouring a high-mannose oligosaccharide is transferred from ER to the Golgi network where it acquires the mannose 6-phosphate (M6P) residues through the sequential action of *N*-acetylglucosaminylphosphotransferase (GnPT) and *N*-acetylglucosamine-1-phosphodiesterase (UCE) (Fig. 3.2.5, pathway 2). M6P residues can bind to M6P receptors allowing the transport of the enzyme to the lysosome. Alternatively, GAA can undergo trimming of mannose, addition of

## acid alpha-glucosidase: structure and function 3.2

sugars and formation of a complex oligosaccharide and is secreted (Fig. 3.2.5, pathway 1). M6P receptors on plasma membrane can bind extracellular enzyme and mediate its internalization and localization to the lysosomes. Thus, N-glycans decorating the seven glycosylation sites of GAA are highly heterogeneous; they are predominantly high-mannose-type, containing typically six to nine mannose residues and one or two M6P residues, but complex-type chains are also present (Mutsaers et al., 1987). The functional relevance of each of the seven glycosylation sites was determined by site-directed mutagenesis (Hermans et al., 1993). Elimination of six of the seven sites does not interfere with enzyme synthesis, lysosomal transport or function; whereas removal of the second glycosylation site at Asn-233 hampers dramatically the formation of mature enzyme. The mutant precursor is synthesized normally and assembled in the ER, but it is not targeted to the lysosomes.

According to CAZy classification, GAA is included in the family GH31, that also contains  $\alpha$ -xylosidases (EC 3.2.1.177), isomaltosyltransferases (EC 2.4.1.-), maltase/glucoamylases (EC 3.2.1.48) and the  $\alpha$ -glucan lyases (EC 4.2.2.13). The GH31 enzymes are *retaining*  $\alpha$ -glycosidases following the classical Koshland double-displacement mechanism (Koshland, 1953) (see section 1.3 for more details) and can be found in a wide range of organisms including archaea, bacteria, plants and animals.

With respect to the functional trait, it is worth to note the striking similarities found between human GAA and the 2 subunits of intestinal brush border sucrase-isomaltase, suggesting that these proteins are derived from the same ancestral gene. The major sequence differences are found in the N-terminal regions accounting, probably, for the different cellular locations of these proteins. Structural conservation of the active site of these enzymes seemed likely because of overlapping substrate specificities (Hoefsloot et al., 1988; Hunziker et al., 1986). The catalytic domain of human GAA is folded into a  $(\beta/\alpha)_8$ -barrel, and the active-site pocket is located at the C-terminal end of the  $\beta$ -strand. Through two different approaches, the substrate analogue and active site-directed inhibitor conduritol  $\beta$ -epoxide (CBE) and site-directed mutagenesis, Hermans and colleagues have identified the catalytic nucleophile

## **3.2 acid alpha-glucosidase: structure and function**

as Asp-518 (Hermans et al., 1991); by contrast, the catalytic proton donor residue, Asp-616, was assigned by similarity (Yoshimizu et al., 2008).

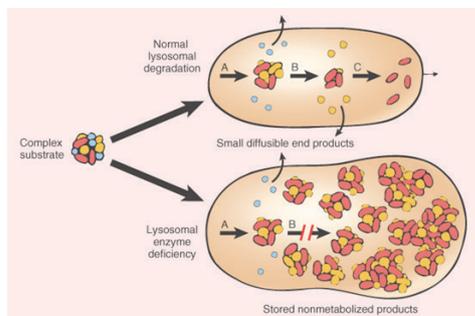
GAA activity is essential for the catabolism of lysosomal deposits of glycogen and the deficiency of this metabolic function is associated to the pathogenesis of a lysosomal storage disease (LSD), known as Pompe disease (see below for a detailed description).

## **3.3 LYSOSOMAL STORAGE DISEASES**

### 3.3 Lysosomal storage diseases

Lysosomes are membrane-bound acidic organelles that contain over 50 different acid hydrolases responsible for the catabolism of a wide range of macromolecules, including glycosphingolipids, glycogen, mucopolysaccharides, oligosaccharides, cholesterol, peptides, and glycoproteins. They also play a role in cell health by digesting viruses, microbes and other organelles.

The deficiency of any individual lysosomal enzyme leads to a lysosomal storage disorder (LSD), which is characterized by the pathological accumulation of the deficient enzyme's substrate (Fig. 3.3.1). Lysosomal storage diseases (LSDs) are a diverse group of hereditary metabolic disorders and, although they are individually rare, their overall prevalence is estimated to be 1 in 8000 births (Meikle et al. 1999).



**Figure 3.3.1:** Schematic illustration of the LSDs pathogenesis. A complex substrate is normally degraded by a series of lysosomal enzymes (A, B, and C) into soluble end products. If there is a deficiency or malfunction of one of the enzymes (e.g., B), catabolism is incomplete and insoluble intermediates accumulate in the lysosomes

The first description of a lysosomal storage disorder was that of Tay-Sachs disease in 1881, even if the lysosome was not discovered until 1955, by Christian De Duve (Appelmans et al., 1955). The first demonstration of a link between an enzyme deficiency and a storage disorder was achieved for Pompe disease by Hers in 1963 (Hers, 1963). To date, over 50 LSDs have been described; they tend to be multisystemic and display a progressive course (Tab. 3.3.1).

Most LSDs are autosomal recessive disorders with the exceptions of Hunter disease (MPS II) and Fabry disease, that show an X-linked inheritance pattern. The pathogenic mutations affect a single gene and can be of various type. Nonsense mutations, deletions, insertions and duplications cause premature termination and/or frame shifts and frequently lead to absent or truncated

## Lysosomal storage diseases 3.3

proteins, which results in a complete loss of functional enzyme. Several splice-site mutations have also been identified and, in some instances, they do not allow the generation of functional mRNA (Polten et al., 1991).

Disease	OMIM	Deficient protein
<b>Mucopolysaccharidosis (MPS)</b>		
MPS I		$\alpha$ -iduronidase
Hurler syndrome	607014	
Scheie syndrome	607016	
MPS II		Iduronate-2-sulfatase
Hunter	309900	
MPS III (Sanfilippo syndrome)		
Type III-A	252900	Heparan N-sulfatase
Type III-B	252920	N-acetyl- $\alpha$ -glucosaminidase
Type III-C	252930	acetyl CoA: $\alpha$ -glucosaminide acetyltransferase
Type III-D	252940	N-acetylglucosamine 6-sulfatase
MPS IV (Morquio syndrome)		
Type IV-A	253000	galactosamine-6-sulfate sulfatase
Type IV-B	253010	$\beta$ -galactosidase
MPS VI (Maroteaux-Lamy syndrome)	253200	Arylsulfatase B
MPS VII (Sly syndrome)	253220	$\beta$ -glucuronidase
MPS IX (Hyaluronidase-Defizienz)	601492	hyaluronidase
<b>Mucopolipidosis (ML)</b>		
ML I	256550	neuraminidase
ML II	252500	N-acetylglucosamine-1-phosphotransferase
ML III (Pseudo-Hurler Syndrome)		N-acetylglucosamine-1-phosphotransferase
Type III-A	252600	
Type III-C	252605	
<b>Sphingolipidosis</b>		
GM1-Gangliosidosis		$\beta$ -galactosidase-1
Type I	230500	
Type II	230600	
Type III	230650	
GM2-Gangliosidosis		
Type I (Tay-Sachs syndrome)	272800	$\beta$ -hexosaminidase A
Type II (Sandhoff syndrome)	268800	$\beta$ -hexosaminidase B
Niemann-Pick disease		sphingomyelinase
Type A	257200	
Type B	607616	
Lipogranulomatosis (Farber disease)	228000	ceramidase
Gaucher disease		$\beta$ -glucocerebrosidase
Type I (adult form)	230800	
Type II (infantile form)	230900	
Type III (juvenile form)	231000	
Fabry disease	301500	$\alpha$ -galactosidase
Metachromatic leukodystrophy	250100	Arylsulfatase A
Krabbe disease	245200	galactosylceramide $\beta$ -galactosidase
<b>Glycogenesis</b>		
Pompe	232300	Acid $\alpha$ -glucosidase
<b>Glycoproteinoses</b>		
alpha-mannosidosis	248500	$\alpha$ -mannosidase
beta-mannosidosis	248510	$\beta$ -mannosidase
Fucosidosis	230000	$\alpha$ -fucosidase
Aspartylglucosaminuria	208400	Aspartylglucosaminidase
Schindler disease	609241	$\alpha$ -N-acetylgalactosaminidase

**Table 3.3.1:** Some known LSDs. OMIM is Online Mendelian Inheritance in Man number

### **3.3 Lysosomal storage diseases**

The effects of missense mutations or small in-frame deletions are less apparent and need molecular biological characterization to understand how they lead to the enzyme deficiency. However, considering the large number of missense mutations identified in LSDs, only a limited number have been investigated from a biochemical point of view. The deleterious effect of a missense mutation is obvious if the affected amino acid is located in the active site (Garman and Garboczi, 2004; Huie et al., 1994) or is important for proper folding. Active-site mutations may lead to a complete loss of enzyme activity, but may also result in alterations of enzyme kinetics (Zhang et al., 2000). Substitution of cysteine residues or charged residues involved in the formation of intramolecular disulphide bonds or salt bridges, respectively, is likely to interfere with proper folding of the enzyme (Hermann et al., 2000; Saarela et al., 2001). Mutations affecting the active centre or intramolecular bonds, however, are in the minority. The effects of most missense mutations are still unknown. Thus, it is difficult to understand why even very conservative amino acid substitutions can severely affect a lysosomal enzyme. In many cases, amino acid substitutions cause misfolding of the enzyme, which results in retention of the mutant polypeptide in the ER and its subsequent proteasomal degradation (Poeppel et al., 2005; Schestag et al., 2002). It has been estimated that degradation is responsible for about 50% of protein defects in all genetic diseases and thus represents the most frequent cause of deficiencies (Carrell and Lomas, 1997). Alternatively, amino acid substitutions allow to pass the ER quality control. One example is given by substitution of proline 426 by leucine (P426L) in arylsulphatase A, one of the most frequent mutation causing Metachromatic leukodystrophy (Polten et al., 1991; Von Bülow et al., 2002).

Since the synthesis and processing of lysosomal hydrolases involve many steps, it is not surprising that the disfunction can originate at different levels: synthesis, folding, activation or targeting. The severity of the phenotype is closely related to the residual enzyme activity. In general, the lower is the residual activity, the earlier is the age of onset, and more severe are the clinical manifestations. There was a critical threshold above of which, enzyme activity is still sufficient to prevent accumulation of substrate (Conzelmann and Sandhoff, 1983).

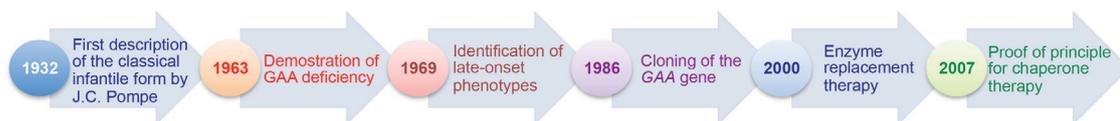
### **Lysosomal storage diseases 3.3**

---

LSDs are not simply a consequence of pure storage, but also result from other complex cell signalling mechanisms, which in turn elicit secondary structural and biochemical events. The molecular pathways through which the storage material causes cellular and organ pathology are largely unknown. A number of secondary biochemical and structural events have been reported. For example raised concentrations of cytokines or chemokines, deriving from macrophage activation, have been found in patients with Gaucher disease (Boot et al., 2004; Hollak et al., 1997). Altered homeostasis of calcium, an important intracellular mediator, appears to play an important role in the sphingolipid storage disorders. There is also evidence that extralysosomal accumulation of substrate may take place and have deleterious effects on trans-membrane and intracellular signalling. For example, in a mouse model of metachromatic leucodystrophy, accumulation of sulphatide was demonstrated in myelin as well in lysosomes. There is also the intriguing possibility that lysosomal storage may adversely affect mitochondrial function (Jolly et al., 2002; Lücke et al., 2004).

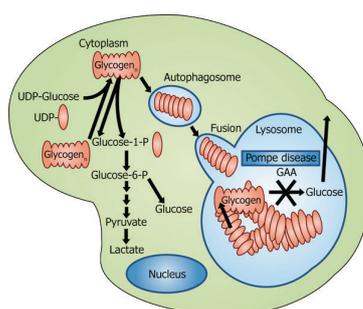
## **3.4 POMPE DISEASE**

Pompe disease (PD, OMIM 232300), also known as glycogen-storage disease type II, glycogenosis type II or acid maltase deficiency, is a pan-ethnic autosomal recessive myopathy with an estimated incidence of 1:40,000 live births. It was described for the first time by the Dutch pathologist Johannes C. Pompe in 1932 (Pompe, 1932) and since then, continuous progress has been made in the understanding of its biochemistry and genetics (Fig.3.4.1)



**Figure 3.4.1:** Milestones in the scientific research on PD

The pathogenesis is associated to mutations in the *GAA* gene, resulting in functional deficiency of acid  $\alpha$ -glucosidase (GAA) with consequent lysosomal accumulation of glycogen in several tissues and organs (Fig. 3.4.2).

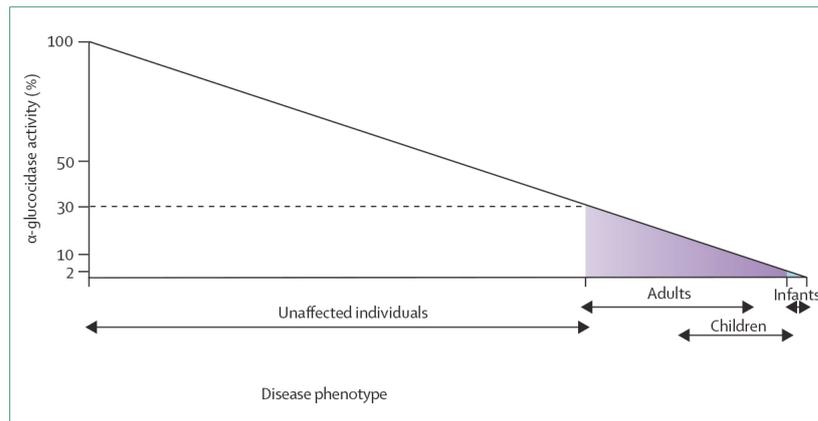


**Figure 3.4.2:** PD pathway

Presently, more than 400 different mutations have been identified (missense, nonsense, deletions, insertions and splice site mutations) and are listed in the PD database of the Erasmus University, Rotterdam, the Netherlands ([http://www.erasmusmc.nl/klinische\\_genetica/research/pompe\\_center/](http://www.erasmusmc.nl/klinische_genetica/research/pompe_center/)). Some of them are recurrent (Hermans et al., 2004) and show an ethnic distribution: the c.525del, p.G309R, p.M519T, p.G219R, p.G828\_N882del, c.525del, and p.E176fsX4 mutations are frequent among Caucasians, the pR854X among Africans, the p.D645E among Asians. In Italy the p.L552P mutation accounts for 8.5% of the alleles in early-onset Italian patients (Pittis et al., 2008), whereas the p.W746X mutation has been found in 10.3% of the alleles in late-onset patients (Montalvo et al., 2006). Other mutations are pan-ethnic, such as the splice site c.-32-13T>G, which is highly prevalent among late-onset PD patients

### 3.4 Pompe disease

(Kroos et al., 2007). The nature and combination of the mutations in the two GAA alleles dictates the degree of enzyme deficiency. Accumulation of lysosomal glycogen starts when GAA activity drops below 30% of average normal activity and the disease phenotype is correlated to the extent of enzyme deficiency: GAA activity may range from complete deficiency (<1%) in the severe forms, to partial (up to 30%) deficiency in milder forms (Fig. 3.4.3) (Hermans et al., 2004).



**Figure 3.4.3:** Model depicting the correlation between residual GAA activity and phenotype

The genotype/phenotype correlation is not always unequivocal: secondary factors can influence substantially the clinical course (De Filippi et al., 2010).

PD is a multisystemic disease with glycogen storage occurring in almost every tissue and cell type. However, clinical manifestations are predominantly related to heart and skeletal muscle involvement. As a consequence, the clinical picture of PD shares common features with that of neuromuscular disorders.

Traditionally, PD has been classified into distinct categories based on the phenotype and the age of onset:

- classic infantile
- juvenile and adult

The infantile form manifests with a severe phenotype, characterized by severe hypertrophic cardiomyopathy, generalized hypotonia, respiratory infections and a rapidly progressive course (van den Hout et al., 2003). Clinical manifestations may appear already in utero, but usually present within the first months of life (1.6-2.0 months) and patients with this PD form rarely survive beyond 1 year of age.

The late-onset juvenile and adult forms emerge between the second and fourth decade and show a more mild clinical presentation (Hagemans et al., 2005).

The cascade of events that starts from glycogen storage and leads to cell death and muscle atrophy and destruction is still poorly understood. For many years, it was assumed that storage *per se* was responsible for cell and tissue damage.

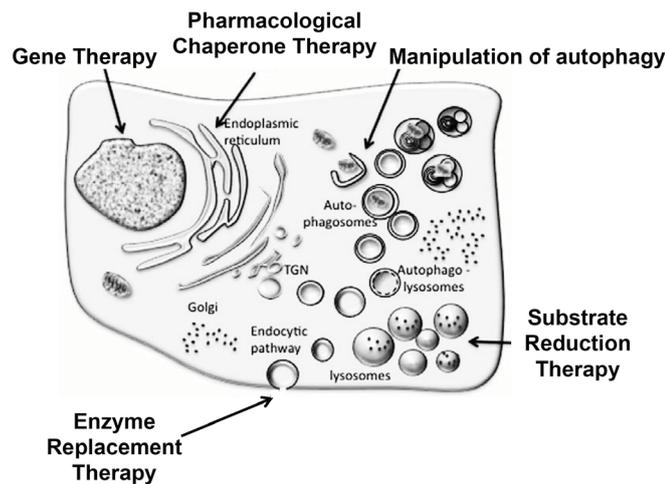
The rupture of glycogen-filled lysosomes was considered the major cause of muscle destruction in PD. Only in the recent years, emerging evidences suggest that secondary factors triggered by storage are likely to concur in tissue damage. The spectrum of such events is wide, including receptor activation by non-physiologic ligands, modulation of receptor responses and signal transduction cascades, activation of inflammatory responses, impaired intracellular trafficking of vesicles, membranes and membrane-bound proteins, impairment of autophagy, and others (Ballabio and Gieselmann, 2009).

The understanding of the role played by the secondary factors in the pathogenetic route of PD has very important clinical implications, specifically for the therapeutic management of the disease.

## **3.5 THERAPEUTIC APPROACHES**

## Therapeutic approaches 3.5

The complexity of PD phenotype and pathophysiology makes therapeutic interventions particularly challenging. Therapies should be directed towards correction of pathology and function in all affected tissues to restore health, or to ameliorate patients' quality of life. The last decade has witnessed extraordinary innovation in the treatment of PD with different approaches now available or under investigation (Fig. 3.5.1). Currently, the only approved therapy for this LSD is the Enzyme Replacement Therapy (ERT) based on intravenous infusions of recombinant GAA (see paragraph 3.5.1 for more details). Nevertheless, this approach has limitations and important issues remain unsolved. Thus, current efforts are being directed towards exploring alternative therapeutic approaches based on innovative strategies, such as Pharmacological Chaperone Therapy (PCT, which will be described in the paragraph 3.5.2), gene therapy, Substrate Reduction Therapy (SRT) and modulation of autophagy (described in the paragraph 3.5.3).



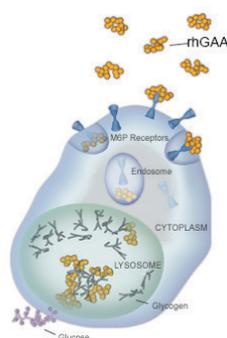
**Figure 3.5.1:** Multiple therapeutic approaches have been used or are under investigation to treat PD, targeting different cellular pathways and functions

### **3.5.1 ENZYME REPLACEMENT THERAPY IN POMPE DISEASE**

### Enzyme replacement therapy in Pompe disease 3.5.1

Enzyme replacement therapy (ERT) with recombinant human GAA (rhGAA) is currently the only approved pharmacologic treatment for PD.

ERT is based on the concept that recombinant lysosomal hydrolases, in most cases enzyme precursors manufactured on large scale in eukaryotic cells, can be administered periodically to patients by an intravenous route. These enzymes are internalized by patients' cells and tissues through the mannose or M6P receptors and are ultimately delivered to lysosomes, where they are activated and restore the function of the defective hydrolases (Fig. 3.5.1.1).



**Figure 3.5.1.1:** Model depicting the cellular uptake of rhGAA

In the 1960s, PD was the first LSD for which attempts at ERT were undertaken by using enzyme preparations from *Aspergillus niger* and human placenta. However, these attempts had no clinical benefit: inappropriate enzyme source and insufficient dosing were the causes of failure (Hug and Schubert, 1967). An important step forward was the knowledge that cell-surface receptors are required for uptake of extracellular glycoproteins via endocytosis. Efficient uptake and targeting to the lysosomes was mediated by the M6P or insulin-like growth factor II receptor (Van der Ploeg et al., 1988). Cloning of the *GAA* gene provided the essential method to explore production of recombinant human enzyme in cell culture and in milk of transgenic animals leading to large-scale production of rhGAA in milk of transgenic rabbits and in CHO cells (Bijvoet et al., 1999; Van Hove et al., 1996; Van Hove et al., 1997; Martiniuk et al., 1986). Both products were tested in a mouse model of PD and showed similar uptake characteristics (Bijvoet et al., 1998), but rhGAA produced in CHO cells has proved to be more effective in clearing glycogen stores in heart and muscle (McVie-Wylie et al., 2008). For these reasons CHO-derived rhGAA (alglucosidase alfa) was selected for future clinical development and registration (Van den Hout et al., 2000; Van den Hout et al., 2004; Klinge et al.,

### **3.5.1 Enzyme replacement therapy in Pompe disease**

2005). In 2006 alglucosidase alfa (Myozyme<sup>®</sup>, Genzyme, Cambridge, MA) received marketing approval by the European Medicine Agency and US Food and Drug Administration. Myozyme is protected by U.S. Patent No. 6,118,045 which expires on August 18, 2018. The therapeutic protocol provides intravenous administration of doses of 20 mg/kg and 40 mg/kg once every 2 weeks. The age at which treatment is started constitutes an important factor that determines efficacy of therapy: early start leads to the best outcome.

ERT in PD is, at present, the only cure and had the great merit of improving survival of the patients and stabilize the disease course. However, it displays several limitations:

- reduced targeting to skeletal muscle, one of the major sites of disease (Raben et al., 2003). Why this happens remains to be fully elucidated, even if several factors have been proposed. Most of the recombinant enzyme (up to 85%) was taken up by liver and only a small fraction reaches heart and even less skeletal muscle. The large mass of this tissue and the relative deficiency of the M6P receptors on muscle cells (Wenk et al., 1991) also contribute to the poor correction of GAA activity. To improve muscle targeting of rhGAA, Zhu and colleagues have proposed a glycoengineered rhGAA, that is chemically modified by conjugation of a synthetic oligosaccharide harbouring M6P residues (Zhu et al., 2005). This modification significantly improved the affinity of the enzyme for the M6P receptors and its uptake by muscle cells. Administration of this carbohydrate-remodeled enzyme to PD mice resulted in an approximately fivefold higher clearance of lysosomal glycogen in muscles when compared to the unmodified enzyme. Nevertheless, the large production of this form of rhGAA is still difficult.
- immune response to rhGAA
- high costs. Initial investments in research and costs related to the production of large amounts of good manufacturing practice recombinant enzymes contribute to keep the prices of ERT extremely high (\$300,000/year/patient)
- burdensome for patients

Therefore, importantly, the ERT drawbacks point to the need for improved therapeutic strategies.

## **3.5.2 PHARMACOLOGICAL CHAPERONE THERAPY**

### **3.5.2 Pharmacological chaperone therapy**

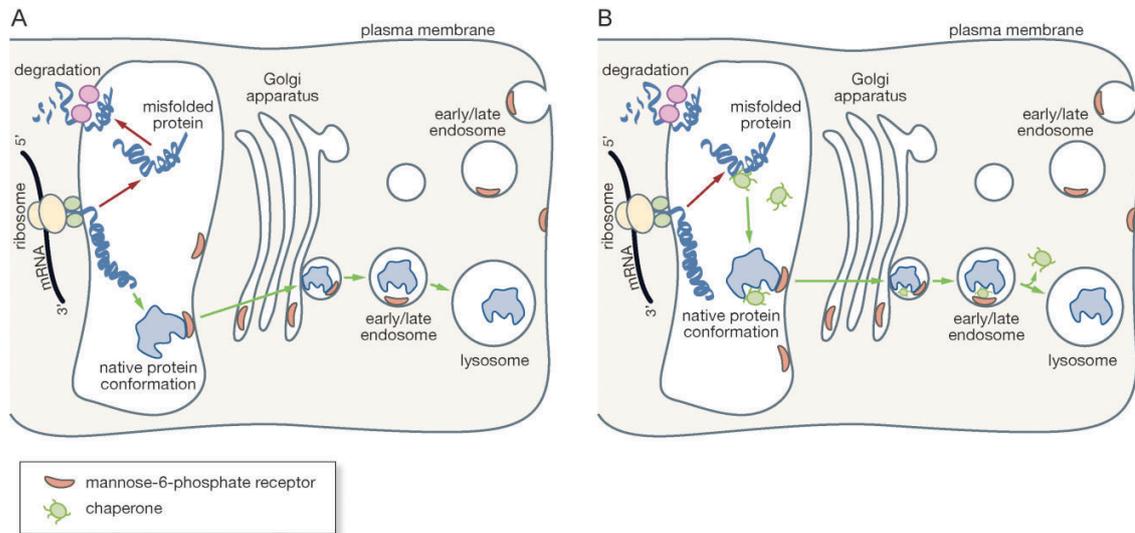
---

Over the past decade, pharmacological chaperone therapy (PCT) has attracted much interest as a potential treatment for many genetic diseases that result from misfolded and/or unstable proteins, including LSDs (Fan, 2008; Parenti, 2009).

Many of the mutations causing human LSDs are missense and may influence mRNA expression, protein folding and stability, intracellular trafficking, substrate binding, catalytic competency, and/or enzyme turnover rate. Folding and maturation of lysosomal enzymes, like many other proteins, are monitored by the quality control system of the ER (Ellgaard and Helenius, 2001). It relies on molecular chaperones and folding factors, such as BiP, calnexin, calreticulin, thiol-disulfide oxidoreductases, and protein disulfide isomerase, that recognize common structural features, such as exposed hydrophobic regions, unpaired cysteine residues and so on, to distinguish stable, native protein conformations from unstable, non-native ones (Anelli and Sitia, 2008; Ushioda et al., 2008). The correctly folded proteins can leave ER and progress through the secretory pathway to the lysosome. If the nascent protein is not properly folded, it is recognized as aberrant by ER quality control system and targeted to the proteasomal degradation (Fig. 3.4.2.1A). ER quality control is highly efficient, but, in some cases, mutant enzymes that retain catalytic activity, or that have only modestly compromised function may be recognized as aberrant and then degraded. Thus, slight modifications in protein stability or conformation, as seen with many lysosomal enzymes carrying missense mutations, may prevent release from the ER and result in premature degradation and a loss of function. PCT is based on the concept that small molecules, such as substrates or active-site directed competitive inhibitors, may assist the folding of mutated enzymes that retain their catalytic activity, and prevent their recognition by the quality control system (Fan and Ishii, 2007). In this way the pharmacological chaperones (PCs) increase intracellular pool of active enzyme, improve trafficking of the protein towards its final destination and partial restore the metabolic functions (Fig. 3.5.2.1B). In addition to lysosomal enzymes, the efficacy of PCs has been shown also for other types of mutated proteins, including G protein-coupled receptors, secreted proteins, transcription factors, ion channels, and transporters, involved in the pathogenesis of diseases such as cystic fibrosis, hypercholesteremia, cataracts, Huntington's, Alzheimer's and

## Pharmacological chaperone therapy 3.5.2

Parkinson's diseases, retinitis pigmentosa, nephrogenic diabetes insipidus, and cancer.



**Figure 3.5.2.1:** **A.** Folding of lysosomal enzymes in absence of PCs; destiny of wild-type protein is indicated by green arrows; while red arrows delineate the route of mutated misfolded enzyme. **B.** Stabilizing effect of PCs on misfolded lysosomal enzyme indicated by green arrows

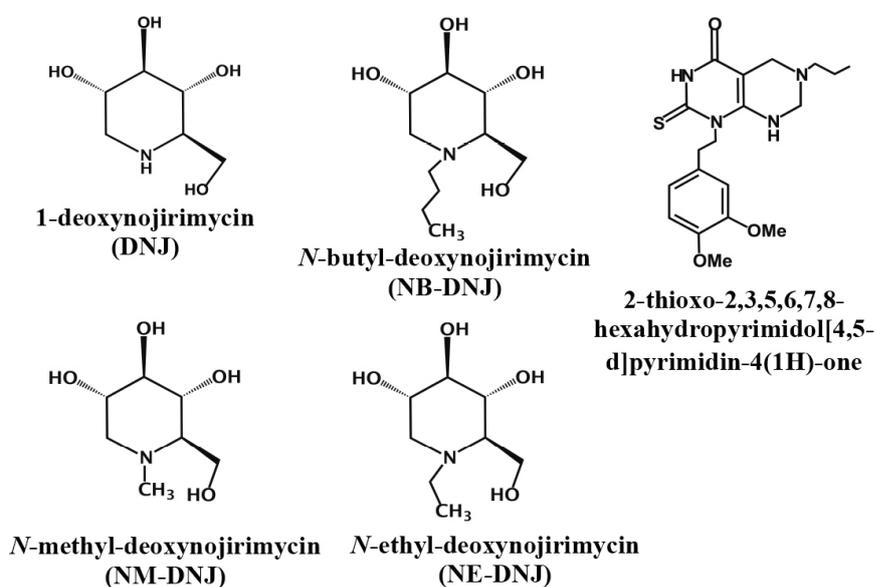
PCs are distinct from chemical chaperones, such as glycerol, dimethylsulfoxide, and trimethylamine N-oxide. Chemical chaperones act non-specifically on many proteins and require very high concentrations to exert their effect. By contrast, PCs are designed to specifically bind the protein of interest, as a consequence therapeutic benefits are achieved at low concentrations.

Several factors make LSDs excellent candidates for PCT. In most LSDs, 1% to 6% of normal activity has been estimated to be sufficient to delay or prevent disease onset or to yield a more mild form of the disease (Schueler et al., 2004). Thus, even a minor increase in enzyme activity obtained by PCT could have a significant impact on substrate levels, and hence disease severity and the rate of disease progression. In addition, the use of PCs has the potential to overcome several ERT limitations such as its impact on the patients' quality of life and high costs. These drugs are low molecular-weight molecules, and thus can be ingested orally with broad biodistribution and do not require invasive life-long infusions, as for ERT.

PCs identified so far interact with the active site and reversibly inhibit enzyme activity. In the case of PD, the imino sugars DNJ and derivatives are currently

### 3.5.2 Pharmacological chaperone therapy

under investigation as PCs (Fig. 3.5.2.2). These compounds contain a piperidine ring with an array of hydroxyl groups that closely resembles the terminal glucose unit of glycogen: this high level of structural similarity allows the binding to the active site and stabilization of GAA, but they also act as reversible competitive inhibitors of enzyme. It has been assumed that, once in the lysosome, the chaperone is displaced from the enzyme due to the excess of natural substrate, with a higher affinity to the catalytic site and that the acidic environment of lysosomes favours the dissociation of the chaperone–enzyme complex. But all these aspects are not completely clear and will need further investigation.



**Figure 3.5.2.2:** Structure of the imino sugars that act as PCs of GAA

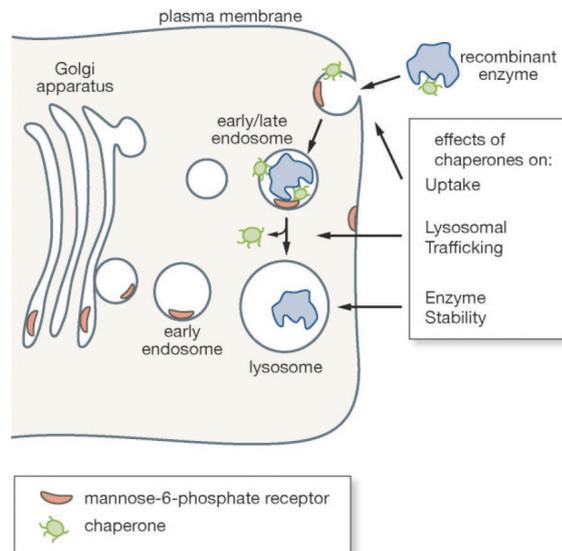
Although the rationale of PCT is clear, the molecular mechanisms underlying the effect of chaperones are still not fully understood.

Not only mutated enzymes, but also recombinant enzymes used for ERT may be prone to mistrafficking and degradation. For instance, in human fibroblasts from PD patients and in muscle cells of a PD mouse model, it has been shown that a fraction of the rhGAA is mistargeted and is ineffective (Cardone et al., 2008; Fukuda et al., 2006). Some evidences indicate that PCs can act also on wild-type enzymes. The mechanisms underlying the effect of chaperones on wild-type enzymes is not clear. The enhancing effects of PCs may take place at different levels (Fig. 3.5.2.3):

- increased uptake

## Pharmacological chaperone therapy 3.5.2

- improved intracellular trafficking to lysosomes
- enhanced enzyme stability



**Figure 3.5.2.3:** Model depicting the synergy between ERT and PCT

In the case of PD, NB-DNJ has proved to be effective to increase stability of rhGAA *in vitro* and, importantly, also to improve the delivery of the recombinant enzyme to the lysosomes. The combination of rhGAA and the PC is more efficient in correcting the enzyme activity than rhGAA alone (Porto et al., 2009). The synergistic effect of PCT and ERT may be particularly useful in patients responding poorly to therapy and in tissues in which sufficient enzyme levels are difficult to obtain. Further investigations are required to address questions as to where do chaperones and enzymes come in contact and why do they confer stability to already properly folded protein.

Unfortunately PCT, like ERT, displays some drawbacks:

- inhibitory action of PCs identified to date (Porto et al., 2012; Yoshimizu et al., 2008)
- only missense mutations are amenable to PCT, accounting for about 15% of mutations leading to PD (Porto et al., 2009)
- adverse effects and high costs of some PCs, like NB-DNJ (Ficicioglu, 2008) and DNJ. A first phase 2 clinical trial, started in the US and sponsored by Amicus Therapeutics, was performed using DNJ as therapeutic agent. The trial was terminated after the appearance of

### **3.5.2 Pharmacological chaperone therapy**

---

adverse events in treated patients and needs careful evaluation of the treatment protocol (Parenti and Andria, 2011).

Hence, future efforts should be directed towards the identification of a second generation of molecules with a good safety profile and better enhancing properties and intracellular distribution. An ideal compound should have a good enhancing activity (Wang et al., 2009) without inhibitory effects, and be effective on a larger number of mutations. To overcome the inhibitory effect associated to the use of compounds directed to active site, one possible scenario could be the identification of molecular chaperones directed to allosteric sites.

### **3.5.3 OTHER THERAPEUTIC APPROACHES**

### **3.5.3 Other therapeutic approaches**

---

#### **Gene therapy**

Gene therapy is based on the use of vectors delivering the therapeutic enzyme to target organs. Different vectors, including Adenoviruses (Ad), adeno-associated (AAV) viruses and lentiviruses, have been tested in PD mouse model. AAV-mediated gene therapy was reported to restore therapeutic levels of GAA activity, substantially reduce glycogen storage and improve muscle morphology both in the heart and in skeletal muscle (Sun et al., 2008). Intravenous administration of an Ad vector encoding GAA resulted in secretion of high levels of the precursor GAA into the plasma of treated animals (Amalfitano et al., 1999). The secreted enzyme was efficiently taken up by the neighboring cells through M6P receptors, resulting in a phenotypic rescue and correction of glycogen storage. Another gene therapy approach is based on the use of a chimeric GAA containing an alternative signal peptide that increased the secretion of GAA from transduced cells and enhanced the receptor-mediated uptake of GAA in striated muscle (Sun et al., 2006). This strategy allowed to reduce significantly glycogen content in striated muscle of PD mice. The lentiviral vector were used to express GAA in hematopoietic stem cells (HSC). PD mouse model transplanted with this genetically engineered HSC, showed a clearance of glycogen in heart, diaphragm, spleen and liver (Van Til et al., 2010). Although these preclinical studies seem to hold great promise, several issues remain to be addressed, such as evasion of immune response, possible toxicity of vectors, choice of the appropriate dosage of viral particles and of the proper route of administration to achieve sustained corrective enzyme levels in target tissues.

#### **Substrate reduction therapy (SRT)**

Substrate reduction therapy (SRT) is aimed to restore the equilibrium between substrate synthesis and degradation, reducing the rate of substrate synthesis, instead of enhancing the activity of the degrading enzymes. Since 2003, this strategy is already in clinical use for mild to moderate forms of Gaucher disease and is based on the use of NB-DNJ (miglustat, Zavesca<sup>®</sup>, Actelion Pharmaceuticals, Allschwil/Basel, Switzerland), a small molecule that inhibits ceramide glucosyltransferase (Platt and Jeyakumar, 2008). A SRT-based approach was also proposed for treatment of PD (Douillard-Guilloux et al., 2008, 2010). The strategy is based on the genetic suppression of glycogen

### **Other therapeutic approaches 3.5.3**

---

synthesis through the use of short hairpin ribonucleic acids (shRNA) targeted to the glycogen synthase. In PD mouse model, this approach allowed to deeply reduce the amount of glycogen in the heart and skeletal muscles, and to significantly decrease lysosomal swelling and autophagic build-up.

#### **Modulation of the autophagy**

Autophagy is a highly regulated mechanism in the lysosomal pathway by which the cells degrade long-lived proteins and damaged or unneeded organelles in the cytoplasm (Wang and Klionsky, 2003). Defects in lysosome function can lead to autophagic stress (Chu, 2006) characterized by accumulation of autophagic intermediates (Kurz et al., 2008). Suppression of autophagy alone or combined with ERT, as alternative treatment for PD, can be beneficial for the restoration of muscle histology and architecture (Raben et al., 2010). Anyway, it is worth to note that the suppression of autophagy in the muscles can theoretically also have unwanted consequences, such as accumulation of dysfunctional mitochondria and oxidative stress (Wu et al., 2009). Hence, when considering the manipulation of autophagy as therapeutic strategy, it is important to find a correct balance.

## **3.6 PURPOSE OF THE WORK**

Acid- $\alpha$ -glucosidase is a glycoside hydrolase whose functional deficiency leads to Pompe disease, a lysosomal storage disorder highly debilitating and life-threatening. Enzyme replacement therapy (ERT) is the only approved therapy, which allows to ameliorate patients' quality of life and slow the clinical course, but it displays several limitations that contribute to reduce its efficacy, in addition to be very expensive and burdensome for patients. In the recent years, Pharmacological chaperone therapy alone or in combination with ERT has proved to be very promising. The pharmacological chaperones (PCs) identified to date are directed to active site of target protein and act as its competitive inhibitors, limiting their chaperoning action. For these reasons, a second generation of PCs with better enhancing effect and effective on a larger number of mutants without inhibitory effect is desirable.

This scenario prompted us to start a project with the goal of identifying new molecular chaperones. The first part of work was focused on the biochemical characterization of rhGAA and on the identification of the parameters to use for testing the efficacy of new PCs. In the second part, we have analysed how the new PCs affect the activity and stability of rhGAA. In particular, the effect on basal activity has allowed to evaluate if they act or not as inhibitors of rhGAA. Instead, the chaperoning action was tested in a cell free system, analyzing the effect of the selected compounds on pH and thermal stability of rhGAA. The study on cell system and *in vivo* were performed by Dr. Parenti's group of TIGEM institute, while the computational biology studies were carried out by Dr. Colombo of Istituto di Chimica del Riconoscimento Molecolare (ICRM, CNR).

## **3.7 EXPERIMENTAL PROCEDURES**

### Reagents

rhGAA (alglucosidase, Myozyme) was from Genzyme, Cambridge, MA. It is a kind gift of the Dr. Giancarlo Parenti and derive from the residual amounts of the reconstituted recombinant enzymes prepared for the treatment of PD patients at the Department of Pediatrics, Federico II University, Naples, Italy. NAC, NAS, NAG, cysteine, serine, glycine, 2-mercaptoethanol, 4-nitrophenyl- $\alpha$ -glucopyranoside (4NP- $\alpha$ -Glc), DNJ, NB-DNJ, and DGJ were from Sigma-Aldrich.

### Enzymatic assay of rhGAA

The standard activity assay of rhGAA was performed on 20 mM 4NP- $\alpha$ -Glc in 100 mM sodium acetate pH 4.0. The reaction was started by adding 5  $\mu$ g of enzyme; after 2 minutes of incubation at 37°C the reaction was blocked in ice by adding 800  $\mu$ l of 1 M sodium carbonate pH 10.2. Suitable blanks, containing all the reagents with the exception of enzyme, were always prepared to take into account the negligible spontaneous hydrolysis of the substrate. Absorbance was measured spectrophotometrically at 420 nm at room temperature and the extinction coefficient to calculate enzymatic units was 17.2 mM<sup>-1</sup> cm<sup>-1</sup>. One enzymatic unit is defined as the amount of enzyme catalyzing the conversion of 1  $\mu$ mol substrate into product in 1 minute, under the indicated conditions. All kinetic data were calculated as the average of at least two experiments and were plotted and refined with the program GraphPad Prism.

The effect of PCs and other molecules on the basal activity of rhGAA was evaluated by performing the standard activity assay in the presence of 0-10 mM of the compounds indicated in the text.

### Effect of pH and temperature on rhGAA activity

The effect of different pHs was determined in 50 mM citrate/phosphate in the range of pH 3.0–7.0. The reaction mixtures, containing 0.75 mg ml<sup>-1</sup> of enzyme, were kept at 37 °C and aliquots were withdrawn at the times indicated (0–48 hours). The residual GAA activity was measured with the standard assay. The effect of PCs and other molecules on the stability of rhGAA at pH 5.0 and 7.0 was determined as described above by adding to

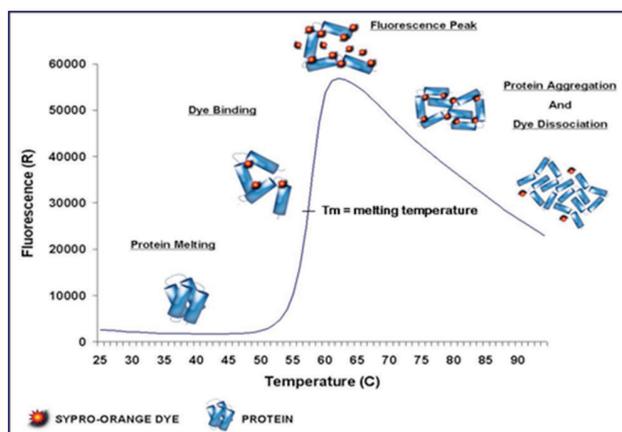
## 3.7 Experimental procedures

the reaction mixtures 0.1, 1 or 10 mM of the different compounds indicated in the text.

To test the effect of temperature, rhGAA ( $0.75 \text{ mg ml}^{-1}$ ) was incubated in 50 mM citrate/phosphate pH 5.0 or 7.0 at 60-65°C. At the times indicated (0–40 min) aliquots were withdrawn and the residual GAA activity was measured with the standard assay.

### Differential scanning fluorimetry of rhGAA

The differential scanning fluorimetry (DSF) (Niesen et al. 2007) was used to evaluate the thermal stability of rhGAA in different conditions. The thermal denaturation was monitored by measuring fluorescence of SYPRO Orange (Invitrogen), a dye that binds to the hydrophobic regions exposed by proteins undergoing thermal unfolding (Fig. 3.7.1).



**Figure 3.7.1:** Schematic representation of thermal shift assay with SYPRO Orange to monitor protein melting profiles

The measurements were performed in neutral pH buffer (25 mM sodium phosphate buffer, 150 mM NaCl, pH 7.4) or acidic pH buffer (25 mM sodium acetate, 150 mM NaCl, pH 5.2). The enzyme (2.5  $\mu\text{g}$ ) was incubated in the absence and in the presence of 5-20 mM of the compounds indicated in the results in a final volume of 25  $\mu\text{l}$ . Thermal stability scans were performed at 1°C/minute in the range 25–95°C in a Real-Time Light Cycler (Biorad, Milan, Italy). SYPRO Orange fluorescence was normalized to maximum fluorescence value within each scan to obtain relative fluorescence. To calculate the melting temperature ( $T_m$ ), the values of fluorescence were processed through an Excel file that determines minimum and maximum intensities within a selected temperature range: these two values delineate

the two asymptotes of the melting transition. The processed data were then fitted to Boltzmann equation implemented in a GraphPad Prism file. The two files used for this calculations are available for download from <ftp://ftp.sgc.ox.ac.uk/pub/biophysics> (Niesen et al., 2007).

### **Inhibition of rhGAA by DNJ**

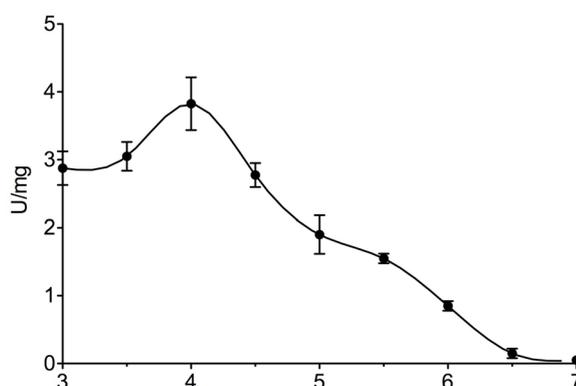
The inhibition of rhGAA was evaluated with variation of the standard activity assay. Briefly, the enzyme (5  $\mu\text{g}$ ) was assayed at various substrate concentrations (0-30 mM 4Np- $\alpha$ -Glc) in presence of 2, 8 or 16  $\mu\text{M}$  DNJ and the activity was measured as described for the standard activity assay.

## **3.8 RESULTS AND DISCUSSION**

### Biochemical characterization of rhGAA

Acid  $\alpha$ -glucosidase (GAA) is a lysosomal glycoside hydrolase belonging to CAZy family GH31 that catalyzes the hydrolysis of both  $\alpha$ -1,4 and  $\alpha$ -1,6 linkages in the natural substrate glycogen. The majority of the biochemical data on wild-type or mutant GAA derives from studies in cell-systems, such as cultured fibroblasts or transiently transfected COS7 and CHO cells (Flanagan et al., 2009; Kakavanos et al., 2006; Okumiya et al., 2007; Parenti et al., 2007; Porto et al., 2009). In the literature, biochemical data on the purified GAA are very limited (Marugan et al., 2010; Yoshimizu et al., 2008); thus, in an attempt to fill this gap we performed a general characterization of the recombinant form of human GAA (rhGAA). The enzyme used in our experiments is the pharmaceutical preparation for ERT known as Myozyme; it was a gift from Dr. Parenti and was obtained from residual solution after clinical infusions.

To determine the pH optimum of the rhGAA activity, we have evaluated the enzymatic hydrolysis of 4NP- $\alpha$ -Glc in the range of pH 3.0–7.0. As expected for a lysosomal enzyme, the maximal activity (3.8 U/mg) was obtained at pH 4.0, then, all the subsequent assays were performed at these conditions.



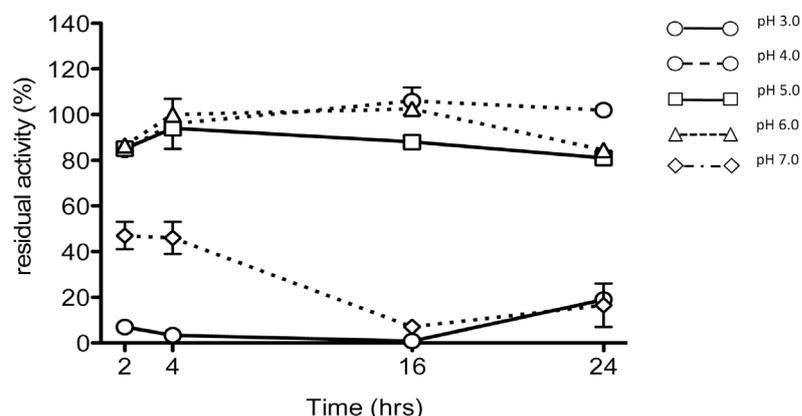
**Figure 3.8.1:** pH profile of rhGAA activity on 4NP- $\alpha$ -Glc

The steady state kinetic constants on 4NP- $\alpha$ -Glc were determined at standard conditions as described in Experimental procedures (section 3.7), and were:  $K_M=6.3\pm 0.9$  mM;  $k_{cat}=7.9\pm 0.3$  s<sup>-1</sup>;  $k_{cat}/K_M=1.26$  s<sup>-1</sup>mM<sup>-1</sup>.

## 3.8 Results and discussion

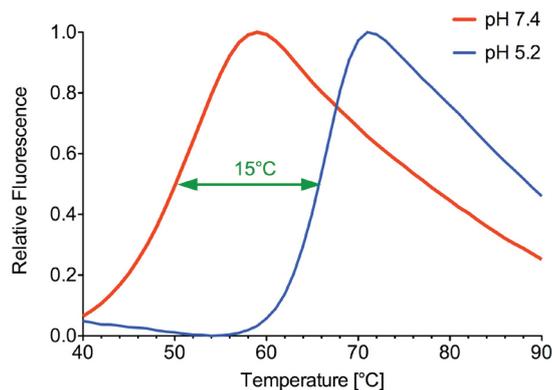
### Identification of the parameters to test the chaperoning action of new ligands

Pharmacological chaperones (PCs) are generally small molecules which have demonstrated to be promising for the treatment of diseases due to protein misfolding (Sampson et al., 2011; Valenzano et al., 2011). The efficacy of PCs is commonly monitored by evaluating the resistance of wild-type enzymes to physical stresses, such as modifications of temperature and pH (Tropak et al., 2007). Resistance to pH variations is of utmost interest, especially for the recombinant enzymes administered in ERT, as neutral pH may be representative of some of the environmental conditions encountered by these enzymes in plasma and in certain cellular compartments. To evaluate how rhGAA activity is affected by pH changes, the enzyme ( $0.75 \text{ mg ml}^{-1}$ ) was incubated in 50 mM citrate phosphate buffer in the range of pH 3.0–7.0 and the residual activity on 4NP- $\alpha$ -Glc was calculated in standard conditions. As expected for an acidophilic lysosomal enzyme, at pH 5.0, rhGAA was highly stable for up to 24 hours. At non-physiological pH representative of non-lysosomal cellular compartments, either extremely acidic (3.0) or neutral (7.0), the enzyme was unstable and rapidly lost its activity (Fig. 3.8.2).



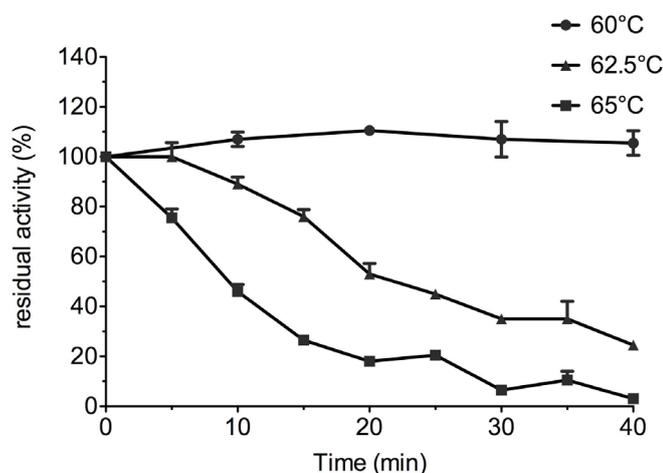
**Figure 3.8.2:** rhGAA stability in the range of pH 3.0-7.0

The thermal stability of the rhGAA structure at pH 5.2 and 7.4 was acquired by differential scanning fluorimetry (DSF) and confirmed to be greater at acidic pH compared to neutral pH. In fact the melting temperature ( $T_m$ ) measured for the enzyme exposed at pH 5.2 is  $65.6 \pm 0.2^\circ\text{C}$  while at pH 7.4 is  $50.2 \pm 0.1^\circ\text{C}$  (Fig. 3.8.3).



**Figure 3.8.3:** Temperature profile of rhGAA at pH 5.2 and 7.4

In addition, the thermal stability of rhGAA activity was monitored by incubating the enzyme ( $0.75 \text{ mg ml}^{-1}$ ) in 50 mM citrate phosphate buffer pH 5.0 in the range of temperature 60–65°C and, at interval times, the residual activity was measured in standard conditions (Fig. 3.8.4). rhGAA is stable for 40 min at 60°C (residual activity≈100%), but its activity is drastically reduced after 20 min at 62.5°C (residual activity≈53%), and 10 min at 65°C (residual activity≈46%).



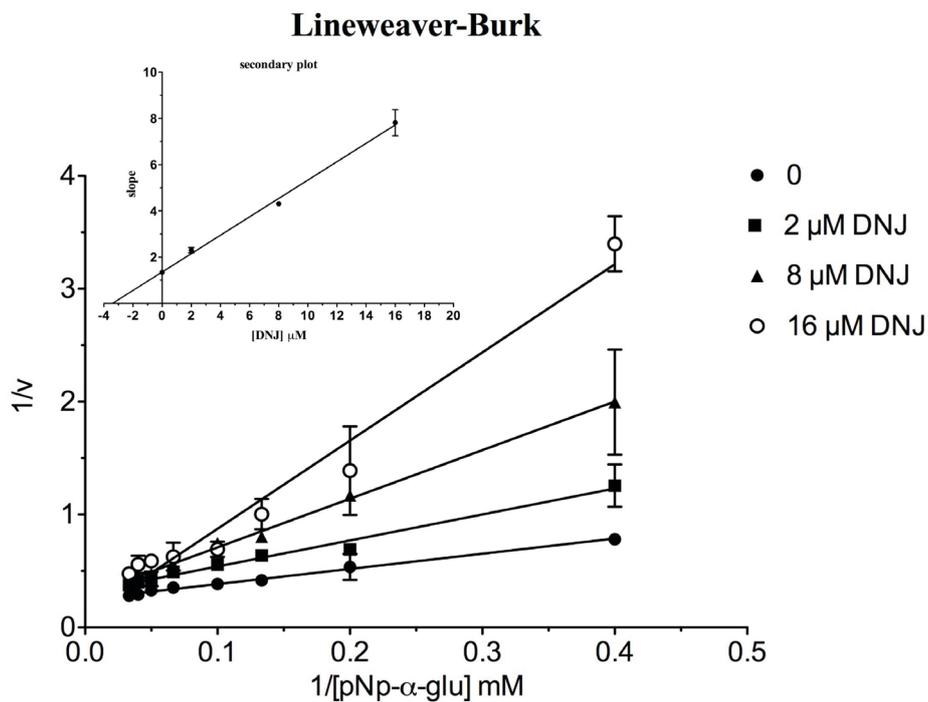
**Figure 3.8.4:** Thermal stability of rhGAA at pH 5.0

We concluded that both temperature and pH can be used as parameters to test the efficacy of new PCs. In particular, we can estimate the chaperoning action of new compounds by evaluating their influence on the stability of enzyme exposed at different temperatures and pHs.

## 3.8 Results and discussion

### DNJ: inhibitory and chaperoning action

1-Deoxynojirimycin (DNJ) is a compound belonging to the class of imino sugars that act as competitive inhibitors of rhGAA (Yoshimizu et al. 2008) and have showed a potential as PCs for PD (Parenti et al. 2007). We evaluated the kinetics of DNJ inhibition by assaying rhGAA in the presence of 2, 8 or 16  $\mu\text{M}$  inhibitor. The initial rate of rhGAA inhibition at various DNJ concentrations was fit to the Lineweaver-Burk (LB) equation (Fig. 3.8.5). As expected, the LB plot showed that the  $V_{\text{max}}$  is unaffected by the presence of DNJ while the effective  $K_{\text{M}}$  on 4NP- $\alpha$ -Glc is increased, confirming a competitive inhibition. The slopes of the LB graphs were then plotted as function of the DNJ concentration with the linear regression analysis (secondary plot). The x-intercept represents the  $K_i$  value (Fig. 3.8.5) of 3.4  $\mu\text{M}$ , which is 7-fold higher than that reported on MU- $\alpha$ -Glc (Yoshimizu et al., 2008), probably due to the different affinity of enzyme for the two synthetic substrates or to the diverse sensitivity of the assay procedures adopted.



**Figure 3.8.5:** Kinetics of rhGAA inhibition by DNJ

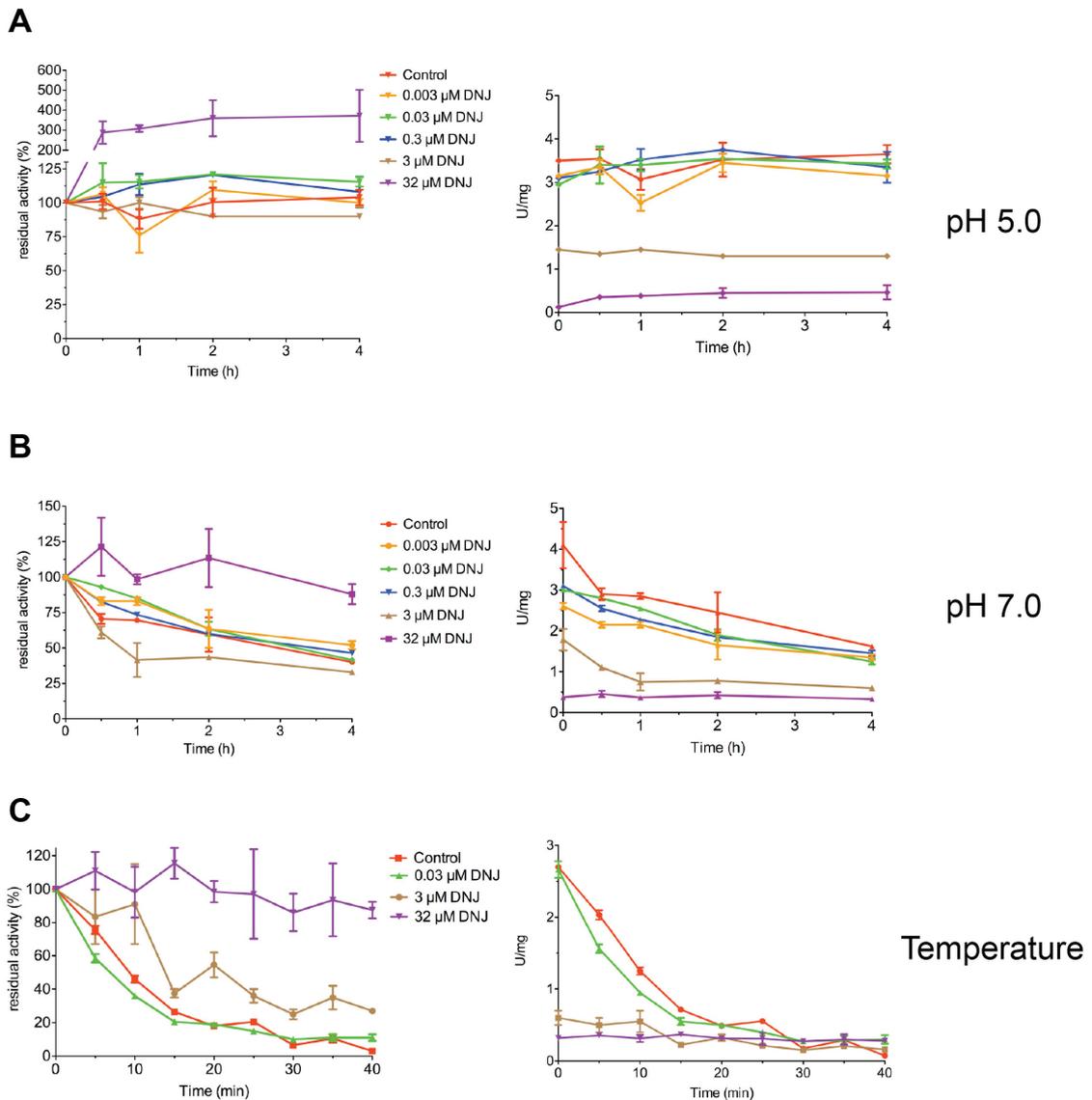
It has been reported that DNJ is able to rescue the activity of some GAA mutants, as observed in fibroblasts deriving from PD patients or in COS7 cells

transfected with GAA mutants (Flanagan et al., 2009; Parenti et al., 2007). Moreover, DNJ can function in synergy with ERT enhancing the stability of the wild-type rhGAA (Flanagan et al., 2009). In fact, DSF analysis indicated that 100  $\mu\text{M}$  DNJ increased the thermal stability of rhGAA at pH 7.4 to a similar level to that observed at pH 5.2 in the absence of DNJ (Flanagan et al., 2009).

On the basis of these observations, we have analysed the effect of DNJ on the activity of rhGAA subjected to various chemical-physical stresses. In particular, we have evaluated the residual activity of rhGAA incubated at pH 5.0 and 7.0 or at 65°C in absence or presence of various concentrations of DNJ (in the range  $3 \times 10^{-3}$  - 32  $\mu\text{M}$ ). As depicted in the Fig. 3.8.6A (left panel), the enzyme maintained its activity for at least 4 hours at pH 5.0 both in absence and in presence of DNJ. In the presence of 32  $\mu\text{M}$  DNJ the residual activity of rhGAA increased by about 3-fold, but at this concentration, which is 10-fold the measured  $K_i$ , DNJ severely inhibited the enzyme (Fig. 3.8.6A, right panel). The stabilizing effect of DNJ on rhGAA is clearly observable at pH 7.0 (Fig. 3.8.6B; left panel). At these conditions, the enzymatic activity of the control was reduced after 4 hours of about 50%, but 32  $\mu\text{M}$  DNJ clearly stabilizes the enzyme (Fig. 3.8.6B; left panel). But, also in this case, remarkable inhibition of DNJ is observed (Fig. 3.8.6B; right panel).

The protective effect of DNJ can be observed also on rhGAA incubated at 65°C. At these conditions, 0.03  $\mu\text{M}$  of the inhibitor had no protective effect while 32  $\mu\text{M}$  DNJ clearly stabilize the enzyme (Fig. 3.8.6C; left panel), even if at the cost of a severe inhibition (Fig. 3.8.6C; right panel). These data confirmed the stabilizing effect of DNJ as observed in cultured fibroblasts and transfected COS7 cells (Flanagan et al., 2009; Parenti et al., 2007). On the other hand it is worth mentioning the inhibitory action exerted by this compound, which could contribute to reduce its chaperoning effect. Additionally, the informations achieved by our experiments support the use of rhGAA stability at various pHs and temperatures as the correct parameters to follow the chaperoning action of new compounds.

## 3.8 Results and discussion



**Figure 3.8.6:** Stability of rhGAA at pH 5.0 (A), pH 7.0 (B) and at 65°C (C) in presence or absence of DNJ (range  $3 \times 10^{-3}$ -32µM). Left panels report on the residual activity of rhGAA while right panels, depicting the specific activity of rhGAA at the various conditions, show the inhibitory effect of the different concentrations of DNJ

### Identification of NAC as new potential PC

PCs identified so far, such as DNJ and derivatives, act as inhibitors of the enzyme target (Valenzano et al., 2011) and are effective in rescuing only some disease-causing missense mutations. For PD, it is possible to speculate that ~10-15% patients may be amenable to PCT (Flanagan et al., 2009). Accordingly, it emerges the need to identify a second generation of PCs with better enhancing effect and without inhibitory action.

## Results and discussion 3.8

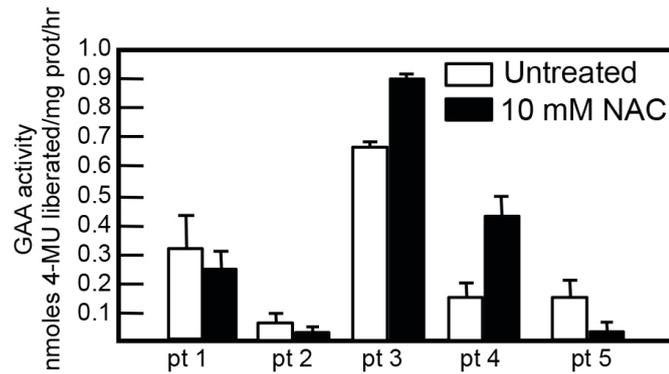
It has been reported that the clinical course of some LSDs is associated with a condition of cellular oxidative stress (Vázquez et al., 2011) and it has been suggested that chaperones might be beneficial for its treatment (Wei et al., 2008). In this framework, our collaborators of Parenti's group tested the effect of some antioxidant drugs (resveratrol, epigallo catechingallate and N-acetylcysteine) on PD fibroblasts (Porto et al., 2012). Among the compounds tested, only N-acetylcysteine (NAC), a known antioxidant currently used as a mucolytic drug, has proved to be effective (Porto et al., 2012). In particular, Parenti's group has investigated the effect of NAC in cultured fibroblasts deriving from five PD patients that carry different mutations and with different phenotypes (Table 3.8.1).

pt No	Phenotype	Genotype
1	Severe	p.W367X / p.G643R
2	Severe	p.H612_D616del- insRGI / p.R375L
3	Intermediate	p.L552P / aberrant splicing
4	Intermediate	c.-32-13T>G - S619N / L552P
5	Intermediate	G549R / aberrant splicing

**Table 3.8.1:** Genotype and phenotype of PD fibroblasts used in this study

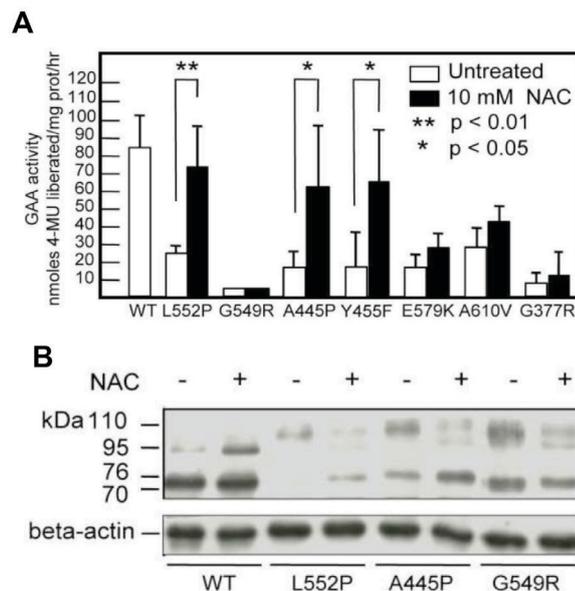
At the concentration of 10 mM, NAC enhanced the residual activity of mutated GAA in fibroblasts from patients 3 and 4 (Fig. 3.8.7).

### 3.8 Results and discussion



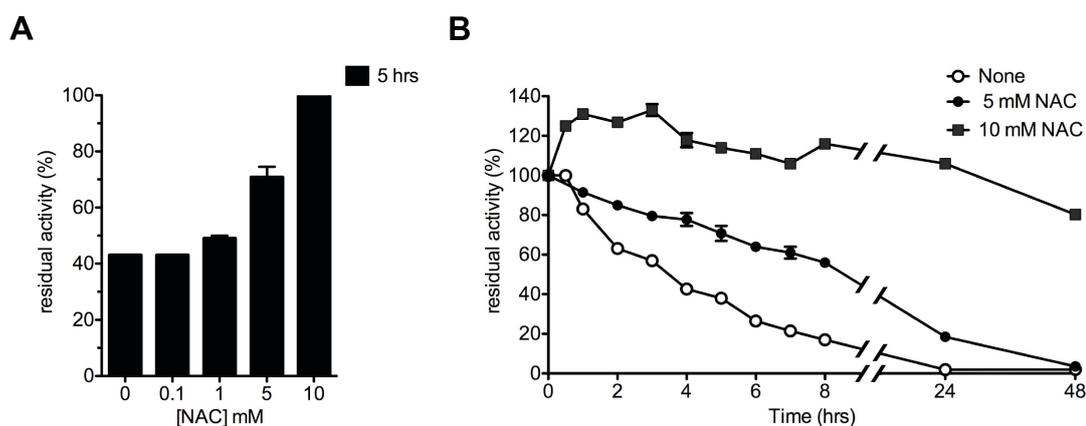
**Figure 3.8.7:** Effect of NAC on the activity of mutated GAA in PD fibroblasts

To determine which mutations are responsive to NAC, Parenti's group expressed a panel of mutated GAA gene constructs in COS7 cells and evaluated the effect of NAC in rescuing the activity of these mutants. As illustrated in Fig. 3.8.8A, NAC rescued the GAA activity of mutants carrying the mutations L552P, A445P, and Y455F. Moreover, the responsive mutants also showed an improved processing into 76 and 70 kDa active isoforms as observed by western blot analysis (Fig. 3.8.8B). Since the GAA precursor is converted into the active forms in the late-endosomal/lysosomal compartment (Wisselaar et al., 1993), this also indicates that NAC improved the trafficking of the enzyme to the lysosomes.



**Figure 3.8.8:** NAC rescued the activity (A) and improved the maturation (B) of mutated GAA expressed in COS7 cells

To assess if the enhancing effect of NAC observed in cell system was due to a correction of oxidative stress or to a direct stabilizing action on GAA, we evaluated the chaperoning action of NAC in a cell-free system by testing the pH and temperature stability of rhGAA as parameters. In particular, we incubated for 5 hrs rhGAA at pH 7.0 in the presence of increasing concentrations of NAC (0.1-10mM). As illustrated in Fig. 3.8.9A, NAC enhanced rhGAA activity at pH 7.0 in a dose-dependent manner and at the concentration of 10 mM about 90% of the activity is maintained even after 48 hours of incubation (Fig. 3.8.9B).



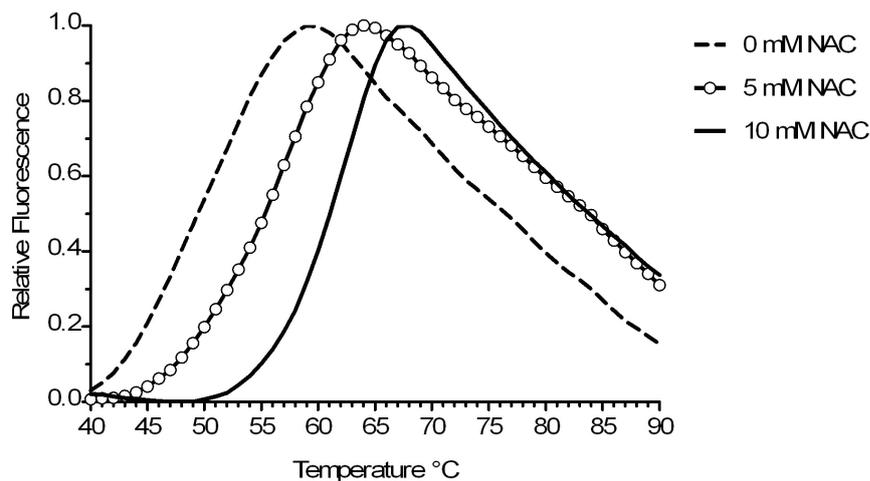
**Figure 3.8.9:** NAC stabilized rhGAA in a dose-dependent manner (A) and the protective effect persisted even after 48 hours at the concentration of 10 mM (B)

The pH of solution containing 10 mM NAC was 7.0, the same value measured in absence of NAC; thus, the stabilization of rhGAA observed in presence of NAC was not due to changes of pH.

NAC demonstrated also to be able to increase the thermal stability of rhGAA structure at pH 7.0. In fact, DSF analysis showed that the  $T_m$  of enzyme co-incubated with 10 mM NAC is  $10.5 \pm 0.6^\circ\text{C}$  higher than  $T_m$  of rhGAA alone (Fig. 3.8.10).

These data suggest that NAC could act as molecular chaperone. To support and validate this hypothesis, and to try to have hints on the stabilization mechanism at molecular level, we tested other compounds.

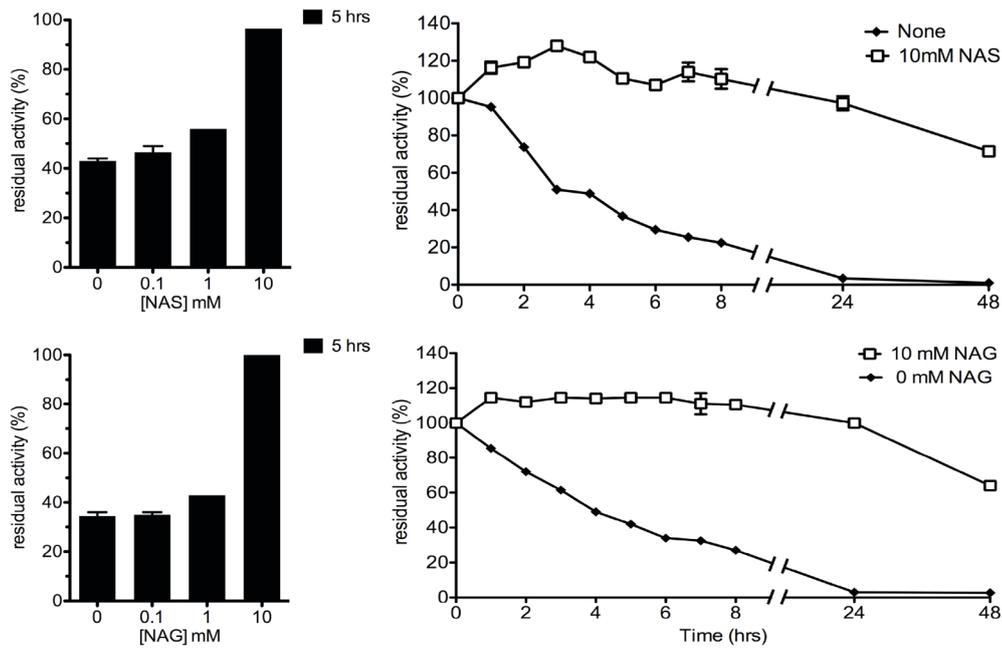
## 3.8 Results and discussion



**Figure 3.8.10:** DSF analysis showing that NAC increased the thermal stability of rhGAA

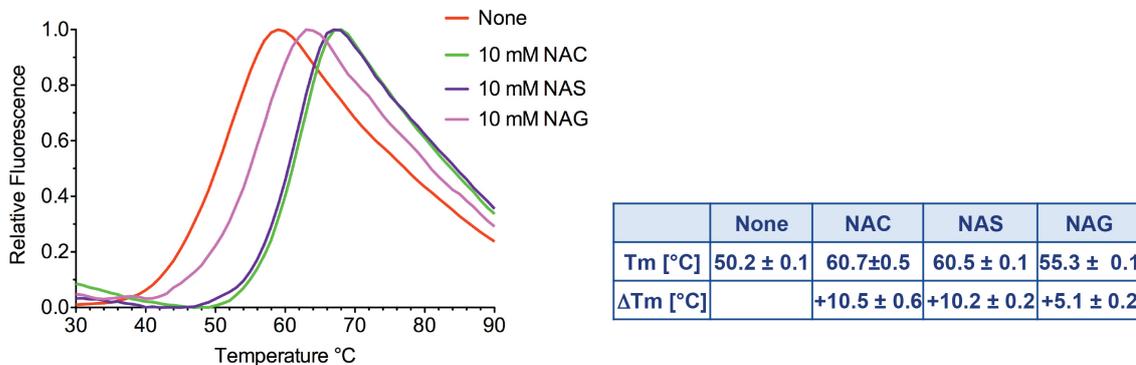
### Understanding the molecular determinants of NAC stabilizing effect

To understand which are the structural features of NAC responsible for its stabilizing action on rhGAA, we tested other compounds. In particular, we evaluated the stabilizing effect of structurally similar acetylated amino acids, such as N-acetyls erine (NAS) and N-acetylglycine (NAG), the non-acetylated counterparts, such as cysteine (Cys), serine (Ser) and glycine (Gly) and the structurally unrelated compound 2-mercaptoethanol (2ME). NAS and NAG behaved as NAC enhancing rhGAA stability at pH 7.0 in a dose-dependent manner and the stabilizing effect persisted even after 48 hours of incubation at the concentration of 10 mM (Fig. 3.8.11).



**Figure 3.8.11:** NAS and NAG stabilized rhGAA in a dose-dependent manner (graphs on the left) and the protective effect persisted even after 48 hours at the concentration of 10 mM (graphs on the right)

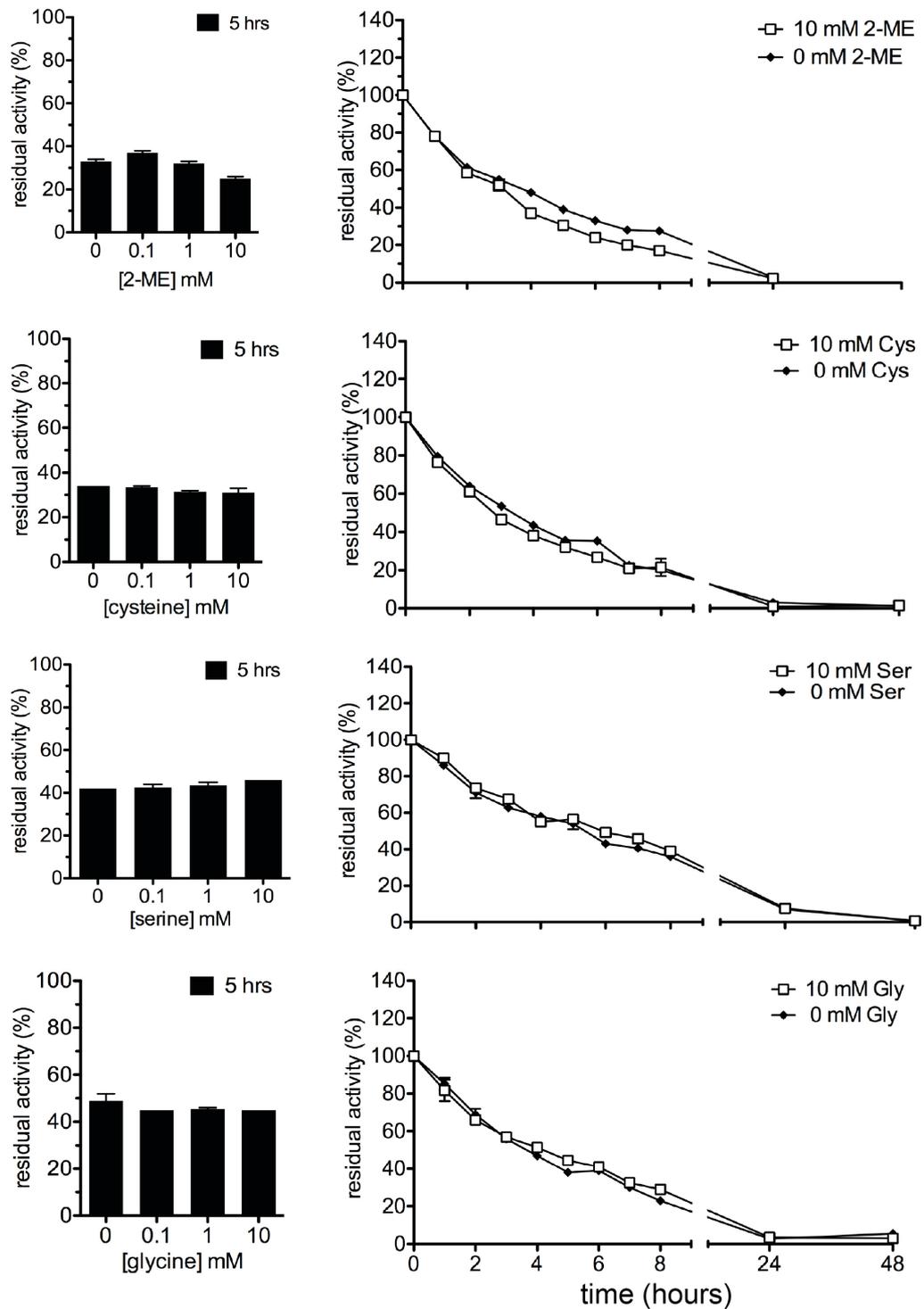
Moreover, DSF analysis showed that NAS and NAG improved the thermal stability of rhGAA structure like NAC (Fig. 3.8.12), increasing the  $T_m$  of the enzyme of  $10.2 \pm 0.2^\circ\text{C}$  and  $5.1 \pm 0.2^\circ\text{C}$ , respectively.



**Figure 3.8.12:** NAS and NAG thermally stabilized rhGAA increasing its melting temperature by  $10.2 \pm 0.2^\circ\text{C}$  and  $5.1 \pm 0.2^\circ\text{C}$ , respectively

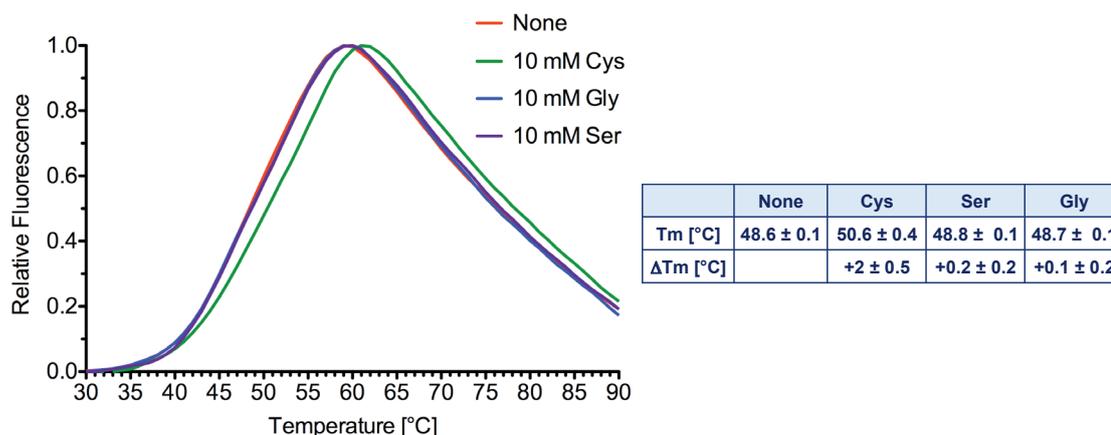
By contrast, the non-acetylated homologs, Cys, Ser and Gly, and the unrelated compound 2ME did not prevent enzyme inactivation at pH 7.0 at all concentrations tested (0.1-10 mM) (Fig. 3.8.13).

### 3.8 Results and discussion



**Figure 3.8.13:** The non-acetylated amino acids Ser, Cys and Gly and 2ME had no effect on rhGAA stability at all conditions tested

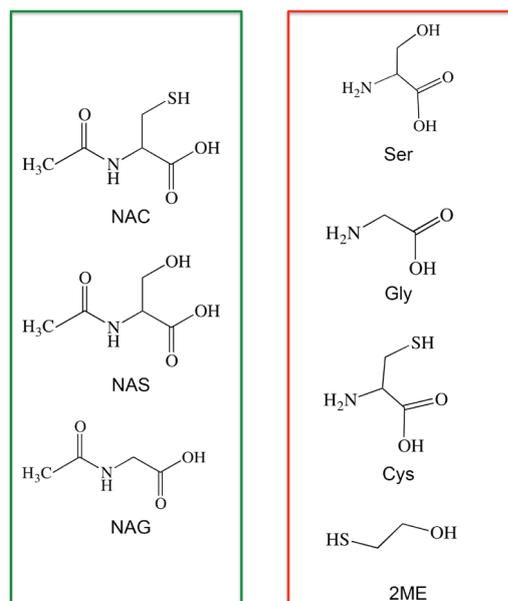
Also the thermal profile of rhGAA at neutral pH resulted unaffected by the non-acetylated amino acids (Fig. 3.8.14).



**Figure 3.8.14:** DSF analysis of rhGAA thermal profile in presence of 10 mM Cys, Ser or Gly

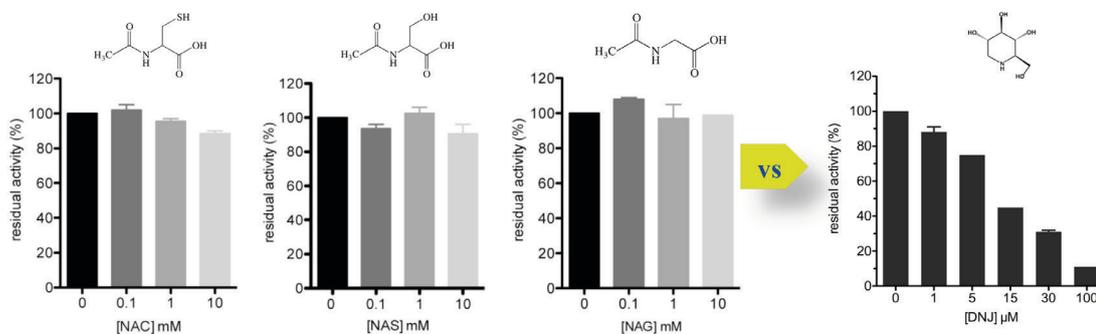
These data allowed us to assert that NAC acts as molecular chaperone rather than as antioxidant. NAS and NAG showed a chaperoning action like NAC, even though they lack the sulfhydryl group that is substituted by an hydroxyl and a hydrogen atom in NAS and NAG, respectively (Fig. 3.8.15). Moreover, 2ME and Cys, which carry a sulfhydryl group (Fig. 3.8.15), did not showed any protective effect on rhGAA. Thus, the stabilizing effect of NAC seems not due to the presence of sulfhydryl group and neither is associated to an antioxidant action. This last concept is also supported by the results achieved by Parenti's group in cell-system, where the other antioxidants studied did not produced any chaperoning effect. Instead, the chaperoning action of NAC seems correlated to the acetyl moiety which is a common feature of NAC, NAS, and NAG, while it is missing in Cys, Ser, and Gly. The identification of chaperoning activity in these molecules was somewhat surprising because they are structurally very different from imino sugars, the only PCs of GAA known so far, which resemble the natural substrates/products of the enzyme.

## 3.8 Results and discussion



**Figure 3.8.15:** Structures of the compounds analyzed. Green box encloses compounds active as molecular chaperones while the red box delimits the inactive molecules

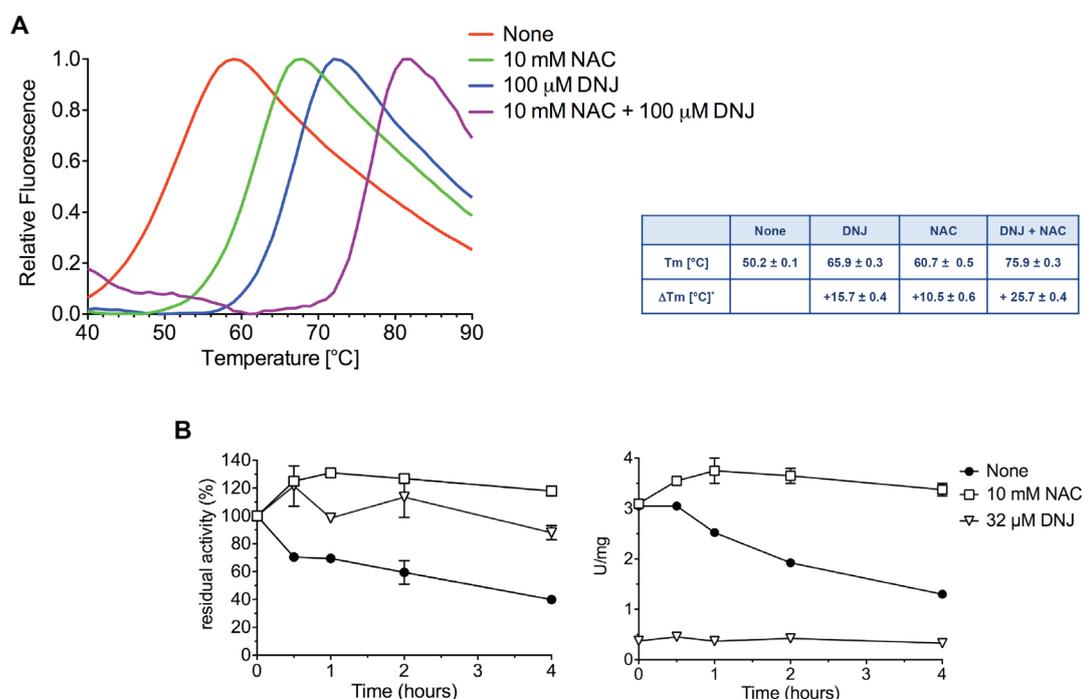
To assess if the new molecular chaperones identified here act as rhGAA inhibitors, we have evaluated the basal activity of enzyme in presence of increasing concentrations of NAC, NAS and NAG (0.1-10 mM). It is worth noting that none of these ligands inhibited rhGAA by concentrations up to 10 mM. In contrast to DNJ whose stabilizing effect produced also a severe inhibition of the  $\alpha$ -glucosidase activity, the new PCs stabilized rhGAA and never interfered with its basal activity at all the conditions tested (Fig. 3.8.16). These data suggest that NAC and the related compounds, NAS and NAG, are the first molecular chaperones that did not interact with the active site of rhGAA.



**Figure 3.8.16:** Analysis of rhGAA basal activity in presence of NAC, NAG and NAS (0.1-10 mM) and DNJ (1-100  $\mu$ M)

### NAC/NAS/NAG and DNJ: comparison and combination

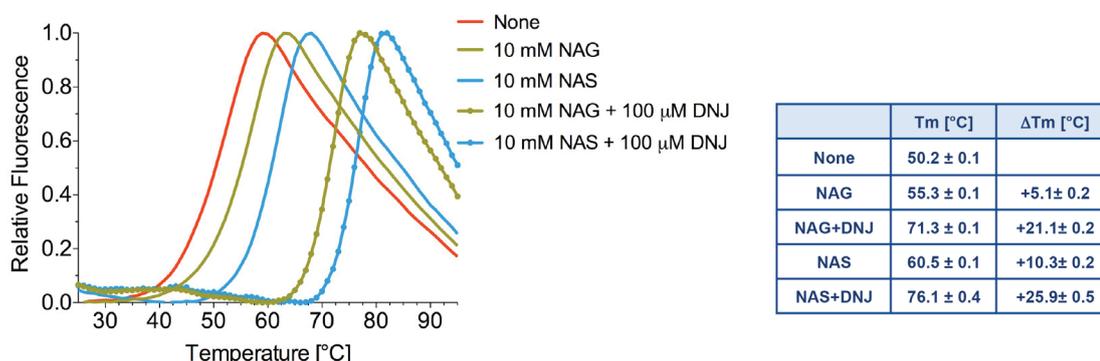
The data acquired in the previous paragraph led us to suppose that the new molecular chaperones and the imino sugars could bind to different domains of rhGAA. To test this hypothesis we evaluated how the combination of NAC and DNJ affects the thermal stability of rhGAA. These compounds individually increased the  $T_m$  of rhGAA at neutral pH, with DNJ alone apparently causing the best shift in  $T_m$  (Fig. 3.8.17A) but at the cost of an almost complete inhibition (Fig. 3.8.17B). Interestingly, the combination of NAC and DNJ resulted in the highest stabilization of the enzyme ( $T_m=75.9\pm 0.3^\circ\text{C}$ ) (Fig. 3.8.17A). The stabilizing effect of the two PCs was additive indicating that NAC and DNJ interact with a non-catalytic and the catalytic domain of rhGAA, respectively.



**Figure 3.8.17:** Effect of the combination of NAC and DNJ on the thermal stability of rhGAA (**A**). Basal activity of rhGAA in presence of 10 mM NAC or 32  $\mu\text{M}$  DNJ (**B**)

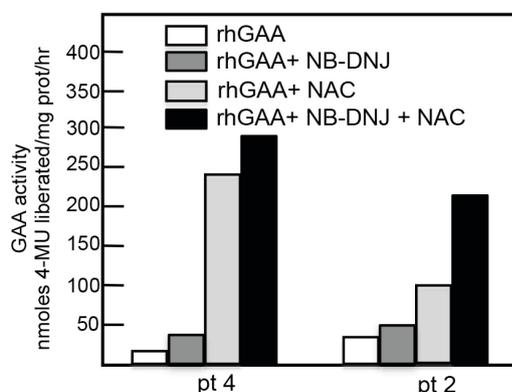
### 3.8 Results and discussion

The cumulative effect was also observed with the combination of NAS or NAG and DNJ (Fig. 3.8.18).



**Figure 3.8.18:** Effect of the combination of NAS or NAG and DNJ on the thermal stability of rhGAA

Moreover, the combination of NAC with the imino sugars gave similar results also in fibroblasts from two PD patients, as observed by Parenti’s group. These cell lines were incubated with rhGAA, with rhGAA plus either NAC or NB-DNJ, and with rhGAA plus the combination of the two chaperones. In both cell lines the combination of NAC and NB-DNJ resulted in the highest enhancement of GAA activity by rhGAA (Fig. 3.8.19).



**Figure 3.8.19:** Combination of NAC and DNJ in PD fibroblasts

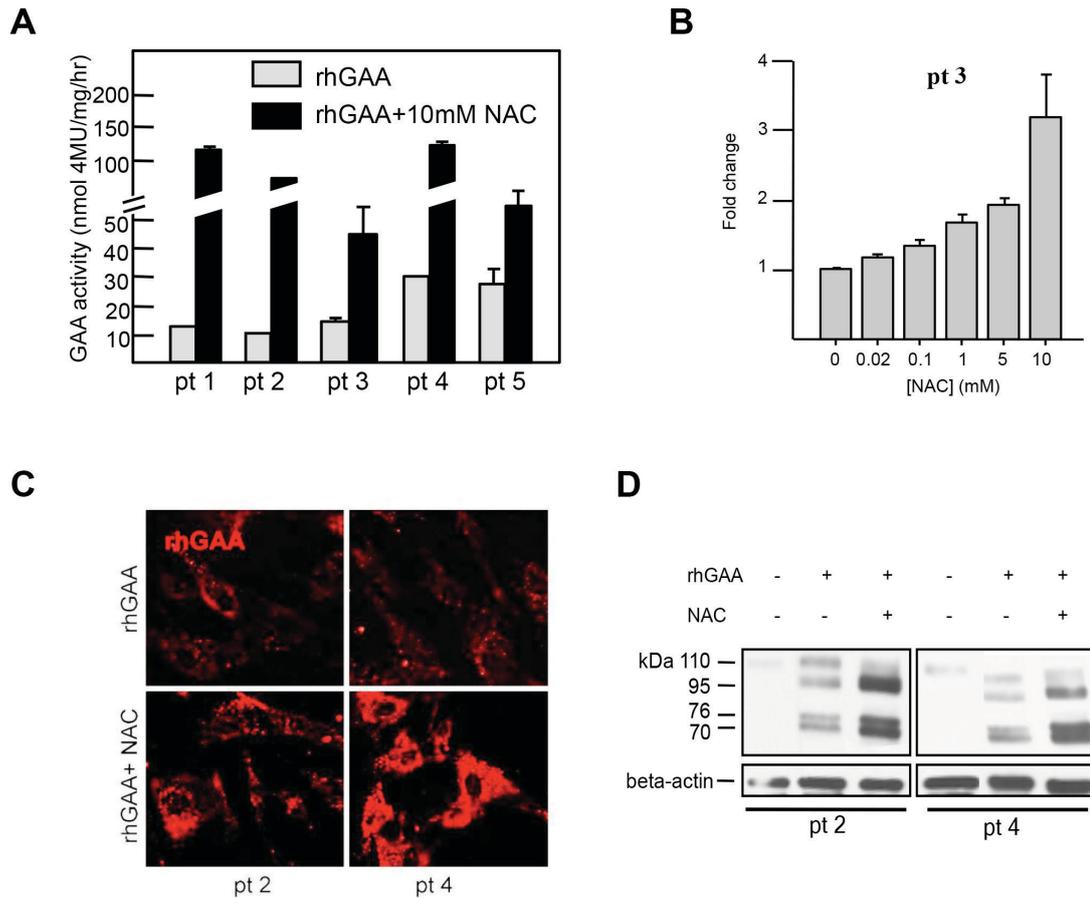
These data support the hypothesis that NAC, NAS, and NAG interact with a protein domain different to that bound by the imino sugars. Accordingly, the combination of the two molecular chaperones might represent an additional advantage for the treatment of PD patients, in order to obtain the best stabilization of rhGAA and the highest synergy with ERT.

### Synergy between NAC and Myozyme in cell system and *in vivo*

The experiments in cell-free system have demonstrated that NAC and the related compounds NAS and NAG enhanced the activity/stability of rhGAA. The synergy between these molecular chaperones and ERT has been validated also in cell-system and *in vivo* by Parenti's group.

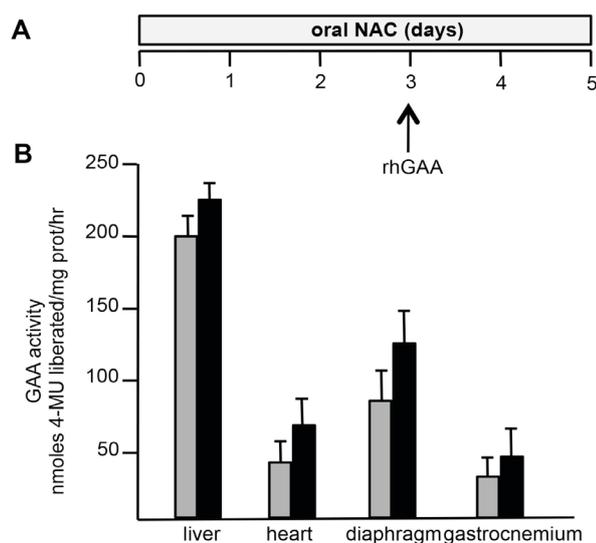
The synergistic effect between rhGAA and NAC was evaluated in fibroblasts from five PD patients. When these cell lines were co-incubated with 50  $\mu$ M rhGAA and 10 mM NAC, the correction of GAA deficiency improved with increments in GAA activity ranging from ~3.7 to 8.7-fold, if compared to the activity of cells treated with rhGAA alone (Fig. 3.8.20A). As already observed in cell free-system (Fig. 3.8.8), also in PD fibroblasts NAC showed a dose-dependent enhancing effect with the maximal increase of 3.3-fold at the concentration of 10 mM (Fig. 3.8.20B). The enhancing effect of GAA activity largely exceeded that due to the rescue of the native mutated enzyme (patients 3 and 4) and was observed also in cell lines non-responsive to NAC alone (patients 1, 2, and 5). These data suggest that NAC acts as chaperone of the wild-type recombinant enzyme also in cell-system. Further indications supporting this hypothesis derive from immunofluorescence analysis and confocal microscopy performed by Parenti's group. This approach allows to detect only the fluorochrome-labeled exogenous enzyme. In the presence of NAC the amounts of fluorochrome-labeled GAA increased compared to cells incubated with fluorescent GAA alone (Figure 3.8.20C). Moreover, western blot analysis (Fig. 3.8.20D), showed increased amounts of GAA active isoforms in the cells treated with NAC and rhGAA, compared to cells treated with the recombinant enzyme alone. Since the GAA precursor is converted into the active forms in the late-endosomal/lysosomal compartment (Wisselaar et al., 1993) this indicates improved lysosomal trafficking of the enzyme.

### 3.8 Results and discussion



**Figure 3.8.20:** Synergy between NAC and rhGAA in PD fibroblasts. Improved GAA deficiency in cell lines co-incubated with rhGAA and NAC (**A**). NAC enhanced rhGAA efficacy in fibroblast of patient 3 in a dose-dependent manner (**B**). Increased amounts of fluorochrome-labelled GAA in the presence of NAC (**C**). Western blot analysis showing an improved processing of GAA precursor in cells treated with rhGAA and NAC

Parenti's group also evaluated the combination of NAC and rhGAA in a mouse model of PD (N Raben et al., 1998). Mice were treated with a single injection of rhGAA at high doses (100mg/kg) in combination with oral NAC for 5 days (Fig. 3.8.21A). Forty-eight hours after rhGAA injection the animals were euthanized and GAA activity was assayed in different tissues. In all tissues examined (liver, heart, diaphragm and gastrocnemium) the combination of NAC and rhGAA resulted in higher levels of GAA compared to rhGAA alone (Fig. 3.8.21B).



**Figure 3.8.21:** Synergy between NAC and rhGAA *in vivo*. Mice were treated with oral NAC for 5 days and received an rhGAA injection on day 3 (A). Improved correction of GAA deficiency in tissues of animals treated with NAC and rhGAA (black bars) compared to the controls treated only with rhGAA (grey bars) (B)

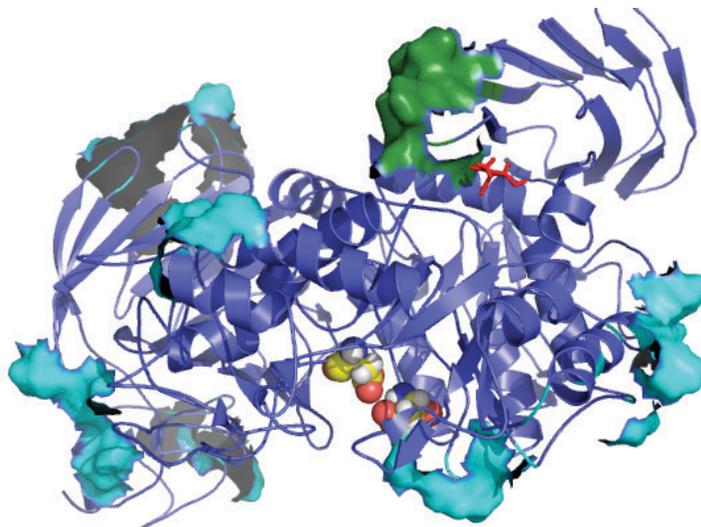
### Modelling the interaction of NAC with the rhGAA

The following computational biology studies, including molecular dynamics simulations and docking experiments, were performed by the group of Dr. Giorgio Colombo of Istituto di Chimica del Riconoscimento Molecolare, Consiglio Nazionale delle Ricerche (ICRM-CNR). The 3D structure of rhGAA is not yet available, so an homology model of GAA was constructed by using the structure of the human maltase-glucoamylase (pdb code 3L4Z) as a template (Sim et al., 2010) (for more details see Porto et al. 2012). Molecular dynamics (MD) simulations in conditions mimicking pH 5.0 and 7.0 showed that the rhGAA structure was more stable at acidic than at neutral pH.

Combining MD simulations with molecular docking experiments, Dr Colombo's group identified the regions of the protein surface that may be most frequently in contact with NAC (Fig. 3.8.22). During the simulation it was observed that NAC bound on specific regions of the protein. A control simulation carried out with non-acetylated Gly in the same conditions showed that this amino acid, which has no chaperoning effect, had a much lower tendency to bind in the vicinity of the surface. Interestingly, one of the identified areas of favourable contact between NAC and GAA partially overlapped the region defined by residues Arg525, Gly793, Glu794, Ser795, Leu796, Glu797, and Gly802 that

### 3.8 Results and discussion

defines a stable pocket, which is at 35 Å distance from the active site (Fig. 3.8.22). This region is conserved among characterized GH31  $\alpha$ -glucosidases from humans and other mammals (Fig. 3.8.23), and is located at the boundary between two domains (one of which contains the active site).



**Figure 3.8.22:** Identification of the areas of favorable NAC contact. The green surface indicated the pocket defined by residues Arg525, Gly793, Glu794, Ser795, Leu796, Glu797, and Gly802. The cyan surfaces indicate alternative NAC contact sites; it is possible to appreciate that they are all superficial and do not define cavities suitable for small-molecule binding. The active site is indicated by yellow van der Waals representations of the atoms

<i>Human_LYAG</i>	485	AATICASSHQFLSTHYNLHNLVGLTEAIAASHRALVKAR	-GTRPFVISRSTFA	535
<i>Bovin_LYAG</i>	541	AATICASSHQFLSTHYDLHNLVGLTEALASHRALVKAR	-GMRPFVISRSTFA	591
<i>Mouse_LYAG</i>	554	AATICASSHQFLSTHYNLHNLVGLTEAIASSRALVKTR	-GTRPFVISRSTFS	604
<i>Human_MGA</i>	565	CKTLCMDAVQHWGKQYDIHNLVGYSMVAATAEAAKT	VFPNKRSEILTRSTFA	616
<i>Human_SUIS</i>	541	SKTICMDAVQNWGKQYDVHSLVGYSMIAATEQAVQK	VFPNKRSEILTRSTFA	592
<i>Human_LYAG</i>	793	GESLEVLERGAYTQVIFLARNTIVNELVRVTS	-EGAGLQLQKVTVLGVATA	843
<i>Bovin_LYAG</i>	847	GESLGVLDGGDYTQLIFLAKNNTFVNKLHVHVS	-EGASLQLRNVTVLGVATA	897
<i>Mouse_LYAG</i>	863	GESLAVLERGAYTLVTFSAKNNTIVNKLVRVTK	-EGAELQLREVTVLGVATA	913
<i>Human_MGA</i>	859	GETKDTVANKVYLLCEF SVTQNRLEVNI	SQSTYKDPNNLAFNEIKILG	909
<i>Human_SUIS</i>	835	GETKDTIQNGNYILYTF SVSNNTLDIVCTHSSYQ	EGTTLAFQTVKILGLTDS	886

Human\_LYAG: human lysosomal alpha-glucosidase P10253 (aa 70-952)  
 Bovin\_LYAG: bovin lysosomal alpha-glucosidase Q9MYM4  
 Mouse\_LYAG: Mouse lysosomal alpha-glucosidase P70699  
 Human\_MGA: Human maltase-glucoamylase, intestinal O43451  
 Human\_SUIS: Human Sucrase-isomaltase, intestinal P14410

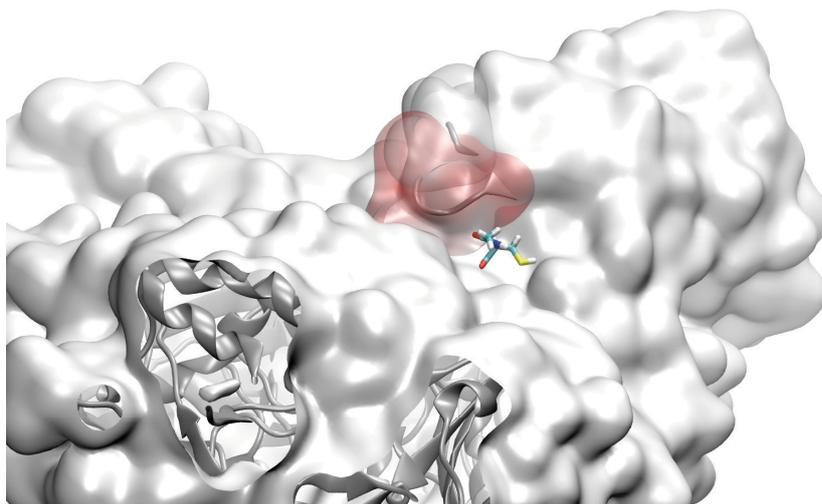
**Figure 3.8.23:** Multiple sequence alignment of GH31  $\alpha$ -glucosidases. Red boxes indicate residues of GAA that are predicted in-silico to interact with NAC (R525; G793; E794; S795; L796; E797; G802)

Remarkably, the binding energies for NAC at this site were comparable to the binding energies calculated for the docking of maltose or DNJ to the GAA active site and similar results were obtained also with NAS and NAG (Tab. 3.8.2).

Molecule	Score	Site
NAC	-4,3342	Allosteric
NAG	-4,0682	Allosteric
NAS	-4,4576	Allosteric
MAL	-4,9417	Active Site
DNJ	-6,6378	Active Site

**Table 3.8.2:** Grid Docking scores (arbitrary units) of acetylated compounds to the allosteric site, and of maltose and DNJ to the active site

Moreover, in the model of the best-ranked pose, the -SH group is directed toward the surface of the protein whereas the acetyl group interacts with the allosteric GAA pocket described above (Fig. 3.8.24) providing further indication that the SH group was not crucial for the binding of the molecule to the enzyme. A control docking run was carried out with DNJ. Importantly, the minimum free energy and most populated docking solution shows DNJ binding preferentially in the active site of GAA, with a lower energy than that measured for the pharmacological chaperone in the allosteric pocket, showing the high affinity of this ligand for the active site and supporting the validity of model used in this study.



**Figure 3.8.24:** Detail of the GAA model with NAC bound to the hot spot: the sulfur atom is highlighted in yellow and the binding site is in red filled spaces

Finally, the GAA-NAC complex refined by MD simulations showed that the presence of NAC at pH 7.0 confers to GAA a local flexibility profile similar to the one of unbound GAA at pH 5.0. Additionally, this effect is not limited to the

### 3.8 Results and discussion

---

immediate vicinity of the NAC-binding site, but extends to distal regions of the protein. As a control, the same procedure was carried out using Gly as a possible ligand, but Gly appears to non-specifically sample many regions of the protein surface with very low affinity. These observations give support to the molecular mechanism of the chaperoning activity of NAC.

Overall the computational data validated the results of *in vitro* experiments concerning both the stabilization through an allosteric pocket rather than active site and the involvement of acetyl group rather than -SH in the binding to the target enzyme.

#### Concluding remarks

We have identified the first molecular chaperones, NAC, NAS and NAG, direct to an allosteric site rather than to active site of GAA. These compounds have proved to enhance the stability/activity of rhGAA at neutral pH in cell-free system, to rescue the activity of some GAA mutants and to improve the efficacy of ERT in cell system and *in vivo*. The synergy between PCT and ERT has important clinical implications. First, it has to consider that in most cases the use of PCT alone, enhancing the endogenous defective enzymes, resulted in minor rescue of enzymatic activity (likely with a modest impact on patients' outcome). By contrast, when the PCs are used in synergy with ERT, the correction of the defective activity is remarkably improved. An additional advantage of the combination of PCT and ERT is that the effect of chaperones is directed towards the wild-type recombinant enzyme and does not depend on the type of mutation carried by patients. This allows to expand the spectrum of treatable patients, including those carrying mutations not or poorly responsive to PCT alone.

The data obtained in PD mice appear promising, although treatment protocols with NAC need to be carefully optimized in studies on a larger number of animals. The concentrations of NAC used in our studies may appear relatively high and difficult to translate into human therapy. However, toxicity of NAC is reported to be low and plasma concentrations in the range of 2-3 mM are usually reached for the treatment of paracetamol overdose (Walsh et al., 1998). Moreover, long-term treatment with NAC doses of 5g/kg/day (higher than doses used in the *in vivo* experiments in our study) has been used in the mouse model

of ethylmalonic encephalopathy and resulted in improved survival of treated animals (Viscomi et al., 2010). It worth mentioning that, for a possible clinical translation of our data, a therapeutic protocol would be based on the administration of the chaperone at the time of ERT infusion, and not on a long-term daily administration of the chaperone. By using such an approach the risk of adverse events would likely be reduced. It is also possible that modifications to NAC might be required to have the stabilizing effect at lower concentrations. The identification of NAC and derivatives, which are structurally very different from the other known PCs identified in PD is promising. In fact, other molecules, whose chaperoning activity cannot be simply inferred from their structure, may be effective in several LSDs, thereby opening new and wider opportunities for the identification of novel therapeutic drugs.

### Future perspectives

NAC, NAS and NAG are the first example of molecular chaperones able to stabilize GAA without interfering with its catalytic activity. The finding that molecules structurally different from the natural substrate can act as molecular chaperones delineates new scenarios in the research of PCs.

Nonetheless, there are still some aspects to be elucidated:

- determination of the site bound by NAC, NAS and NAG
- characterization of the molecular mechanism through which the new PCs exert their chaperoning action
- identification of new PCs effective on a larger number of GAA mutants and with a better affinity for enzyme. These last issues could be addressed by using NAC, NAS and NAG as lead molecules and improving their chaperoning action, for instance through their structural variations.

## **3.9 REFERENCES**

- Amalfitano A, McVie-Wylie AJ, Hu H, Dawson TL, Raben N, Plotz P, Chen YT. 1999. Systemic correction of the muscle disorder glycogen storage disease type II after hepatic targeting of a modified adenovirus vector encoding human acid-alpha-glucosidase. *Proceedings of the National Academy of Sciences of the United States of America* **96**:8861–8866.
- Anelli T, Sitia R. 2008. Protein quality control in the early secretory pathway. *The EMBO journal* **27**:315–27.
- Appelmans F, Wattiaux R, De Duve C. 1955. Tissue fractionation studies. 5. The association of acid phosphatase with a special class of cytoplasmic granules in rat liver. *The Biochemical journal* **59**:438–445.
- Ballabio A, Gieselmann V. 2009. Lysosomal disorders: from storage to cellular damage. *Biochimica et biophysica acta* **1793**:684–696.
- Bijvoet A, Kroos M, Pieper F, Van der Vliet M, De Boer H, Van der Ploeg A, Verbeet M, Reuser A. 1998. Recombinant human acid alpha-glucosidase: high level production in mouse milk, biochemical characteristics, correction of enzyme deficiency in GSDII KO mice. *Human molecular genetics* **7**:1815–1824.
- Bijvoet A, Van Hirtum H, Kroos M, Van de Kamp E, Schoneveld O, Visser P, Brakenhoff J, Weggeman M, Van Corven E, Van der Ploeg A, Reuser A. 1999. Human acid alpha-glucosidase from rabbit milk has therapeutic effect in mice with glycogen storage disease type II. *Human Molecular Genetics* **8**:2145–2153.
- Boot RG, Verhoek M, De Fost M, Hollak CEM, Maas M, Bleijlevens B, Van Breemen MJ, Van Meurs M, Boven LA, Laman JD, Moran MT, Cox TM, Aerts JMFG. 2004. Marked elevation of the chemokine CCL18/PARC in Gaucher disease: a novel surrogate marker for assessing therapeutic intervention. *Blood* **103**:33–39.
- Cardone M, Porto C, Tarallo A, Vicinanza M, Rossi B, Polishchuk E, Donaudy F, Andria G, De Matteis MA, Parenti G. 2008. Abnormal mannose-6-phosphate receptor trafficking impairs recombinant alpha-glucosidase uptake in Pompe disease fibroblasts. *PathoGenetics* **1**:6.
- Carrell RW, Lomas DA. 1997. Conformational disease. *Lancet* **350**:134–138.
- Chambers JP, Williams JC. 1983. Acid alpha-glucosidase from human heart. *Enzyme* **29**:109–119.
- Chen Y, Amalfitano A. 2000. Towards a molecular therapy for glycogen storage disease type II (Pompe disease). *Molecular medicine today* **6**:245–251.
- Chu CT. 2006. Autophagic stress in neuronal injury and disease. *Journal of neuropathology and experimental neurology* **65**:423–432.

### 3.9 References

---

- Conzelmann E, Sandhoff K. 1983. Partial enzyme deficiencies: residual activities and the development of neurological disorders. *Developmental neuroscience* **6**:58–71.
- Douillard-Guilloux G, Raben N, Takikita S, Batista L, Caillaud C, Richard E. 2008. Modulation of glycogen synthesis by RNA interference: towards a new therapeutic approach for glycogenosis type II. *Human molecular genetics* **17**:3876–3886.
- Douillard-Guilloux G, Raben N, Takikita S, Ferry A, Vignaud A, Guillet-Deniau I, Favier M, Thurberg BL, Roach PJ, Caillaud C, Richard E. 2010. Restoration of muscle functionality by genetic suppression of glycogen synthesis in a murine model of Pompe disease. *Human molecular genetics* **19**:684–696.
- Ellgaard L, Helenius A. 2001. ER quality control: towards an understanding at the molecular level. *Current opinion in cell biology* **13**:431–437.
- Fan JQ. 2008. A counterintuitive approach to treat enzyme deficiencies: use of enzyme inhibitors for restoring mutant enzyme activity. *Biological chemistry* **389**:1–11.
- Fan JQ, Ishii S. 2007. Active-site-specific chaperone therapy for Fabry disease. Yin and Yang of enzyme inhibitors. *The FEBS journal* **274**:4962–4971.
- Ficicioglu C. 2008. Review of miglustat for clinical management in Gaucher disease type 1. *Therapeutics and clinical risk management* **4**:425–431.
- De Filippi P, Ravaglia S, Bembi B, Costa A, Moglia A, Piccolo G, Repetto A, Dardis A, Greco G, Ciana G, Canevari F, Danesino C. 2010. The angiotensin-converting enzyme insertion/deletion polymorphism modifies the clinical outcome in patients with Pompe disease. *Genetics in medicine : official journal of the American College of Medical Genetics* **12**:206–211.
- Flanagan JJ, Rossi B, Tang K, Wu X, Mascioli K, Donaudy F, Tuzzi MR, Fontana F, Cubellis MV, Porto C, Benjamin E, Lockhart DJ, Valenzano KJ, Andria G, Parenti G, Do H V. 2009. The pharmacological chaperone 1-deoxyojirimycin increases the activity and lysosomal trafficking of multiple mutant forms of acid alpha-glucosidase. *Human mutation* **30**:1683–1692.
- Fukuda T, Ahearn M, Roberts A, Mattaliano RJ, Zaal K, Ralston E, Plotz PH, Raben N. 2006. Autophagy and mistargeting of therapeutic enzyme in skeletal muscle in Pompe disease. *Molecular therapy : the journal of the American Society of Gene Therapy* **14**:831–839.
- Garman SC, Garboczi DN. 2004. The molecular defect leading to Fabry disease: structure of human alpha-galactosidase. *Journal of molecular biology* **337**:319–335.

- Hagemans MLC, Winkel LPF, Van Doorn PA, Hop WJC, Loonen MCB, Reuser AJJ, Van der Ploeg AT. 2005. Clinical manifestation and natural course of late-onset Pompe's disease in 54 Dutch patients. *Brain: a journal of neurology* **128**:671–677.
- Hasilik A, Neufeldg EF. 1980. Biosynthesis of Lysosomal Enzymes in Fibroblasts. *The Journal of biological chemistry* **25**:4937–4945.
- Hermann S, Schestag F, Polten A, Kafert S, Penzien J, Zlotogora J, Baumann N, Gieselmann V. 2000. Characterization of four arylsulfatase A missense mutations G86D, Y201C, D255H, and E312D causing metachromatic leukodystrophy. *American journal of medical genetics* **91**:68–73.
- Hermans MMP, Krooss MA, Beeurnens J Van, Oostras BA, ReusersII AJJ. 1991. Human Lysosomal alpha-Glucosidase. *Biochemistry* **266**:13507–13512.
- Hermans MMP, Wisselaar H a, Kroos M a, Oostra B a, Reuser a J. 1993. Human lysosomal alpha-glucosidase: functional characterization of the glycosylation sites. *The Biochemical journal* **289**:681–6.
- Hermans MMP, Van Leenen D, Kroos M a, Beesley CE, Van Der Ploeg AT, Sakuraba H, Wevers R, Kleijer W, Michelakakis H, Kirk EP, Fletcher J, Bosshard N, Basel-Vanagaite L, Besley G, Reuser AJJ. 2004. Twenty-two novel mutations in the lysosomal alpha-glucosidase gene (GAA) underscore the genotype-phenotype correlation in glycogen storage disease type II. *Human mutation* **23**:47–56.
- Hers HG. 1963. alpha-Glucosidase deficiency in generalized glycogenstorage disease (Pompe's disease). *The Biochemical journal* **86**:11–16.
- Hoefsloot LH, Hoogeveen-Westerveld M, Kroos MA, Van Beeumen J, Reuser AJ, Oostra BA. 1988. Primary structure and processing of lysosomal alpha-glucosidase; homology with the intestinal sucrase-isomaltase complex. *The EMBO journal* **7**:1697–1704.
- Hoefsloot LH, Hoogeveen-Westerveld, Marianne Reuser AJJ, Ostra BA. 1990. Characterization of the human lysosomal alpha-glucosidase gene **272**:493–497.
- Hollak CE, Evers L, Aerts JM, Van Oers MH. 1997. Elevated levels of M-CSF, sCD14 and IL8 in type 1 Gaucher disease. *Blood cells, molecules & diseases* **23**:201–212.
- Hug G, Schubert WK. 1967. Lysosomes in type II glycogenosis. Changes during administration of extract from *Aspergillus niger*. *The Journal of cell biology* **35**:C1–6.

### 3.9 References

---

- Huie ML, Hirschhorn R, Chen AS, Martiniuk F, Zhong N. 1994. Mutation at the catalytic site (M519V) in glycogen storage disease type II (Pompe disease). *Human mutation* **4**:291–293.
- Hunziker W, Spiess M, Semenza G, Lodish HF. 1986. The sucrase-isomaltase complex: Primary structure, membrane-orientation, and evolution of a stalked, intrinsic brush border protein. *Cell* **46**:227–234.
- Jeffrey PL, Brown DH, Brown BI. 1970a. Studies of lysosomal alpha-glucosidase. II. Kinetics of action of the rat liver enzyme. *Biochemistry* **9**:1416–1422.
- Jeffrey PL, Brown DH, Brown BI. 1970b. Studies of lysosomal alpha-glucosidase. I. Purification and properties of the rat liver enzyme. *Biochemistry* **9**:1403–1415.
- Jolly RD, Brown S, Das AM, Walkley SU. 2002. Mitochondrial dysfunction in the neuronal ceroid-lipofuscinoses (Batten disease). *Neurochemistry international* **40**:565–571.
- Kakavanos R, Hopwood JJ, Lang D, Meikle PJ, Brooks DA. 2006. Stabilising normal and mis-sense variant alpha-glucosidase. *FEBS letters* **580**:4365–4370.
- Klinge L, Straub V, Neudorf U, Schaper J, Bosbach T, Görlinger K, Wallot M, Richards S, Voit T. 2005. Safety and efficacy of recombinant acid alpha-glucosidase (rhGAA) in patients with classical infantile Pompe disease: results of a phase II clinical trial. *Neuromuscular disorders : NMD* **15**:24–31.
- Koshland DE. 1953. Stereochemistry and the mechanism of enzymatic reactions. *Biological Reviews* **28**:416–436.
- Kroos MA, Pomponio RJ, Hagemans ML, Keulemans JLM, Phipps M, DeRiso M, Palmer RE, Ausems MGEM, Van der Beek NAME, Van Diggelen OP, Halley DJJ, Van der Ploeg AT, Reuser AJJ. 2007. Broad spectrum of Pompe disease in patients with the same c.-32-13T->G haplotype. *Neurology* **68**:110–115.
- Kurz T, Terman A, Gustafsson B, Brunk UT. 2008. Lysosomes and oxidative stress in aging and apoptosis. *Biochimica et biophysica acta* **1780**:1291–1303.
- Lücke T, Höppner W, Schmidt E, Illsinger S, Das AM. 2004. Fabry disease: reduced activities of respiratory chain enzymes with decreased levels of energy-rich phosphates in fibroblasts. *Molecular genetics and metabolism* **82**:93–97.
- Martiniuk F, Mehler M, Pellicer A, Tzall S, La Badie G, Hobart C, Ellenbogen A, Hirschhorn R. 1986. Isolation of a cDNA for human acid alpha-glucosidase and detection of genetic heterogeneity for mRNA in three alpha-

- glucosidase-deficient patients. *Proceedings of the National Academy of Sciences of the United States of America* **83**:9641–9644.
- Martiniuk F, Bodkin M, Tzall S, Hirschhorn R. 1991. Isolation and partial characterization of the structural gene for human acid alpha glucosidase. *DNA and cell biology* **10**:283–292.
- Marugan JJ, Zheng W, Motabar O, Southall N, Goldin E, Sidransky E, Aungst RA, Liu K, Sadhukhan SK, Austin CP. 2010. Evaluation of 2-thioxo-2,3,5,6,7,8-hexahydropyrimido[4,5-d]pyrimidin-4(1H)-one analogues as GAA activators. *European journal of medicinal chemistry* **45**:1880–97.
- Matsui H, Sasaki M, Takemasa E, Kaneta T, Chiba S. 1984. Kinetic studies on the substrate specificity and active site of rabbit muscle acid alpha-glucosidase. *Journal of biochemistry* **96**:993–1004.
- McVie-Wylie AJ, Lee KL, Qiu H, Jin X, Do H, Gotschall R, Thurberg BL, Rogers C, Raben N, O'Callaghan M, Canfield W, Andrews L, McPherson JM, Mattaliano RJ. 2008. Biochemical and pharmacological characterization of different recombinant acid alpha-glucosidase preparations evaluated for the treatment of Pompe disease. *Molecular genetics and metabolism* **94**:448–55.
- Meikle PJ, Hopwood JJ, Clague AE, Carey WF. 1999. Prevalence of lysosomal storage disorders. *JAMA : the journal of the American Medical Association* **281**:249–54.
- Montalvo ALE, Bembi B, Donnarumma M, Filocamo M, Parenti G, Rossi M, Merlini L, Buratti E, De Filippi P, Dardis A, Stroppiano M, Ciana G, Pittis MG. 2006. Mutation profile of the GAA gene in 40 Italian patients with late onset glycogen storage disease type II. *Human mutation* **27**:999–1006.
- Moreland RJ, Jin X, Zhang XK, Decker RW, Albee KL, Lee KL, Cauthron RD, Brewer K, Edmunds T, Canfield WM. 2005. Lysosomal acid alpha-glucosidase consists of four different peptides processed from a single chain precursor. *The Journal of biological chemistry* **280**:6780–91.
- Murray AK, Brown BI, Brown DH. 1978. The molecular heterogeneity of purified human liver lysosomal alpha-glucosidase (acid alpha-glucosidase). *Archives of biochemistry and biophysics* **185**:511–24.
- Mutsaers JH, Van Halbeek H, Vliegenthart JF, Tager JM, Reuser AJ, Kroos M, Galjaard H. 1987. Determination of the structure of the carbohydrate chains of acid alpha-glucosidase from human placenta. *Biochimica et biophysica acta* **911**:244–51.
- Niesen FH, Berglund H, Vedadi M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols* **2**:2212–21.

### 3.9 References

---

- Nilssen O, Berg T, Riise HMF, Ramachandran U, Evjen G, Hansen GM, Malm D, Tranebjaerg L, Tollersrud OK. 1997. -Mannosidosis:Functional Cloning of the Lysosomal -Mannosidase cDNA and Identification of a Mutation in Two Affected Siblings. *Human Molecular Genetics* **6**:717–726.
- Okumiya T, Kroos MA, Vliet L Van, Takeuchi H, Van der Ploeg AT, Reuser AJJ. 2007. Chemical chaperones improve transport and enhance stability of mutant alpha-glucosidases in glycogen storage disease type II. *Molecular genetics and metabolism* **90**:49–57.
- Oude Elferink RP, Brouwer-Kelder EM, Surya I, Strijland A, Kroos M, Reuser AJ, Tager JM. 1984. Isolation and characterization of a precursor form of lysosomal alpha-glucosidase from human urine. *European journal of biochemistry / FEBS* **139**:489–95.
- Parenti G, Zuppaldi A, Gabriela Pittis M, Rosaria Tuzzi M, Annunziata I, Meroni G, Porto C, Donaudy F, Rossi B, Rossi M, Filocamo M, Donati A, Bembi B, Ballabio A, Andria G. 2007. Pharmacological enhancement of mutated alpha-glucosidase activity in fibroblasts from patients with Pompe disease. *Molecular therapy: the journal of the American Society of Gene Therapy* **15**:508–14.
- Parenti G. 2009. Treating lysosomal storage diseases with pharmacological chaperones : from concept to clinics. *Molecular Medicine* **1**:268–279.
- Parenti G, Andria G. 2011. Pompe Disease: From New Views on Pathophysiology to Innovative Therapeutic Strategies. *Current Pharmaceutical Biotechnology* **12**:902–915.
- Pittis MG, Donnarumma M, Montalvo ALE, Dominissini S, Kroos M, Rosano C, Stroppiano M, Bianco MG, Donati MA, Parenti G, D'Amico A, Ciana G, Di Rocco M, Reuser A, Bembi B, Filocamo M. 2008. Molecular and functional characterization of eight novel GAA mutations in Italian infants with Pompe disease. *Human mutation* **29**:E27–36.
- Platt FM, Jeyakumar M. 2008. Substrate reduction therapy. *Acta paediatrica (Oslo, Norway : 1992). Supplement* **97**:88–93.
- Poeppel P, Habetha M, Marcão A, Büssow H, Berna L, Gieselmann V. 2005. Missense mutations as a cause of metachromatic leukodystrophy. Degradation of arylsulfatase A in the endoplasmic reticulum. *The FEBS journal* **272**:1179–88.
- Polten A, Fluharty AL, Fluharty CB, Kappler J, Von Figura K, Gieselmann V. 1991. Molecular basis of different forms of metachromatic leukodystrophy. *The New England journal of medicine* **324**:18–22.
- Pompe JC. 1932. Over idiopathische hypertrophie van het hart. *Ned. Tijdschr. Geneesk.* **76**:304–312.

- Porto C, Cardone M, Fontana F, Rossi B, Tuzzi MR, Tarallo A, Barone MV, Andria G, Parenti G. 2009. The pharmacological chaperone *N*-butyldeoxynojirimycin enhances enzyme replacement therapy in Pompe disease fibroblasts. *Molecular therapy : the journal of the American Society of Gene Therapy* **17**:964–71.
- Porto C, Ferrara MC, Meli M, Acampora E, Avolio V, Rosa M, Cobucci-Ponzano B, Colombo G, Moracci M, Andria G, Parenti G. 2012. Pharmacological Enhancement of  $\alpha$ -Glucosidase by the Allosteric Chaperone *N*-acetylcysteine. *Molecular therapy* **20**:2201–11.
- Raben N, Nagaraju K, Lee E, Kessler P, Byrne B, Lee L, LaMarca M, King C, Ward J, Sauer B, Plotz P. 1998. Targeted disruption of the acid alpha-glucosidase gene in mice causes an illness with critical features of both infantile and adult human glycogen storage disease type II. *The Journal of biological chemistry* **273**:19086–92.
- Raben N, Danon M, Gilbert AL, Dwivedi S, Collins B, Thurberg BL, Mattaliano RJ, Nagaraju K, Plotz PH. 2003. Enzyme replacement therapy in the mouse model of Pompe disease. *Molecular genetics and metabolism* **80**:159–69.
- Raben N, Schreiner C, Baum R, Takikita S, Xu S, Xie T, Myerowitz R, Komatsu M, Van der Meulen JH, Nagaraju K, Ralston E, Plotz PH. 2010. Suppression of autophagy permits successful enzyme replacement therapy in a lysosomal storage disorder-murine Pompe disease. *Autophagy* **6**:1078–89.
- Reuser AJ, Kroos MA, Ponne NJ, Wolterman RA, Loonen MC, Busch HF, Visser WJ, Bolhuis PA. 1984. Uptake and stability of human and bovine acid alpha-glucosidase in cultured fibroblasts and skeletal muscle cells from glycogenosis type II patients. *Experimental cell research* **155**:178–89.
- Saarela J, Laine M, Oinonen C, Von Schantz C, Jalanko A, Rouvinen J, Peltonen L. 2001. Molecular pathogenesis of a disease: structural consequences of aspartylglucosaminuria mutations. *Human molecular genetics* **10**:983–95.
- Sampson HM, Robert R, Liao J, Matthes E, Carlile GW, Hanrahan JW, Thomas DY. 2011. Identification of a NBD1-binding pharmacological chaperone that corrects the trafficking defect of F508del-CFTR. *Chemistry & biology* **18**:231–42.
- Scheibe R, Wenzel KW, Hofmann E. 1985. Purification and some properties of lysosomal alpha-glucosidase of rat liver. *Biomedica biochimica acta* **44**:1279–86.
- Schestag F, Yaghootfam A, Habetha M, Poeppel P, Dietz F, Klein RA, Zlotogora J, Gieselmann V. 2002. The functional consequences of mis-

### 3.9 References

---

- sense mutations affecting an intra-molecular salt bridge in arylsulphatase A. *The Biochemical journal* **367**:499–504.
- Schueler UH, Kolter T, Kaneski CR, Zirzow GC, Sandhoff K, Brady RO. 2004. Correlation between enzyme activity and substrate storage in a cell culture model system for Gaucher disease. *Journal of inherited metabolic disease* **27**:649–58.
- Sim L, Willemsma C, Mohan S, Naim HY, Pinto BM, Rose DR. 2010. Structural basis for substrate selectivity in human maltase-glucoamylase and sucrase-isomaltase N-terminal domains. *The Journal of biological chemistry* **285**:17763–70.
- Sun B, Zhang H, Benjamin DK, Brown T, Bird A, Young SP, McVie-Wylie A, Chen Y-T, Koeberl DD. 2006. Enhanced efficacy of an AAV vector encoding chimeric, highly secreted acid alpha-glucosidase in glycogen storage disease type II. *Molecular therapy: the journal of the American Society of Gene Therapy* **14**:822–30.
- Sun B, Young SP, Li P, Di C, Brown T, Salva MZ, Li S, Bird A, Yan Z, Auten R, Hauschka SD, Koeberl DD. 2008. Correction of multiple striated muscles in murine Pompe disease through adeno-associated virus-mediated gene therapy. *Molecular therapy: the journal of the American Society of Gene Therapy* **16**:1366–71.
- Tashiro K, Iwamasa T, Kato H, Ogata S, Anai M. 1986. Purification and Characterization of Two Components of Acid {alpha}-Glucosidase from Pig Liver. *J. Biochem.* **99**:693–701.
- Tropak MB, Blanchard JE, Withers SG, Brown ED, Mahuran D. 2007. High-throughput screening for human lysosomal beta-N-acetyl hexosaminidase inhibitors acting as pharmacological chaperones. *Chemistry & Biology* **14**:153–164.
- Ushioda R, Hoseki J, Araki K, Jansen G, Thomas DY, Nagata K. 2008. ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER. *Science (New York, N.Y.)* **321**:569–72.
- Valenzano KJ, Khanna R, Powe AC, Boyd R, Lee G, Flanagan JJ, Benjamin ER. 2011. Identification and characterization of pharmacological chaperones to correct enzyme deficiencies in lysosomal storage disorders. *Assay and drug development technologies* **9**:213–235.
- Van den Hout H, Reuser AJ, Vulto AG, Loonen MC, Cromme-Dijkhuis A, Van der Ploeg AT. 2000. Recombinant human alpha-glucosidase from rabbit milk in Pompe patients. *Lancet* **356**:397–8.
- Van den Hout HM, Hop W, Van Diggelen OP, Smeitink J, Smit GP, Poll-The BT, Bakker HD, Loonen MC, De Klerk JB, Reuser AJJ, Van der Ploeg AT.

2003. The natural course of infantile Pompe's disease: 20 original cases compared with 133 cases from the literature. *Pediatrics* **112**:332–340.
- Van den Hout JMP, Kamphoven JHJ, Winkel LPF, Arts WFM, De Klerk JBC, Loonen MCB, Vulto AG, Cromme-Dijkhuis A, Weisglas-Kuperus N, Hop W, Van Hirtum H, Van Diggelen OP, Boer M, Kroos MA, Van Doorn PA, Van der Voort E, Sibbles B, Van Corven EJJM, Brakenhoff JPJ, Van Hove J, Smeitink JAM, De Jong G, Reuser AJJ, Van der Ploeg AT. 2004. Long-term intravenous treatment of Pompe disease with recombinant human alpha-glucosidase from milk. *Pediatrics* **113**:e448–457.
- Van der Ploeg AT, Loonen MC, Bolhuis PA, Busch HM, Reuser AJ, Galjaard H. 1988. Receptor-mediated uptake of acid alpha-glucosidase corrects lysosomal glycogen storage in cultured skeletal muscle. *Pediatric research* **24**:90–4.
- Van Hove JL, Yang HW, Wu JY, Brady RO, Chen YT. 1996. High-level production of recombinant human lysosomal acid alpha-glucosidase in Chinese hamster ovary cells which targets to heart muscle and corrects glycogen accumulation in fibroblasts from patients with Pompe disease. *Proceedings of the National Academy of Sciences of the United States of America* **93**:65–70.
- Van Hove JL, Yang HW, Oliver LM, Pennybacker MF, Chen YT. 1997. Purification of recombinant human precursor acid alpha-glucosidase. *Biochemistry and molecular biology international* **43**:613–623.
- Van Til NP, Stok M, Aerts Kaya FSF, De Waard MC, Farahbakhshian E, Visser TP, Kroos MA, Jacobs EH, Willart MA, Van der Wegen P, Scholte BJ, Lambrecht BN, Duncker DJ, Van der Ploeg AT, Reuser AJJ, Verstegen MM, Wagemaker G. 2010. Lentiviral gene therapy of murine hematopoietic stem cells ameliorates the Pompe disease phenotype. *Blood* **115**:5329–37.
- Vázquez MC, Del Pozo T, Robledo FA, Carrasco G, Pavez L, Olivares F, González M, Zanlungo S. 2011. Alteration of gene expression profile in niemann-pick type C mice correlates with tissue damage and oxidative stress. *PloS one* **6**:e28777.
- Viscomi C, Burlina AB, Dweikat I, Savoiaro M, Lamperti C, Hildebrandt T, Tiranti V, Zeviani M. 2010. Combined treatment with oral metronidazole and N-acetylcysteine is effective in ethylmalonic encephalopathy. *Nature medicine* **16**:869–71.
- Von Bülow R, Schmidt B, Dierks T, Schwabauer N, Schilling K, Weber E, Usón I, Von Figura K. 2002. Defective oligomerization of arylsulfatase a as a cause of its instability in lysosomes and metachromatic leukodystrophy. *The Journal of biological chemistry* **277**:9455–9461.

### 3.9 References

---

- Walsh TS, Hopton P, Philips BJ, Mackenzie SJ, Lee A. 1998. The effect of N-acetylcysteine on oxygen transport and uptake in patients with fulminant hepatic failure. *Hepatology (Baltimore, Md.)* **27**:1332–40.
- Wang CW, Klionsky DJ. 2003. The molecular mechanism of autophagy. *Molecular medicine (Cambridge, Mass.)* **9**:65–76.
- Wang GN, Reinkensmeier G, Zhang SW, Zhou J, Zhang LR, Zhang LH, Butters TD, Ye XS. 2009. Rational design and synthesis of highly potent pharmacological chaperones for treatment of N370S mutant Gaucher disease. *Journal of medicinal chemistry* **52**:3146–9.
- Wei H, Kim SJ, Zhang Z, Tsai PC, Wisniewski KE, Mukherjee AB. 2008. ER and oxidative stresses are common mediators of apoptosis in both neurodegenerative and non-neurodegenerative lysosomal storage disorders and are alleviated by chemical chaperones. *Human molecular genetics* **17**:469–77.
- Wenk J, Hille A, Von Figura K. 1991. Quantitation of Mr 46000 and Mr 300000 mannose 6-phosphate receptors in human cells and tissues. *Biochemistry international* **23**:723–31.
- Wisselaar HA, Kroos MA, Hermans MM, Van Beeumen J, Reuser AJ. 1993. Structural and functional changes of lysosomal acid alpha-glucosidase during intracellular transport and maturation. *The Journal of biological chemistry* **268**:2223–31.
- Wu JJ, Quijano C, Chen E, Liu H, Cao L, Fergusson MM, Rovira II, Gutkind S, Daniels MP, Komatsu M, Finkel T. 2009. Mitochondrial dysfunction and oxidative stress mediate the physiological impairment induced by the disruption of autophagy. *Aging* **1**:425–37.
- Yoshimizu M, Tajima Y, Matsuzawa F, Aikawa SI, Iwamoto K, Kobayashi T, Edmunds T, Fujishima K, Tsuji D, Itoh K, Ikekita M, Kawashima I, Sugawara K, Ohyanagi N, Suzuki T, Togawa T, Ohno K, Sakuraba H. 2008. Binding parameters and thermodynamics of the interaction of imino sugars with a recombinant human acid alpha-glucosidase (alglucosidase alfa): insight into the complex formation mechanism. *Clinica chimica acta; international journal of clinical chemistry* **391**:68–73.
- Zhang S, Bagshaw R, Hilson W, Oho Y, Hinek A, Clarke JT, Callahan JW. 2000. Characterization of beta-galactosidase mutations Asp332-->Asn and Arg148-->Ser, and a polymorphism, Ser532-->Gly, in a case of GM1 gangliosidosis. *The Biochemical journal* **348 Pt 3**:621–32.
- Zhu Y, Li X, Vie-wylie AMC, Jiang C, Thurberg BL, Raben N, Mattaliano RJ, Cheng SH. 2005. Carbohydrate-remodelled acid  $\alpha$ -glucosidase with higher affinity for the cation-independent mannose 6-phosphate receptor demonstrates improved delivery to muscles of Pompe mice. *New York* **628**:619–628.

## LIST OF PUBLICATIONS

The results exposed in chapter 2 are part of the following manuscript in preparation:

- Maria Carmina Ferrara, Beatrice Cobucci-Ponzano, Andrea Carpentieri, Bernard Henrissat, Mosè Rossi, Angela Amoresano, and Marco Moracci. **The identification and molecular characterization of the first archaeal exo-N-acetyl- $\beta$ -glucosaminidase demonstrates that family GH116 can be structured in three subfamilies**

Chapter 3 reports on results included in the following published article:

- Caterina Porto, Maria Carmina Ferrara, Massimiliano Meli, Emma Acampora, Valerio Avolio, Margherita Rosa, Beatrice Cobucci-Ponzano, Giorgio Colombo, Marco Moracci, Generoso Andria, and Giancarlo Parenti. (2012) **Pharmacological Enhancement of  $\alpha$ -Glucosidase by the Allosteric Chaperone N-acetylcysteine.** *Molecular therapy* **20**:2201–11.

Moreover the results exposed in chapter 3 are protected by provisional Patent US 61/656,246 “Allosteric chaperones for use in the treatment of Pompe disease”, inventors Andria Generoso, Parenti Giancarlo, Marco Moracci, Beatrice Cobucci-Ponzano, Maria Carmina Ferrara, Caterina Porto.

## CONFERENCES

### 2012

- *N-acetylcysteine as potential therapeutic agent in the treatment of Pompe disease*, Porto C., **Ferrara M.C.**, Meli M., Acampora E., Avolio V., Rosa M., Cobucci-Ponzano B., Colombo G., Moracci M., Andria G., Parenti G., PhD student conference, MRC laboratory for Molecular Cell Biology and Institute of Protein Biochemistry, 4-6 November, Vico Equense, Naples, Italy.
- *Metagenomic analysis of extreme environments in astrobiology: understanding the evolution of primitive life*, Strazzulli A., Giglio R., **Ferrara M.C.**, Zhou Y., Xu J., Cobucci-Ponzano B., Contursi P. and Moracci M., VI Congresso Nazionale dell'Italian Society for space biomedicine and biotechnology, 25-27 October, Brindisi, Italy.
- *New extremophilic biocatalysts for the biotransformation of plant biomasses*, Moracci M., Cobucci Ponzano B., Morana A., Ionata E., Strazzulli A., Giglio R., **Ferrara M.C.**, Del Monaco G., Maurelli L., Marcolongo L., Iacono R., Masturzo G., La Cara F., Italian Forum on industrial Biotechnology and Bioeconomy, 23-24 October, Milan, Italy.
- *Glycosylation in the crenarchaeon *Sulfolobus solfataricus**, **Ferrara M.C.**, Cobucci-Ponzano B., Carpentieri A., Amoresano A. and Moracci M., IV Workshop of the Italian Astrobiology Society, 19-21 September, Perugia, Italy.
- *Metagenomic analysis of an extreme biotope*, Giglio R., **Ferrara M.C.**, Zhou Y., Xu J., Cobucci-Ponzano B., Contursi P. and Moracci M., IV Workshop of the Italian Astrobiology Society, 19-21 September, Perugia, Italy.
- *Identification of novel allosteric chaperones as potential therapeutic agents in the treatment of Pompe disease*, **Ferrara M.C.**, Porto C., Meli M., Acampora E., Avolio V., Rosa M., Cobucci-Ponzano B., Colombo G., Andria G., Parenti G. and Moracci M., 56th National Meeting of the Italian Society of Biochemistry and Molecular Biology, 26-29 September, Chieti, Italy.

- *Archaeal glycosylation: deciphering the machinery in Sulfolobus solfataricus*, **Ferrara M.C.**, Cobucci-Ponzano B., Carpentieri A., Amoresano A. and Moracci M., 9th International Congress on Extremophiles, 10-13 September, Seville, Spain.
- *Pharmacological enhancement of alpha-glucosidase*, **Ferrara M.C.**, Porto C., Meli M., Acampora E., Avolio V., Rosa M., Cobucci-Ponzano B., Colombo G., Moracci M., Andria G., Parenti G., XI National Biotechnology Congress, 27-29 June, Varese, Italy.

## 2011

- *Analysis of the glycome and identification of the enzymes involved in the glycosylation of Sulfolobus solfataricus*, Cobucci-Ponzano B., **Ferrara M.C.**, Strazzulli A., Contursi P., Fiorentino G., Limauro D., Carpentieri A., Amoresano A., Rossi M., Bartolucci S. and Moracci M., Thermophiles, 11-16 September, Big Sky, Montana, USA.
- *Archaeal glycosylation: analysis of the glycome and identification of the enzymes involved*, **Ferrara M.C.**, Cobucci-Ponzano B., Contursi P., Fiorentino G., Limauro D., Carpentieri A., Amoresano A., Rossi M., Bartolucci S. and Moracci M., European Carbohydrate Symposium (Eurocarb 16), 3-7 July, Sorrento, Naples, Italy.
- *Glycosylation in the crenarchaeon Sulfolobus solfataricus*, **Ferrara M.C.**, Cobucci-Ponzano B., Contursi P., Fiorentino G., Limauro D., Bartolucci S. and Moracci M., Carbohydrate Bioengineering Meeting (CBM9), 15-18 May, Lisbon, Portugal.

## 2010

- *Beta-Glycosyl azide donors for the preparation of alpha-glycosynthases*, Cobucci-Ponzano B., Zorzetti C., Strazzulli A., **Ferrara M.C.**, Bedini E., Corsaro M.M., Sulzenbacher G., Rossi M., Moracci M., Gordon Research Conference on Glycobiology, 8-13 May, Lucca, Italy.
- *Carbohydrate active enzymes in glicobiology: pharmaceutical targets and new tools for the synthesis of innovative carbohydrate-based drugs*, Cobucci-Ponzano B., Strazzulli A., **Ferrara M.C.**, Zorzetti C., Bedini E.,

Corsaro M.M., Donaudy F., Parenti G., Rossi M. and Moracci M., DSV conference, 11-12 October, Rome, Italy.

## **ACKNOWLEDGMENTS**

This thesis marks the end of a nice and fruitful period spent in the laboratory of Glycobiology at the Institute of Protein Biochemistry, CNR, where I have carried out my PhD project.

First, a special thank you to my supervisor Dr. Marco Moracci for giving me the opportunity to perform my formation activities in his laboratory, for his constant presence, availability and remarkable scientific support throughout the development of my PhD project, for teaching me to be a good scientist and for strengthening my passion for research, and also for his nice discussions in scientific and non-scientific matters.

I would thank Dr. Beatrice Cobucci-Ponzano not only for her noticeable scientific assistance, but especially for her friendly presence in each situation and for teaching me to be strong and decisive in the achieving the objectives.

Thank you also to my colleague and friend Dr. Andrea Strazzulli for his "productive criticism" on every thing concerning the science and the life inside and outside of the lab.

I would thank the two special undergraduate students, Roberta Iacono and Giuseppe Masturzo for their friendly support and for sharing good times during my last year of PhD, and all my previous and present colleagues.

A very special thanks to Andrea for supporting and encouraging all my choices in scientific and private life and for his lovely patience.

Thank you to my family for their constant and tireless support in all situations I have faced in my life.

I would thank all my friends for all the happy moments spent together.

# Pharmacological Enhancement of $\alpha$ -Glucosidase by the Allosteric Chaperone *N*-acetylcysteine

Caterina Porto<sup>1,2</sup>, Maria C Ferrara<sup>3</sup>, Massimiliano Meli<sup>4</sup>, Emma Acampora<sup>2</sup>, Valeria Avolio<sup>2</sup>, Margherita Rosa<sup>2</sup>, Beatrice Cobucci-Ponzano<sup>3</sup>, Giorgio Colombo<sup>4</sup>, Marco Moracci<sup>3</sup>, Generoso Andria<sup>2</sup> and Giancarlo Parenti<sup>1,2</sup>

<sup>1</sup>Telethon Institute of Genetics and Medicine (TIGEM), Naples, Italy; <sup>2</sup>Department of Pediatrics, Federico II University, Naples, Italy; <sup>3</sup>Institute of Protein Biochemistry (IBP), Consiglio Nazionale delle Ricerche (CNR), Naples, Italy; <sup>4</sup>Istituto Chimica del Riconoscimento Molecolare, Consiglio Nazionale delle Ricerche (CNR), Milan, Italy

Pompe disease (PD) is a metabolic myopathy due to the deficiency of the lysosomal enzyme  $\alpha$ -glucosidase (GAA). The only approved treatment for this disorder, enzyme replacement with recombinant human GAA (rhGAA), has shown limited therapeutic efficacy in some PD patients. Pharmacological chaperone therapy (PCT), either alone or in combination with enzyme replacement, has been proposed as an alternative therapeutic strategy. However, the chaperones identified so far also are active site-directed molecules and potential inhibitors of target enzymes. We demonstrated that *N*-acetylcysteine (NAC) is a novel allosteric chaperone for GAA. NAC improved the stability of rhGAA as a function of pH and temperature without disrupting its catalytic activity. A computational analysis of NAC–GAA interactions confirmed that NAC does not interact with GAA catalytic domain. NAC enhanced the residual activity of mutated GAA in cultured PD fibroblasts and in COS7 cells overexpressing mutated GAA. NAC also enhanced rhGAA efficacy in PD fibroblasts. In cells incubated with NAC and rhGAA, GAA activities were 3.7–8.7-fold higher than those obtained in cells treated with rhGAA alone. In a PD mouse model the combination of NAC and rhGAA resulted in better correction of enzyme activity in liver, heart, diaphragm and gastrocnemius, compared to rhGAA alone.

Received 12 May 2012; accepted 3 July 2012; advance online publication 18 September 2012. doi:10.1038/mt.2012.152

## INTRODUCTION

Pompe disease (PD, OMIM 232300) is an inborn metabolic disorder caused by the functional deficiency of  $\alpha$ -glucosidase (GAA, acid maltase, E.C.3.2.1.20), an acid glycoside hydrolase involved in the lysosomal breakdown of glycogen. GAA deficiency results in glycogen accumulation in lysosomes and in secondary cellular damage, through mechanisms not fully understood.<sup>1,2</sup> Although GAA deficiency in PD is generalized, muscles are particularly vulnerable to glycogen storage. The disease manifestations are thus predominantly related to the involvement of cardiac and skeletal

muscles. The phenotypic spectrum of the disease is wide and varies from a devastating classical infantile-onset form, to attenuated late-onset phenotypes. The manifestations related to progressive muscle hypotonia, which cause severe motor impairment and eventually respiratory failure, impact severely on the health of PD patients.<sup>1,3</sup>

Like for several other lysosomal storage diseases, an enzyme replacement therapy (ERT) with recombinant human  $\alpha$ -glucosidase (rhGAA) has become available for PD in the early 2000s. ERT was shown to improve patients' survival and function and to stabilize the disease course.<sup>4–8</sup> However, despite treatment, some patients experience limited clinical benefit or show signs of disease progression and it is clear that reaching therapeutic concentrations of the recombinant enzyme in skeletal muscle is particularly challenging.<sup>9</sup>

Several factors concur in limiting therapeutic success of ERT, including the age at start of treatment,<sup>10,11</sup> the immunological and cross-reactive material status of patients,<sup>12</sup> the preferential uptake of rhGAA by liver and the insufficient targeting of the enzyme to muscles,<sup>13</sup> the relative deficiency of the mannose-6-phosphate receptor in muscle cells,<sup>14</sup> and the "build up" of the autophagic compartment observed in myocytes.<sup>15,16</sup> In addition, studies in other lysosomal storage diseases treatable by ERT, such as Gaucher disease (due to the deficiency of  $\beta$ -glucocerebrosidase) and Fabry disease (due to  $\alpha$ -galactosidase A deficiency), point to the role of factors intrinsically related to the recombinant enzymes used for ERT, and suggest that these enzymes may be relatively unstable when exposed to stresses, like non-acidic pH, during their transit to lysosomes.<sup>17,18</sup>

In the recent years, pharmacological chaperone therapy (PCT) has been proposed as a strategy to increase physical stability of recombinant enzymes and to enhance the therapeutic action of ERT. This approach, based on using small-molecule ligands that increase stability of mutated proteins and prevent their degradation, was first designed for the treatment of diseases due to protein misfolding.<sup>19</sup> Recent studies, however, have shown that chaperones are not only able to rescue misfolded defective proteins, but may also potentiate the effects of the wild-type recombinant enzymes used for ERT.<sup>20</sup> We and others have provided preclinical evidence supporting this concept in two relatively prevalent lysosomal disorders, PD,<sup>21</sup> and Fabry disease.<sup>18,22</sup> In both disorders, when recombinant enzymes were administered to mutant

**Correspondence:** Giancarlo Parenti, Telethon Institute of Genetics and Medicine (TIGEM), Via P. Castellino 111, 80131, Naples, Italy. E-mail: (parenti@tigem.it) or (parenti@unina.it)

fibroblasts in combination with the chaperone molecules *N*-butyl-deoxyojirimycin (NB-DNJ) and 1-deoxy-galactonojirimycin, respectively, the lysosomal trafficking, the maturation and the intracellular activity of the enzymes improved. A similar effect was also obtained in cultured macrophages treated with the recombinant  $\beta$ -glucocerebrosidase (the enzyme used for ERT in Gaucher disease), in the presence of the chaperone isofagomine.<sup>23</sup>

Although small-molecule chaperones have several advantages, compared to ERT, in terms of biodistribution, oral availability, reduced impact on patients' quality of life, a major reason of concern for the clinical use of these drugs is that the chaperones so far identified for the treatment of lysosomal storage diseases are active site-directed molecules and are reversible competitive inhibitors of the target enzymes.<sup>24</sup> Therefore, the identification of second-generation chaperones that protect the enzymes from degradation without interfering with its activity may be advantageous. Extensive search for new chaperones is currently being done by high-throughput screenings with chemical libraries.<sup>25,26</sup>

Here we report on the identification of *N*-acetylcysteine (NAC), a known pharmaceutical drug, as a novel allosteric chaperone for GAA. The strategy used for the identification and characterization of NAC's effects was based on the combination of biochemical studies, both in cell-free and in cellular systems, and a computational analysis of NAC-GAA interactions. We found that this drug stabilizes wild-type GAA at nonacidic pH, enhances the residual activity of mutated GAA and improves the efficacy of rhGAA used for ERT in this disease. This novel chaperone does not interact with the catalytic domain of GAA, and consequently is not a competitive inhibitor of the enzyme.

## RESULTS

### NAC improves rhGAA stability *in vitro*

Resistance of wild-type enzymes to physical stresses, such as modifications of temperature and pH, is commonly taken as an indicator to monitor the efficacy of pharmacological chaperones.<sup>24</sup> rhGAA 6.8  $\mu$ mol/l was stable at pH 5.0 for up to 24 hours while rapidly lost its activity at nonphysiological pH (3.0 and 7.0), representative of nonlysosomal cellular compartments (Figure 1a). Coincubation with 10 mmol/l NAC (compound 1 in Supplementary Figure S1) rescued rhGAA activity at pH 7.0 in a dose-dependent manner (Figure 1b), and persisted even after 48 hours of incubation (Figure 1c). The same concentration of NAC also thermally stabilized 0.95  $\mu$ mol/l rhGAA increasing by  $10.5 \pm 0.5^\circ\text{C}$  its melting temperature ( $T_m$ ) (Figure 1d).

The related amino acids *N*-acetylserine and *N*-acetylglycine (NAS and NAG, compounds 2 and 3 in Supplementary Figure S1) also behaved as NAC by inducing remarkable stabilization of rhGAA at pH 7.0 (Supplementary Figure S2a–b), while the nonacetylated homologs cysteine, serine, and glycine and 2-mercaptoethanol (4–7 in Supplementary Figure S1) did not prevent enzyme inactivation (Supplementary Figure S2c–f), suggesting that the stabilizing effect was due to the presence of the acetyl rather than to the sulfidryl group.

### Modeling the interaction of NAC with the rhGAA protein

To rationalize the interaction of NAC with the enzyme, a homology model of rhGAA was built based on its amino acid sequence and the 3D-structure of the highly homologous human enzyme

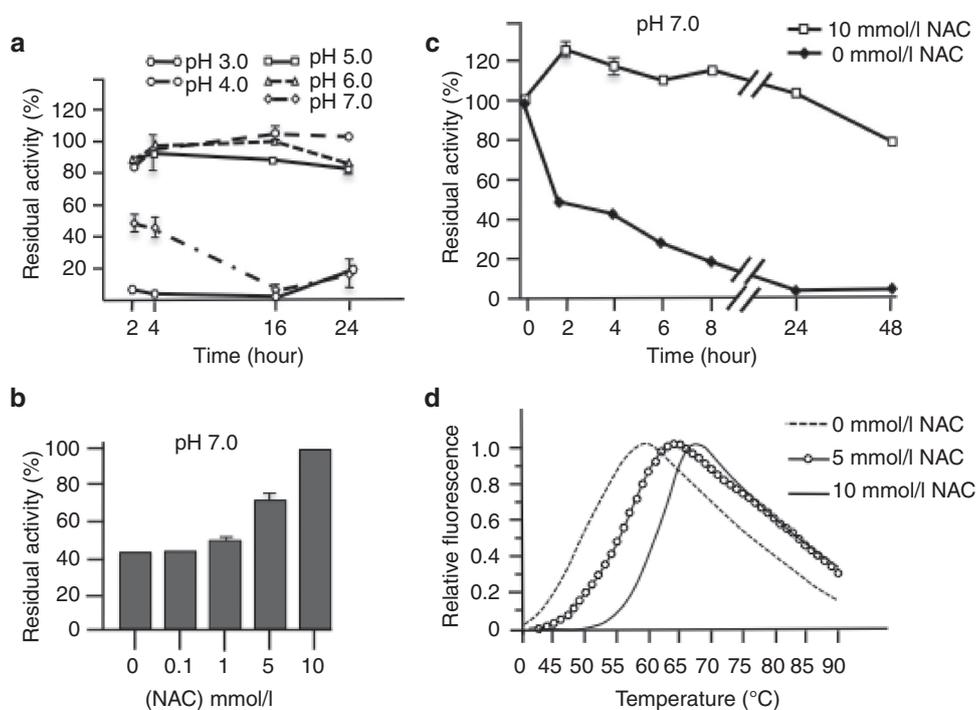
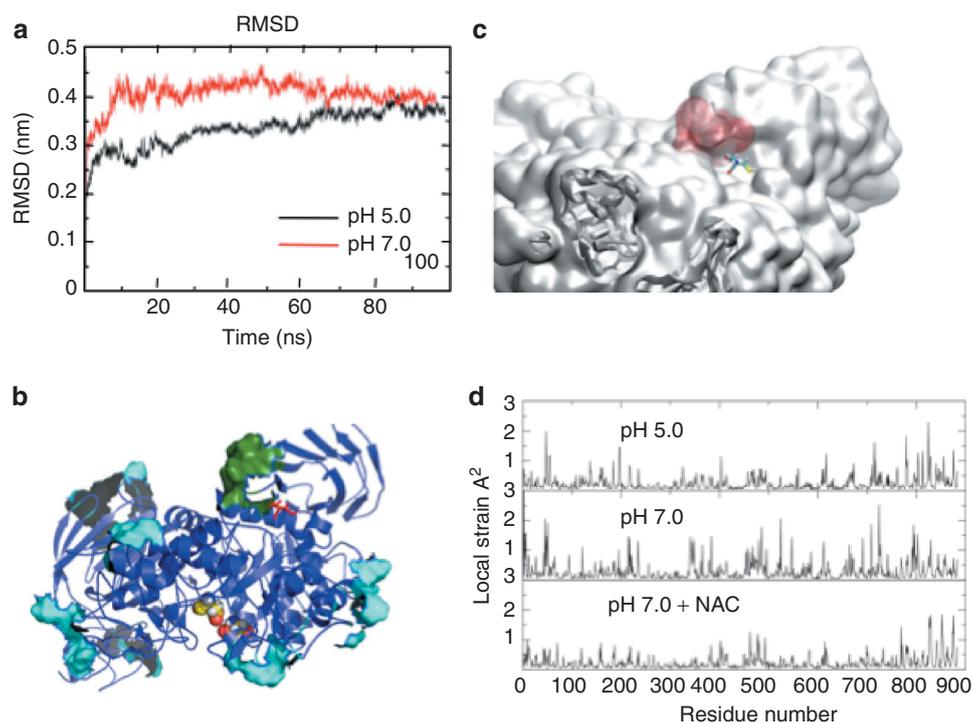


Figure 1 Characterization of recombinant human  $\alpha$ -glucosidase (rhGAA) *in vitro* and analysis of pharmacological chaperones stabilization. (a) rhGAA was stable at pH 5.0 for up to 24 hours while rapidly lost its activity at pH 3.0 and 7.0. (b) Coincubation with 10 mmol/l *N*-acetylcysteine (NAC) rescued rhGAA activity at pH 7.0 in a dose-dependent manner, and (c) persisted even after 48 hours of incubation. (d) NAC thermally stabilized rhGAA increasing by  $10.5 \pm 0.5^\circ\text{C}$  its melting temperature.



**Figure 2** Modeling the structural determinants of  $\alpha$ -glucosidase (GAA) stability and mapping the *N*-acetylcysteine (NAC)-binding sites. **(a)** Comparison of the molecular dynamics (MD) simulations of the GAA model at pH 5.0 (red) and 7.0 (black). **(b)** Identification of the NAC clustering regions: the surfaces indicate the areas of favorable NAC contact calculated from the simulations of GAA in the presence of 10 copies of NAC. The green surface indicated the surface identified as a possible allosteric-binding site as explained in the results (combined analysis of NAC-binding regions, changes in the dynamics of the residues in response to pH, SiteMap screening to define the suitability of the pocket to host a small molecule). The cyan surfaces indicate alternative NAC contact sites. It is possible to appreciate that they are all superficial and do not define cavities suitable for small-molecule binding. The active site is indicated by yellow van der Waals representations of the atoms. **(c)** Detail of the GAA model with NAC bound to the hot spot: the sulfur atom is highlighted in yellow and the binding site is in red filled spaces. **(d)** The residue-based profile of average strain calculated over the simulation time for the three GAA conditions; from top to bottom: isolated protein at pH 5.0, 7.0, and 7.0 plus NAC.

maltase-glucoamylase (pdb code 3L4Z) showing 44% identity over the aligned 868 residues of GAA. Molecular dynamics (MD) simulations in conditions mimicking pH 5.0 and 7.0 showed that the rhGAA structure was more stable at pH 5.0 than at neutral pH (**Figure 2a**) and the analysis of the protein's pH-dependent geometric strain<sup>27</sup> further confirmed that rhGAA was, in general, more rigid at pH 5.0 than at pH 7.0 (**Supplementary Figure S3a**). Then, we searched for sites (hot spots) of the protein whose local flexibility and structural organization changed upon passing from pH 5.0 to 7.0 (**Supplementary Figure S3b**). Projecting the results on the 3D structural model, the region defined by residues Arg525, Gly793, Glu794, Ser795, Leu796, Glu797, and Gly802 emerged as a possible target (**Figure 2b**). This region, which is conserved among characterized GH31  $\alpha$ -glucosidases from humans and other mammals (**Supplementary Figure S3c**), is located at the boundary between two domains (one of which contains the active site) defining a stable pocket, which is at 35 Å distance from the active site (**Supplementary Figure S3d**). Thus, a putative ligand, stabilizing the interdomain interface, might confer to rhGAA at neutral pH the same properties showed by the apoenzyme at acidic pH.

### NAC-binding mechanism and binding site

To identify the regions of the protein surface that may be most frequently in contact with NAC, we combined MD simulations with

molecular docking experiments. In a first, “coarse-grained” exploration of binding spots on the protein surface, multiple copies of NAC were simulated in the presence of GAA. The resulting density maps report on the areas of GAA where NAC clusters most favorably. NAC was observed multiple times during the simulation to bind with favorable interactions on specific regions of the protein. A control simulation was carried out with nonacetylated Gly in the same conditions and the results showed that this amino acid, which has no chaperoning effect, had a much lower tendency to bind in the vicinity of the surface. The list of residues defining the surface on which NAC makes favorable contacts is reported in **Supplementary Table S1**. Interestingly, one of the identified areas of favorable contact between NAC and GAA partially overlapped with the hotspot region located at the boundary between the catalytic domains and the flanking one identified through geometric strain analysis (**Figure 2b**). The list of identified possible binding regions was further filtered by the application of the “Site Map” function of the Maestro suite of programs, which aims to identify possible sites on a protein that fulfill structural and functional requirements optimal for binding a small, drug-like molecule thus being identified as druggable sites. The application of this further filter on rhGAA surface returned the pH responsive hot spot region identified earlier as the only one showing the stereochemical properties needed by a putative receptor site to bind a small-molecule (**Figure 2b**).

A refined docking analysis using the Grid program from the Maestro suite was next carried (see Materials and Methods section) to investigate the pose and the energy of NAC in the putative allosteric site. Interestingly, the binding energies for NAC at this site were comparable to the binding energies calculated for the docking of maltose or DNJ to the active site of GAA and similar results were obtained also with NAS and NAG (Supplementary Table S2). Moreover, in the model of the best-ranked pose, the -SH group is directed toward the surface of the protein whereas the acetyl group interacts with the allosteric GAA pocket described above (Figure 2c), providing further indication that the SH group was not crucial for the binding of the molecule to the enzyme. A control docking run was carried out with DNJ. DNJ was left free to move on the whole surface of GAA, without imposing restraints. Importantly, the minimum free energy and most populated docking solution shows DNJ binding preferentially in the active site of GAA, with a lower energy than that measured for the pharmacological chaperone in the allosteric pocket, showing the high affinity of this ligand for the active site and supporting the validity of our model (see Supplementary Table S2).

Finally, the GAA–NAC complex was refined by MD simulations as described above for the apoenzyme at the two different

pHs. Interestingly, the presence of NAC at pH 7.0 confers to GAA a local flexibility profile similar to the one of unbound GAA at pH 5.0, thereby suggesting that NAC may actually act as a pharmacological chaperone of the enzyme (Figure 2d). This effect is not limited to the immediate vicinity of the NAC-binding site, but extends to distal regions of the protein. As a control, the same procedure was carried out using Gly as a possible ligand, but the resulting maps did not show any specific feature and Gly appears to nonspecifically sample many regions of the protein surface with very low affinity.

NAC and related compounds, binding reversibly to allosteric sites remote from the catalytic site, may act by blocking conformational fluctuations leading to a destabilized state of GAA, thus rescuing its functional state. This suggests that these molecules would be the first pharmacological chaperones that do not work as competitive inhibitors of the enzyme, expanding the molecular diversity of GAA pharmacological chaperones beyond imino sugar molecules. The modeling data were further supported by *in vitro* experiments showing that different concentrations of NAC and derivatives (up to 10 mmol/l) did not affect rhGAA activity (Figure 3a). Instead DNJ acted as a competitive inhibitor with a  $K_i$  of 3.4  $\mu$ mol/l (Figure 3b).

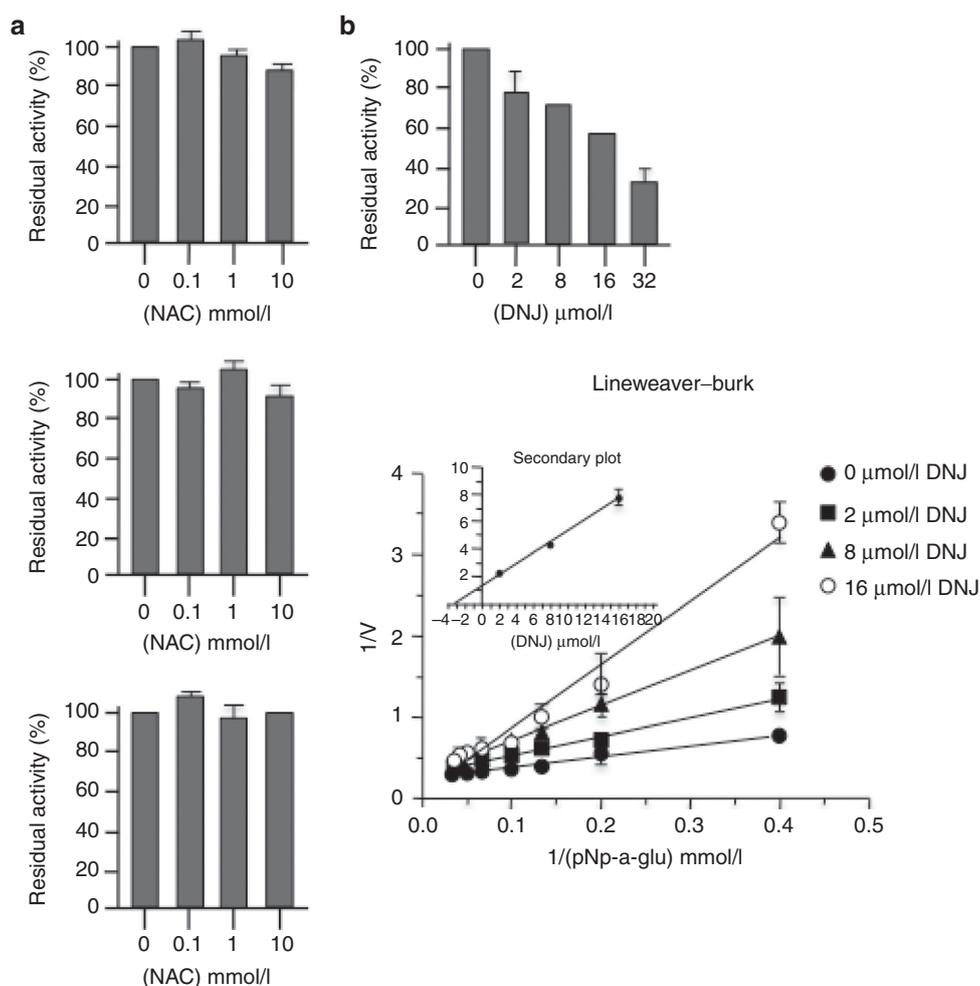


Figure 3 *N*-acetylcysteine (NAC) and related compounds are not  $\alpha$ -glucosidase (GAA) inhibitors. GAA activity was measured in the presence of different concentrations of NAC and of the related compounds NAS and NAG, ranging from 0.1 to 10 mmol/l. (a) None of the compounds tested inhibited recombinant human  $\alpha$ -glucosidase (rhGAA) activity, indicating that these compounds are not competitive inhibitors of the enzyme. (b) Instead DNJ acted as a competitive inhibitor with a  $K_i$  of 3.4  $\mu$ mol/l.

## NAC rescues mutated GAA in PD fibroblasts and transfected COS7 cells

We investigated the effect of NAC in cultured fibroblasts from five PD patients that carry different mutations and with different phenotypes (Table 1). The concentration of 10 mmol/l used for this purpose was in the same range as that used in studies on the antioxidant effect of NAC.<sup>28,29</sup>

NAC enhanced the residual activity of mutated GAA in fibroblasts from patients 3 and 4 (Figure 4a). Both patients carried on one allele the mutation p.L552P, previously reported to be responsive to NB-DNJ and DNJ.<sup>30,31</sup>

The response of individual mutations to NAC was further evaluated by expressing a panel of mutated GAA gene constructs in COS7 cells (Figure 4b). The mutated constructs were chosen to be representative of both imino sugar-responsive and non-responsive mutations, in order to compare the chaperoning profile of NAC with that of imino sugars. The mutations p.L552P, p.A445P, and p.Y455F showed significant enhancement of GAA activity in the

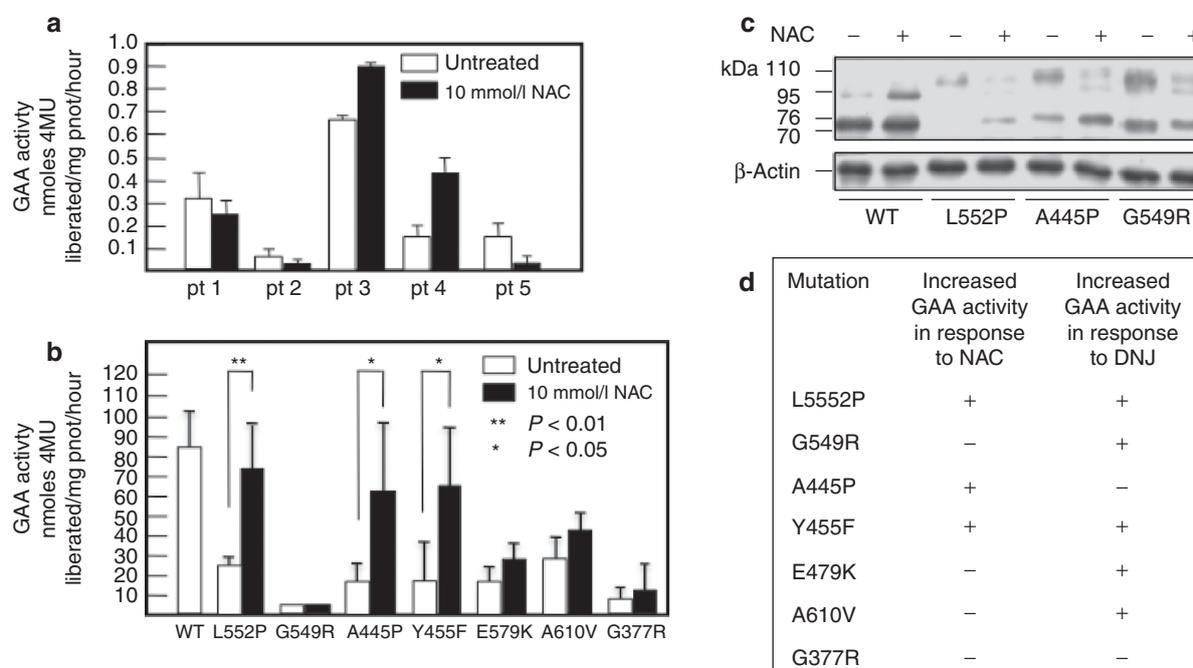
presence of 10 mmol/l NAC. The enhancement of enzyme activity for responsive mutations paralleled the increase in the amounts of the 76 and 70 kDa active isoforms of GAA on western blot analysis. Figure 4c shows a western blot analysis of COS7 cells overexpressing two of the responsive (p.L552P, p.A445P) and one non-responsive (p.G549R) mutation. For this latter mutation no change was seen in the amounts of the GAA active isoforms, already detectable in the absence of NAC, as previously reported.<sup>31</sup> These results suggest that NAC has a different chaperoning profile compared to the active site-directed chaperones DNJ and NB-DNJ (Figure 4d).

## NAC enhances rhGAA efficacy in PD fibroblasts and in the mouse model of PD

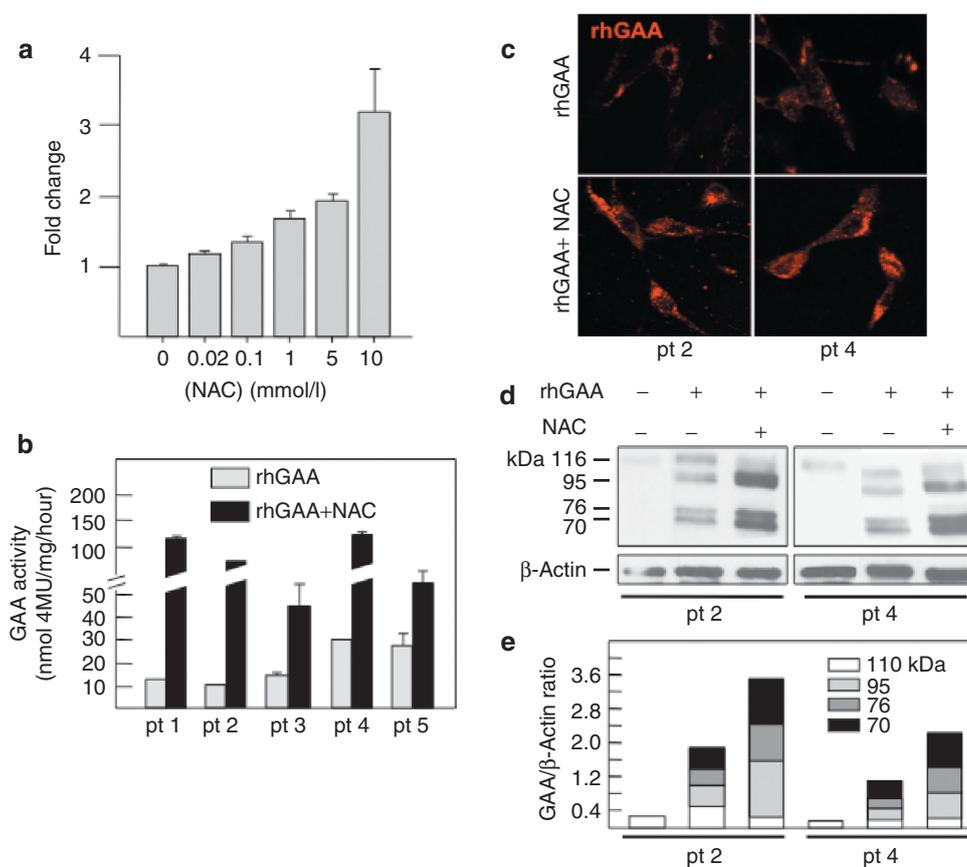
We tested whether NAC is able to synergistically enhance the efficacy of rhGAA. In fibroblasts from patient 3 coadministration of rhGAA and NAC (0.02–10 mmol/l) resulted in improved GAA activity with a dose-dependent effect (Figure 5a). Increases of 1.3-, 1.7- and 2.0-fold were already observed at NAC concentrations

**Table 1** PD fibroblast cell lines studied

pt No	Phenotype	Genotype	Average GAA residual activity <sup>a</sup>	Studies in which the same cell line was used
1	Severe	p.W367X/p.G643R	0.20	Parenti <i>et al.</i> <sup>30</sup> ; Cardone <i>et al.</i> <sup>32</sup> ; Porto <i>et al.</i> <sup>21</sup>
2	Severe	p.H612_D616del-insRGI/p.R375L	0.14	Not reported in previous studies
3	Intermediate	p.L552P/aberrant splicing	0.69	Rossi <i>et al.</i> 2007; Cardone <i>et al.</i> <sup>32</sup> ; Porto <i>et al.</i> <sup>21</sup>
4	Intermediate	c.-32-13T>G - S619N/L552P	0.15	Not reported in previous studies
5	Intermediate	G549R/aberrant splicing	0.16	Rossi <i>et al.</i> 2007; Cardone <i>et al.</i> <sup>32</sup> ; Porto <i>et al.</i> <sup>21</sup>



**Figure 4** Effect of *N*-acetylcysteine (NAC) on the residual activity of mutated  $\alpha$ -glucosidase (GAA) in fibroblasts and COS7 cells. **(a)** Incubation of Pompe disease (PD) fibroblast with 10 mmol/l NAC resulted in enhanced residual GAA activity in 2 of the 5 cell lines (from patients 3 and 4). The response of mutated GAA to NAC was also evaluated by expressing mutated GAA gene constructs in COS7 cells. **(b)** The mutations p.L552P, p.A445P, and p.Y455F, showed increases of GAA activity in the presence of NAC, and increased amounts of 76 and 70 kDa active GAA isoforms by western blot **(c)**. These results indicate that NAC (10 mmol/l) has a different chaperoning profile, in terms of enhancement of GAA activity, compared to the imino sugar chaperone *N*-butyl-deoxyojirimycin (NB-DNJ) (20–100  $\mu$ mol/l).<sup>31</sup>



**Figure 5** Synergy between *N*-acetylcysteine (NAC) and recombinant human  $\alpha$ -glucosidase (rhGAA) in Pompe disease (PD) fibroblasts. **(a)** The efficacy of rhGAA was enhanced by different concentrations (0.02–10 mmol/l) of NAC in patient 3, showing a dose-dependent effect. Five PD fibroblast cell lines were incubated with 50  $\mu$ mol/l rhGAA in the absence and presence of 10 mmol/l NAC. **(b)** In all cell lines coincubation of rhGAA with the chaperone resulted in an improved correction of GAA deficiency, with increases in GAA activity ranging from ~3.7–8.7-fold the activity of cells treated with rhGAA alone. **(c)** In the presence of NAC also the amount of fluorochrome-labeled GAA increased, compared to cells incubated with fluorescent GAA alone. **(d)** The amounts of GAA polypeptides in the cells treated with NAC and rhGAA were also increased, compared to cells treated with the recombinant enzyme alone. Density scans of GAA bands show a relative increase of mature 76 and 70 kDa GAA isoforms in the presence of NAC **(e)**.

of 0.1, 1, and 5 mmol/l, respectively, with a maximal increase of 3.3-fold at 10 mmol/l. We then incubated five PD fibroblast cell lines with 50  $\mu$ mol/l rhGAA in the absence and in the presence of 10 mmol/l NAC. The efficacy of rhGAA in correcting the enzymatic deficiency varied among the different cell lines, as already reported in previous papers,<sup>22,32</sup> possibly due to individual factors implicated in the uptake and intracellular trafficking of the recombinant enzyme in each cell line. However, in all cell lines tested, coincubation of rhGAA with NAC resulted in an improved correction of GAA deficiency, with increases in GAA activity ranging from ~3.7–8.7-fold compared to the activity of cells treated with rhGAA alone (**Figure 5b**).

The enhancing effect largely exceeded that due to the rescue of the native mutated enzyme (patients 3 and 4) and was observed also in non-responsive cell lines (patients 1, 2, and 5). We also observed an increase in the amounts of fluorochrome-labeled GAA in the presence of the chaperone NAC, compared to cells incubated with fluorescent GAA alone (**Figure 5c**). Using this approach only the fluorescent exogenous enzyme is detectable and variations in the intensity of fluorescence only reflect the effects on the recombinant enzyme. The combination of these results

supports the concept that the enhancing effect of chaperones is directed towards the wild-type recombinant enzyme.

A western blot analysis (**Figure 5d**) and the quantitative analysis of each band (**Figure 5e**) showed increased amounts of GAA-related polypeptides in the cells treated with NAC and rhGAA, compared to cells treated with the recombinant enzyme alone. The processing of rhGAA into the active isoforms was also improved in the presence of NAC with a relative increase of the intermediate (95 kDa) and mature (76–70 kDa) GAA molecular forms, compared to the 110 kDa precursor. Since the GAA precursor is converted into the active forms in the late-endosomal/lysosomal compartment,<sup>33</sup> this indicates improved lysosomal trafficking of the enzyme.

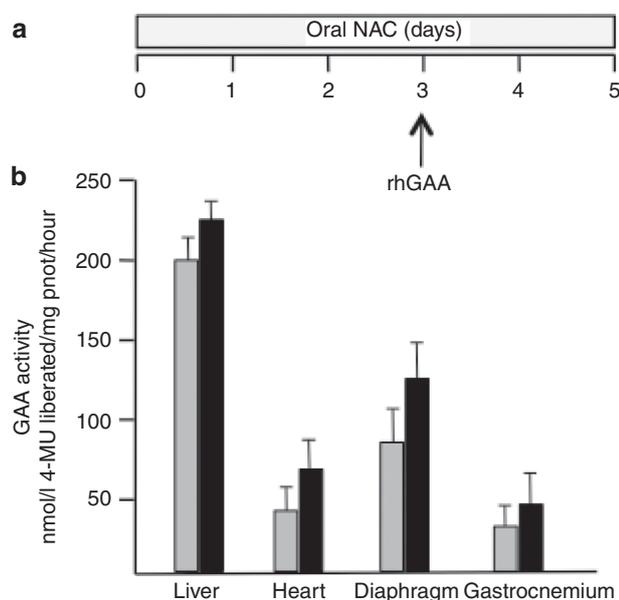
Other antioxidant drugs (resveratrol, epigallo catechin-gallate) did not enhance rhGAA in PD cultured fibroblasts (**Supplementary Figure S4**). These results, together with the analysis of NAC–GAA interaction, with the data in cell-free systems, and with the lack of effect on another lysosomal glycosidase,  $\alpha$ -galactosidase A that is defective in Fabry disease and belongs, like GAA, to the GH-D superfamily (**Supplementary Figure S5**), exclude that the effect of NAC is due to its antioxidant properties.

We also tested the combination of NAC and rhGAA in a mouse model of PD.<sup>34</sup> Mice were treated with a single injection of rhGAA at high doses (100 mg/kg) in combination with oral NAC for 5 days (Figure 6a). Mice treated with the recombinant enzyme alone were used as controls. Forty-eight hours after rhGAA injection the animals were euthanized and GAA activity was assayed in different tissues. In all tissues examined (liver, heart, diaphragm and gastrocnemium) the combination of NAC and rhGAA resulted in higher levels of GAA compared to rhGAA alone (Figure 6b).

### Combined effect of NAC and imino sugar chaperones

A corollary of the fact that NAC and imino sugar chaperones interact with different protein domains, is that their effect may be cumulated. This hypothesis was supported by the results of thermal denaturation of rhGAA performed in the presence of NAC, of the imino sugar DNJ, and of both molecules. Both NAC and DNJ increased rhGAA thermal stability, with DNJ causing the best shift in  $T_m$  ( $65.9 \pm 0.3^\circ\text{C}$ ). The combination of NAC and DNJ ( $T_m = 75.9 \pm 0.3$ ), however, resulted in the highest stability of the enzyme (Figure 7a).

A similar cumulative effect was also seen in PD fibroblasts from patients 2 and 4 incubated with rhGAA. The cells were incubated with rhGAA, with rhGAA plus either NAC or NB-DNJ, and with rhGAA plus the combination of the two chaperones. In both cell lines the combination of NAC and NB-DNJ resulted in the greater enhancement of GAA activity by rhGAA (Figure 7b). This might represent an additional advantage for the treatment of patients, in order to obtain the best stabilization of rhGAA and the highest synergy with ERT.

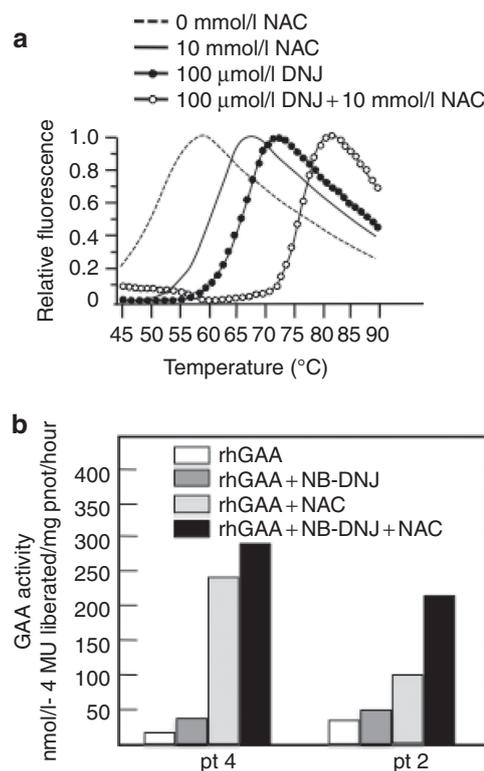


**Figure 6** Synergy between *N*-acetylcysteine (NAC) and recombinant human  $\alpha$ -glucosidase (rhGAA) *in vivo*. (a) Mice were treated with oral NAC for 5 days and received an rhGAA injection on day 3. Animal treated with rhGAA alone were used as controls. In all tissues examined (liver, heart, diaphragm, and gastrocnemium) the combination of NAC and rhGAA (black bars) resulted in higher GAA enzyme activity compared to (b) rhGAA alone (grey bars).

### DISCUSSION

PCT is an attractive strategy to improve enzyme stability and is rapidly evolving towards clinical applications. However, concerns have been raised on its clinical translation. First, known chaperones are active site-directed molecules and thus potential inhibitors of target enzymes.<sup>24</sup> In addition, it has been shown that chaperones are effective in rescuing only some disease-causing missense mutations, mainly located in specific enzyme domains, and are thus potentially effective in a limited number of patients. For PD, it is possible to speculate that ~10–15% patients may be amenable to PCT.<sup>31</sup>

These problems can be addressed by the identification of novel and allosteric noninhibitory chaperones. In this study, we have shown that NAC, and the related compounds NAS and NAG, are able to stabilize GAA without interfering with its activity and have a different chaperoning profile, compared to known chaperones. NAC is a known antioxidant that was evaluated in our laboratory, together with other related drugs (resveratrol, epigallo catechin-gallate) in PD fibroblasts for possible effects on rhGAA intracellular trafficking. The characterization of NAC's mechanism of action on rhGAA, however, indicated that molecular interactions with the enzyme, rather than the antioxidant effect, were responsible



**Figure 7** Combination of the effect of *N*-acetylcysteine (NAC) with imino sugar chaperones. Thermal stability scans of recombinant human  $\alpha$ -glucosidase (rhGAA) were performed in the absence and in the presence of NAC or DNJ. Both chaperones increased thermal stability of rhGAA, with *N*-butyl-deoxynojirimycin (NB-DNJ) resulting in the best shift in  $T_m$  of the enzyme. (a) The combination of DNJ and NAC resulted in further increase of the enzyme stability. Pompe disease (PD) fibroblasts from patients 2 and 4 were treated with rhGAA, with rhGAA plus either NAC or NB-DNJ, and with rhGAA plus the combination of the two chaperones. In both cell lines the combination of NAC and NB-DNJ resulted in the highest enhancement of GAA activity by rhGAA (b).

for rhGAA stabilization and that the other antioxidants studied did not stabilize the enzyme. This was somewhat surprising because NAC is structurally very different from the imino sugars, the only known pharmacological chaperones of GAA so far, that resemble the natural substrates/products of the enzyme.

We showed that NAC improved stability of GAA in response to physical stresses such as temperature and pH changes. Resistance to pH variations is of interest, as neutral pH may be representative of some of the environmental conditions encountered by recombinant enzymes in plasma and in certain cellular compartments. It has been shown that pH induces conformational changes in other lysosomal enzymes, such as  $\beta$ -glucocerebrosidase,<sup>35</sup> and that binding with pharmacological chaperones may favor the most stable conformations of the enzyme.

Here, NAC, and the structurally related amino acids NAS and NAG, prevented the loss of GAA activity as a function of pH and increased the enzyme thermal stability with no inhibition, while the nonacetylated counterparts had no effect. To rationalize these data at molecular level, we analyzed the conformational dynamics of a 3D model of GAA through MD simulations. Although a high-resolution 3D-structure would certainly be the most accurate tool to perform these studies, structural models based on X-ray crystallography are not available for many glycosidases similar to GAA. According to these simulations, interaction with NAC favored a more stable conformation of the enzyme. Interestingly, the integrated analysis of the results of GAA simulations carried out in different conditions showed that a possible binding site for NAC could be located at 35 Å distance from the active site, thereby validating the allosteric stabilization observed *in vitro*. The binding site was defined based on the combined characterization of the variations in the local dynamics of specific protein sites in response to different pH conditions and to the presence/absence of possible small-molecule ligands. Such analyses were next combined to the investigation of the druggability of possible identified binding pockets. The integrated analysis of the data eventually returned one solution that we defined as the putative allosteric pocket. Docking of NAC and its analogues to the allosteric pocket helped define the most favorable pose for the chaperone in complex with the protein. A control docking run with DNJ proved that this molecule strongly favors binding to the active site (**Supplementary Table S2**).

These observations were largely confirmed by the analyses performed in cellular systems (fibroblasts from PD patients and COS7 cells overexpressing mutated GAA constructs). In COS7 cells that overexpress mutated GAA incubation with NAC resulted in significantly increased residual GAA activity for three of the seven mutations studied. Remarkably, the chaperoning profile of NAC showed differences compared to that of NB-DNJ and DNJ. This may translate into an expansion of the number of chaperone-responsive mutations, and should be further investigated in large-scale studies. It would be tempting to envisage that preliminary screenings *in vitro* on a number of chaperones would allow personalization of treatment protocols aimed at obtaining the greatest beneficial effect in different PD patients.

NAC also increased the efficacy of rhGAA in correcting the enzyme defect in PD fibroblasts. Compared to the effect of NAC, and of chaperones in general, on the mutated enzymes, this effect

holds greater promise for the cure of patients affected by PD, and possibly of other lysosomal disorders. It should be considered that, while the enhancement of endogenous defective enzymes by chaperones in most cases resulted in minor changes in terms of residual activity (likely with a modest impact on patients' outcome), the synergy of these drugs with ERT induced in PD fibroblasts remarkable increases in specific activity. In this study, coadministration of NAC and rhGAA resulted in complete correction of the enzymatic defect. An additional advantage of the combination of PCT and ERT is that the effect of chaperones is directed towards the wild-type recombinant enzyme and does not depend on the type of mutation carried by patients. Thus, any patient on ERT may benefit from the coadministration of PCT.

The results that support a synergy between chaperones and recombinant enzymes have important clinical implications and may translate into improved clinical efficacy of ERT. The data obtained in our proof-of-concept *in vivo* experiments in PD mice appear promising in this respect, although treatment protocols with NAC need to be carefully optimized in studies on a larger number of animals. The concentrations used in our studies may appear relatively high and difficult to translate into human therapy. However, toxicity of NAC is reported to be low and plasma concentrations in the range of 2-3 mmol/l can be reached for the treatment of paracetamol overdose.<sup>36</sup> Long-term treatment with NAC doses of 5 g/kg/day (higher than those used in the *in vivo* experiments our study) has been used in the mouse model of ethylmalonic encephalopathy and resulted in improved survival of treated animals.<sup>37</sup> NAC therapy was also used in patients affected by this metabolic disorder, albeit at lower doses, and again resulted in improved clinical and neurological course and in improvement of biochemical markers. In addition it must be taken into account that, for a possible clinical translation of our data, a therapeutic protocol would be based on the administration of the chaperone at the time of ERT infusion, and not on a long-term daily administration of the chaperone. By using such an approach the risk of adverse events would likely be reduced. It is also possible that modifications to NAC might be required to have the same stabilizing effect at lower concentrations.

It should be mentioned that clinical trials based on the combination of imino sugar competitive chaperones and ERT are already in progress (see trials NCT00214500 and NCT01196871 at <http://clinicaltrials.gov>; Telethon foundation trial GUP09017, <http://www.telethon.it/ricerca-progetti/progetti-finanziati>).

The identification of NAC and derivatives, which are structurally very different from the other known pharmacological chaperones identified in PD is promising. In fact, other molecules, whose chaperoning activity cannot be simply inferred from their structure, may be effective in several lysosomal storage disease, thereby opening new and wider opportunities for the identification of novel therapeutic drugs.

## MATERIALS AND METHODS

**Fibroblast cultures.** Fibroblasts from PD and Fabry disease patients were derived from skin biopsies performed for diagnostic purposes after obtaining the informed consent from patients or from their parents. All cell lines were grown at 37°C with 5% CO<sub>2</sub> in Dulbecco's modified Eagle's medium (Invitrogen, Grand Island, NY) and 10% fetal bovine serum

(Sigma-Aldrich, St Louis, MO), supplemented with 100 U/ml penicillin and 100 mg/ml streptomycin.

**Reagents.** rhGAA (alugcosidase, Myozyme) and rh- $\alpha$ -Gal A (agalsidase- $\beta$ , Fabrazyme, for the experiments shown in **Supplementary Figure S5**) were from Genzyme, Cambridge, MA. As source of enzyme we used the residual amounts of the reconstituted recombinant enzymes prepared for the treatment of PD and Fabry patients at the Department of Pediatrics, Federico II University, Naples, Italy. NAC, NAS, NAG, cysteine, serine, glycine, 2-mercaptoethanol, 4-nitrophenyl- $\alpha$ -glucopyranoside (4NP-Glc), DNJ, NB-DNJ, and 1-deoxy-galactonojijirimycin were from Sigma-Aldrich.

The rabbit anti-GAA primary antibody used for immunofluorescence and western blot analysis was a gift from Dr Bruno Bembi and Dr Andrea Dardis, Centro di Coordinamento Regionale per le Malattie Rare, Udine, Italy; the anti- $\beta$ -actin mouse monoclonal antibody was from Sigma-Aldrich. The anti-rabbit secondary antiserum were from Molecular Probes, Eugene, OR; horseradish peroxidase-conjugated anti-rabbit or anti-mouse immunoglobulin G were from Amersham, Freiburg, Germany. Labeling of rhGAA was performed using the Alexa Fluor 546 labeling kit (Molecular Probes) according to the manufacturer instructions.

**Enzyme characterization.** The standard activity assay of rhGAA was performed in 200  $\mu$ l by using 5  $\mu$ g of enzyme at 37 °C in 100 mmol/l sodium acetate pH 4.0 and 20 mmol/l 4NP-Glc. The reaction was started by adding the enzyme; after suitable incubation time (1-2 minutes) the reaction was blocked by adding 800  $\mu$ l of 1 mol/l sodium carbonate pH 10.2. Absorbance was measured at 420 nm at room temperature, the extinction coefficient to calculate enzymatic units was 17.2 mmol/l<sup>-1</sup>  $\times$  cm<sup>-1</sup>. One enzymatic unit is defined as the amount of enzyme catalyzing the conversion of 1  $\mu$ mol substrate into product in 1 minute, under the indicated conditions.

**Effect of pH on rhGAA.** The effect of different pHs on the rhGAA stability was measured by preparing reaction mixtures containing 0.75 mg ml<sup>-1</sup> of enzyme in the presence of 50 mmol/l citrate/phosphate (pH 3.0–7.0) at a certain pH. rhGAA mixture was kept at 37 °C and aliquots were withdrawn at the times indicated (0–24 hours) and the residual GAA activity was measured with the standard assay.

To test the effect on the pH stability of rhGAA of pharmacological chaperones and other molecules and to test the effect of chaperones on rhGAA activity, experiments were performed as described above by adding to the reaction mixtures 0.1, 1, 10 mmol/l of the different compounds indicated in the text.

**Analysis of thermal stability of rhGAA.** Thermal stability scans of rhGAA were performed as described in ref. 38; briefly, 2.5  $\mu$ g of enzyme were incubated in the absence and in the presence of NAC and DNJ, 10 and 0.1 mmol/l, respectively, SYPRO Orange dye, and sodium phosphate 25 mmol/l buffer, 150 mmol/l NaCl, pH 7.4 or sodium acetate 25 mmol/l buffer, 150 mmol/l NaCl, pH 5.2. Thermal stability scans were performed at 1 °C/minute in the range 25–95 °C in a Real-Time Light Cycler (Biorad, Milan, Italy). SYPRO Orange fluorescence was normalized to maximum fluorescence value within each scan to obtain relative fluorescence.

**Modeling and molecular dynamics studies of GAA.** The atomic structural model of GAA was constructed via homology modeling. The homology module of the ICM software (Molsoft, San Diego, CA; www.molsoft.com) was used for this purpose. The structure of the human maltase-glucoamylase,<sup>39</sup> showing 44% identity over the aligned 868 residues of GAA, was used as a template. After alignment of the target sequence with the related 3D structure the modeling proceeded through the following steps: (i) a starting model was created based on the homologous structure with the conserved portion fixed and the nonconserved portion having standard covalent geometry and free torsion angles; (ii) the Biased Probability Monte Carlo procedure was applied to search the subspaces of either all the nonconservative side-chain torsion angles or torsion angles in a loop backbone and surrounding side chains. The Biased Probability

Monte Carlo procedure globally optimizes the energy function consisting of ECEPP/3 and solvation energy terms. Next, the structure was refined through energy minimization with the Gromos96 force-field and the good quality of the model was assessed using ICM, the WHAT\_CHECK module of the program WHAT\_IF and the program PROCHECK. This model was compared with previous modeling studies and the results appeared to be consistent with what reported by other authors.<sup>31,40</sup>

The structure obtained after the modeling step was subjected to MD simulation analysis. Three sets of MD simulations were run using the same protocol: one simulation for the isolated apo-GAA, one for GAA bound to NAC, and one for GAA surrounded by multiple copies of NAC or NAG. pH variations were mimicked by protonating ionizable groups according to their pK<sub>a</sub> values.

In each MD run, the system was solvated in a tetrahedral solvation box with a size allowing the presence of a 1.2-nm layer of water on each side of the protein. All simulations and the analysis of the trajectories were performed using the 4.0.3 version of the GROMACS software package<sup>41</sup> using the GROMOS96 force field<sup>42,43</sup> and the SPC water model.<sup>44</sup> In the case of the simulation with multiple copies of NAC or NAG, 10 molecules of co-solute were randomly placed in the simulation box at a distance of 1.0 nm from GAA.

Each system was first energy relaxed with 2,000 steps of steepest descent energy minimization followed by another 2,000 steps of conjugate gradient energy minimization, in order to remove possible bad contacts from the initial structures. Next, the systems were first equilibrated by 50 ps of MD runs with position restraints on the protein and ligand to allow relaxation of the solvent molecules. These first equilibration runs were followed by other 50 ps runs without position restraints on the solute. The first 20 ns of each trajectory were not used in the subsequent analysis in order to minimize convergence artifacts. Equilibration of the trajectories was checked by monitoring the equilibration of the root mean square deviation with respect to the initial structure, and of the internal protein energy.

Production runs span 100 ns each system studied. The electrostatic term was described by using the particle mesh Ewald algorithm. The LINCS<sup>45</sup> algorithm was used to constrain all bond lengths. For the water molecules the SETTLE algorithm<sup>46</sup> was used. A dielectric permittivity,  $\epsilon = 1$ , and a time step of 2 fs were used. All atoms were given an initial velocity obtained from a Maxwellian distribution at the desired initial temperature of 300K. The density of the system was adjusted performing the first equilibration runs at NPT condition by weak coupling to a bath of constant pressure ( $P_0 = 1$  bar, coupling time  $\hat{\sigma}_p = 0.5$  ps).<sup>47</sup> In all simulations the temperature was maintained close to the intended values by weak coupling to an external temperature bath with a coupling constant of 0.1 ps. The proteins and the rest of the system were coupled separately to the temperature bath.

The structural stability was initially monitored by analyzing the time evolution of structural parameters such as the root mean square deviation of each conformation visited from the starting structure.

The mechanistic determinants of GAA structural stability at different pHs was obtained by analyzing the protein's pH-dependent geometric strain.<sup>27</sup> This quantity reports on the variations of the local deformations of the residues' contact networks as a consequence of the changes in their respective local structural environment. This parameter thus reflects the extent of conformational changes throughout the protein structure in response to specific perturbations.

**Analysis of internal dynamic properties from MD trajectories.** For each MD trajectory we computed the regions in which the geometric deformation accumulates. Such quantity, defined as geometric strain accumulates can be characterized by defining the strain of a given amino acid,  $i$ , as

$$p(i) = \sum A_{ij} f(\langle d_{ij} \rangle)$$

where  $j$  runs over all protein amino acids, and  $f$  is a sigmoidal function that restricts the contribution to the sum to amino acids that are within

about 5 Å from amino acid  $i$ :  $f(x) = [1 - \tanh(x-5)]/2$ , where  $x$  is expressed in Ångströms. Regions that respond differently to a certain perturbation (such as the pH change, the presence of a certain ligand) are characterized by different values of  $p$  because in the course of the dynamical evolution their local network of contacts appreciably changes due to the relative motion of neighboring substructures.

To identify the regions of the protein surface that may be most frequently in contact with NAC, we combined MD simulations with molecular docking experiments. The density maps report the areas on the protein where NAC molecules cluster for most of the time. It is worth mentioning that this simulation should be intended as a coarse-grained exploration of the surface rather than a docking calculation. We ran an additional 100 ns MD simulation with one copy of GAA surrounded by 10 copies of NAC molecules, initially placed at random positions in the simulation box.

In order to generate a refined and optimized model of the possible NAC–GAA complex, the surface of the protein was further scanned for possible small-molecule binding sites with the “Site Map” function of the Maestro suite of programs. This function identifies possible sites on a protein that have the structural and functional requirements that are optimal for binding a small molecule. The pH responsive hot spot region identified earlier was thus used as a target for a set of unrestrained docking calculations, using a large receptor 3-D grid of 20 Å of length on each side around the hot spot, which also included the residues that resulted to contact NAC in the aforementioned MD exploration.

**Effect of NAC in COS7 cells and PD fibroblasts.** Cultured PD fibroblasts and COS7 cells were treated with 10mmol/l NAC for 24 hours before being harvested and used for GAA assay. COS cells were transfected with mutated GAA gene constructs as indicated in refs. 30,31.

GAA activity was assayed by using the fluorogenic substrate 4-methylumbelliferyl- $\alpha$ -D-glucopyranoside (Sigma-Aldrich) as described in ref. 22. Briefly, 25  $\mu$ g of protein were incubated with the fluorogenic substrate (2 mmol/l) in 0.2 mol/l acetate buffer, pH 4.0, for 60 minutes in incubation mixtures of 100  $\mu$ l. The reaction was stopped by adding 700  $\mu$ l of glycine-carbonate buffer, pH 10.7. Fluorescence was read at 365 nm (excitation) and 450 nm (emission) on a Turner Biosystems Modulus fluorometer. Protein concentration in cell homogenates was measured by the Bradford assay (Biorad, Hercules, CA).

**Western blot analysis.** To study GAA immunoreactive material, fibroblast extracts were subjected to western blot analysis as described in ref. 21. The cells were harvested, washed in phosphate-buffered saline, resuspended in water, and disrupted by five cycles of freeze-thawing. Equal amounts (20  $\mu$ g protein) of fibroblast extracts were subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis and proteins were transferred to PVD membrane (Millipore, Billerica, MA). An antihuman GAA antiserum was used as primary antibody to detect GAA polypeptides; to detect  $\beta$ -actin, a monoclonal mouse antibody was used. Immunoreactive proteins were detected by chemiluminescence (ECL; Amersham). To quantitate the amounts of GAA-related bands western blots were analyzed by ImageJ.

**Immunofluorescence analysis and confocal microscopy.** To study the distribution of AlexaFluor546-labeled GAA, PD fibroblasts grown on coverslips were fixed using methanol, permeabilized using 0.1% saponin and locked with 0.01% saponin, 1% fetal bovine serum diluted in phosphate-buffered saline for 1 hour. The cells were incubated with the primary antibodies, with secondary antibodies in blocking solution and then mounted with vectashield mounting medium (Vector Laboratories, Burlingame, CA). Samples were examined with a Zeiss LSM 5–10 laser scanning confocal microscope.

**In vivo studies.** Animal studies were performed according to the European Union Directive 86/609, regarding the protection of animals used for experimental purposes. PD mice were housed in the TIGEM animal facility. Every procedure on the mice was performed with the aim of ensuring

that discomfort, distress, pain, and injury would be minimal. Mice were euthanized following avertin anesthesia by cervical dislocation. To study the effects of the combination of NAC and rhGAA the animals were allowed to drink 138 mmol/l NAC in water *ad libitum* (4.2 g/kg/day) for 5 days. On day 3 the animals received a single rhGAA injection. On day 5 the animals were euthanized and the tissues were homogenized with a tissue lyser and GAA activity was measured as described in ref. 21.

## SUPPLEMENTARY MATERIAL

**Figure S1.** Molecular structures of the compounds analyzed in this study.

**Figure S2.** Effect of NAS, NAG, and nonacetylated amino acids on rhGAA at different concentrations (0, 0.1, 1, 10 mmol/l) and up to 48 hours of incubation.

**Figure S3.** The analysis of the root mean square fluctuation (RMSF) per residue showed that the protein residues experience lower structural fluctuations at acidic pH than at neutral pH, suggesting a higher structural stability at the lower pH condition (data not shown).

**Figure S4.** Effect of the antioxidants epigallo catechingallate (EGCG) and resveratrol on the efficacy of rhGAA in cultured PD fibroblasts (patient 3).

**Figure S5.** Molecular modeling data suggest that the interactions of NAC occur at a specific protein domain that shows sequence homology within other glycosidases belonging to family GH31.

**Table S1.** The lists of residues reported in the table have been calculated by running a simulation of GAA in the presence of 10 NAC molecules starting from random positions in the simulation box.

**Table S2.** Grid Docking scores of acetylated compounds to the allosteric site, and of DNJ to the active site.

## ACKNOWLEDGMENTS

This work was supported by the Telethon Foundation, Rome, Italy [grant TGPMT4TELD to G.P.], by Programma Operativo Nazionale (PON) 01\_00862, by the Agenzia Spaziale Italiana [project MoMa n.1/014/06/0] and by the project “Nuove glicosidasi d’interesse terapeutico nella salute umana” within the exchange programme Galileo 2009-10 of the Università Italo-Francese (M.C.F., B.C.-P., and M.M.). We thank Dr Graciana Diez-Roux (Telethon Institute of Genetics and Medicine) for critically reviewing the manuscript and for helpful discussion. The study was performed in Naples, Italy.

## REFERENCES

- van der Ploeg, AT and Reuser, AJ (2008). Pompe’s disease. *Lancet* **372**: 1342–1353.
- Shea, L and Raben, N (2009). Autophagy in skeletal muscle: implications for Pompe disease. *Int J Clin Pharmacol Ther* **47** Suppl 1: S42–S47.
- Parenti, G and Andria, G (2011). Pompe disease: from new views on pathophysiology to innovative therapeutic strategies. *Curr Pharm Biotechnol* **12**: 902–915.
- Van den Hout, JM, Kamphoven, JH, Winkel, LP, Arts, WF, De Klerk, JB, Loonen, MC *et al.* (2004). Long-term intravenous treatment of Pompe disease with recombinant human alpha-glucosidase from milk. *Pediatrics* **113**: e448–e457.
- Van den Hout, H, Reuser, AJ, Vulto, AG, Loonen, MC, Cromme-Dijkhuis, A and Van der Ploeg, AT (2000). Recombinant human alpha-glucosidase from rabbit milk in Pompe patients. *Lancet* **356**: 397–398.
- Kishnani, PS, Corzo, D, Nicolino, M, Byrne, B, Mandel, H, Hwu, WL *et al.* (2007). Recombinant human acid [alpha]-glucosidase: major clinical benefits in infantile-onset Pompe disease. *Neurology* **68**: 99–109.
- Strothotte, S, Strigl-Pill, N, Grunert, B, Kornblum, C, Eger, K, Wessig, C *et al.* (2010). Enzyme replacement therapy with alglucosidase alfa in 44 patients with late-onset glycogen storage disease type 2: 12-month results of an observational clinical trial. *J Neurol* **257**: 91–97.
- van der Ploeg, AT, Clemens, PR, Corzo, D, Escolar, DM, Florence, J, Groeneveld, GJ *et al.* (2010). A randomized study of alglucosidase alfa in late-onset Pompe’s disease. *N Engl J Med* **362**: 1396–1406.
- Schooser, B, Hill, V and Raben, N (2008). Therapeutic approaches in glycogen storage disease type II/Pompe Disease. *Neurotherapeutics* **5**: 569–578.
- Chien, YH, Lee, NC, Thurberg, BL, Chiang, SC, Zhang, XK, Keutzer, J *et al.* (2009). Pompe disease in infants: improving the prognosis by newborn screening and early treatment. *Pediatrics* **124**: e1116–e1125.
- Kishnani, PS, Corzo, D, Leslie, ND, Gruskin, D, Van der Ploeg, A, Clancy, JP *et al.* (2009). Early treatment with alglucosidase alpha prolongs long-term survival of infants with Pompe disease. *Pediatr Res* **66**: 329–335.
- Kishnani, PS, Goldenberg, PC, DeArme, SL, Heller, J, Benjamin, D, Young, S *et al.* (2010). Cross-reactive immunologic material status affects treatment outcomes in Pompe disease infants. *Mol Genet Metab* **99**: 26–33.

13. Raben, N, Danon, M, Gilbert, AL, Dwivedi, S, Collins, B, Thurberg, BL *et al.* (2003). Enzyme replacement therapy in the mouse model of Pompe disease. *Mol Genet Metab* **80**: 159–169.
14. Wenk, J, Hille, A and von Figura, K (1991). Quantitation of Mr 46000 and Mr 300000 mannose 6-phosphate receptors in human cells and tissues. *Biochem Int* **23**: 723–731.
15. Fukuda, T, Ahearn, M, Roberts, A, Mattaliano, RJ, Zaal, K, Ralston, E *et al.* (2006). Autophagy and mistargeting of therapeutic enzyme in skeletal muscle in Pompe disease. *Mol Ther* **14**: 831–839.
16. Raben, N, Baum, R, Schreiner, C, Takikita, S, Mizushima, N, Ralston, E *et al.* (2009). When more is less: excess and deficiency of autophagy coexist in skeletal muscle in Pompe disease. *Autophagy* **5**: 111–113.
17. Xu, YH, Ponce, E, Sun, Y, Leonova, T, Bove, K, Witte, D *et al.* (1996). Turnover and distribution of intravenously administered mannose-terminated human acid beta-glucosidase in murine and human tissues. *Pediatr Res* **39**: 313–322.
18. Benjamin, ER, Khanna, R, Schilling, A, Flanagan, JJ, Pellegrino, LJ, Brignol, N *et al.* (2012). Co-administration with the pharmacological chaperone AT1001 increases recombinant human α-galactosidase A tissue uptake and improves substrate reduction in Fabry mice. *Mol Ther* **20**: 717–726.
19. Fan, JQ (2008). A counterintuitive approach to treat enzyme deficiencies: use of enzyme inhibitors for restoring mutant enzyme activity. *Biol Chem* **389**: 1–11.
20. Parenti, G (2009). Treating lysosomal storage diseases with pharmacological chaperones: from concept to clinics. *EMBO Mol Med* **1**: 268–279.
21. Porto, C, Cardone, M, Fontana, F, Rossi, B, Tuzzi, MR, Tarallo, A *et al.* (2009). The pharmacological chaperone N-butyldeoxyjirimycin enhances enzyme replacement therapy in Pompe disease fibroblasts. *Mol Ther* **17**: 964–971.
22. Porto, C, Pisani, A, Rosa, M, Acampora, E, Avolio, V, Tuzzi, MR *et al.* (2012). Synergy between the pharmacological chaperone 1-deoxygalactonojirimycin and the human recombinant alpha-galactosidase A in cultured fibroblasts from patients with Fabry disease. *J Inher Metab Dis* **35**: 513–520.
23. Shen, JS, Edwards, NJ, Hong, YB and Murray, GJ (2008). Isogomine increases lysosomal delivery of exogenous glucocerebrosidase. *Biochem Biophys Res Commun* **369**: 1071–1075.
24. Valenzano, KJ, Khanna, R, Powe, AC, Boyd, R, Lee, G, Flanagan, JJ *et al.* (2011). Identification and characterization of pharmacological chaperones to correct enzyme deficiencies in lysosomal storage disorders. *Assay Drug Dev Technol* **9**: 213–235.
25. Tropak, MB, Blanchard, JE, Withers, SG, Brown, ED and Mahuran, D (2007). High-throughput screening for human lysosomal beta-N-Acetyl hexosaminidase inhibitors acting as pharmacological chaperones. *Chem Biol* **14**: 153–164.
26. Zheng, W, Padia, J, Urban, DJ, Jadhav, A, Goker-Alpan, O, Simeonov, A *et al.* (2007). Three classes of glucocerebrosidase inhibitors identified by quantitative high-throughput screening are chaperone leads for Gaucher disease. *Proc Natl Acad Sci USA* **104**: 13192–13197.
27. Morra, G, Potestio, R, Micheletti, C and Colombo, G (2012). Corresponding functional dynamics across the Hsp90 Chaperone family: insights from a multiscale analysis of MD simulations. *PLoS Comput Biol* **8**: e1002433.
28. Lamers, ML, Almeida, ME, Vicente-Manzanares, M, Horwitz, AF and Santos, MF (2011). High glucose-mediated oxidative stress impairs cell migration. *PLoS ONE* **6**: e22865.
29. Rieber, M and Rieber, MS (2003). N-Acetylcysteine enhances UV-mediated caspase-3 activation, fragmentation of E2F-4, and apoptosis in human C8161 melanoma: inhibition by ectopic Bcl-2 expression. *Biochem Pharmacol* **65**: 1593–1601.
30. Parenti, G, Zuppaldi, A, Gabriela Pittis, M, Rosaria Tuzzi, M, Annunziata, I, Meroni, G *et al.* (2007). Pharmacological enhancement of mutated alpha-glucosidase activity in fibroblasts from patients with Pompe disease. *Mol Ther* **15**: 508–514.
31. Flanagan, JJ, Rossi, B, Tang, K, Wu, X, Mascioli, K, Donaudy, F *et al.* (2009). The pharmacological chaperone 1-deoxyjirimycin increases the activity and lysosomal trafficking of multiple mutant forms of acid alpha-glucosidase. *Hum Mutat* **30**: 1683–1692.
32. Cardone, M, Porto, C, Tarallo, A, Vicinanza, M, Rossi, B, Polishchuk, E *et al.* (2008). Abnormal mannose-6-phosphate receptor trafficking impairs recombinant alpha-glucosidase uptake in Pompe disease fibroblasts. *Pathogenetics* **1**: 6.
33. Wisselaer, HA, Kroos, MA, Hermans, MM, van Beeumen, J and Reuser, AJ (1993). Structural and functional changes of lysosomal acid alpha-glucosidase during intracellular transport and maturation. *J Biol Chem* **268**: 2223–2231.
34. Raben, N, Nagaraju, K, Lee, E, Kessler, P, Byrne, B, Lee, L *et al.* (1998). Targeted disruption of the acid alpha-glucosidase gene in mice causes an illness with critical features of both infantile and adult human glycogen storage disease type II. *J Biol Chem* **273**: 19086–19092.
35. Lieberman, RL, Wustman, BA, Huertas, P, Powe, AC Jr, Pine, CW, Khanna, R *et al.* (2007). Structure of acid beta-glucosidase with pharmacological chaperone provides insight into Gaucher disease. *Nat Chem Biol* **3**: 101–107.
36. Walsh, TS, Hopton, P, Philips, BJ, Mackenzie, SJ and Lee, A (1998). The effect of N-acetylcysteine on oxygen transport and uptake in patients with fulminant hepatic failure. *Hepatology* **27**: 1332–1340.
37. Viscomi, C, Burlina, AB, Dweikat, I, Savoirdo, M, Lamperti, C, Hildebrandt, T *et al.* (2010). Combined treatment with oral metronidazole and N-acetylcysteine is effective in ethylmalonic encephalopathy. *Nat Med* **16**: 869–871.
38. Niesen, FH, Berglund, H and Vedadi, M (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**: 2212–2221.
39. Sim, L, Willemsma, C, Mohan, S, Naim, HY, Pinto, BM and Rose, DR (2010). Structural basis for substrate selectivity in human maltase-glucoamylase and sucrase-isomaltase N-terminal domains. *J Biol Chem* **285**: 17763–17770.
40. Tajima, Y, Matsuzawa, F, Aikawa, S, Okumiya, T, Yoshimizu, M, Tsukimura, T *et al.* (2007). Structural and biochemical studies on Pompe disease and a “pseudodeficiency of acid alpha-glucosidase”. *J Hum Genet* **52**: 898–906.
41. Hess, B, Kutzner, C, Van Der Spoel, D, Lindahl, E (2008) GROMACS 4: Algorithms for highly efficient load-balanced, scalable molecular simulation. *J Chem Theory Comput* **4**: 435–447.
42. van Gunsteren, WF, Daura, WFX and Mark, AE (1998) GROMOS force field. *Encyclopedia of Computational Chemistry* **2**: 1211–1216.
43. Scott, WRP, Hünenberger, PH, Tironi, IG, Mark, AE, Billeter, SR, Fennen, J *et al.* (1999). The GROMOS biomolecular simulation program package. *J Phys Chem A* **103**: 3596–3607.
44. Berendsen HJ, Grigera JR, Straatsma TP (1987) The missing term in effective pair potentials. *J Phys Chem* **91**: 6269–6271.
45. Hess, E, Bekker, H, Berendsen, HJC and Fraaije, JGEM (1997). LINCS: A linear constraint solver. *J Comput Chem* **18**: 1463–1472.
46. Miyamoto, S, Kollmann, PA (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comp Chem* **13**: 952–962.
47. Berendsen, HJ, Postma, JP, van Gunsteren, WF, DiNola, A, Haak, JR (1984). Molecular dynamics with coupling to an external bath. *J Chem Phys* **81**: 3684.