UNIVERSITÀ DEGLI STUDI DI NAPOLI

FEDERICO II

DOCTORATE SCHOOL IN BIOLOGY

Cycle XXXIV

THE ROLE OF MISSENSE MUTATIONS IN DISEASES: TWO SIDES OF A COIN

Supervisor

Prof.ssa Maria Vittoria Cubellis

Coordinator

Prof. Sergio Esposito

***Ph.D.* student**

Maria Monticelli

ACADEMIC YEAR 2020/2021

A coloro che mi hanno guidata verso la complessità,

insegnandomi ad amarla.

"Ad ogni passo ci vien ricordato che noi non dominiamo la natura come un conquistatore domina un popolo straniero soggiogato, che non la dominiamo come chi è estraneo ad essa, ma che noi le apparteniamo con carne e sangue e cervello e viviamo nel suo grembo: tutto il nostro dominio sulla natura consiste nella capacità, che ci eleva al di sopra delle altre creature, di conoscere le sue leggi e di impiegarle in modo appropriato."

(F. Engels)

INDEX

1. ABSTRACT

1.1 English version

Missense mutations play a central role in non-physiological conditions.

On the one hand, they represent the molecular basis of many genetic diseases, particularly rare diseases; on the other hand, common missense variants in the general population can influence the course of different diseases.

The present *Ph.D.* project aimed at deepening this topic by using different approaches and techniques.

Whole-exome sequencing analysis (WES) was employed to establish the genetic causes of four cases of Beckwith-Wiedemann syndrome associated with Multi-Locus Imprinting Disturbance, and *in silico* analysis of the identified mutations of interest was performed to figure out their role on the mutant protein, PADI6.

In silico docking was used as a starting point for rational drug discovery to find pharmacological chaperones (PC) for the phosphomannomutase-2 deficiency (PMM2-CDG). *In vitro* expression of recombinant proteins and their analysis *via* canonical methods based on spectrofluorimetry or innovative techniques based on ^{31}P-NMR, allowed the validation of *in silico* results. A potential PC for PMM2-CDG was identified this way, namely β-glucose-1,6-bisphosphate.

Cell-based techniques represented the chosen method for drug repurposing. In particular, the possibility of potentiating approved pharmacological chaperones by modulation of proteostasis was evaluated for Fabry-Anderson disease. Aspirin turned out to enhance and prolong the PC-induced protein stabilization.

COVID-19 outbreak represented the starting point to investigate the role of common missense variants in influencing various conditions.

Bioinformatics analysis of literature was employed to identify common variants in the general population potentially associated with SARS-CoV*. WES analysis on patients with different clinical outcomes confirmed the prediction. In particular, one variant was identified in *TMPRSS2* that showed a deleterious effect on the protease and a protective effect on the patients, especially in some age- and sex-related sub-groups.

1.2 Italian version

Le mutazioni missenso hanno un ruolo di primo piano in situazioni non fisiologiche. Da un lato, rappresentano le basi molecolari di molte malattie genetiche, in particolare malattie rare; dall'altro, varianti missenso comuni nella popolazione possono influenzare il decorso di varie malattie.

Questo progetto di dottorato ha avuto come scopo l'approfondimento di quest'argomento attraverso l'impiego di diversi approcci e tecniche.

Il sequenziamento dell'esoma (WES) è stato impiegato per stabilire le cause genetiche di quattro casi di pazienti affetti da sindrome di Beckwith-Wiedemann associata a disturbi dell'imprinting multi-locus (MLID), e successive analisi *in silico* sono state effettuate per comprendere il ruolo delle mutazioni di interesse identificate sulla proteina mutante, PADI6.

Il docking *in silico* è servito come primo step per l'approccio razionale alla ricerca di chaperones farmacologici (PC) per il deficit di fosfomannomutasi-2 (PMM2-CDG). L'espressione *in vitro* delle proteine ricombinanti e la loro analisi attraverso metodi canonici, basati sulla spettrofluorimetria, e tecniche innovative, basate sulla ^{31}P-NMR, hanno consentito la validazione dei risultati ottenuti *in silico*. In questo modo, un potenziale PC per la PMM2-CDG è stato identificato: il β-glucosio-1,6-bisfosfato.

L'utilizzo di test su cellule ha rappresentato il metodo d'elezione per il riposizionamento di farmaci. In particolare, è stata valutata la possibilità di potenziare lo chaperon in uso per la malattia Fabry-Anderson, attraverso l'utilizzo di modulatori della proteostasi. L'aspirina ha mostrato la capacità di potenziare e prolungare gli effetti stabilizzanti indotti dal PC.

La pandemia COVID-19 è poi servita da spunto per studiare il ruolo di varianti missenso comuni nell'influenzare varie condizioni. Nello specifico, un'analisi della letteratura condotta per via bioinformatica ha consentito di identificare varianti comuni nella popolazione generale e potenzialmente associate a SARS-CoV*. L'analisi di esomi di pazienti con differenti quadri clinici ha confermato la predizione. In particolare, è stata identificata una variante in *TMPRSS2* che ha un effetto deleterio sulla proteina ed un effetto protettivo sui pazienti, in particolare su alcuni sottogruppi legati all'età ed al genere.

2. AIM OF THE WORK

The present thesis is organized in the form of a paper collection; thus, published or submitted papers are presented after a general introduction. The narrative thread of this *Ph.D.* project is represented by missense mutations, faced from two different sides: 1. as causative of genetic diseases; 2. as influencing the acquisition or the course of other diseases.

The scheme in Figure 2.1 synthetically describes the evolution of the project and the connections between the different topics.

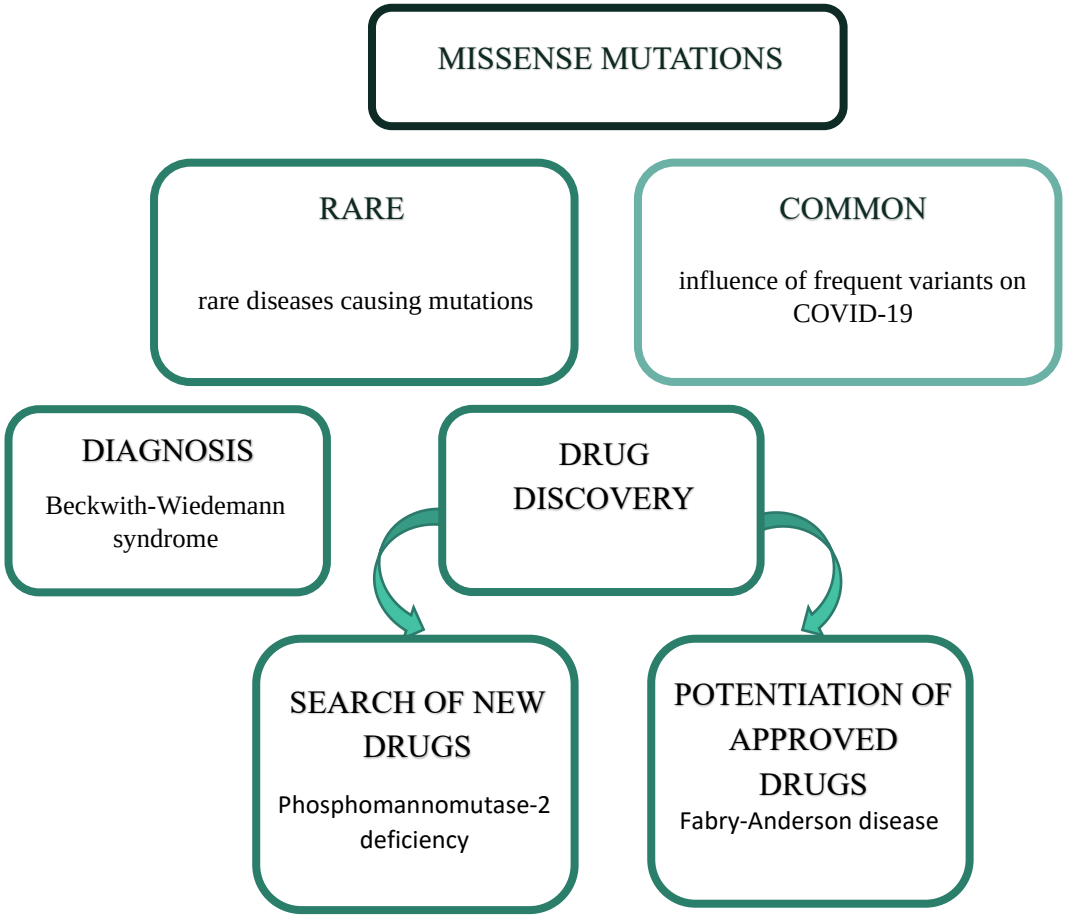

Figure 2.1: project description

The perspective of deepening the role of common variants in influencing the acquisition or the course of other diseases was challenged during the pandemic emergence of COVID-19. The identification of risk factors for the severe course of the infection represented an early challenge for the scientific community, but the role of host genetic factors in influencing the course of the disease is still being clarified.

During the first pandemic wave, we explored this topic starting with bioinformatics analysis, which allowed us to identify thirteen common variants in the general population associated with the term "SARS-CoV*" in literature. For five of these variants, we were able to prove a correlation between the frequency in different populations and the incidence/mortality rates of the associated countries. The results of this research were published in (Monticelli, Mele, Andreotti, et al., 2021) and are here presented in chapter four.

A collaboration with the GEN-COVID Multicenter Study and the group led by Pf. A. Renieri from the Dept. of Medical Biotechnologies of the University of Siena, allowed us to validate those results by analyzing whole-exome sequencing data from patients classified according to the severity of the disease. In particular, the common variant p.Val197Met in TMPRSS2 emerged as a predictive factor to identify patients at risk of a severe course of COVID-19, especially among two groups: those of young men and elderly women. These results were published in (Monticelli, Mele, Benetti, et al., 2021) and are presented in chapter five.

The core of the project is the research on rare diseases caused by missense mutations through *in vitro* and *in silico* approaches, for diagnosis and drug discovery.

Whole-exome sequencing (WES) is among the most powerful new generation tools to identify mutations possibly causing genetic diseases. The bioinformatics and biochemistry supports are then central to strengthen the hypothesis and analyze identified mutations concerning their effects on the protein. This approach was applied in a collaboration with the group led by Pf. A. Riccio from the Dept. of Ambiental, Biological and Pharmaceutical Sciences and Technologies (DiSTABiF) of the University of Caserta, that provided the WES of four patients clinically diagnosed with the Beckwith-Wiedemann syndrome (BWS) and those of their parents. We were able to identify five novel pathogenic variants in the gene *PADI6*, encoding a component of the subcortical maternal complex. These maternal loss-of-function variants were causative of BWS with multi-locus imprinting disturbances. The results of this collaboration were published in (Cubellis et al., 2020) and represent the object of chapter six.

About diseases caused by missense mutations, one of the most interesting therapeutic approaches is the use of the so-called “pharmacological chaperones” (PCs). PCs are small molecules specific for a target protein, able to stabilize its native form. In the last twenty years, there has been a lot of interest generated by research on PCs, as many have been recently approved. We performed a deep search of the literature and provided a comprehensive review of the topic, published as (Liguori et al., 2020) and presented in chapter seven.

The possibility to use a PC as a therapeutical approach was faced for a rare disease belonging to the Congenital Disorders of Glycosylation (CDG): the deficiency of phosphomannomutase-2, also known as PMM2-CDG. This disease has currently no therapeutic approach available. We explored the characteristics of a potential PC that is analogous to the substrate, namely β-glucose-1,6-bisphosphate. We started from *in silico* experiments, then moved to *in vitro* validation of its usefulness. These results were published in (Monticelli et al., 2019) and are discussed in chapter eight. PCs designed from analogous of the substrate are easy to find and have usually good stabilizing properties but represent enzymatic inhibitors, and this is a controversial side-effect; they can be identified as “first-generation PCs”. This is the case of the PC currently used for the cure of Fabry-Anderson disease, a lysosomal storage disease, also known as alpha-galactosidase deficiency. The PC is 1-deoxygalactonojirimycin (DGJ), which requires an intermittent regimen to minimize side effects. We were inspired by the recent approval by the FDA of combined therapy for cystic fibrosis, we explored the possibility of boosting the stabilization by DGJ combining it with other drugs. Thus, the dosage would be reduced, and the inhibitory effect would be minimized. A powerful combined therapy was obtained with acetylsalicylic acid, that strongly enhanced alpha-galactosidase (AGAL) in cells in combination with DGJ and was shown to prolong the PC effects. Since drug discovery for rare diseases is a difficult, long, and expensive process, the possibility of drug repositioning is very attractive and cost-effective. The results of this research were submitted as Monticelli et al. to the International Journal of Molecular Sciences (EISSN 1422-0067) and are currently under review; they represent the object of chapter nine.

3. INTRODUCTION

In the present thesis, the role of missense mutations in diseases has been addressed from different points of view and using different approaches and techniques, *in vitro* and *in silico*. Diagnosis and drug discovery for rare diseases met the analysis of common variants' roles in a viral pandemic, such as COVID-19.

This introduction retraces the different aspects, highlighting the essential issues to be deepened in the next chapters. Following a general discussion about protein folding and protein destabilization by missense mutations (p. 3.1), the paragraphs continue according to the order of the papers collected. Thus, the first question to be faced is the role common variants in the course of COVID-19 (p. 3.2). Next, the definition of "rare disease" and its related implications (p. 3.3) precedes the highlights about the three rare diseases object of the thesis: the Beckwith-Wiedemann syndrome (p. 3.4), Phosphomannomutase-2 deficiency (p. 3.6) and Fabry-Anderson disease (p. 3.7). Also, information about common approaches for the treatment of rare disease are discussed (p. 3.5).

3.1 **Protein folding and missense mutations**

The three-dimensional structure of an active protein is called the "native state". The information related to the three-dimensional structure is contained in the protein primary sequence, *i.e.* the amino acid sequence.

Small proteins can reach their native state through a two-state transition, including a disordered denatured state and the ordered native state. Larger proteins undergo a more complex folding process, which often implies the activity of molecular chaperones (Baker, 2000).

Protein stability is thermodynamically measured as ΔG (free energy variation) and is defined as the ratio of folded to unfolded molecules. Folded and unfolded states' free energies are marginally different, and this implies that folded proteins are marginally stable (typically, ΔG values are in the range -3 to -10 kcal mol^{-1}) (DePristo et al., 2005). The transition from the unfolded to the native state includes a passage through a higher-free-energy transition state (Baker, 2000). Folding is essential for the correct protein function. Its leading biological role is well represented by the big cell machinery dedicated to protein folding and unfolded and aggregated proteins degradation (Molinari, 2007).

Missense mutations occur when a single base-pair change causes the substitution of an amino acid in the protein (NIH, n.d.). The effect of a mutation is measured as $\Delta\Delta G$, which is the difference between the wild-type and the mutant free energy variations. Negative values of $\Delta\Delta G$ are characteristics of stabilizing mutations, while positive values of $\Delta\Delta G$ stands for destabilizing mutations (DePristo et al., 2005).

3.2 The role of common variants in the course of COVID-19

COVID-19 is a viral infection caused by SARS-CoV-2, that causes fever, cough, shortness of breath, dyspnea, and muscle pain, among the most frequent symptoms. SARS-CoV-2 belongs to the *Betacoronavirus* genus; it was identified in Wuhan, Hubei Province, China, in December 2019, and rapidly spread all over the world, leading to the declaration of a pandemic emergence by the World Health Organization in March 2020. The variety of symptoms and the variability of the response to the infection led the scientific community to put a strong effort into the identification of possible risk factors associated with a more severe outcome. Age and co-morbidities were identified at the very beginning of the pandemic emergency. In particular, cardiovascular diseases, diabetes, and chronic lung diseases represent major risk factors for COVID-19 (Fang et al., 2020; Rod et al., 2020; Zhou et al., 2021). Besides this, different countries experienced different impacts by COVID-19, particularly during the so-called “first-wave”. A recent paper clearly described the role that different factors, such as socio-behavioral habits, political management, co-existence with other viruses, could have in influencing the pandemic course (Yamamoto & Bauer, 2020). Among these, particular relevance should be given to host genetic factors. It must be noticed that the influence of genetic factors has been hypothesized since the beginning but represented the most difficult to be proved, due to the factors listed above. Moreover, the globalized world we live in certainly complicates the association between population variants reported in literature and countries, thus the identification of genetic factors would require the analysis of patients’ samples.

We faced this challenge at the very beginning of the emergence when data coming from all over the world were not biased by pharmaceuticals variables, such as that

of vaccinated people percentages. We were initially interested in the identification of common variants with different frequencies among the populations, that could represent the basis for the various experiences of the countries. We were able to identify different variants of interest with bioinformatics analysis. As discussed above, the analysis of patients' samples could represent the only method to unbiased the eventual effect. We were able to validate the data obtained from literature by analyzing the exomes of 1177 Italian patients, thus confirming the important role of a *TMPRSS2* variant in influencing the course of COVID-19.

3.3 Rare diseases and their implications: from the diagnosis to the treatment

A rare disease, also known as an orphan disease, affects a small number of people compared to the general population. European Union considers as rare a disease affecting fewer than 5 people in 10000 (European Commission, n.d.), but different definitions exist worldwide (Orphan Drug Act, 1983). The global average prevalence has been indicated in 40 cases per 100000 people (Richter et al., 2015). The number of rare diseases is estimated to be comprised among 5000 and 8000, with 7000 representing the most commonly cited number (Haendel et al., 2020).

Diseases caused by mutations in a single gene are defined as Mendelian diseases (Alkuraya, 2021). Almost 80% of rare diseases are genetic diseases, and nearly all of them are Mendelian (Rahit & Tarailo-Graovac, 2020). The common aspects are related to the difficulties experienced in the diagnosis, treatments and management, and the absence of cures.

The diagnosis of a rare disease is a real challenge, that implies a co-operation of factors, including technologies, experience, and expertise (Salvatore et al., 2020). Many patients affected by a rare disease remain undiagnosed for years, in so far that the National Institutes of Health of USA (NIH) launched the Undiagnosed Diseases Program (UDP) in 2008, including different participant countries (Taruscio et al., 2015).

In 2010 the first case of a Mendelian disorder diagnosed by the use of whole-exome sequencing (WES) was reported (Ng et al., 2010). The combination between the draft genome and the next-generation sequencing technologies has been described

as a “game-changer” in the field of Mendelian disease. This is true for diagnosis, but also for drug discovery and, most importantly, for the development of personalized medicine (Alkuraya, 2021).

The absence of a cure is a common aspect among rare diseases, with few exceptions. Also, some patients can be treated by symptomatic treatments, which do not represent a cure.

Drug discovery approaches can be classified into two major branches: *de novo* and drug repositioning. *De novo* drug discovery is a time-consuming and expensive process, that requires many years before approval and drug commercialization. For these reasons, pharmaceutical industries are usually not very interested in investing their financial resources for this approach in the field of rare diseases.

Drug repositioning is very interesting to overcome some of the problems related to *de novo* discovery. It consists of the use of previously approved drugs for new therapeutic purposes (Ashburn & Thor, 2004). Compared to *de novo* drug discovery, drug repositioning reduces the time between “bench and bedside” and the research-related costs; most important, it reduces the risk of failure, from more than 95% to around 45% (Ashburn & Thor, 2004; Cha et al., 2018; Nosengo, 2016; Wouters et al., 2020).

3.4 Multi-locus imprinting disturbance in the Beckwith-Wiedemann syndrome: the role of the subcortical maternal complex

Genomic imprinting represents the epigenetic marking of parental origin; it results in monoallelic expression (Monk et al., 2019).

Beckwith-Wiedemann syndrome (BWS) is a growth disorder characterized by different symptoms and a generalized overgrowth. Imprinting defects in at least one of two domains usually represent the molecular basis of BWS: the locus including the *H19* and *IGF2* genes and the locus including the *KCNQ1OT1* and *CDKN1C* genes. Typically, the first locus shows gain-of-methylation while the second one shows loss-of-methylation.

In addition to these defects, almost one-third of BWS patients show methylation defects to a greater extent, a condition known as Multi-Locus Imprinting Disturbance (MLID). The causes are generally unknown, but in some cases,

mutations in maternal-effect or zygotic mutations have been identified. Many of these genes are components of the Subcortical Maternal Complex (SCMC), a multiprotein complex expressed in oocytes and early embryos (Bebbere et al., 2016), which plays a fundamental role in the oocyte-to-embryo transition. It is composed of at least *NLRP2*, *NLRP5*, *OOEP*, *TLE6*, *PADI6*, and *KHDC3L* (Demond et al., 2019).

Whether the SCMC affects imprinting is currently unknown. One of its members, *KHDC3L*, has been recently indicated as necessary for de novo methylation in oocytes (Demond et al., 2019).

Very little is known about the SCMC.

Among its components, PADI6 has been recently identified as part of the complex, but the catalytic activity of PADI6 has not yet been identified. Notwithstanding this, most of the *PADI6* mutations so far identified and associated with MLID or infertility, affect the arginine deiminase domain.

We enriched the group of *PADI6* mutations associated with MLID identifying five new pathogenic variants in three different families, all of them localized in the arginine deiminase domain, and suggested that maternal PADI6 is necessary for methylation maintenance in early embryos.

3.5 **Pharmacological chaperones and small molecules for the treatment of genetic diseases**

The term "pharmacological chaperon" indicates a small molecule able to specifically bind a target protein and stabilize its native form; it was introduced by (Morello et al., 2000). The attention towards this kind of drug has enormously increased in the last twenty years since they represent an easy approach for the treatment of diseases caused by missense mutations (Liguori et al., 2020).

The peculiar characteristic of pharmacological chaperones, that differentiate them from other non-specific small molecules, is the specificity of the binding to a target protein. Osmolytes, hydrophobic chaperones and proteasome inhibitors are different types of small molecules used at the same scope but acting generically on unstable proteins. The effect of all these molecules is the increase in the intracellular concentration of mutant proteins. The mechanism of action of PCs is to raise the

difference between the free energy of the folded and unfolded states by lowering the first one (Liguori et al., 2020).

Pharmacological chaperones (PCs) typically have side effects that depend upon their discovery approach. The easiest way to find a new ligand specific for a protein of interest is starting from the structure of its natural ligands. The design of a drug based on a natural protein-ligand usually leads to an active site ligand. The main consequence of this approach is that the drug is also an enzymatic activity inhibitor. These kinds of drugs can be referred to as "first-generation" pharmacological chaperones.

One of the recent methods used to overcome this issue is the search of allosteric ligands, which can be achieved by bioinformatics techniques, such as molecular docking, as described in (Citro, Peña-García, et al., 2016).

Techniques for the screening of PCs include *in vitro* tests and cell-based assays. Among the most interesting approaches, the Thermal Shift Assay (TSA) should be named. TSA is a fluorescence-based method that offers many advantages over different techniques, such as calorimetry, circular dichroism, limited proteolysis, and so on. It is a fast high-throughput method that allows the processing of many samples with a limited amount of protein. Among the limitations associated with this technique, the necessity of pure proteins is of utmost importance. The best results would be obtained performing the assay directly on the mutants of interest; anyway, when this is not possible, the eventual stabilization of the wt-protein usually represents a good index of possible action on the mutants (Andreotti et al., 2015).

3.6 **PMM2-CDG: a rare disease without a chaperon**

Phosphomannomutase-2 deficiency is the most common of congenital disorders of glycosylation, also known as PMM2-CDG or Jaeken disease. It was firstly described by Pf. Jaak Jaeken in the late 1970s, and since then almost 1000 patients have been identified worldwide (Jaeken et al., 1997; Van Schaftingen & Jaeken, 1995). It has a very wide phenotypic spectrum, with a prevalent involvement of the central nervous system (Grünewald, 2009; Monin et al., 2014; Serrano et al., 2015).

PMM2-CDG is caused by mutations in the gene *PMM2*, encoding the enzyme phosphomannomutase-2. Cytogenetic location of PMM2 is 16p13.2 (Bjursell et al., 1998) and PMM2-CDG is transmitted as autosomal recessive. A total absence of PMM2 activity is not compatible with life. As for the other CDGs, PMM2-CDG has a low rate of homozygosity (Kane et al., 2016); often, patients are compound heterozygotes with at least one hypomorphic mutation (Grünewald, 2009).

PMM2 is a 28 kDa protein and its active form is dimeric. It acts as a phosphomannomutase, thus it catalyzes the inter-conversion of mannose-6-phosphate (Man-6-P) and mannose-1-phosphate (Man-1-P), as shown in Figure 3.1.

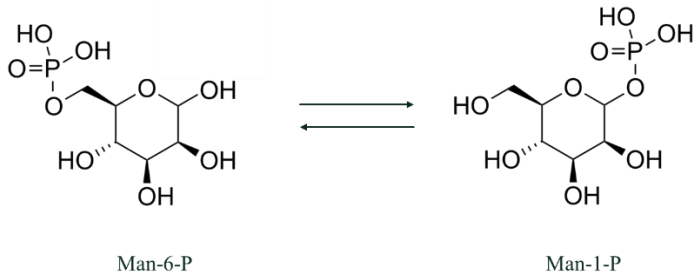

Figure 3.1: reaction catalyzed by phosphomannomutase-2

Man-1-P is the precursor to GDP-mannose, which is essential for the $Man_5GlcNAc_2$-PP-dolichol (Aebi, 2013; Maeda & Kinoshita, 2008).

The mechanism of action of PMM2 requires the activation by a bisphosphate sugar, either mannose-1,6-bisphosphate (Man-1,6-P) or glucose-1,6-bisphosphate (Glc-1,6-P).

Phosphomannomutase-2 has a paralogous enzyme in humans, which is phosphomannomutase-1 (PMM1). They share 63% of identity and derive from a duplication event that happened very early in vertebrate evolution (Quental et al., 2010).

Both the enzymes have phosphomannomutase activity, but PMM1 has also phosphatase activity on bisphosphate sugars. In particular, phosphatase activity is strongly enhanced in the presence of inosine monophosphate (IMP) [42]. No syndrome is associated with PMM1 mutations. In humans, *PMM1* is expressed in many tissues, especially in the heart, brain, and liver; the tissue distribution of

phosphomannomutase expression is reminiscent of the pathology in patients with PMM2-CDG syndrome (Matthijs et al., 1997).

Given these information and associated with the lack of compensation by wt-PMM1 in PMM2-CDG patients, one hypothesis has been advanced, that PMM1 could even counteract the low PMM2 activity in patients by hydrolyzing its bisphosphate activators.

Since the CNS is prevalently involved in the disease, PMM2-CDG cannot benefit from Enzyme Replacement Therapy. Therapies with small molecules would be ideal, as far as they are proved to cross the blood-brain barrier.

According to the hypothesis regarding the role of PMM1, the use of bisphosphate sugars as PCs would be useful to partially restore PMM2 activity in cells. For these reasons, we explored the possibility of using the β-anomer of glucose-1,6-bisphosphate for a therapeutic approach, to exploit its stabilizing properties and avoid the hydrolysis by PMM1.

3.7 **Fabry disease: what if pharmacological chaperones are not enough?**

Fabry-Anderson disease is a lysosomal storage disease caused by mutations in the gene *GLA*, located at Xq22.1, firstly described in 1898 by a German dermatologist (Kint, 1970). Fabry disease has a large phenotypic and genotypic spectrum, ranging from severe to mild cases, whose onset is during adulthood. More than 700 variants have been reported for *GLA* so far, with no prevalent mutations. Most of them are missense mutations occurring in the hydrophobic core, thus causing a destabilization.

GLA encodes alpha-galactosidase A (AGAL), a homodimeric enzyme that hydrolyses the terminal galactose from glycosphingolipids; its main substrate is globotriaosylceramide (Gb3, figure 3.2). The mature lysosomal form is 46 kDa, but it is synthesized as a 50 kDa precursor (ER and Golgi forms) (Lemansky et al., 1987). Patients affected by Fabry disease face a Gb3 accumulation in lysosomes. Anyway, it has been demonstrated that lyso-Gb3, the deacylated form of Gb3, is a good biomarker for the disease (Nowak et al., 2017).

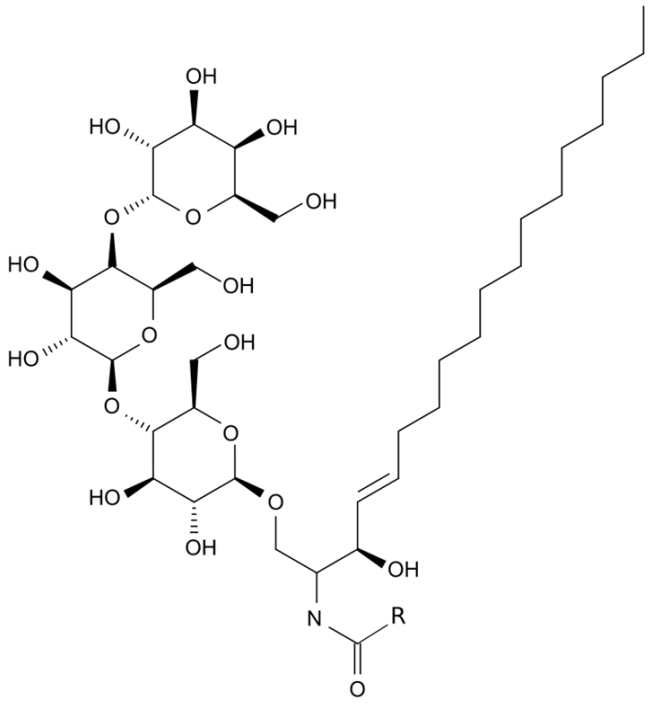

Figure 3.2: globotriaosylceramide (Gb3), the substrate for AGAL

Differently from what was discussed above for PMM2-CDG, Fabry disease has currently different therapeutic options. Mainly, patients can be treated with Enzyme Replacement Therapy (ERT), thanks to the lack of involvement of CNS, or with a pharmacological chaperone. Two different formulations of the recombinant enzyme are available, namely Fabrazyme® from Sanofi and Replagal® from Shire Pharmaceuticals. The approved PC is commercialized as Galafold® from Amicus Therapeutics; it is 1-deoxygalactonojirimycin (DGJ), an imino-sugar analogous of the substrate. It acts as a classical PC, binding the active site, stabilizing the protein but inhibiting its enzymatic activity. Patients are treated with an intermittent regimen, and the therapy would have two alternating phases. In the first one, *i.e.* in the presence of DGJ, AGAL would be stabilized but inhibited; in the second one, *i.e.* in the absence of DGJ, AGAL would be active but slowly degraded.

The response to DGJ is mutant-specific, thus the choice for the therapy requires a personalized analysis and must consider inter-assay variability on DGJ amenability (Cimmaruta et al., 2017; Citro, Cammisa, et al., 2016; Lukas et al., 2020). To help physicians in the choice of eligible patients, an online tool has been developed by

our research group, based on the comparison of data present in the literature (Cammisa et al., 2013).

The limitations to the use of PCs are well-known. Recently, the Food And Drug Administration (FDA) approved a combined therapy for cystic fibrosis (Ridley & Condren, 2020), that potentiates the beneficial effects of the drugs already used for the disease.

We advanced the hypothesis that a combined therapy approach would be useful in the case of Fabry disease, to potentiate DGJ usefulness and reduce its side effects. Such an approach had been proposed by our collaborator Dr. Jan Lukas and co-workers from the Department of Neurology of the University Medical Center Rostock (Seemann et al., 2020). The very innovative approach that we present is the possibility to modulate AGAL quantity inside cells combining DGJ with a very common use drug: acetylsalicylic acid (aspirin).

4 WHY DOES SARS-COV-2 HIT IN DIFFERENT WAYS? HOST GENETIC FACTORS CAN INFLUENCE THE ACQUISITION OR THE COURSE OF COVID-19

European Journal of Medical Genetics 64 (2021) 104227

Contents lists available at ScienceDirect

European Journal of Medical Genetics

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

Review

Why does SARS-CoV-2 hit in different ways? Host genetic factors can influence the acquisition or the course of COVID-19

Maria Monticelli [a], Bruno Hay Mele [b], Giuseppina Andreotti [c,**], Maria Vittoria Cubellis [a,c,*], Guglielmo Riccio [d]

[a] Department of Biology, Università Federico II, 80126, Napoli, Italy

[b] Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121, Napoli, Italy

[c] Istituto di Chimica Biomolecolare –CNR, 80078, Pozzuoli, Italy

[d] Scuola di Specializzazione in Pediatria, Università degli Studi di Trieste, 34127, Trieste, Italy

* Corresponding author. Department of Biology, Università Federico II, 80126, Napoli, Italy.

** Corresponding author.

E-mail addresses: maria.monticelli@unina.it (M. Monticelli), bruno.haymele@szn.it (B.H. Mele), gandreotti@icb.cnr.it (G. Andreotti), cubellis@unina.it (M.V. Cubellis), gugli.riccio@libero.it (G. Riccio).

ABSTRACT

The identification of high-risk factors for the infection by SARS-CoV-2 and the negative outcome of COVID-19 is crucial. The genetic background of the host might account for individual responses to SARS-CoV-2 infection besides age and comorbidities. A list of candidate polymorphisms is needed to drive targeted screens, given the existence of frequent polymorphisms in the general population.

We carried out text mining in the scientific literature to draw up a list of genes referable to the term "SARS- CoV*". We looked for frequent mutations that are likely to affect protein function in these genes. Ten genes, mostly involved in innate immunity, and thirteen common variants were identified, for some of these, the involvement in COVID-19 is supported by publicly available epidemiological data. We looked for available data on the population distribution of these variants and we demonstrated that the prevalence of five of them, Arg52Cys (rs5030737), Gly54Asp (rs1800450) and Gly57Glu (rs1800451) in *MBL2*, Ala59Thr (rs25680) in *CD27*, and Val197Met (rs12329760) in *TMPRSS2*, correlates with the number

of cases and/or deaths of COVID-19 observed in different countries. The association of the *TMPRSS2* variant provides epidemiological evidence of the usefulness of transmembrane protease serine 2 inhibitors for the cure of COVID-19. The identified genetic variants represent a basis for the design of a cost-effective assay for population screening of genetic risk factors in the COVID-19 pandemic.

Introduction

Many people affected by COVID-19 develop fever, cough, fatigue, shortness of breath, muscular pain, and headache. In the most severe cases medical conditions, such as respiratory failure, occur and eventually lead to death.

Age and comorbidities, such as cardiovascular diseases, diabetes, and chronic lung diseases, are known risk factors for worse outcomes (Fang et al., 2020). Furthermore, altered biomarkers, namely C-reactive Protein, high-sensitivity troponin T, prothrombin time, fibrin degradation products, D-Dimer, and CD4+ count, help differentiate patients who will develop severe symptoms from those who will only be mildly affected (Zheng et al., 2020). These factors alone are not sufficient to identify all subjects at high risk, and on top of that, they tend to overlap. Genetic factors, on the other hand, also influence risks associated with COVID-19 but are more likely to be independent of age-related comorbidities. Specific studies on the subject are very few and in most cases gene+*s have been selected regardless of the existence of frequent polymorphisms in the general population, lowering their potential as markers for large-scale genetic screenings (Hussain et al., 2020).

SARS-CoV-2 and SARS-CoV belong to the Betacoronavirus genus (Chen et al., 2020) and share the same cellular receptor, the angiotensin-converting enzyme 2 (ACE2), and very similar nucleotide sequences (Yan et al., 2020).

In this paper, we carried out a bioinformatic analysis to predict polymorphisms that could either influence the acquisition or the course of COVID-19. Text mining was carried out to extract a list of human genes that have been associated with the term "SARS-CoV*" in the literature. We highlighted the polymorphisms of these genes that are very frequent in the general population and are likely to affect the function of the mutated protein. We hope that our prediction can guide and speed up

experimental tests and help to find those high-risk subjects that have not been recognized by age, co-morbidities, or biomarkers so far.

Question: which host genetic factors are likely to influence sensitivity to COVID-19?

High variability of symptoms and outcomes characterizes SARS-CoV-2 infection. The phenotypic spectrum of the disease ranges from asymptomatic individuals to patients requiring ventilation due to severe pneumonia (García, 2020). Different hypotheses have been put forward, the role of the host's genetic background is currently emerging (Godri Pollitt et al., 2020; Latini et al., 2020; Strafella et al., 2020; Zhang, S.-Y. et al., 2020). Data from a genome-wide association (GWAS) study involving 1980 patients with COVID-19 and 2381 healthy control participants in Spain and Italy have been recently published (Group, 2020). Two loci were strongly associated with susceptibility to SARS-CoV-2, one on chromosome 9 covering the ABO blood group and another one on chromosome 3 covering six genes, *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, and *XCR1*.

We would like to contribute to gene-specific candidate-driven studies by suggesting missense variants that can cause severe symptoms or facilitate the infection and expand the list identified by GWAS studies. We evaluated missense mutations that are frequent in the general population and occur in genes that do not cause haploinsufficiency. We selected those that are deleterious for the protein product, and we excluded those associated with blood groups. A deleterious mutation on a single allele might not be enough to cause a life-threatening disease and hence it could be relatively frequent but might confer special sensitivity or resistance to viral infection. We excluded variants occurring in promoters, introns, or untranslated regions because their deleteriousness is less predictable. Hence our study will provide a subset of the genetic variants that are likely to affect the sensitivity to infection or the course of COVID-19.

Procedure to identify candidate variants

We obtained from the EXome Aggregation Consortium (EXAC) a list of human genes (14995) with a pLI <0.9. pLI is a value that measures the probability of being

loss-of-function intolerant (pLI). Genes with a pLI<0.9 have a low probability of causing haploinsufficiency when affected by the severe mutation. We developed a simple pipeline to use such a list of genes and corresponding proteins as input and look within the Scopus database for papers containing any of these protein names and the term "SARS-CoV*" in the article's title, abstract, or keywords. The script, which is written in R language and exploits the rscopus package, is provided in Supplementary File S1. We excluded the genes covered by the GWAS analysis (Group, 2020) and obtained a "SARS-CoV* referable genes" list (106 genes; Supplementary File S2).

We searched within GnomAD for frequent variants, missense or loss- of-function, in "SARS-CoV* referable genes", setting the threshold to allele frequency at 1%. Lastly, we used Polyhen2 (http://genetics.bwh. harvard.edu/pph2/), to exclude benign mutations. Frequent missense variants were found in 56 genes and frequent deleterious variants in 10 genes. The scheme of the search is provided in Fig. 4.1.

The list of identified variants with their frequencies is reported in Table 4.1. We observed that the candidate variants were mostly in genes involved in innate immune defense. References to databases and programs utilized in this analysis are provide as a supplementary file S3.

Correlation between allele frequencies and COVID-19 cases and deaths per population

Data on the prevalence in different populations were available for six of our candidate variants. The datasets relative to rs12329760 (Val197Met, *TMPRSS2*), rs25680 (Ala59Thr, *CD27*), and rs3775291 (Leu412Phe, *TLR3*) were found in ALFRED; the data relative to the *MBL2* variants rs5030737 (Arg52Cys), rs1800450 (Gly54Asp), and rs1800451 (Gly57Glu) were reported by Verdu et al. Data on the population distribution of the remaining six variants were not available. Notwithstanding the limitations associated with publicly available data on the number of cases and deaths per country that are influenced by socio-economic factors, we found interesting results for five variants, namely Arg52Cys (rs5030737), Gly54Asp (rs1800450), and Gly57Glu (rs1800451) in *MBL2*, Val197Met (rs12329760) in *TMPRSS2* and Ala59Thr (rs25680) in *CD27*.

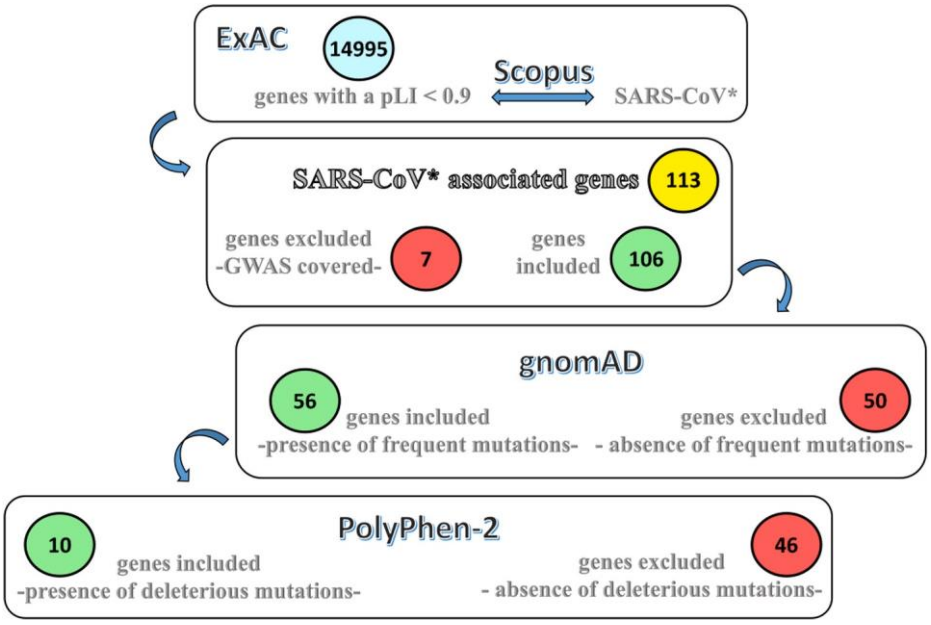

Fig. 4.1. Procedure to identify candidate variants. The EXome Aggregation Consortium (EXAC) provided a list of human genes and annotated them with a value that measures the probability of being loss-of- function intolerant (pLI). The names of the genes were associated to the names of the corresponding proteins. Such list was used as the input to look within the Scopus database for papers containing any of the protein names and the term "SARS-CoV*" in the article's title, abstract or keywords. "SARS- CoV referable genes", encompassing frequent variants, missense or loss of function, were identified setting the threshold to allele frequency to 1%, and excluding variants that did not pass quality control or occur in non-canonical transcripts. The genes covered by the GWAS analysis (Group, 2020) were excluded. Among these genes, those encompassing frequent deleterious variants were identified.

Table 4.1

Missense variants that might influence the acquisition or the course of COVID-19.

Gene name	Protein name/Mim phenotype	Reference SNP ID number	Amino acid change	Variant frequency	GWAS *p*-value
ACE	Angiotensin-converting enzyme	rs3730025	Tyr244Cys	0.0107	
AHSG	Alpha-2-HS-glycoprotein #203650	rs140827890	Ala164Thr	0.0114	0.05
CD27	CD27 antigen #615122	rs25680	Ala59Thr	0.1912	
GFAP	Glial fibrillary acidic protein #203450	rs59291670	Asp157Asn	0.11	
IFI30	Gamma-interferon-inducible lysosomal thiol reductase	rs11554159	Arg76Gln	0.2281	
MBL2	Mannose-binding protein C #614372	rs1800450	Gly54Asp	0.141	
		rs5030737	Arg52Cys	0.0558	
		rs1800451	Gly57Glu	0.0318	
MEFV	Pyrin #134610 #249100	rs3743930	Glu148Gln	0.0708	
TLR2	Toll-like receptor 2 #607948	rs5743704	Pro631His	0.0281	0.02
		rs5743708	Arg753Gln	0.0176	
TLR3	Toll-like receptor 3 #609423	rs3775291	Leu412Phe	0.2789	0.03
TMPRSS2	Transmembrane protease serine 2	rs12329760	Val197Met	0.2452	0.01528

Table 4.1: Likely deleterious frequent variants (allele frequency >1%) in genes referable to the term "SARS-CoV*" in the literature. Variant frequency is referred to gnomAD total exome frequency obtained from wAnnovar annotation. For variants covered in the GWAS study, GCST90000255 Fisher *p*-values are provided.

MBL2 encodes a mannose-binding protein C that binds mannose, fucose, and N-acetylglucosamine on different microorganisms and activates the lectin complement pathway. The deficiency of mannose- binding protein C (MIM

#614372) is associated with increased susceptibility to infection (Degn et al., 2011). MBL2 binds and activates the lectin complement pathway in viral infections such as HIV (Garred et al., 1997) and influenza (Kase et al., 1999). We tested three frequent missense mutations in *MBL2*: Gly54Asp (rs1800450), Arg52Cys (rs5030737), and Gly57Glu (rs1800451). All these mutations affect the ability of the protein to bind carbohydrate surfaces and MASP-2, the mannan-binding lectin serine protease 2 which activates the complement system (Turner, 2003); the lowest interaction is observed with Arg52Cys (Larsen et al., 2004). In a study conducted in China, it was proven that Gly54Asp (rs1800450) is significantly associated with susceptibility to SARS-CoV infection but not with disease severity (Ip et al., 2005; Zhang et al., 2005); unfortunately, the other two variants were not tested. We found that the allele frequency per population of Arg52Cys (rs5030737) correlates positively with the number of cases per country [*p*-value 0.0394] of COVID-19, but not with the number of deaths (Fig. 4.2, panels A and B), while the allele frequency per population of Gly54Asp (rs1800450) correlates positively both with the number of cases [*p*-value 0.0018] and the number of deaths [*p*-value 0.02] (Fig. 4.2, panels C and D). This finding supports the hypothesis that mannose-binding protein C plays a protective role and that MBL2 inactivation is a risk factor for SARS-CoV-2 infection, and that the immune response plays a role in the course of the disease. Unexpectedly, the allele frequency per population of Gly57Glu (rs1800451) was found to be negatively correlated with the number of cases [*p*-value 0.0122] (Fig. 4.3, panel A). The correlation with the number of deaths was significant when the data from Iran were excluded [*p*-value 0.0241] (Fig. 4.3, panel B).

TMPRSS2 encodes transmembrane protease serine 2. Mouse models have shown that SARS, MERS CoVs and Influenza viruses use this protease to cleave the Spike protein and achieve viral internalization (Iwata-Yoshikawa et al., 2019). A negative correlation between the allele frequency of Val197Met and the number of cases [*p*-value 0.0012] and deaths [*p*-value 0.0057] per country (Fig. 4.4 panels A and B) was found. We observed the number of cases and deaths as a function of prevalence clusters in two groups and the high mortality group is associated with a low prevalence of the variant, indicating that genetic *TMPRSS2* inactivation is a protective factor against SARS-CoV-2 infection and progression towards severe

cases. A lower frequency of Val197Met in Italian COVID-19 patients compared to the allelic frequency of this variant in GnomAD for the EUR reference population has recently been reported (Latini et al., 2020). Val197Met is a SNP covered by the GWAS study (Group, 2020) and it is significantly associated with susceptibility to (p-value < 0.05). TMPRSS2 is a druggable protein that can be inhibited by camostat mesylate, a drug approved in Japan, by bromhexine and nafamostat. Clinical trials are ongoing for their usage in COVID-19 (ClinicalTrials.gov Bromexhine: NCT04355026, NCT04273763, NCT04340349; Camostat: NCT04321096, NCT04355052, NCT04353284, NCT04338906, NCT04470544; Nafamostat: NCT04352400; NCT04473053). Our data provide epidemiologic evidence of the role of the protease in SARS-CoV-2 infection and encourages the usage of protease inhibitors for the therapy (Hoffmann et al., 2020).

CD27 receptor is thought to have an essential role in connecting the CD4 and CD8 branches of the immune system. Unbalances between the two branches has been associated with a worse prognosis in SARS-CoV infections (Li et al., 2008). Known pathological mutations of CD27 exist, and they are associated with cytokine storms (Amberger et al., 2009) and hemophagocytic lymphohistiocytosis (HLH) like syndromes that frequently characterize severe SARS-CoV-2 infections (Shoenfeld, 2020). We observed a statistically significant negative correlation between the allele frequency of Ala59Thr and the number of deaths per country [p-value 0.0337] (Fig. 4.5, panel B), but not with the numbers of cases (Fig. 4.5 panel A).

Description of the other genes potentially related to COVID-19 response

As discussed above, the association between six inactivating variants (Table 4.1) and SARS-CoV-2 susceptibility could not be demonstrated because of the absence of data on population distribution. Nevertheless, literature data support a possible role of some of them in COVID-19.

The role of ACE in innate immunity has been reviewed (Bernstein et al., 2018). A very frequent intronic insertion/deletion I/D polymorphism influences ACE expression (Rigat et al., 1990). A recent paper has suggested the usage of ACE I/D polymorphisms for the identification of high-risk COVID-19 patients (Delanghe et al., 2020). The hypothesis that variants of ACE are associated with COVID-19 risks

is very interesting and is in line with several observations. Although ACE2, rather than ACE, has been identified to be one of the receptors of the SARS-CoV-2, it has been proposed that the unbalance between ACE and ACE2 expression leads to lung injury (Guo et al., 2020). Children, who have higher levels of ACE than adults, are much less affected by COVID-19 (Guo et al., 2020). ACE inhibitors have been proven to upregulate ACE2, and some authors speculate a possible role of these extremely commonly used drugs in infection.

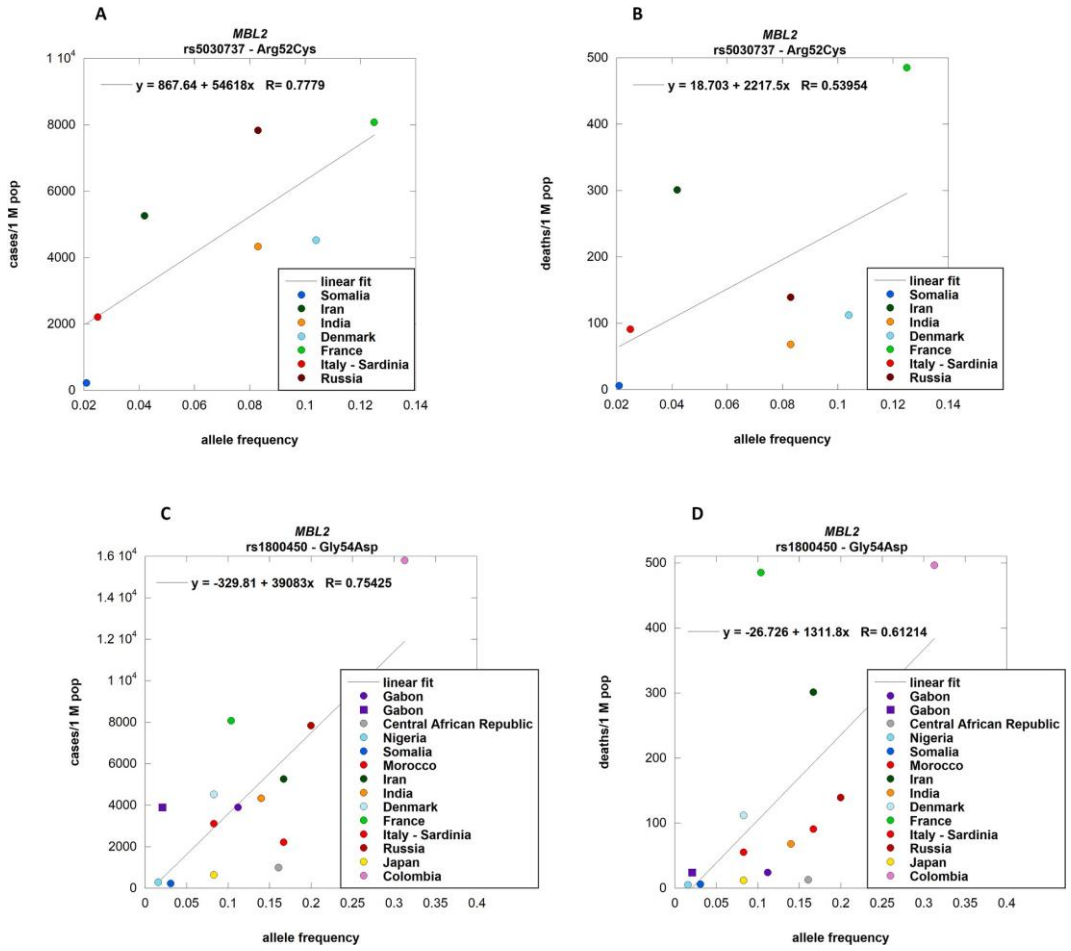

Fig. 4.2. *MBL2* allele frequencies per population correlation with COVID-19 cases and deaths. rs5030737 (Arg52Cys) allele frequencies in different countries from Verdu et al. were correlated with the respective number of cases (panel A; *p*-value 0.0394) and deaths (panel B; n. s.) of COVID-19. rs1800450 (Gly54Asp) allele frequencies were correlated both with cases (panel C; *p*-value 0.0018) and deaths (panel D; *p*-value 0.0199) of COVID-19 per country. Different symbols relate to different populations associated with the same country [Purple circle: GBP; purple square: GBB.] Details about populations and their association to countries are provided in supplementary file S1.

(For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Their role in COVID-19 disease, however, is very controversial, because upregulation of *ACE2* may have a protective role against lung inflammation and fibrosis (Guo et al., 2020). Publications on ACE I/D polymorphisms and SARS-CoV are of little help because they provide different results (Chan et al., 2005; Itoyama et al., 2004). A final word on this subject could come from a very recent paper by Richards and coworkers (Butler-Laporte et al., 2020). Although the authors did not include Tyr244Cys (rs3730025) in their study, they found that variants which decrease the expression of *ACE* do not increase COVID-19 susceptibility or severity.

AHSG encodes a protein known as human fetuin which is thought to be an immune modulator, mainly acting on cytokine release from macrophages. Macrophages play a central role in the IL1, 6, TNF alfa, INF gamma dependent cytokine storm. Its deletion has already been linked to worse outcomes in SARS-CoV infected patients (Zhu et al., 2011). The SNP Ala164Thr is covered by the GWAS study (Group, 2020) and significantly associates with susceptibility to COVID-19 (*p*-value 0.05). The role of *GFAP* in COVID-19 might be prevalently associated to the severity of the disease and in particular to the development of neuro- logical symptoms. *GFAP* is a TGF-beta induced gene that participates in the ERG dependent TGF-beta triggering of fibrosis consequent to ROS production. It has been proven that SARS-CoV viruses possess a deubiquitinating protein named PLpro that stimulates such pathway, inducing lung fibrosis in infected patients. Pulmonary fibrosis has been described in cases of SARS-CoV infections, but so far it has not been described to be a relevant feature of the disease (Yang et al., 2006). Very little is known so far about COVID-19 pathology, and future studies could demonstrate the importance of fibrosis since it is a central feature in uncontrolled activation of macrophage-mediated inflammation.

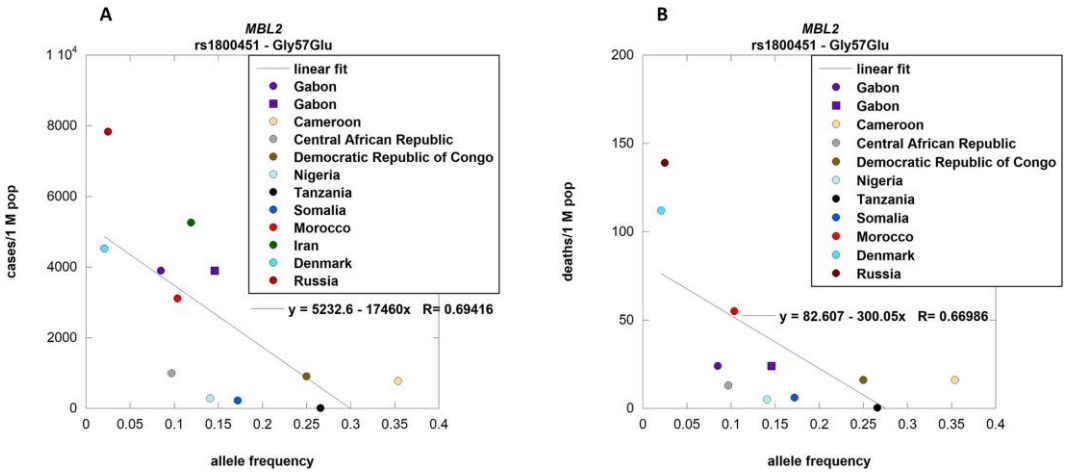

Fig. 4.3. *MBL2* allele frequencies per population correlation with COVID-19 cases and deaths. rs1800451 (Gly57Glu) allele frequencies were correlated with cases (panel A; *p*-value 0.0122) and deaths (panel B; *p*-value 0.0241) of COVID-19 per country. The outsider Iran was excluded to observe a significant correlation with the number of deaths. Different symbols relate to different populations associated with the same country [Purple circle: GBP; purple square: GBB.] Details about populations and their association to countries are provided in supplementary file S1. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

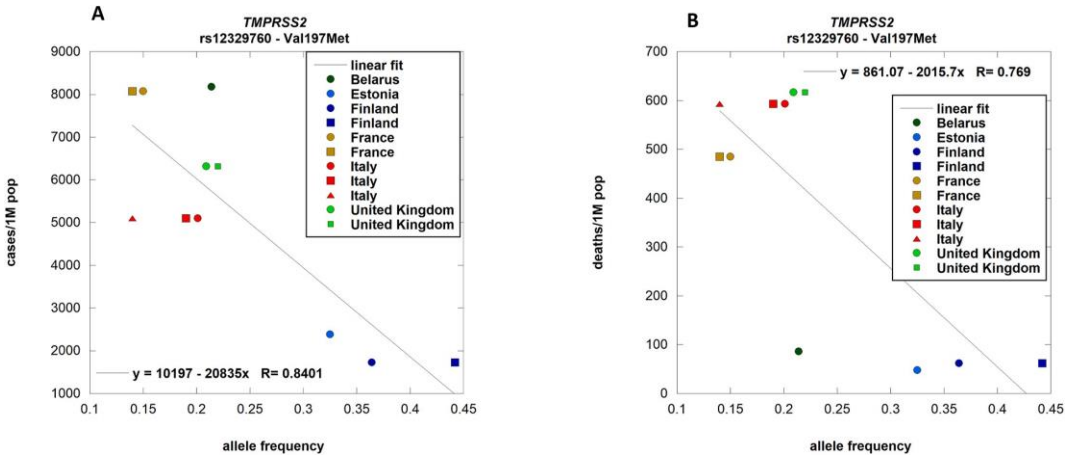

Fig. 4.4. *TMPRSS2* allele frequencies per population correlation with COVID-19 cases and deaths.

rs12329760 (Val197Met) allele frequencies in different countries from ALFRED were correlated with the respective number of cases (panel A; *p*-value 0.0012) and deaths (panel B; *p*-value 0.0057) of COVID-19. Different symbols relate to different populations associated with the same country [Blu circle: SA004049R; blu square: SA004617S. Light brown circle: SA001504K; light brown

square: SA001503J. Red circle: SA004057Q; red square: SA002255O; red triangle SA001505L. Light green circle: SA004050J; light green square: SA001508O.] Details about populations and their association to countries are provided in supplementary file S1. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

IFI30 (Gamma-interferon-inducible lysosomal thiol reductase, GILT) has a role in protecting against the internalization of SARS-CoV viruses in the lung tissue (Chen, D. et al., 2019). Polymorphisms in this gene may be correlated with a predisposition to infection by SARS-CoV-2, as well as with a worse prognosis due to higher initial viral charge. The viral charge has, in fact, already been proven to be linked to worse clinical outcomes (Chen, D. et al., 2019).

MEFV (Pyrin) is associated to Familial Mediterranean Fever whose symptoms, fever, abdominal/chest pain, elevation of C-reactive protein, and leucocytosis, overlap in part with those of COVID-19. Deleterious variants in this gene might influence the severity of the infection. *MEFV* is thought to be involved in the NLRP3 dependant IL-1 beta modulation (Manukyan and Aminov, 2016). The capability of SARS-CoV viruses to activate the NLRP3 inflammasome has already been described (Chen, I.-Y. et al., 2019). There are some known pathological mutations of *MEFV* that cause autoinflammatory disease due to a IL1, 6, TNF- alfa INF-gamma cytokine storm (Manukyan and Aminov, 2016).

Toll like receptors activate Type-I interferon (alpha and beta) response that enhance innate antiviral responses such as the limitation of cell replication, the induction of apoptosis, and the synthesis of enzymes with direct antiviral effect (Kawasaki and Kawai, 2014). Some viruses have been described to elude the Type I interferon response by shifting the reply towards the Type II (gamma) interferon branch, which favors viral diffusion through tissues by activating a cytokine storm (McGonagle et al., 2020).

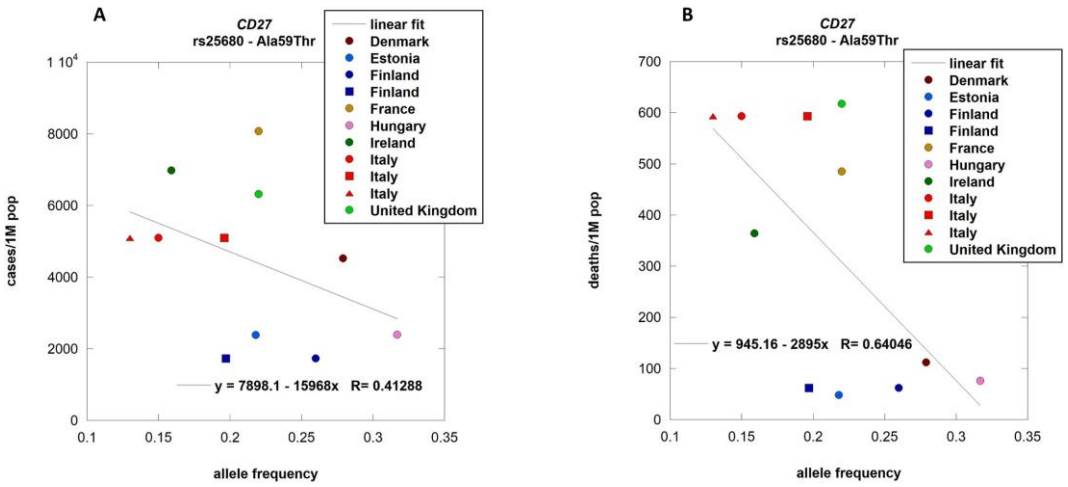

Fig. 4.5. *CD27* allele frequencies per population correlation with COVID-19 cases and deaths. rs25680 (Ala59Thr) allele frequencies in different countries from ALFRED were correlated with the respective number of cases (panel A; n.s.) and deaths (panel B; *p*-value 0.0337) of COVID-19. Different colors relate to different countries while different symbols relate to different populations associated with the same country [Blue circle: SA004377V; blue square: SA004049R. Red circle: SA002255O; red square: SA004057Q; red triangle SA001505L.] Details about populations and their association to countries are provided in supplementary file S1. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Toll-like receptor 2 interacts with TICAM2 (Stack et al., 2014) that in turn contributes to SARS-CoV pathogenesis in mice (Gralinski et al., 2017). Recent papers on *TLR2* have linked the function of this gene to viral as well as to the already known bacterial targets (Carty and Bowie, 2010). TLR2 might bind SARS-CoV-2 proteins and have a key role in directing the immune response to this pathogen. Arg753Gln (rs5743708) and Pro631His (rs5743704) in TLR2 are relatively frequent in the European population, approximately 3%. A few papers confirm the role of Arg753Gln (rs5743708) in respiratory diseases (Patarˇci´c et al., 2015) (Smelaya et al., 2016). The effects of Arg753 substitution by glutamine on TLR2 were proven *in vitro*. The mutation impairs agonist-induced phosphorylation hetero-dimerization with TLR6, and recruitment of myeloid differentiation primary response protein 88 (MYD88) and MyD88 adapter-like (MAL) (Xiong et al., 2012). The SNP

Arg753Gln (rs5743708) is covered in the GWAS study (Group, 2020) and it is associated with COVID-19 significantly (*p*-value < 0.05). The role in the susceptibility to viral infection is confirmed by the finding that intranasal administration of a TLR2/6 agonist reduces upper respiratory tract viral shedding in an animal model (Proud et al., 2020).

Toll-like receptor 3 is known to be involved in the innate immune response to SARS-CoV infections (Totura et al., 2015). Its role in the new epidemics is supported by the fact that rare loss-of-function variants were found among male patients with life-threatening COVID-19 pneumonia more frequently than in mildly affected controls (Zhang, Q. et al., 2020). One frequent missense mutation in *TLR3*, Leu412Phe (rs3775291) was associated with several viral diseases (El-Bendary et al., 2018; Fischer et al., 2018; Grygorczuk et al., 2017; Studzin´ska et al., 2017). It has an allele frequency ranging from 25% to 35% worldwide except for African populations, where it is 6%. Although it is necessary to be extremely cautious when considering the causes of the low numbers of reported cases from Africa, nonetheless it is tempting to associate it to the low incidence of rs3775291. Another indication in favour of the role of rs3775291 in the susceptibility to SARS-CoV-2 comes from the GWAS study (Group, 2020) where a significative difference (*p*-value < 0.03) was observed between affected people and healthy controls. When we analysed the epidemiological data of general European populations, as we did for variants in *TMPRSS2* (Fig. 4.4) or *CD27* (Fig. 4.5), we did not find significant correlation (data not shown), but it would be interesting to look in a cohort of males.

Four deleterious frequent variation were found in the genes identified by GWAS analysis (Group, 2020), although these SNPs were not covered by the study.

Thr199Met in the gene *SLC6A20* encoding Sodium- and chloride- dependent transporter XTRP3 and Asp182Asn in the gene *LZTFL1* encoding Leucine zipper transcription factor-like protein 1 are relatively frequent in the European population but not in African and East Asian populations.

GluE994Lys and Arg1140Gln are in the gene *FYCO1* encoding FYVE and coiled-coil domain-containing protein 1 that participate in the transportation and maturation of autophagosomes.

Conclusion

COVID-19 pandemic upset the world. The fast spread of SARS-CoV-2 through over the world and some peculiar novelties of the virus, such as the high number of asymptomatic individuals, rendered it a difficult challenge from different aspects. A strong effort was put in the discovery of risk factors associated with a more serious phenotype. Besides this, different countries experienced different impacts caused by SARS-CoV-2. As well described by Yamamoto and Bauer (2020), many factors contribute to these differences, e.g. socio-behavioral habits, political management, co-existence with other viruses, genetic factors. Genetic factors can play a major role in defining the outcome of the infection. This correlation is hard to prove due to all the side-factors described above, that certainly influence the official data from the different countries. Moreover, the globalized world we live in certainly complicates the association between population variants and countries.

Nevertheless, our data showed that a statistically significant correlation could be proved for some common variants in three genes, namely *MBL2*, *TMPRSS2*, and *CD27*, with the number of cases or deaths observed per country.

These data suggest that the genetic background is of utmost importance in the evaluation of COVID-19 susceptibility and that the discovery of target genes could be useful in the treatment and prevention of the infection. Moreover, the evaluation of the role of these polymorphisms might be necessary for the management of very common drugs during a pandemic.

Acknowledgements

This work is dedicated to our friend and colleague Maria Malanga.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmg.2021.104227.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Amberger, J., Bocchini, C.A., Scott, A.F., Hamosh, A., 2009. McKusick's online Mendelian inheritance in man (OMIM®). Nucleic Acids Res. 37 (Suppl. l_1), D793–D796.

Bernstein, K.E., Khan, Z., Giani, J.F., Cao, D.-Y., Bernstein, E.A., Shen, X.Z., 2018.

Angiotensin-converting enzyme in innate and adaptive immunity. Nat. Rev. Nephrol. 14 (5), 325.

Butler-Laporte, G., Nakanishi, T., Mooser, V., Renieri, A., Amitrano, S., Zhou, S., Chen, Y., Forgetta, V., Richards, J.B., 2020. The effect of angiotensin-converting enzyme levels on Covid-19 susceptibility and severity: a Mendelian randomization study. Int. J. Epidemiol.

Carty, M., Bowie, A.G., 2010. Recent insights into the role of Toll-like receptors in viral infection. Clin. EXp. Immunol. 161 (3), 397–406.

Chan, K.A., Tang, N.L., Hui, D.S., Chung, G.T., Wu, A.K., Chim, S.S., Chiu, R.W., Lee, N.,

Choi, K., Sung, Y.M., 2005. Absence of association between angiotensin converting enzyme polymorphism and development of adult respiratory distress syndrome in patients with severe acute respiratory syndrome: a case control study. BMC Infect. Dis. 5 (1), 26.

Chen, D., Hou, Z., Jiang, D., Zheng, M., Li, G., Zhang, Y., Li, R., Lin, H., Chang, J., Zeng, H., 2019a. GILT restricts the cellular entry mediated by the envelope glycoproteins of SARS-CoV, Ebola virus and Lassa fever virus. Emerg. Microb. Infect. 8 (1), 1511–1523.

Chen, I.-Y., Moriyama, M., Chang, M.-F., Ichinohe, T., 2019b. Severe acute respiratory syndrome coronavirus viroporin 3a activates the NLRP3 inflammasome. Front.

Microbiol. 10, 50.

Chen, Y., Liu, Q., Guo, D., 2020. Emerging coronaviruses: genome structure, replication, and pathogenesis. J. Med. Virol. 92 (4), 418–423.

Degn, S.E., Jensenius, J.C., Thiel, S., 2011. Disease-causing mutations in genes of the complement system. Am. J. Hum. Genet. 88 (6), 689–705.

Delanghe, J., Speeckaert, M., De Buyzere, M., 2020. The host's angiotensin-converting enzyme polymorphism may explain epidemiological findings in COVID-19 infections. Clin. Chim. Acta 505, 192–193.

El-Bendary, M., Neamatallah, M., Elalfy, H., Besheer, T., Elkholi, A., El-Diasty, M., Elsareef, M., Zahran, M., El-Aarag, B., Gomaa, A., 2018. The association of single nucleotide polymorphisms of Toll-like receptor 3, Toll-like receptor 7 and Toll-like receptor 8 genes with the susceptibility to HCV infection. Br. J. Biomed. Sci. 75 (4), 175–181.

Fang, L., Karakiulakis, G., Roth, M., 2020. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? The Lancet. Respir. Med. 8 (4), e21.

Fischer, J., Koukoulioti, E., Schott, E., Fülo¨p, B., Heyne, R., Berg, T., van Bo¨mmel, F., 2018. Polymorphisms in the toll-like receptor 3 (TLR3) gene are associated with the natural course of hepatitis B virus infection in Caucasian population. Sci. Rep. 8 (1), 1–8.

García, L.F., 2020. Immune response, inflammation, and the clinical spectrum of COVID-19. Front. Immunol. 11, 1441.

Garred, P., Madsen, H.O., Balslev, U., Hofmann, B., Pedersen, C., Gerstoft, J., Svejgaard, A., 1997. Susceptibility to HIV infection and progression of AIDS in relation to variant alleles of mannose-binding lectin. Lancet 349 (9047), 236–240.

Godri Pollitt, K.J., Peccia, J., Ko, A.I., Kaminski, N., Dela Cruz, C.S., Nebert, D.W., Reichardt, J.K., Thompson, D.C., Vasiliou, V., 2020. COVID-19 vulnerability: the potential impact of genetic susceptibility and airborne transmission. Hum. Genom. 14, 1–7.

Gralinski, L.E., Menachery, V.D., Morgan, A.P., Totura, A.L., Beall, A., Kocher, J., Plante, J., Harrison-Shostak, D.C., Scha¨fer, A., de Villena, F.P.-M., 2017. Allelic variation in the toll-like receptor adaptor protein ticam2 contributes to

SARS- coronavirus pathogenesis in mice. G3: Genes, Genomes, Genetics 7 (6), 1653–1663.

Group, S.C.-G., 2020. Genomewide association study of severe Covid-19 with respiratory failure. N. Engl. J. Med. 383 (16), 1522–1534.

Grygorczuk, S., Parczewski, M., S´wierzbin´ska, R., Czupryna, P., Moniuszko, A., Dunaj, J.,

Kondrusik, M., Pancewicz, S., 2017. The increased concentration of macrophage migration inhibitory factor in serum and cerebrospinal fluid of patients with tick-borne encephalitis. J. Neuroinflammation 14 (1), 126.

Guo, J., Huang, Z., Lin, L., Lv, J., 2020. Coronavirus disease 2019 (covid-19) and cardiovascular disease: a viewpoint on the potential influence of angiotensin-converting enzyme inhibitors/angiotensin receptor blockers on onset and severity of severe acute respiratory syndrome coronavirus 2 infection. Journal of the American Heart Association 9 (7), e016219.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.-H., Nitsche, A., 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell.

Hussain, M., Jabeen, N., Raza, F., Shabbir, S., Baig, A.A., Amanullah, A., Aziz, B., 2020. Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. J. Med. Virol.

Ip, W.E., Chan, K.H., Law, H.K., Tso, G.H., Kong, E.K., Wong, W.H., To, Y.F., Yung, R.W.,

Chow, E.Y., Au, K.L., 2005. Mannose-binding lectin in severe acute respiratory syndrome coronavirus infection. JID (J. Infect. Dis.) 191 (10), 1697–1704.

Itoyama, S., Keicho, N., Quy, T., Phi, N.C., Long, H.T., Van Ban, V., Ohashi, J., Hijikata, M., Matsushita, I., Kawana, A., 2004. ACE1 polymorphism and progression of SARS. Biochem. Biophys. Res. Commun. 323 (3), 1124–1129.

Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., Nagata, N., 2019. TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. J. Virol. 93 (6).

Kase, T., Suzuki, Y., Kawai, T., Sakamoto, T., Ohtani, K., Eda, S., Maeda, A., Okuno, Y., Kurimura, T., Wakamiya, N., 1999. Human mannan-binding lectin inhibits the infection of influenza A virus without complement. Immunology 97 (3), 385.

Kawasaki, T., Kawai, T., 2014. Toll-like receptor signaling pathways. Front. Immunol. 5, 461.

Larsen, F., Madsen, H.O., Sim, R.B., Koch, C., Garred, P., 2004. Disease-associated mutations in human mannose-binding lectin compromise oligomerization and activity of the final protein. J. Biol. Chem. 279 (20), 21302–21311.

Latini, A., Agolini, E., Novelli, A., Borgiani, P., Giannini, R., Gravina, P., Smarrazzo, A., Dauri, M., Andreoni, M., Rogliani, P., 2020. COVID-19 and genetic variants of protein involved in the SARS-CoV-2 entry into the host cells. Genes 11 (9), 1010.

Li, C.K.-f., Wu, H., Yan, H., Ma, S., Wang, L., Zhang, M., Tang, X., Temperton, N.J., Weiss, R.A., Brenchley, J.M., 2008. T cell responses to whole SARS coronavirus in humans. J. Immunol. 181 (8), 5490–5500.

Manukyan, G., Aminov, R., 2016. Update on pyrin functions and mechanisms of familial Mediterranean fever. Front. Microbiol. 7, 456.

McGonagle, D., Sharif, K., O'Regan, A., Bridgewood, C., 2020. Interleukin-6 use in COVID-19 pneumonia related macrophage activation syndrome. Autoimmun. Rev. 102537.

Patarˇci´c, I., Gelemanovi´c, A., Kirin, M., Kolˇci´c, I., Theodoratou, E., Baillie, K.J., De Jong, M.D., Rudan, I., Campbell, H., Polaˇsek, O., 2015. The role of host genetic factors in respiratory tract infectious diseases: systematic review, meta-analyses and field synopsis. Sci. Rep. 5, 16119.

Proud, P.C., Tsitoura, D., Watson, R.J., Chua, B.Y., Aram, M.J., Bewley, K.R., Cavell, B.E., Cobb, R., Dowall, S., Fotheringham, S.A., 2020. Prophylactic intranasal administration of a TLR2/6 agonist reduces upper respiratory tract viral shedding in a SARS-CoV-2 challenge ferret model. EBioMedicine 63, 103153.

Rigat, B., Hubert, C., Alhenc-Gelas, F., Cambien, F., Corvol, P., Soubrier, F., 1990. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene

accounting for half the variance of serum enzyme levels. J. Clin. Invest. 86 (4), 1343–1346.

Shoenfeld, Y., 2020. Corona (COVID-19) time musings: our involvement in COVID-19 pathogenesis, diagnosis, treatment and vaccine planning. Autoimmun. Rev.

Smelaya, T.V., Belopolskaya, O.B., Smirnova, S.V., Kuzovlev, A.N., Moroz, V.V., Golubev, A.M., Pabalan, N.A., Salnikova, L.E., 2016. Genetic dissection of host immune response in pneumonia development and progression. Sci. Rep. 6, 35021.

Stack, J., Doyle, S.L., Connolly, D.J., Reinert, L.S., O'Keeffe, K.M., McLoughlin, R.M., Paludan, S.R., Bowie, A.G., 2014. TRAM is required for TLR2 endosomal signaling to type I IFN induction. J. Immunol. 193 (12), 6090–6102.

Strafella, C., Caputo, V., Termine, A., Barati, S., Gambardella, S., Borgiani, P., Caltagirone, C., Novelli, G., Giardina, E., Cascella, R., 2020. Analysis of ACE2 Genetic Variability Among Populations Highlights A Possible Link with COVID19- Related Neurological Complications.

Studzin´ska, M., Jabłon´ska, A., Wi´sniewska-Ligier, M., Nowakowska, D., Gaj, Z., Le´snikowski, Z.J., Wo´zniakowska-Gęsicka, T., Wilczyn´ski, J., Paradowska, E., 2017. Association of TLR3 L412F polymorphism with cytomegalovirus infection in children. PloS One 12 (1), e0169420.

Totura, A.L., Whitmore, A., Agnihothram, S., Sch¨afer, A., Katze, M.G., Heise, M.T., Baric, R.S., 2015. Toll-like receptor 3 signaling via TRIF contributes to a protective innate immune response to severe acute respiratory syndrome coronavirus infection. mBio 6 (3).

Turner, M.W., 2003. The role of mannose-binding lectin in health and disease. Mol. Immunol. 40 (7), 423–429.

Xiong, Y., Song, C., Snyder, G.A., Sundberg, E.J., Medvedev, A.E., 2012. R753Q polymorphism inhibits Toll-like receptor (TLR) 2 tyrosine phosphorylation, dimerization with TLR6, and recruitment of myeloid differentiation primary response protein 88. J. Biol. Chem. 287 (45), 38327–38337.

Yamamoto, N., Bauer, G., 2020. Apparent difference in fatalities between Central Europe and East Asia due to SARS-COV-2 and COVID-19: four hypotheses for possible explanation. Med. Hypotheses 144, 110160.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., Zhou, Q., 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 367 (6485), 1444–1448.

Yang, Z., Suzuki, R., Daniels, S.B., Brunquell, C.B., Sala, C.J., Nishiyama, A., 2006. NG2 glial cells provide a favorable substrate for growing axons. J. Neurosci. 26 (14), 3829–3839.

Zhang, H., Zhou, G., Zhi, L., Yang, H., Zhai, Y., Dong, X., Zhang, X., Gao, X., Zhu, Y., He, F., 2005. Association between mannose-binding lectin gene polymorphisms and susceptibility to severe acute respiratory syndrome coronavirus infection. J. Infect. Dis. 192 (8), 1355–1361.

Zhang, Q., Bastard, P., Liu, Z., Le Pen, J., Moncada-Velez, M., Chen, J., Ogishi, M., Sabli, I.K., Hodeib, S., Korol, C., 2020a. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science 370 (6515).

Zhang, S.-Y., Zhang, Q., Casanova, J.-L., Su, H.C., 2020b. Severe COVID-19 in the young and healthy: monogenic inborn errors of immunity? Nat. Rev. Immunol. 20 (8), 455–456.

Zheng, Y., Xu, H., Yang, M., Zeng, Y., Chen, H., Liu, R., Li, Q., Zhang, N., Wang, D., 2020. Epidemiological characteristics and clinical features of 32 critical and 67 noncritical cases of COVID-19 in Chengdu. J. Clin. Virol. 104366.

Zhu, X., Wang, Y., Zhang, H., Liu, X., Chen, T., Yang, R., Shi, Y., Cao, W., Li, P., Ma, Q., 2011. Genetic variation of the human α-2-Heremans-Schmid glycoprotein (AHSG) gene associated with the risk of SARS-CoV infection. PloS One 6 (8), e23730.

5. PROTECTIVE ROLE OF A *TMPRSS2* VARIANT ON SEVERE COVID-19 OUTCOME IN YOUNG MALES AND ELDERLY WOMEN

genes

MDPI

Communication

Protective Role of a *TMPRSS2* Variant on Severe COVID-19 Outcome in Young Males and Elderly Women

Maria Monticelli [1], Bruno Hay Mele [1,2], Elisa Benetti [3], Chiara Fallerini [3,4], Margherita Baldassarri [3,4], Simone Furini [3], Elisa Frullanti [3,4], Francesca Mari [3,4,5], GEN-COVID Multicenter Study †, Giuseppina Andreotti [6,*], Maria Vittoria Cubellis [1,6,*] and Alessandra Renieri [3,4,5]

[1] Department of Biology, Università Federico II, 80126 Napoli, Italy; maria.monticelli@unina.it (M.M.); bruno.haymele@szn.it (B.H.M.)

[2] Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Napoli, Italy

[3] Med Biotech Hub and Competence Center, Department of Medical Biotechnologies, University of Siena, 53100 Siena, Italy; elisa.benetti@dbm.unisi.it (E.B.); fallerini2@unisi.it (C.F.); margherita.baldassarri@dbm.unisi.it (M.B.); simone.furini@unisi.it (S.F.); elisa.frullanti@dbm.unisi.it (E.F.); francesca.mari@unisi.it (F.M.); alessandra.renieri@unisi.it (A.R.)

[4] Medical Genetics, Department of Medical Biotechnologies, University of Siena, 53100 Siena, Italy

[5] Genetica Medica, Azienda Ospedaliero-Universitaria Senese, 53100 Siena, Italy

[6] Istituto di Chimica Biomolecolare—CNR, 80078 Pozzuoli, Italy

***** Correspondence: gandreotti@icb.cnr.it (G.A.); cubellis@unina.it (M.V.C.); Tel.: +39-081-8675241 (G.A.); +39-081-679152 (M.V.C.)

† Collaborators of the GEN-COVID Multicenter Study are listed in the Appendix A.

Abstract

The protease encoded by the *TMPRSS2* gene facilitates viral infections and has been implicated in the pathogenesis of SARS-CoV-2. We analyzed the TMPRSS2 sequence and correlated the protein variants with the clinical features of a cohort of 1177 patients affected by COVID-19 in Italy. Nine relatively common variants (allele frequency > 0.01) and six missense variants which may affect the protease activity according to PolyPhen-2 in HumVar-trained mode were identified. Among them, p.V197M (p.Val197Met) (rs12329760) emerges as a common variant that has a deleterious effect on the protease and a protective effect on the patients. Its role appears particularly relevant in two subgroups of patients—young males and

elderly women—and among those affected by co-morbidities, where the variant frequency is higher among individuals who were mildly affected by the disease and did not need hospitalization or oxygen therapy than among those more severely affected, who required oxygen therapy, ventilation or intubation. This study provides useful information for the identification of patients at risk of developing a severe form of COVID-19 and encourages the usage of drugs affecting the expression of *TMPRSS2* or inhibiting protein activity.

Keywords: COVID-19; TMPRSS2; V197M; missense mutation; Whole-Exome Sequencing (WES)

Introduction

The gene *TMPRSS2* encodes a plasma membrane-anchored serine protease (TMPS2_HUMAN) that activates several endogenous substrates such as pro-hepatocyte growth factor [1], PAR-2 [2], matriptase/ST14 [3], and ACE2 [4]. In addition to this, it is essential for the pathogenesis of several human viruses, coronavirus 229E [5], Middle East respiratory syndrome coronavirus [6], influenza A viruses [7], and parainfluenza viruses [8].

The protein is composed of a small cytoplasmic region (aa 1–121), a transmembrane helix (aa 122–142), and an extracellular region (aa 143–529). In turn, the last one is composed of three domains, an LDL-receptor-like domain (aa 149–186), a SRCR-like domain (aa 187–279), and a peptidase domain (aa 293–526).

TMPRSS2 is expressed in the prostate glands predominantly, but also in type II pneumocytes in the lung, nasal goblet secretory cells, and small intestine [9–12]. It is regulated by androgenic hormones *in vivo* [1].

Genetic determinants of susceptibility and/or severity of COVID-19 have been sought in *TMPRSS2* in predictive [13–16] as well as in Whole-Exome Sequencing (WES) studies [17,18]. An analysis extended to large cohorts confirms the protective role of a common polymorphism p.V197M (rs12329760) in *TMPRSS2* and, in particular, its effect in young males and elderly women.

Materials and Methods

Whole-Exome Sequencing (WES) data derived from the GEN-COVID Multicenter Study [19] were analyzed. Odds ratio (OR) for different contingency tables were calculated using unconditional maximum likelihood estimation (Wald's method); confidence intervals (CI) at 95% were calculated using the normal approximation. Independence was tested using a Chi-squared test. All calculations were performed within the R environment for statistical computing, using the OR function from the epitools package. R scripts are provided in Supplementary file S1.

The sequence of human TMPS2_HUMAN was obtained from the UniProt database [20]. Structural templates were searched using the SWISS-MODEL server [21] and a model of the region spanning aa 187 to 526 was built using the structure of human hepsin [22]. A protein model was visualized using PyMol [23]. The solvent accessibility and the effect of mutations on protein stability were calculated for variants occurring in the region spanning aa 187 to 526 using SDM [24].

PolyPhen-2 [25], which is a program that can graduate the severity of missense variants [26], was used to identify mutations with a high impact on TMPS2_HUMAN.

Results

We analyzed the WES data of 1177 patients affected by COVID-19 and selected according to the following inclusion criteria: (i) endotracheal intubation; (ii) CPAP/BiPAP ventilation; (iii) oxygen therapy; (iv) hospitalized without oxygen support; and (v) not hospitalized. In this cohort, we identified 52 variants in *TMPRSS2*, nine of which are relatively common (allele frequency > 0.01). *TMPRSS2* encodes two isoforms that differ for 37 aa at the amino-terminus and both can activate respiratory viruses [27]. In this paper, the numbering of the longest isoform is used to identify variants in coding regions. The frequencies of two common mutations, one missense and one synonymous, p.V197M (rs12329760) and p.G296G (rs2298659), correlate with COVID-19 severity (Figure 5.1 and Table 5.1).

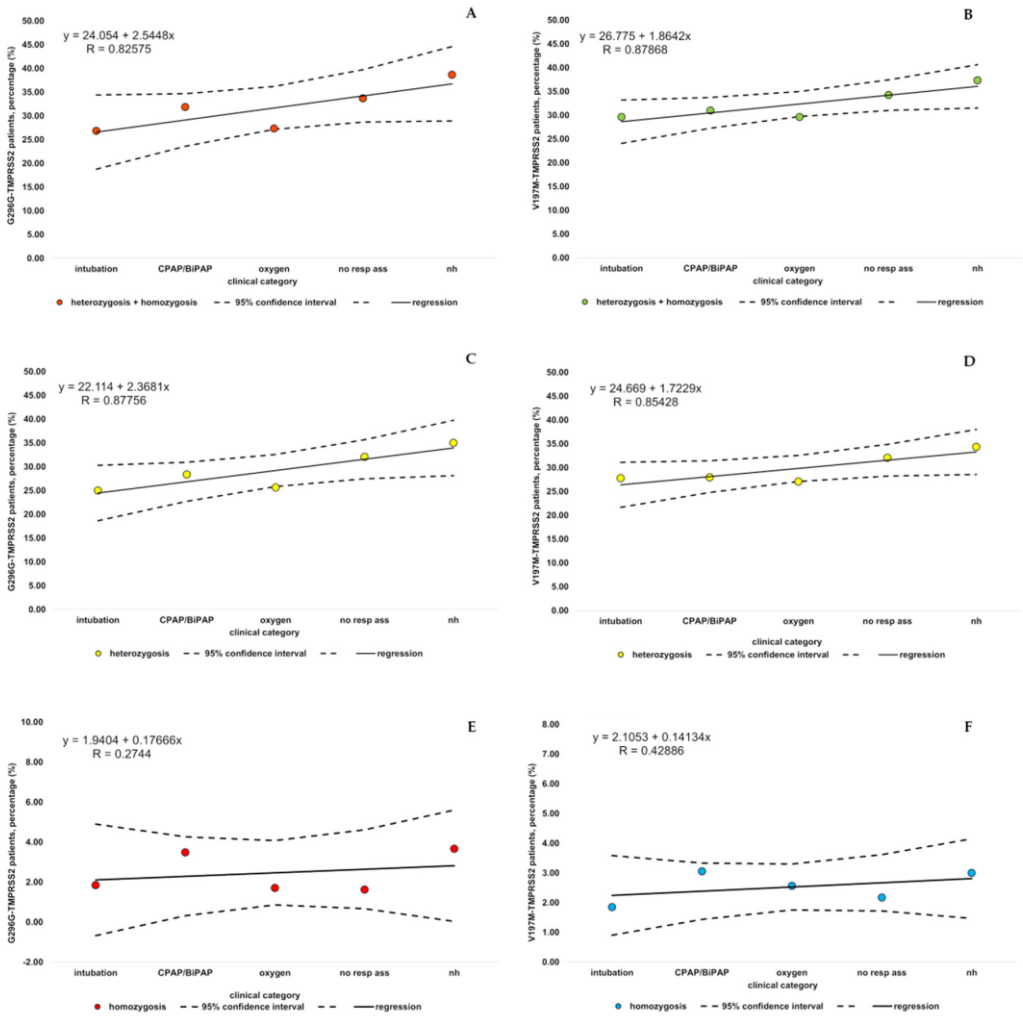

Figure 5.1. Prevalence of p.G296G or p.V197M in COVID-19 patients subdivided by different clinical outcomes. The percentages of carriers (A,B), heterozygous (C,D) and homozygous (E,F) individuals for p.G296G (A,C,E) or p.V197M (B,D,F) in each clinical category, *i.e.*, patients who needed intubation, ventilation (CPAP/BiPAP), or oxygen therapy, patients who did not need oxygen therapy (no resp support) or were not hospitalized (nh), are reported. The 95% confidence interval around the regression line is calculated as the product of the regression standard error and the value of the 0.975-th quantile of a t distribution with 3 degrees of freedom.

The percentage of carriers in severely (intubation + CPAP/BiPAP + oxygen) or mildly (no resp support + nh) affected patients differs significantly and is higher in the latter group (p.V197M p = 0.0289, OR = 0.7601, LC = 0.5941, HC = 0.9724;

p.G296G p = 0.0039, OR = 0.6947, LC = 0.5422 HC = 0.8899; Table 5.1).

(A) p.G296G			
	Mild	**Severe**	**Marginal_Rows**
Wild-type	306 (63%)	490 (71%)	796
G296G	178 (37%)	198 (29%)	376
marginal_cols	484	688	1172
(B) p.V197M			
	Mild	**Severe**	**Marginal_Rows**
Wild-type	313 (64%)	482 (70%)	795
V197M	176 (36%)	206 (30%)	382
marginal_cols	489	688	1177

Table 5.1. Contingency tables of *TMPRSS2* mutations in the two categories of patients.
Patients were divided in two categories according to the severity of the disease: mild (including not hospitalized and no respiratory assistance required) and severe (including oxygen therapy, ventilation, or intubation required).

Both SNPs are covered by the GWAS study of Severe COVID-19 with Respiratory Failure and are significantly associated with susceptibility (p.V197M p = 0.0153 OR = 0.8497, p.G296G p = 0.0462 OR = 0.8730) [28].

Several lines of evidence support the hypothesis that p.V197M affects the stability and/or the function of the protease. It is predicted to be deleterious by SIFT [29] and PolyPhen-2 [25], in both predictive modes, HumDiv and HumVar (Supplementary file S2).

p.V197M does not occur in the catalytic domain but in the SRCR domain that is likely needed for protein-protein interaction [30,31]. The region spanning aa 187 to 526 can be built by homology modeling. The side chain of Valine197 is buried (Figure 5.2), and solvent accessibility (%) 15.4 and its substitution by Methionine affects the stability of the protein (pseudo$\Delta\Delta$G= −2.00 kcal/mol).

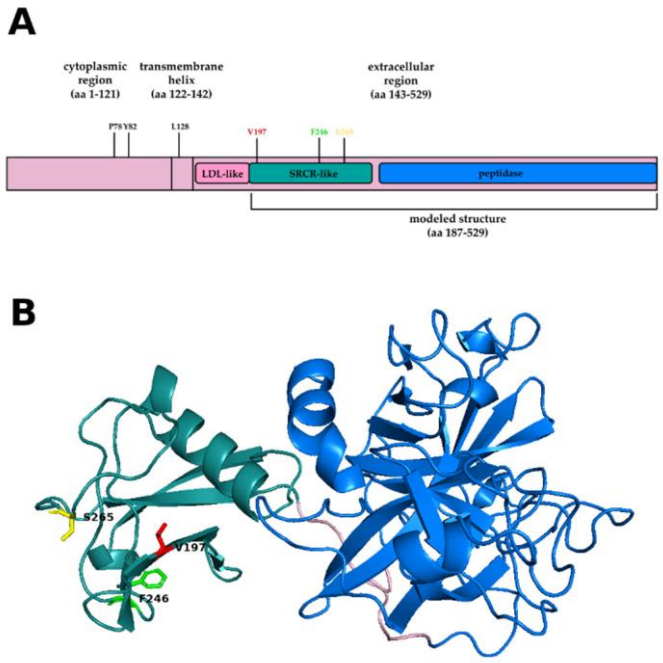

Figure 5.2. TMPRSS2 structure. A linear representation of TMPRSS2 is shown in panel (A). The probably and possibly damaging mutations identified by the HumVar-trained model in wAnnovar are highlighted. The model of the region spanning from aa 187 to 526 is shown as a cartoon in panel (B). The SRCR-like domain is in deep teal cyan, peptidase domain is in marine blue, and the linker is in light pink. V197 (red), F246 (green), and S265 (yellow) belong to the SRCR-like domain.

On the other side, the synonymous variant p.G296G, which corresponds to NM_001135099:exon9:c.C888T, is not annotated as an eQTL and does not fall at the exon- intron junction. We hypothesize that the missense mutation is protective and is often associated in cis with the synonymous variant, which has little effect *per se*.

The frequency of the other common variants does not correlate with the disease severity (Supplementary file S3). This is expected since most of them are synonymous or intron variants. The only missense mutation is p.G8V that is predicted to be benign by SIFT [29] and PolyPhen-2 [25] in both predictive modes, HumDiv and HumVar (Supplementary file S2).

It has been observed and it is commonly accepted that age, sex, and co-morbidities influence the severity of COVID-19. We divided patients according to sex and age (below and above median values) and the severity of the disease. The number of carriers of p.V197M (heterozygous + homozygous individuals) among mild and

severe cases differs significantly in two sub-cohorts, those of young males (p = 0.0200, OR = 0.5804, LC = 0.3663, HC = 0.9197) and elderly women (p = 0.0347, OR = 0.5346, LC = 0.2977, HC = 0.9601) (Table 5.2). To check whether slightly altering the age frame may flip the p-value to non-significant, we tested different ages and found that, provided that the boundary between young and old patients falls between 52 and 66 years for males and between 54 and 70 years for females, the protective effect remains significant for young males and elderly women (Supplementary Figure S1).

We restricted the two groups of patients and analyzed young males and elderly women who were affected by co-morbidities such as diabetes and/or hypertension and/or obesity or other diseases (young males affected by co-morbidities p = 0.0057, OR = 0.2969, LC = 0.1254, HC = 0.7027; elderly women affected by co-morbidities p = 0.0391, OR = 0.4667, LC = 0.2262, HC = 0.9627) (Table 5.3). The results hold significance when we consider the first quartile of the male or the last quartile of the female population (very young males in Q1 affected by co-morbidities p = 0.0343, OR = 0.2121, LC = 0.0485, HC = 0.9281; very elderly women in Q4 affected by co-morbidities p = 0.0090, OR = 0.2455, LC = 0.0812, HC = 0.7424) (Table 5.4).

Data in the two sub-cohorts of elderly males and young women are shown in Supplementary file S4. The subdivision of patients was carried out starting from the complete data set reporting the age, sex, and co-morbidities of individual patients. It is provided in Supplementary file S5.

Table 5.2. Contingency tables of V197M TMPRSS2 in young males and elderly women.

(A) Male, Young (Age ≤ 60)			
	Mild	**Severe**	**Marginal_Rows**
Wild-type	102 (63%)	140 (75%)	242
Val197Met	59 (37%)	47 (25%)	106
marginal_cols	161	187	348
(B) Female, Elderly (Age ≥ 58)			
	Mild	**Severe**	**Marginal_Rows**
Wild-type	37 (57%)	131 (71%)	168
Val197Met	28 (43%)	53 (29%)	81

marginal_cols	65	184	249

Patients were divided according to sex and age (above and below median age) and the severity of the disease.

Table 5.3. Contingency tables of V197M TMPRSS2 in young males and elderly women affected by co-morbidities.

(A) Male with Co-Morbidities, Young (Age ≤ 60)			
	Mild	**Severe**	**Marginal_Rows**
Wild-type	16 (50%)	65 (77%)	81
Val197Met	16 (50%)	19 (23%)	35
marginal_cols	32	84	116
(B) Female with Co-Morbidities, Elderly (Age ≥ 58)			
	Mild	**Severe**	**Marginal_Rows**
Wild-type	21 (54%)	105 (71%)	126
Val197Met	18 (46%)	42 (29%)	60
marginal_cols	39	147	186

Patients with co-morbidities were divided according to sex and age (above and below median age) and the severity of the disease.

Table 5.4. Contingency tables of V197M TMPRSS2 in very young males and very elderly women affected by co-morbidities.

(A) Male, Very Young (Age ≤ 50)			
	Mild	**Severe**	**Marginal_Rows**
Wild-type	7 (39%)	12 (75%)	19
Val197Met	11 (61%)	4 (25%)	15
marginal_cols	18	16	34
(B) Female, Very Elderly (Age ≥ 71)			
	Mild	**Severe**	**Marginal_Rows**
Wild-type	6 (37.5%)	66 (71%)	72
Val197Met	10 (62.5%)	27 (29%)	37
marginal_cols	16	93	109

Patients were divided according to sex and age (quartiles) and the severity of the disease.

It is not surprising that the effects of a deleterious mutation in *TMPRSS2* are seen in both sexes because the expression of the gene is only slightly higher in males [13]. The difference in the times of life can be explained because androgens and estrogens have opposite effects on gene expression, as proved by the data collected from the Expression Atlas [32] (Supplementary file S6). Young males who do not carry p.V197M are at risk because of high testosterone levels. Particularly relevant is the risk of wild-type elderly women who are not protected by estrogens. Androgenic hormones decline with age less rapidly than estradiol after menopause and this effect might explain the risk in elderly females and the protective role of the variant.

Other rare missense mutations are found in the Italian cohort in heterozygosity (Supplementary file S2). We will discuss the germline mutations that can affect the protein. A few fall in the region spanning aa 187 to 526 that can be modeled by homology, thus precluding a comparative analysis based on structural effects. PolyPhen-2 [25] uses sequence conservation to predict deleterious effects and two databases for training and testing predictions. The HumDiv model is trained using Mendelian disease variants vs. divergence from close mammalian homologs of human proteins (≥95% sequence identity). HumVar is trained using all human variants associated with some disease (except cancer mutations) or loss of activity/function vs. common (minor allele frequency > 1%) human polymorphism with no reported association with a disease of other effects.

The HumVar-trained model, which is best suited for distinguishing mutations with drastic effects, predicts that only one rare variant, p.L128P (rs147711290), has a strong effect (probably damaging, D, supplementary file S2) on the protein, perhaps because it occurs in the transmembrane anchor where proline destabilizes the helix [33]. p.L128P, which is relatively frequent only among Ashkenazy Jews (0.2%), was found in the Italian cohort in a single patient on oxygen therapy. The HumVar-trained model predicts that p.S265I and p.F246I (rs150554820), which occur on the surface of SRCR-like domain (aa 187–279) (Figure 5.2), p.P78L (rs138651919) and p. Y82D (rs201679623), which are located in the poorly characterized cytoplasmic region, have a moderate effect (possibly damaging, P, supplementary file S2) that is not sufficient to protect the patients. The effects of p.S265I and

p.F246I (rs150554820) were calculated using the model shown in Figure 5.2. Both mutations affect relatively buried amino acids, S265 and F246 (Solvent accessibility (%) 11.5 and 1.2, respectively), but have a very mild effect on protein stability (S265I and F246I pseudo$\Delta\Delta G$ = −0.29 kcal/mol and −0.07 kcal/mol, respectively).

These variants are not enriched in the mildly affected patients but were observed in severely affected ones too, 5 and 6, respectively.

Discussion and Conclusions

From our analysis, based on the WES of a large cohort of Italian patients, a variant in *TMPRSS2* emerges as a predictive factor to identify patients at risk of a severe course of COVID-19.

The allele frequency of p.V197M in our study is 0.18, in line with what was already reported (0.17 [13]). This value is well below the one reported in GnomAD for Eastern Asians (0.38), Finns (0.37), and Africans (0.29). Among Non-Finnish Europeans, the allele frequency of p.V197M is 0.23 with an apparent gradient from North to South, Northern Sweden 0.29, Estonia 0.31 (Genetic variation in the Estonian population), UK 0.21 (UK 10K study—Twins), and Spain 0.17 (Medical Genome Project healthy controls from Spanish population). It is highly suggestive correlating the low frequency of V197M with the high impact that the first wave of the epidemics had in Italy.

Other missense mutations are not enriched in the mildly affected patients, possibly because their effect on the protein is weak.

The protein encoded by *TMPRSS2* belongs to a family of membrane proteases, some of which promote SARS-CoV-2 infection [4,5,11]. No common deleterious missense mutations are found in these genes but for a variant in *TMPRSS4*, p.P413L, which is frequent among Latinos and mixed Americans (0.1259) but not among Europeans (0.002). The fact that a missense mutation with a destabilizing effect on the protein product, such as p.V197M, influences the course of COVID-19 implicitly suggests that *TMPRSS2* can be targeted for therapies either reducing its expression or inhibiting its protein product.

Sexual hormones can be used for therapeutic purposes [34,35]. Quite interestingly, and in line with our results, Seeland et al. reported that: “their retrospective study of hormone therapy in female COVID-19 patients shows that the fatality risk for women > 50 years receiving estradiol therapy (user group) is reduced by more than 50%; the OR was 0.33, 95% CI (0.18, 0.62) and the hazard ratio (HR) was 0.29, 95% CI (0.11,0.76). For younger, pre-menopausal women (15–49 years), the risk of COVID-19 fatality is the same irrespective of estradiol treatment, probably because of higher endogenous estradiol levels” [36].

To obtain more targeted effects, specific inhibitors at the protein level could be considered since they could have a protective effect analogous to that exerted by the mutation. Preliminary clinical data concerning Camostat [37,38] and Nafamostat [39,40] have been published. Their identification was the result of reposition which is a useful approach to reduce the time and costs of drug development [41] and has been largely employed during the emergency posed by COVID-19 [42,43].

In vitro Camostat and Nafamostat bind and inhibit TMPS2_HUMAN with great affinity, IC50 6.2 nM and 0.27 nM, respectively, but are not specific [44] and are not devoid of side effects in the patients [45]. *In silico* docking experiments have been carried out to find other inhibitors of TMPS2_HUMAN [46], but *in vitro* validation of the hits has not yet been carried out.

The effect of V197M on TMPS2_HUMAN was predicted by several authors [13–18]. WES analysis conducted on a large cohort of patients proves that the variant has indeed a statistically significant protective role in COVID-19. We do hope that our study will not only help to identify at risk patients, especially among elderly women, but also encourage the development of drugs for their treatment.

During the revision of this paper, we became aware that an independent research group proved the protective role of *TMPRSS2* variants in COVID-19 in the general population [47].

Supplementary Materials: The following are available online at www.mdpi.com/2073- 4425/12/4/596/s1, Supplementary Tables S1: R scripts; Supplementary Tables S2: TMPRSS2 variants in COVID-19 patients’ cohort;

Supplementary Tables S3: contingency tables of TMPRSS2 common variants with no correlation; Supplementary Tables S4: contingency tables of old males and young women; Supplementary Tables S5: complete data set reporting the age, sex, and co-morbidities of individual patients; Supplementary Tables S6: differential expression data for TMPRSS2 from Expression Atlas. Supplementary Figure S1: influence of the boundary age on the statistical significance of p.V197M protective effect.

Author Contributions: Conceptualization, M.V.C., G.A., and A.R.; validation, M.M. and B.H.M.; formal analysis, E.B., C.F., and S.F.; resources, GEN_COVID Multicenter Study; data curation, M.V.C., G.A., A.R., and F.M.; writing—original draft preparation, M.V.C., G.A., M.B., and E.F. All authors have read and agreed to the published version of the manuscript.

Funding: MIUR project "Dipartimenti di Eccellenza 2018–2020" to Department of Medical Biotechnologies University of Siena, Italy (Italian D.L. n.18 17 March 2020). Private donors for COVID-19 research. "Bando Ricerca COVID-19 Toscana" project to Azienda Ospedaliero- Universitaria Senese. Charity fund 2020 from Intesa San Paolo dedicated to the project N. B/2020/0119 "Identificazione delle basi genetiche determinanti la variabilità clinica della risposta a COVID-19 nella popolazione italiana".

Institutional Review Board Statement: The GEN-COVID study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the University Hospital of Siena Ethical Review Board (protocol code 16929, dated 16 March 2020). This observational study has been inserted in www.clinicaltrial.gov on 16 September 2020 (NCT04549831).

Informed Consent Statement: As part of GEN-COVID Multicenter Study written informed consent was obtained from all individuals who contributed samples and data. Detailed clinical and laboratory characteristics (data), specifically related to COVID-19, were collected for all subjects.

Data Availability Statement: The data and samples referenced here are housed in the GEN-COVID Patient Registry and the GEN-COVID Biobank and are available for consultation. You may contact Alessandra Renieri (e-mail: alessandra.renieri@unisi.it).

Acknowledgments: This study is part of the GEN-COVID Multicenter Study, https://sites.google.com/dbm.unisi.it/gen-covid (accessed on 19 April 2021) the Italian multicenter study aimed at identifying the COVID-19 host genetic bases from 8 April 2020. Specimens were provided by the COVID-19 Biobank of Siena, which is part of the Genetic Biobank of Siena, member of BBMRI-IT, of Telethon Network of Genetic Biobanks (project no. GTB18001), of EuroBioBank (http://www.eurobiobank.org/ (accessed on 19 April 2021)), and of RD-Connect. We thank the CINECA consortium for providing computational resources and the Network for Italian Genomes (NIG) http://www.nig.cineca.it (accessed on 19 April 2021) for its support (from July 27th, 2020 the GEN-COVID section is also available at http://nigdb.cineca.it/ (accessed on 19 April 2021)). We thank private donors for the support provided to A.R. (Department of Medical Biotechnologies, University of Siena, 53100 Siena, Italy) for the COVID-19 host genetics research project (D.L n.18 of 17 March 2020). We also thank the COVID-19 Host Genetics Initiative (https://www.covid19hg.org/ (accessed on 19 April 2021)) from April 2020, MIUR project "Dipartimenti di Eccellenza 2018-2020" to the Department of Medical Biotechnologies University of Siena, Italy and "Bando Ricerca COVID-19 Toscana" project to Azienda Ospedaliero-Universitaria Senese. We also thank Intesa San Paolo for the 2020 charity fund dedicated to the project N. B/2020/0119 "Identificazione delle basi genetiche determinanti la variabilità clinica della risposta a COVID-19 nella popolazione italiana".

Conflicts of Interest: The authors declare no conflict interests.

The list of collaborators of GEN-COVID Multicenter Study is available online as **Appendix A**, at https://www.mdpi.com/2073-4425/12/4/596.

References

1. Lucas, J.M.; Heinlein, C.; Kim, T.; Hernandez, S.A.; Malik, M.S.; True, L.D.; Morrissey, C.; Corey, E.; Montgomery, B.; Mostaghel, E.; et al. The Androgen-Regulated Protease TMPRSS2 Activates a Proteolytic Cascade Involving Components of the Tumor Microenvironment and Promotes Prostate Cancer Metastasis. Cancer Discov. 2014, 4, 1310–1325, doi:10.1158/2159-8290.CD-13- 1010.

2. Wilson, S.; Greer, B.; Hooper, J.; Zijlstra, A.; Walker, B.; Quigley, J.; Hawthorne, S. The Membrane-Anchored Serine Protease, TMPRSS2, Activates PAR-2 in Prostate Cancer Cells. Biochem. J. 2005, 388, 967–972, doi:10.1042/BJ20041066.

3. Ko, C.J.; Huang, C.C.; Lin, H.Y.; Juan, C.P.; Lan, S.W.; Shyu, H.Y.; Wu, S.R.; Hsiao, P.W.; Huang, H.P.; Shun, C.T.; et al. Androgen-Induced TMPRSS2 Activates Matriptase and Promotes Extracellular Matrix Degradation, Prostate Cancer Cell Invasion, Tumor Growth, and Metastasis. Cancer Res. 2015, 75, 2949–2960, doi:10.1158/0008-5472.CAN-14-3297.

4. Heurich, A.; Hofmann-Winkler, H.; Gierer, S.; Liepold, T.; Jahn, O.; Pohlmann, S. TMPRSS2 and ADAM17 Cleave ACE2 Differentially and Only Proteolysis by TMPRSS2 Augments Entry Driven by the Severe Acute Respiratory Syndrome Coronavirus Spike Protein. J. Virol. 2014, 88, 1293–1307, doi:10.1128/jvi.02202-13.

5. Bertram, S.; Dijkman, R.; Habjan, M.; Heurich, A.; Gierer, S.; Glowacka, I.; Welsch, K.; Winkler, M.; Schneider, H.; Hofmann- Winkler, H.; et al. TMPRSS2 Activates the Human Coronavirus 229E for Cathepsin-Independent Host Cell Entry and Is Expressed in Viral Target Cells in the Respiratory Epithelium. J. Virol. 2013, 87, 6150–6160, doi:10.1128/jvi.03372-12.

6. Shirato, K.; Kawase, M.; Matsuyama, S. Middle East Respiratory Syndrome Coronavirus Infection Mediated by the Transmembrane Serine Protease TMPRSS2. J. Virol. 2013, 87, 12552–12561, doi:10.1128/jvi.01890-13.

7. Garten, W.; Braden, C.; Arendt, A.; Peitsch, C.; Baron, J.; Lu, Y.; Pawletko, K.; Hardes, K.; Steinmetzer, T.; Böttcher- Friebertshäuser, E. Influenza Virus

Activating Host Proteases: Identification, Localization and Inhibitors as Potential Therapeutics. Eur. J. Cell Biol. 2015, 94, 375–383.

8. Abe, M.; Tahara, M.; Sakai, K.; Yamaguchi, H.; Kanou, K.; Shirato, K.; Kawase, M.; Noda, M.; Kimura, H.; Matsuyama, S.; et al. TMPRSS2 Is an Activating Protease for Respiratory Parainfluenza Viruses. J. Virol. 2013, 87, 11930–11935, doi:10.1128/jvi.01490- 13.

9. Vaarala, M.H.; Porvari, K.S.; Kellokumpu, S.; Kyllönen, A.P.; Vihko, P.T. Expression of Transmembrane Serine Protease TMPRSS2 in Mouse and Human Tissues. J. Pathol. 2001, 193, 134–140, doi:10.1002/1096-9896(2000)9999:9999<::AID- PATH743>3.0.CO;2-T.

10. Chen, Y.W.; Lee, M.S.; Lucht, A.; Chou, F.P.; Huang, W.; Havighurst, T.C.; Kim, K.; Wang, J.K.; Antalis, T.M.; Johnson, M.D.; et al. TMPRSS2, a Serine Protease Expressed in the Prostate on the Apical Surface of Luminal Epithelial Cells and Released into Semen in Prostasomes, Is Misregulated in Prostate Cancer Cells. Am. J. Pathol. 2010, 176, 2986–2996, doi:10.2353/ajpath.2010.090665.

11. Zang, R.; Castro, M.F.G.; McCune, B.T.; Zeng, Q.; Rothlauf, P.W.; Sonnek, N.M.; Liu, Z.; Brulois, K.F.; Wang, X.; Greenberg, H.B.; et al. TMPRSS2 and TMPRSS4 Promote SARS-CoV-2 Infection of Human Small Intestinal Enterocytes. Sci. Immunol. 2020, 5, eabc3582, doi:10.1126/sciimmunol.abc3582.

12. Ziegler, C.G.K.; Allon, S.J.; Nyquist, S.K.; Mbano, I.M.; Miao, V.N.; Tzouanas, C.N.; Cao, Y.; Yousif, A.S.; Bals, J.; Hauser, B.M.; et al. SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. Cell 2020, 181, 1016–1035.e19, doi:10.1016/j.cell.2020.04.035.

13. Asselta, R.; Paraboschi, E.M.; Mantovani, A.; Duga, S. ACE2 and TMPRSS2 Variants and Expression as Candidates to Sex and Country Differences in COVID-19 Severity in Italy. Aging (Albany NY) 2020, 12, 10087–10098, doi:10.18632/aging.103415.

14. Jeon, S.; Blazyte, A.; Yoon, C.; Ryu, H.; Jeon, Y.; Bhak, Y.; Bolser, D.; Manica, A.; Shin, E.-S.; Cho, Y.S.; et al. Ethnicity-Dependent Allele Frequencies Are Correlated with COVID-19 Case Fatality Rate. Res. Sq. 2020, 1, doi:10.21203/rs.3.rs-37798/v1.

15. Paniri, A.; Hosseini, M.M.; Akhavan-Niaki, H. First Comprehensive Computational Analysis of Functional Consequences of TMPRSS2 SNPs in Susceptibility to SARS-CoV-2 among Different Populations. J. Biomol. Struct. Dyn. 2020, 1–18, doi:10.1080/07391102.2020.1767690.

16. Senapati, S.; Kumar, S.; Singh, A.K.; Banerjee, P.; Bhagavatula, S. Assessment of Risk Conferred by Coding and Regulatory Variations of TMPRSS2 and CD26 in Susceptibility to SARS-CoV-2 Infection in Human. J. Genet. 2020, 99, 1–5, doi:10.1007/s12041-020-01217-7.

17. Latini, A.; Agolini, E.; Novelli, A.; Borgiani, P.; Giannini, R.; Gravina, P.; Smarrazzo, A.; Dauri, M.; Andreoni, M.; Rogliani, P.; et al. COVID-19 and Genetic Variants of Protein Involved in the SARS-CoV-2 Entry into the Host Cells. Genes 2020, 11, 1010, doi:10.3390/genes11091010.

18. Al-Mulla, F.; Mohammad, A.; Al Madhoun, A.; Haddad, D.; Ali, H.; Eaaswarkhanth, M.; John, S.E.; Nizam, R.; Channanath, A.; Abu-Farha, M.; et al. ACE2 and FURIN Variants Are Potential Predictors of SARS-CoV-2 Outcome: A Time to Implement Precision Medicine against COVID-19. Heliyon 2021, 7, e06133, doi:10.1016/j.heliyon.2021.e06133.

19. Daga, S.; Fallerini, C.; Baldassarri, M.; Fava, F.; Valentino, F.; Doddato, G.; Benetti, E.; Furini, S.; Giliberti, A.; Tita, R.; et al. Employing a Systematic Approach to Biobanking and Analyzing Clinical and Genetic Data for Advancing COVID-19 Research. Eur. J. Human Genet. 2021;1-15. doi:10.1038/s41431-020-00793-7.

20. Bateman, A. UniProt: A Worldwide Hub of Protein Knowledge. Nucleic Acids Res. 2019, 47, D506–D515, doi:10.1093/nar/gky1049.

21. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; De Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. Nucleic Acids Res. 2018, 46, W296–W303, doi:10.1093/nar/gky427.

22. Herter, S.; Piper, D.E.; Aaron, W.; Gabriele, T.; Cutler, G.; Cao, P.; Bhatt, A.S.; Choe, Y.; Craik, C.S.; Walker, N.; et al. Hepatocyte Growth Factor Is a Preferred in Vitro Substrate for Human Hepsin, a Membrane-Anchored Serine

Protease Implicated in Prostate and Ovarian Cancers. Biochem. J. 2005, 390, 125–136, doi:10.1042/BJ20041955.

23. DeLano, W.L. The PyMOL Molecular Graphics System; Version 2.3; Schrödinger LLC: New York, NY, USA, 2020.

24. Pandurangan, A.P.; Ochoa-Montaño, B.; Ascher, D.B.; Blundell, T.L. SDM: A Server for Predicting Effects of Mutations on Protein Stability. Nucleic Acids Res. 2017, 45, W229–W235, doi:10.1093/nar/gkx439.

25. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A Method and Server for Predicting Damaging Missense Mutations. Nat. Methods 2010, 7, 248–249.

26. Cimmaruta, C.; Citro, V.; Andreotti, G.; Liguori, L.; Cubellis, M.V.; Hay Mele, B. Challenging Popular Tools for the Annotation of Genetic Variations with a Real Case, Pathogenic Mutations of Lysosomal Alpha-Galactosidase. BMC Bioinf. 2018, 19, 39–46, doi:10.1186/s12859-018-2416-7.

27. Zmora, P.; Moldenhauer, A.S.; Hofmann-Winkler, H.; Pöhlmann, S. TMPRSS2 Isoform 1 Activates Respiratory Viruses and Is Expressed in Viral Target Cells. PLoS ONE 2015, 10, e0138380, doi:10.1371/journal.pone.0138380.

28. The Severe Covid-19 GWAS Group; Ellinghaus, D.; Degenhardt, F.; Bujanda, L.; Buti, M.; Albillos, A.; Invernizzi, P.; Fernández, J.; Prati, D.; Baselli, G.; et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N. Engl. J. Med. 2020, 383, 1522–1534, doi:10.1056/NEJMoa2020283.

29. Ng, P.C.; Henikoff, S. SIFT: Predicting Amino Acid Changes That Affect Protein Function. Nucleic Acids Res. 2003, 31, 3812–3814, doi:10.1093/nar/gkg509.

30. Resnick, D.; Pearson, A.; Krieger, M. The SRCR Superfamily: A Family Reminiscent of the Ig Superfamily. Trends Biochem. Sci.
1994, 19, 5–8, doi:10.1016/0968-0004(94)90165-1.

31. Hohenester, E.; Sasaki, T.; Timpl, R. Crystal Structure of a Scavenger Receptor Cysteine-Rich Domain Sheds Light on an Ancient Superfamily. Nat. Struct. Biol. 1999, 6, 228–232, doi:10.1038/6669.

32. Papatheodorou, I.; Moreno, P.; Manning, J.; Fuentes, A.M.P.; George, N.; Fexova, S.; Fonseca, N.A.; Füllgrabe, A.; Green, M.; Huang, N.; et al. Expression

Atlas Update: From Tissues to Single Cells. Nucleic Acids Res. 2020, 48, D77–D83, doi:10.1093/nar/gkz947.

33. Hofmann, K.; Stoffel, W. TMbase: A Database of Membrane Spanning Protein Segments. Biol. Chem. 1993, 374, 166.

34. Cattrini, C.; Bersanelli, M.; Latocca, M.M.; Conte, B.; Vallome, G.; Boccardo, F. Sex Hormones and Hormone Therapy during Covid-19 Pandemic: Implications for Patients with Cancer. Cancers 2020, 12, 2325, doi:10.3390/cancers12082325.

35. Salciccia, S.; Del Giudice, F.; Eisenberg, M.L.; Mastroianni, C.M.; De Berardinis, E.; Ricciuti, G.P.; Maggi, M.; Sciarra, A. Androgen-Deprivation Therapy and SARS-Cov-2 Infection: The Potential Double-Face Role of Testosterone. Ther. Adv. Endocrinol. Metab. 2020, 11, 11.

36. Seeland, U.; Coluzzi, F.; Simmaco, M.; Mura, C.; Bourne, P.E.; Heiland, M.; Preissner, R.; Preissner, S. Evidence for Treatment with Estradiol for Women with SARS-CoV-2 Infection. BMC Med. 2020, 18, 1–9, doi:10.1186/s12916-020-01851-z.

37. Breining, P.; Frølund, A.L.; Højen, J.F.; Gunst, J.D.; Staerke, N.B.; Saedder, E.; Cases-Thomas, M.; Little, P.; Nielsen, L.P.; Søgaard, O.S.; et al. Camostat Mesylate against SARS-CoV-2 and COVID-19—Rationale, Dosing and Safety. Basic Clin. Pharmacol. Toxicol. 2021, 128, doi:10.1111/bcpt.13533.

38. Hoffmann, M.; Hofmann-Winkler, H.; Smith, J.C.; Krüger, N.; Sørensen, L.K.; Søgaard, O.S.; Hasselstrøm, J.B.; Winkler, M.; Hempel, T.; Raich, L.; et al. Camostat Mesylate Inhibits SARS-CoV-2 Activation by TMPRSS2-Related Proteases and Its Metabolite GBPA Exerts Antiviral Activity. bioRxiv 2020, 1, doi:10.1101/2020.08.05.237651.

39. Jang, S.; Rhee, J.Y. Three Cases of Treatment with Nafamostat in Elderly Patients with COVID-19 Pneumonia Who Need Oxygen Therapy. Int. J. Infect. Dis. 2020, 96, 500–502, doi:10.1016/j.ijid.2020.05.072.

40. Takahashi, W.; Yoneda, T.; Koba, H.; Ueda, T.; Tsuji, N.; Ogawa, H.; Asakura, H. Potential Mechanisms of Nafamostat Therapy for Severe COVID-19 Pneumonia with Disseminated Intravascular Coagulation. Int. J. Infect. Dis. 2021, 102, 529–531, doi:10.1016/j.ijid.2020.10.093.

41. Mele, B.H.; Citro, V.; Andreotti, G.; Cubellis, M.V. Drug Repositioning Can Accelerate Discovery of Pharmacological Chaperones. Orphanet J. Rare Dis. 2015, 10, 55, doi:10.1186/s13023-015-0273-2.

42. Khadka, S.; Yuchi, A.; Shrestha, D.B.; Budhathoki, P.; Al-Subari, S.M.M.; Alhouzani, Z.T.M.; Butt, A.I. Repurposing Drugs for COVID-19: An Approach for Treatment in the Pandemic. Altern. Ther. Health Med. 2020, 26, 100-107.

43. Tworowski, D.; Gorohovski, A.; Mukherjee, S.; Carmi, G.; Levy, E.; Detroja, R.; Mukherjee, S.B.; Frenkel-Morgenstern, M. COVID19 Drug Repository: Text-Mining the Literature in Search of Putative COVID19 Therapeutics. Nucleic Acids Res. 2021, 49, D1113–D1121, doi:10.1093/nar/gkaa969.

44. Shrimp, J.H.; Kales, S.C.; Sanderson, P.E.; Simeonov, A.; Shen, M.; Hall, M.D. An Enzymatic TMPRSS2 Assay for Assessment of Clinical Candidates and Discovery of Inhibitors as Potential Treatment of COVID-19. ACS Pharmacol. Transl. Sci. 2020, 3, 997– 1007, doi:10.1021/acsptsci.0c00106.

45. Okajima, M.; Takahashi, Y.; Kaji, T.; Ogawa, N.; Mouri, H. Nafamostat Mesylate-Induced Hyperkalemia in Critically Ill Patients with COVID-19: Four Case Reports. World J. Clin. Cases 2020, 8, 5320–5325, doi:10.12998/wjcc.v8.i21.5320.

46. Singh, N.; Decroly, E.; Khatib, A.M.; Villoutreix, B.O. Structure-Based Drug Repositioning over the Human TMPRSS2 Protease Domain: Search for Chemical Probes Able to Repress SARS-CoV-2 Spike Protein Cleavages. Eur. J. Pharm. Sci. 2020, 153, 105495, doi:10.1016/j.ejps.2020.105495.

47. Andolfo, I.; Russo, R.; Lasorsa, V.A.; Cantalupo, S.; Rosato, B.E.; Bonfiglio, F.; Frisso, G.; Abete, P.; Cassese, G.M.; Servillo, G.; et al. Common variants at 21q22.3 locus influence MX1 and TMPRSS2 gene expression and susceptibility to severe COVID-19. Iscience 2021, 102322, doi:10.1016/j.isci.2021.102322.

6. LOSS-OF-FUNCTION MATERNAL-EFFECT MUTATIONS OF *PADI6* ARE ASSOCIATED WITH FAMILIAL AND SPORADIC BECKWITH-WIEDEMANN SYNDROME WITH MULTI-LOCUS IMPRINTING DISTURBANCE

Cubellis et al. Clinical Epigenetics (2020) 12:139
https://doi.org/10.1186/s13148-020-00925-2

Clinical Epigenetics

RESEARCH Open Access

Loss-of-function maternal-effect mutations of *PADI6* are associated with familial and sporadic Beckwith-Wiedemann syndrome with multi-locus imprinting disturbance

Maria Vittoria Cubellis[1], Laura Pignata[2], Ankit Verma[2,3], Angela Sparago[2], Rosita Del Prete[2], Maria Monticelli[1], Luciano Calzari[4], Vincenzo Antona[5], Daniela Melis[6], Romano Tenconi[7], Silvia Russo[4], Flavia Cerrato[2*] and Andrea Riccio[2,3*]

* Correspondence: flavia.cerrato@unicampania.it; andrea.riccio@unicampania.it

[2] Department of Environmental Biological and Pharmaceutical Sciences and Technologies (DiSTABiF), Università degli Studi della Campania "Luigi Vanvitelli", Caserta, Italy

Full list of author information is available at the end of the article

Abstract

Background: PADI6 is a component of the subcortical maternal complex, a group of proteins that is abundantly expressed in the oocyte cytoplasm, but is required for the correct development of early embryo. Maternal-effect variants of the subcortical maternal complex proteins are associated with heterogeneous diseases, including female infertility, hydatidiform mole, and imprinting disorders with multi-locus imprinting disturbance. While the involvement of PADI6 in infertility is well demonstrated, its role in imprinting disorders is less well established.

Results: We have identified by whole-exome sequencing analysis four cases of Beckwith-Wiedemann syndrome with multi-locus imprinting disturbance whose mothers are carriers of *PADI6* variants. *In silico* analysis indicates that these variants result in loss of function, and segregation analysis suggests they act as either recessive or dominant- negative maternal-effect mutations. Genome-wide methylation analysis revealed heterogeneous and extensively altered methylation profiles of imprinted loci in the patients, including two affected sisters, but not in their healthy siblings.

Conclusion: Our results firmly establish the role of PADI6 in imprinting disorders. We report loss-of-function maternal-effect variants of *PADI6* that are associated with heterogeneous multi-locus imprinting disturbances in the progeny. The rare finding of two siblings affected by Beckwith-Wiedemann syndrome suggests that in some cases, familial recurrence risk of these variants may be high. However, the heterogeneous phenotypes of the other pedigrees suggest that altered oocyte PADI6 function results in stochastic maintenance of methylation imprinting with unpredictable consequences on early embryo health.
Keywords: Multi-locus imprinting disturbance, PADI6, Beckwith-Wiedemann syndrome, Genomic imprinting, DNA methylation, Maternal-effect variants, Subcortical maternal complex, Infertility

Background

Differential DNA methylation between the maternally and paternally inherited chromosomes controls the monoallelic and parent-of-origin dependent expression of the imprinted genes, a group of about 100 loci with important roles in fetal growth, metabolism, and behavior [1]. The organization in clusters or domains of the imprinted genes enables their fine-tuned and coordinated regulation by cis-acting control regions. In particular, each imprinting cluster typically harbors one germline DMR (gDMR), in which differential allelic methylation is acquired during gametogenesis and is maintained throughout development in somatic cells, escaping the methylation reprogramming occurring in early embryogenesis. Currently, 38 gDMRs have been identified in the human genome and the majority of them are methylated on the maternal chromosome. Differential methylation of the gDMR is required for maintaining the parent-of-origin dependent expression of the entire imprinted gene cluster.

Genetic and epigenetic defects altering the expression of imprinted genes are associated with 12 rare clinical conditions, known as imprinting disorders [2]. Among this group of pathologies, the Beckwith-Wiedemann syndrome (BWS, OMIM #130650, prevalence of 1:10,340 live births [3]) is characterized by macrosomia, macro- glossia, abdominal wall defects, neonatal hypoglycemia, lateralized overgrowth, and predisposition to Wilms tumor and other embryonal

cancers [4]. At the molecular level, this disorder is caused by molecular defects affecting one or both of two imprinting domains located at chromosome 11p15.5: the telomeric domain that includes the *H19* and *IGF2* genes and is controlled by the *H19*/*IGF2*:IG-DMR, and the centromeric domain that includes the *KCNQ1OT1* and *CDKN1C* genes and is controlled by the *KCNQ1OT1*:TSS-DMR. The molecular defects of BWS are as follows: loss of methylation (LOM) of the *KCNQ1OT1*:TSS-DMR that is present in 50% of cases, mosaic segmental paternal unidisomy of chromosome 11p15 (upd(11)pat) in 20%, gain of methylation (GOM) of the *H19*/*IGF2*:IG-DMR in 5–10%, and maternal loss-of-function mutations of *CDKN1C* gene in 5% [4].

A subset of patients with imprinting disorders display methylation changes in addition to that of the locus that is normally associated with the clinical phenotype. This condition is known as multi-locus imprinting disturbance (MLID) and appears particularly frequent (one third of the cases according to ref. [4]) in BWS with *KCNQ1OT1*:TSS-DMR LOM [5, 6]. Its etiology is unknown in the majority of the cases [7]. However, in a number of families, causative loss-of-function mutations affecting either zygotic or oocyte-derived trans-acting factors have been found in the patients with MLID or in their mothers, respectively [1]. In particular, zygotic mutations affecting the zinc-finger *ZFP57* gene have been identified in the transient neonatal diabetes type 1 (TNDM1) [8], while maternal-effect mutations have been reported in healthy women with offspring with MLID and variable imprinting disorders including BWS, or women with reproductive problems, such as infertility, recurrent pregnancy loss, or hydatidiform mole [7]. Maternal-effect genes are abundantly expressed in oocytes but exert their phenotypic effect in the offspring, being critical for correct development of early embryos. Most of the maternal-effect genes associated with MLID or hydatidiform mole encode for components of the subcortical maternal complex (SCMC) that is localized in the periphery of the oocyte cytoplasm and early embryo until the blastocyst stage [9]. Among these genes, more compelling pieces of evidence have been provided for association of NLRP2, NLRP5, NLRP7, and KHDC3L, while the role of PADI6, OOEP, and TLE6 in the etiology of imprinting disorders is not definitely established [7]. Although several important functions have been attributed to the SCMC during

oocyte-to-embryo transition and early development (for a review, see [9]), the mechanisms through which this protein complex influences methylation imprinting remain elusive.

PADI6 encodes a member of the peptidylarginine deiminase (PAD) family, a class of enzymes converting arginine residues to citrulline. All mammalian PADs, including PADI6, share 70–95% identity in their amino acid sequence including two immunoglobulin- like domains at the N terminus and a highly conserved C-terminal domain that harbors the active site [10]. PADI6 is the only PAD for which no enzymatic activity has been detected *in vitro* [11]. Mouse and human studies demonstrate that *PADI6* is highly expressed in oocytes and early embryos, where it co-localizes with the other components of the SCMC [12, 13]. PADI6 is required for the formation of the oocyte lattices that are believed to work as ribosomal storage for early embryo [14]. Indeed, the development of Padi6-null embryos is arrested at the 2-cell stage, and their ribosomal components, *de novo* protein synthesis, and embryonic genome activation are impaired [15]. In humans, homozygous loss-of- function mutations of *PADI6* are associated with female infertility and hydatidiform mole (Additional file 1: Table S1) [13, 16–19]. In addition, biallelic or monoallelic missense *PADI6* variants have been linked with three cases of Silver-Russell syndrome and one case of BWS [20].

Other important aspects needing clarification concern the phenotypes associated with MLID and their risk of recurrence in the families in which maternal-effect variants segregate. A few familial cases with MLID have been described so far, and highly variable expressivity and incomplete penetrance of the phenotype have been reported [20–22]. In some cases, siblings with different imprinting disorders or complex phenotypes, or pedigrees with multiple pregnancy losses and imprinting disorders were observed [20, 23].

Here, we describe four BWS cases with MLID that are associated with maternal-effect loss-of-function variants of *PADI6*. Two of the probands are siblings, indicating that maternal variants of *PADI6* may act as trans-acting mutations in either familial or sporadic BWS.

Results

Clinical cases

The four patients described in this study were referred to our laboratory with clinical diagnosis of BWS [4]. Family 1 included two sisters with typical BWS features. The older one (III-1, Fig. 6.1a) was a 4-year-old girl born from non-consanguineous parents. She was born pre- term (32 weeks of gestation) because of maternal pre-eclampsia and polyhydramnios, with birth weight of 1980 g (50–70th centile), length of 44 cm (70–90th centile), and cranial circumference of 31 cm (75–90th centile). At birth, she presented breathing and feeding difficulties, placental hyperplasia, and macroglossia. Hypoglycemia and anemia occurred in the perinatal period. After birth, patent foramen ovale was observed, with closure at 18 months. She also showed visual defects. Clinical examination at three and a half years revealed the presence of lateralized overgrowth of the body with the right side bigger than the left side (1 cm dysmetry), angiomas, naevus flammeus, umbilical hernia, and diastasis recti. According to the recent BWS consensus, these features correspond to a score of 8 points that is consistent with the clinical diagnosis of classical BWS [4]. Her sister (III-2, Fig. 6.1a) was a 1-year-old child, born at 39 weeks of gestation with placental hyperplasia and polyhydramnios. At birth, her weight was 3000 g (25– 50th centile), length 48 cm (25th centile), and cranial circumference 32 cm (3–10th centile). Typical BWS features were observed, including neonatal hypoglycemia, macroglossia, naevus flammeus, and umbilical hernia. Clinical examination at 6 months revealed the presence of angiomas and patent foramen ovale. Overall, a clinical score of 6 was reached, confirming the clinical diagnosis of classical BWS, also in this case. No clinical signs of BWS were reported in the parents (II-1; II-2), grandparents (I-1 and I-2), and uncles (II-3 and II-4).

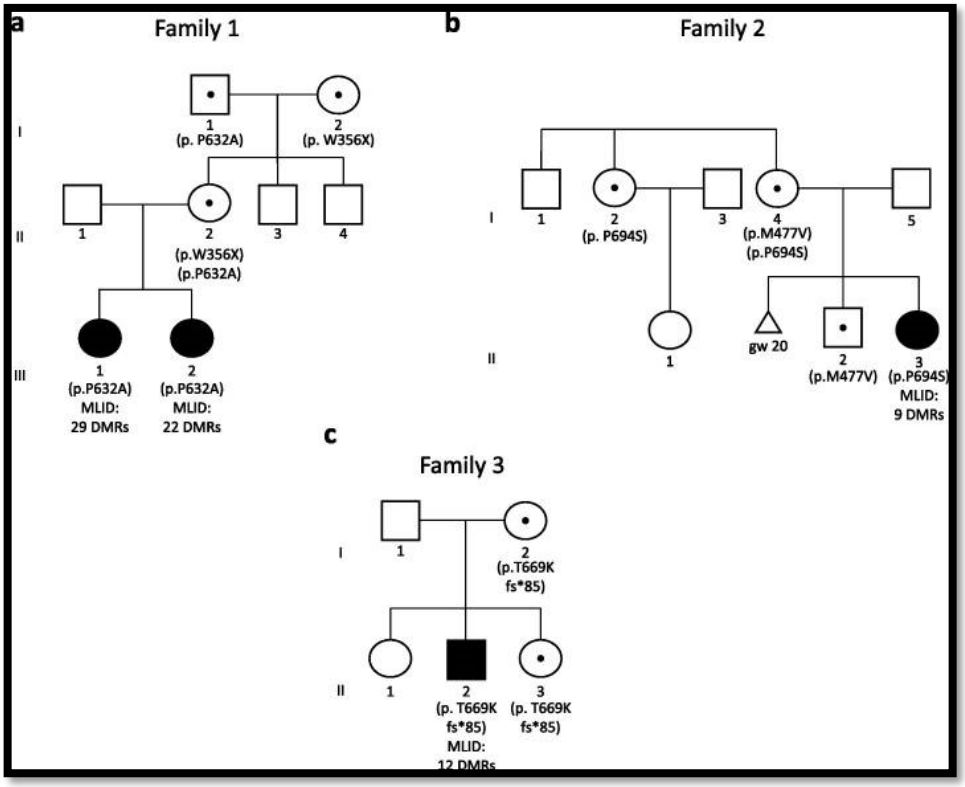

Fig. 6.1 Genetic characterization of the three families under study. Pedigree and corresponding *PADI6* variants of families 1 (a), 2 (b), and 3 (c). Black filled symbol represents individual with BWS features, and black central dot unaffected carriers of *PADI6* variants. Weeks of gestation (gw) are reported for the aborted fetus in family 2. MLID status and the number of DMRs with altered methylation are also reported

The proband of family 2 (II-3, Fig. 6.1b) was a 33-year- old lady born at 35 weeks of gestation from healthy unrelated parents. Birth weight was 2950 g (75–90th centile), birth length 51 cm (90–95th centile), and head circumference 34 cm (75–90th centile). She also presented diastasis recti, macroglossia, umbilical hernia, ear lobe creases, naevus flammeus, facial dysmorphism, organomegaly, renal dysplasia, and asymmetry of the chest, which allowed the clinical diagnosis of classical BWS (points = 8). Overgrowth started at the age of 6 months, no lateralized overgrowth and no feeding difficulties were observed, and normal developmental milestones were achieved. The patient was followed annually for the first 12 years of life with clinical evaluation and abdominal ultrasound. Last evaluation was at the age of 30 years. Intellectual development was normal, and she obtained a university degree. The proband's mother had a healthy phenotype, but

reported a spontaneous pregnancy loss at 20 weeks of gestation when she was 28 years old. At birth of the first son, she was 30 years old and her husband 32 years old. The proposita was born 5 years later. No conception difficulty was reported for the couple, and no clinical signs of BWS were observed in the other family members (Fig. 6.1b).

The proband of family 3 (II-2, Fig. 6.1c) was an 18-year- old male born from healthy non-consanguineous parents. He was born at 38 weeks of gestation, with birth weight of 5530 g (> 95th centile) and birth length of 54 cm (> 95th centile). During the postnatal period, he was affected by convulsions. Clinical examination at 1 year revealed the presence of macroglossia, prognathism, lateralized overgrowth, cryptorchidism, and patent foramen ovale, which together with macrosomia correspond to a clinical score of 5 that is sufficient for the clinical diagnosis of classical BWS. During follow-up, delay of language development and dorsal scoliosis were observed. More recently, because of recurrent occurrence of headache, the patient was subjected to magnetic resonance imaging, and Arnold-Chiari I malformation was diagnosed. The other family members showed no feature of BWS (Fig. 6.1c).

DNA methylation analysis

Molecular testing for BWS was performed by DNA methylation analysis of the 11p15.5 region [4]. DNA methylation was measured by bisulfite treatment of peripheral blood leukocyte (PBL) DNA followed by pyrosequencing. *KCNQ1OT1*:TSS-DMR LOM and normal methylation of the *H19*/*IGF2* IG-DMR were found in all four cases, confirming the diagnosis of BWS (Additional file 2: Figure S1). Since MLID is frequently associated with *KCNQ1OT1*:TSS-DMR LOM, we extended the DNA methylation analysis to five further DMRs that are frequently hypomethylated in BWS patients [5]. The results demonstrated that all four patients had at least one additional hypomethylated DMR, thus confirming the molecular diagnosis of MLID (Additional file 2: Figure S1). To better characterize the methylation abnormalities of our patients, we determined their genome-wide methylation profiles in PBL DNAs by employing the Illumina

Infinium methylation EPIC array, and compared them with those of their siblings and mothers, and 12 age-matched controls. After quality control filtering, ~ 736,000 CpG sites in each sample were retained. All these datasets were proven to be comparable after batch adjustment by principal component analysis (PCA, Additional file 3: Figure S2), and no evident difference was observed in the bimodal distribution of their genomic DNA methylation (Additional file 4: Figure S3). However, when only the CpGs underlying imprinted DMRs were considered, the four patients clustered separately from the controls and their healthy relatives, indicating that different methylation levels were present in imprinted loci (Fig. 6.2). To better compare the methylation profiles, we generated a heatmap (Fig. 6.3) depicting the methylation levels of all known imprinted DMRs [25] that were covered by at least three CpGs in our dataset. A striking difference was seen in the methylation of multiple DMRs between the four patients and controls. The methylation level of several DMRs of the patients differed in a statistically significant manner and was generally hypomethylated with respect to the controls, while only a few DMRs differed moderately in the control individuals from their average (Fig. 6.3 and Additional file 5: TableS2). In addition, the number and methylation level of the affected DMRs differed among the patients and even between affected siblings. Overall, the probands of family 1 had the most dramatic changes: III-1 and III-2 showed the most intense methylation changes and the highest number of affected DMRs. Overall, the most frequently affected loci in addition to *KCNQ1OT1* were *GNAS*, *MCTS2P*, *NHP2L1*, *PPIEL*, *DIRAS*, *ZNF331*, *IGF1R*, *ERLIN*, and *WRB*. Hypomethylation was observed in both maternally (e.g., *KCNQ1OT1*, *GNAS*, *MCTS2P*) and paternally methylated (e.g., *MEG3*) gDMRs. The paternally methylated secondary DMRs *GNAS*-NESP, *ZNF597*-TSS, and *ZDBF2*/*GPR1* were found hypermethylated, likely as consequence of hypomethylation of their respective maternally methylated gDMRs. A slight hypomethylation of the *KCNQ1OT1*:TSS-DMR was also evident in the mother of the family 3 proband (I-2 in Fig. 6.3 and Additional file 2: Fig S1c).

Identification of maternal *PADI6* variants

To investigate whether familial MLID was associated with genetic variants, whole-exome sequencing (WES) was performed on the DNAs of the probands and their parents of family 1. After filtering and exclusion of frequent variants, no damaging variant that was present in homozygosity or compound heterozygosity in both the probands was identified.

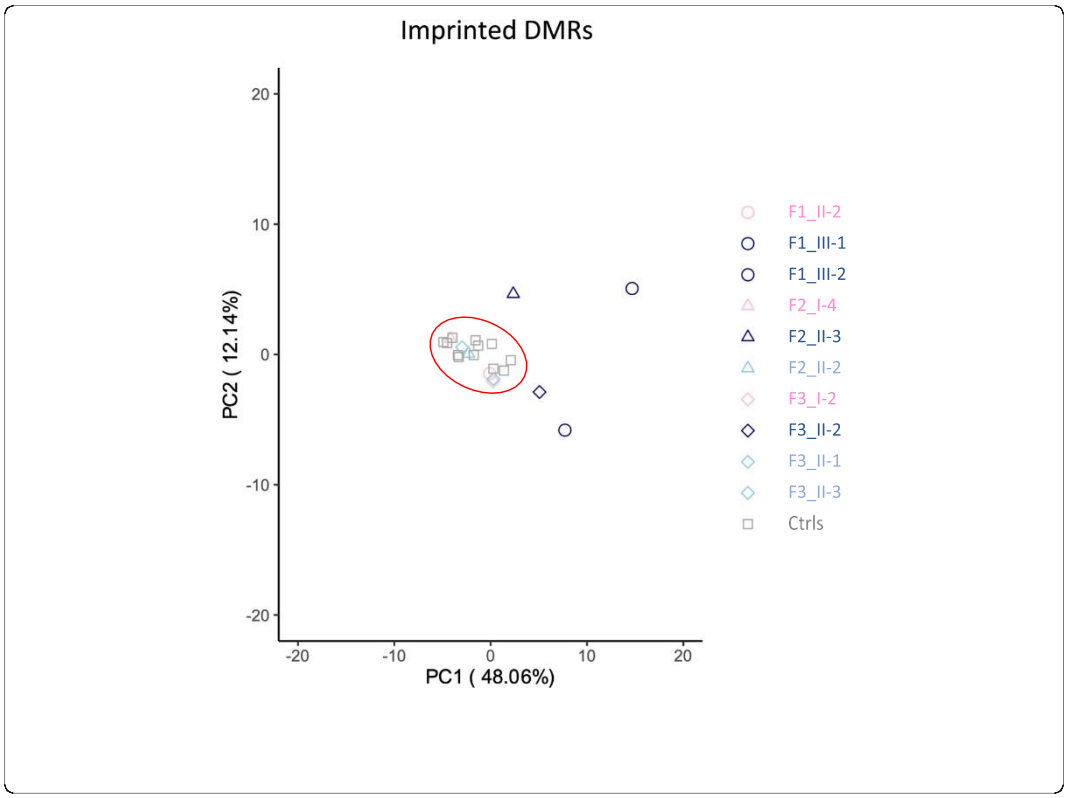

Fig. 6.2 PCA of imprinted DMR methylation. DNA methylation of the imprinted DMRs of the four probands, their mothers and siblings, and 12 age-matched healthy controls was compared by PCA. Symbol code: circles represent members of family 1, triangles members of family 2, rhombi members of family 3, and squares the controls. Color code: probands are represented in dark blue, mothers in pink, siblings in light blue, controls in gray.

Then, we looked for rare homozygous and compound heterozygous variants in their mother (Additional file 6: Table S3). According to the score of the mutation prediction tools (e.g., PolyPhen-2 and SIFT) of these variants and available information on the affected proteins, we identified *PADI6* as the best candidate gene

causative of MLID in this family. Two single-nucleotide variants (SNVs) of *PADI6* were found in compound heterozygosity in the mother of the probands (II-2). The identified variants were a novel stop-gain mutation (chr1, g.17718714 G → A, p.W356X) in exon 10 and an extremely rare missense mutation (chr1, g.17727743 C → G, p.P632A) in exon 17, respectively (Table 6.1). Both variants affected the arginine deiminase domain of PADI6 (Fig. 6.4). P632 was conserved in orthologous proteins, and the P632A variant was predicted to be damaging by PolyPhen-2 and deleterious by SIFT (Table 6.1). Segregation of the variants was demonstrated by exome-seq and Sanger sequencing (Additional file 6: Table S3, Additional file 7: Figure S4a). II-2 inherited the missense variant from her father (I-1) and the stop-gain mutation from her mother (I-2). The variants were not present in the patients' father (II-1), and only the missense variant was inherited by the probands. The presence of biallelic damaging variants in the probands' mother indicates that they act as loss-of-function maternal-effect mutations, as previously suggested [20].

To look for further maternal-effect variants associated with MLID, WES was performed on I-4 of family 2 and I-2 of family 3. Two extremely rare missense variants (chr1, g.17727929 C → T, p.P694S; chr1, g.17721538 A → G, p.M477V) of *PADI6* were identified in compound heterozygosity in the former woman and a novel heterozygous frameshift variant (GRCh37/hg19, chr1, g.17727855 C → -, p.T669K fs*85) also in *PADI6* in the latter (Additional file 6: Table S3 and Table 6.1). Both P694 and M477 are conserved in orthologous proteins, and P694S is predicted to be damaging by PolyPhen-2 and deleterious by SIFT, but is considered neutral on the protein structure by the prediction tool SDM, while M477V is predicted to be benign by PolyPhen-2 and tolerated by SIFT, but internally located (accessible surface area Å2 = 3.2) and destabilizing the protein structure by SDM (Table 6.1). As with the variants of family 1, also those identified in families 2 and 3 affect the arginine deiminase domain of PADI6 (Fig. 6.4). Segregation analysis demonstrated that M477V was inherited by the healthy proband's brother (II-2) and P694S by the proband (II-3) and her maternal aunt (I-2), in family 2 (Additional file 7: Figure S4b). In family 3, the frameshift variant pT669K fs*85 was inherited by the proband (II-2) and his healthy youngest sister (II-3, Additional file 7: Figure S4c). The frameshift T669Kfs*85 modifies the carboxy-terminal of PADI6, the last

residues being substituted by a longer peptide (Fig. 6.5a). Since this *PADI6* variant was present in heterozygosity in I-2 of family 3, we wondered if this mutation could have a dominant effect. We generated a 3D structural model of human PADI6 by aligning it to the paralogous PADI4 that is known to be a dimer [26]. On the basis of this model, we observed that the carboxy-terminal part of PADI6 was not involved in dimerization; hence, the formation of heterodimers with dominant-negative effect in the heterozygous patient is possible (Fig. 6.5b).

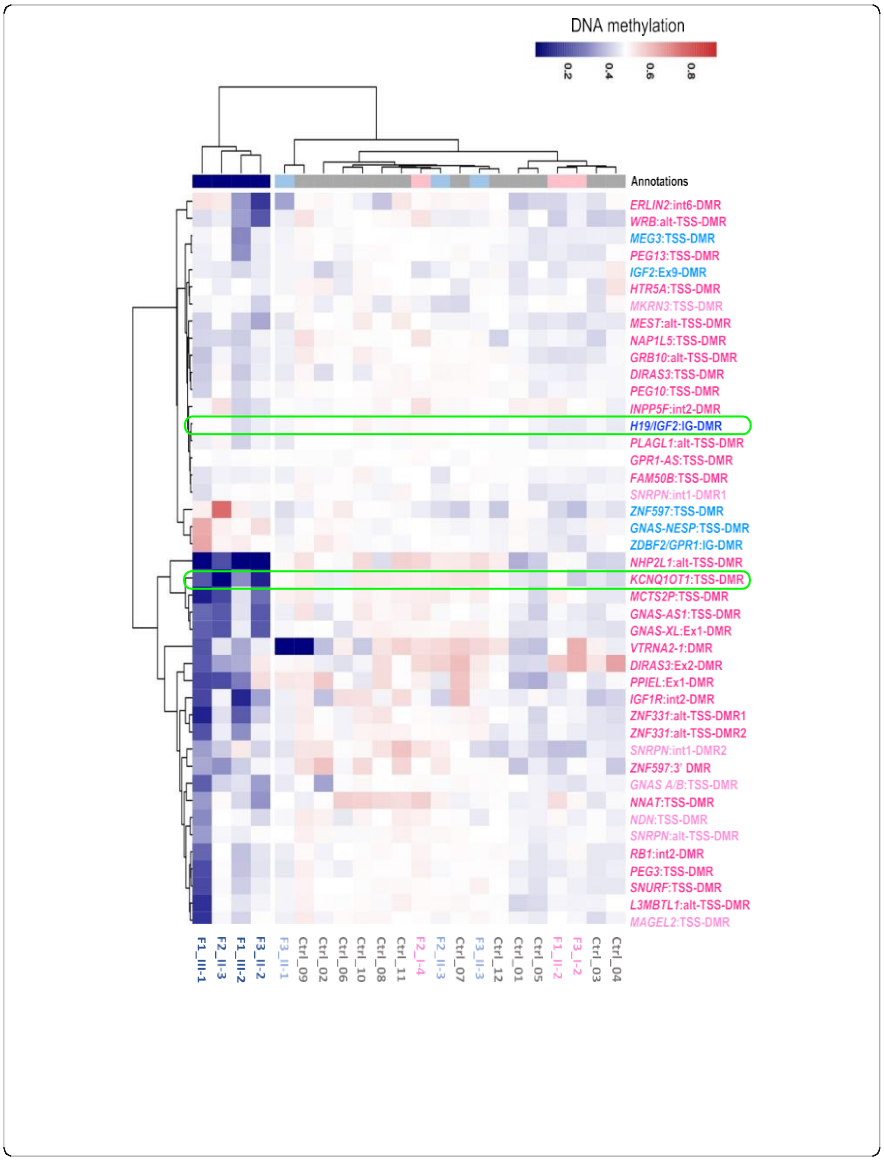

Fig. 6.3 DNA methylation analysis of imprinted DMRs. Heatmap showing imprinted DMR methylation levels of the probands of families 1, 2, and 3 (F1_III-1, F1_III-2, F2_II-3, F3_II-2); their mothers (F1_II-2, F2_I-4, F3_I-2) and siblings (F2_II-2, F3_II-1, F3_II-3); and 12 age-

matched control individuals, organized by hierarchical clustering. Clustering is based on CpG methylation levels of 736 probes overlapping with 43 imprinted DMRs, containing at least three CpGs. Maternally methylated germline DMRs are in dark pink; maternally methylated secondary DMRs are in light pink. Paternally methylated germline DMRs are in dark blue; paternally methylated secondary DMRs are in light blue. The *KCNQ1OT1*:TSS-DMR and the *H19*/*IGF2*:IG-DMR diagnostic of BWS are highlighted in green. The *VTRNA2-1* DMR has been reported to be polymorphic [24].

Table 6.1: List of the *PADI6* pathogenic variants and prediction of their effect on the protein structure and function.

Position	Alleles	Variant ID	Allele frequency	AA	Molecular effect	Conservation in orthologous proteins	Prediction				Genotype	Family
							SIFT	Polyphen-2	SDM (ΔΔ*G*)	Accessible surfaces area		
chr1 17718714	G > A	–	-	Trp356Ter	Nonsense	–	–	–	–	-	Comp het	1
chr1 17727743	C > G	rs755260464[a]	0.000004012	Pro632Ala	Missense	Yes	Deleterious	Possibly damaging	Stabilizing (0.600 kcal/mol)	$Å^2$ = 0.61		
chr1:17721538	A > G	rs761556429[b]	0.000004008	Met477Val	Missense	Yes	Tolerated	Benign	Destabilizing (− 0.400 kcal/	$Å^2$ = 3.2	Comp	2
chr1:17727929	C > T	rs1368496637[c]	0.000008	Pro694Ser	Missense	Yes	Deleterious	Possibly damaging	Neutral (0.000 kcal/mol)	$Å^2$ = 61.7		
chr1:17727854-17727856	Del C	–	-	Thr669Lys	fs*85	–	–	–	–	-	Het	3

Allele frequency was worldwide, as reported by GnomeAD Comp het compound heterozygous, Het heterozygous

[a]Observed only in heterozygosity in a Swedish male

[b]Observed only in heterozygosity in a Southern European male

[c]Observed only in two males in heterozygosity

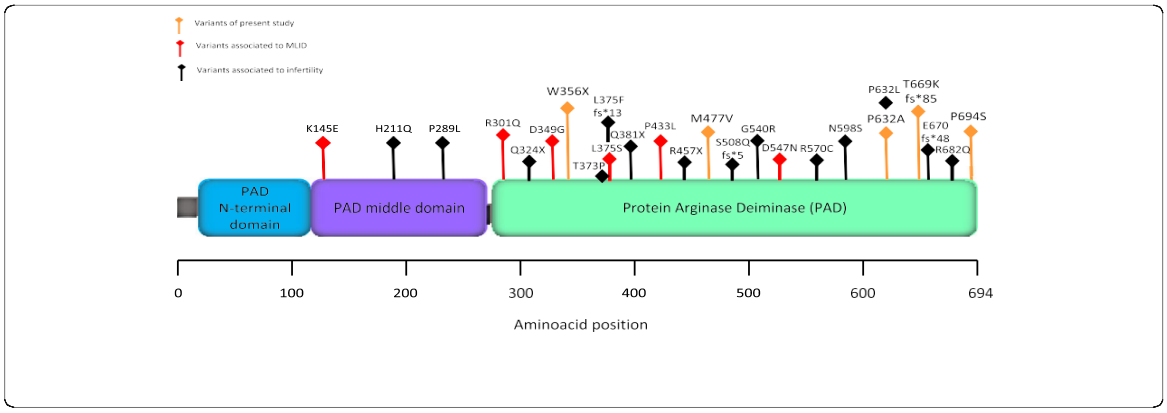

Fig. 6.4 Position of the pathogenic variants in the PADI6 protein. The variants associated with infertility of the carrier women are represented with black rhombi, and variants associated with MLID in the offspring with orange or red rhombi, if described in present or previous studies, respectively.

Discussion

DNA methylation defects of imprinting disorders arise in early embryogenesis, often as consequence of genetic mutations that act either *in cis* or *in trans*. In the latter case, MLID generally occurs in the patients [1]. Recently, a number of maternal-effect variants mostly affecting the components of the oocyte SCMC have been associated with MLID. However, the presence of heterogeneous phenotypes in the offspring of the carrier mothers makes prediction of the clinical condition and determination of recurrence risks very challenging [7]. Several studies have indicated *PADI6* variants as the cause of female infertility [16–19], but its role in imprinting disorders is less well established [20]. We report on four cases of BWS with MLID, which arise from maternal-effect loss-of-function *PADI6* mutations, and characterize their genomic methylation profiles. Our results firmly establish the role of *PADI6* in imprinting disorders and provide novel information on the recurrence risk and the extent by which these variants impact DNA methylation.

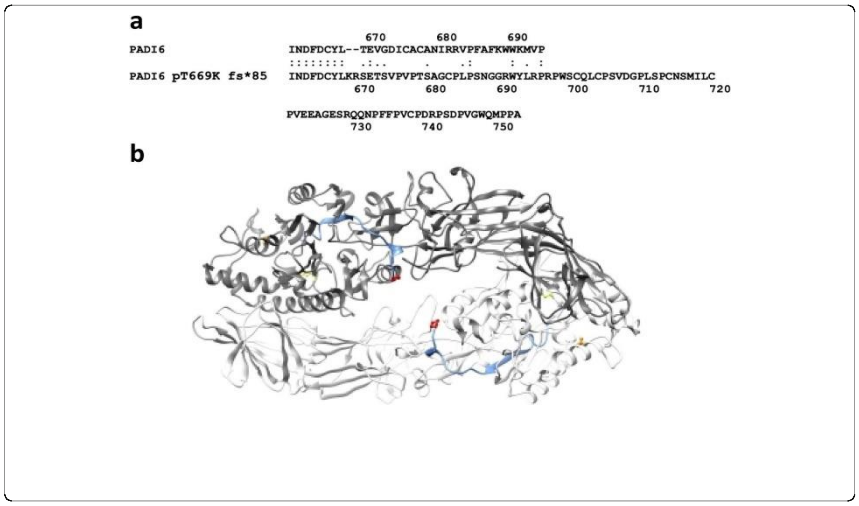

Fig. 6.5 Effect of the frameshift variant on PADI6 protein structure. a Alignment of the C-terminal residues of the wild-type PADI6 with those of the protein resulting from the T669Kfs*85 variant. b Model of human wild-type PADI6 in dimeric form. The amino acids that are colored are the sites were mutations occur. The model is represented as a cartoon, and the two chains are colored in light or dark gray, respectively. The 26-aa carboxy-terminal region of the protein that is substituted by a longer peptide in the frameshift variant T669Kfs*85 is colored in blue; the side chains of M477 (yellow), P632 (orange), and P694 (red) are shown as balls-and-sticks.

The most compelling evidence of the association of *PADI6* with BWS and MLID as maternal-effect gene is provided by our family 1. In this case, the healthy mother of two girls with BWS carries a truncating mutation and a deleterious missense variant in compound heterozygosity. The truncating variant is present also in the maternal grandmother, indicating that monoallelic inactivation of *PADI6* is not sufficient to cause the BWS phenotype in the progeny. Consistent with this observation, two damaging *PADI6* mutations have also been identified in the proband's mother of family 2. In this case, one of the variants (M477V) was predicted as benign by PolyPhen-2 but destabilizing by a structure prediction tool, and it occurs at a site that is completely conserved in orthologous PADI6 proteins and is buried in the protein structure (Table 6.1). Differently from the previous cases, the frameshift *PADI6* variant affects the proband's mother of family 3 in only one allele. This variant removes the last 26 residues of the wild-type protein and substitutes them with an 85-aa-long peptide. By generating an *in silico* model, we predicted that if PADI6 dimerizes as the paralogous PADI4, this mutant protein

would be able to heterodimerize with the wild-type isoform and possibly act as dominant-negative mutant in the heterozygous patient.

The progenies of the carrier mothers of families 2 and 3 included both affected and healthy siblings. Also, the proband's mother of family 3 showed slight *KCNQ1OT1* hypomethylation suggesting maternal inheritance of the frameshift variant, but she had no feature of BWS. Incomplete penetrance of the clinical phenotype associated with maternal variants of *PADI6* is also evident in two previously reported pedigrees [20] and is consistent with the findings on other SCMC members [20, 23]. This could be due to partial loss of protein activity resulting in mosaic and heterogeneous alteration of DNA methylation in early embryos. Therefore, although the presence of two affected siblings in family 1, and a BWS patient and a miscarriage in family 2, suggests that in some cases recurrence risk of maternal SCMC variants may be high, this is not always the case, and analysis of larger cohorts is needed for better predictions.

Our genome-wide methylation analysis revealed that the MLID associated with maternal *PADI6* variants may involve many imprinted loci. Strikingly, in the most intensely affected patient (III-1 of family 1), up to 50% of the investigated DMRs showed abnormal methylation. Interestingly, the methylation profile differed among patients and was extremely heterogeneous even in the affected siblings of family 1. Both maternally methylated and paternally methylated imprinted DMRs were found hypomethylated in the patients, suggesting that the methylation defects arise post-fertilization. Some secondary DMRs were hypermethylated, consistent with the described hierarchical control by their respective gDMR [1]. Also, both hypo- and hypermethylation were present at variable level, indicating that the methylation disturbances were mosaic.

It is possible that the methylation abnormalities of loci other than *KCNQ1OT1* may modify the typical clinical picture of BWS [22, 27–29]. Consistent with this hypothesis, we observe some clinical features (e.g., absence of macrosomia at birth) that are infrequent in BWS. It is also worth to mention that three of the four MLID cases with *PADI6* variants described so far displayed a Silver- Russell syndrome (SRS) phenotype [20]. These patients have *H19*/*IGF2*:IG-DMR LOM, and some clinical features (feeding difficulties and body asymmetry) overlapping with the

cases of the present study. This demonstrates that maternal *PADI6* variants may result in a so far unpredictable wide spectrum of DNA methylation profiles and clinical phenotypes in the offspring (Additional file 1: Table S1). These variants add further to the complexity of phenotypes associated with MLID and maternal variants so far (Additional file 8: Table S4). The great majority of these maternal variants affect SCMC protein members. However, investigation on larger cohorts is needed to establish how frequent these variants are and how large the spectra of the involved genes and phenotypes are.

Human *PADI6* variants have been associated with female infertility and hydatidiform mole [13, 16–19]. It is possible that the reproductive outcome is correlated with the severity of the *PADI6* variants. Indeed, infertility is more frequently associated with truncating variants completely disrupting the protein function or missense variants that according to SDM strongly destabilize protein structure (Additional file 1: Table S1). Conversely, MLID is more frequently associated with missense variants only partially reducing protein function (Additional file 1: Table S1). MLID-associated variants are generally non-destabilizing (e.g., P632A and P694S) or only moderately destabilizing (e.g., M477V), or a destabilizing variant is in compound heterozygosity with a non-destabilizing variant (Table 6.1 and Additional file 1: Table S1). Both P632A and P694S may have functional effect, because they are included in regions that are very conserved in all PAD proteins. In particular, P694 that is the last residue of the protein is part of the consensus sequence FxxWxMxP of the human C-terminome, which is known to have a functional role [30]. Concerning P632 variants, in addition to P632A we identified in MLID patients, the variant P632L has been associated with female infertility [18]. Since P632 is poorly exposed to solvent and buried inside the protein, substitution of proline with alanine may be less deleterious than substitution with leucine that has a longer side chain. In conclusion, severe loss-of-function variants may lead to female infertility, while hypomorphic variants may cause MLID. Intriguingly, in the present study, the relatively more severe maternal *PADI6* genotype (a truncating variant and a functionally damaging missense variant) is associated with a rare pedigree with two affected sib- lings, suggesting that severity of PADI6 variants may also influence recurrence risk.

It is unknown how *PADI6* or the other SCMC members affect methylation imprinting. It has been recently demonstrated that *KHDC3L* that is associated with recurrent hydatidiform mole is necessary for de novo methylation in human oocytes [31]. *NLRP2*, *NLRP5*, and *PADI6* that are associated with MLID and mosaic hypo-methylation of both maternal and paternal gDMRs more likely affect post-zygotic imprinting maintenance. Although it is unclear if PADI6 has catalytic activity, most of the pathogenic missense variants of *PADI6* identified so far affect its arginine deiminase domain (Fig. 6.4), indicating an important role of this domain for protein function. It is possible that imprinting disturbances arise as consequence of defective storage, de novo synthesis, cellular localization, or inappropriate degradation of several maternal-zygotic factors including those controlling DNA methylation maintenance in early embryos. Our study suggests that hypomorphic *PADI6* mutations cause MLID, while complete loss-of-function variants lead to female infertility. Thus, further functional studies reproducing the hypomorphic variants that are present in the MLID patients may help clarifying the role of SCMC in the maintenance of epigenome integrity.

Conclusions

In summary, we investigated the etiology of MLID cases associated with BWS. Using WES analysis, we identified several loss-of-function variants of *PADI6* that segregated with MLID and BWS, as maternal-effect mutations. Further, we showed that the genomic methylation profiles of the MLID patients with *PADI6* variants can be extensively altered but also very heterogeneous. Overall, our data suggest that maternal *PADI6* is necessary for methylation maintenance in early embryos and that its deficiency results in stochastic imprinted disturbances that may variably impact the health of the progeny.

Methods

Patient cohort

The four patients described in this study were part of a cohort of 72 BWS patients with *KCNQ1OT1*:TSS-DMR LOM. Of these, 18 (25%) were demonstrated to have MLID after testing 7 DMRs by pyroseq, COBRA, or MS- MLPA. In four pedigrees

we have sent out for exome sequencing so far, maternal *NLRP5* variants were identified in one family [22] and *PADI6* variants in the other 3 families (present study).

Ethics

Genetic analyses were performed after written informed consent was obtained from the patients or patients' parents. The research work was carried out in accordance with the ethical principles and the Italian legislation. The study was approved by the Ethical Committee of the University of Campania "Luigi Vanvitelli" (Naples, Italy; approval number: 1135, 13 October 2016).

Genomic DNA extraction

Genomic DNA was isolated from peripheral blood lymphocytes (PBLs) by using the salting-out procedure ([32]; Nucleic acids research). Nucleic acid concentration was determined by using NanoDrop spectrophotometer (Termo Fisher Scientific, Italy).

DNA methylation analysis

Locus-specific analysis by pyrosequencing

One and a half micrograms of PBL DNA was treated with sodium bisulfite by using the EpiTect Bisulfite kit (Qiagen- Italia, Milan, Italy) following the manufacturer's protocol. About 200 ng of converted DNA was amplified by using the PyroMark PCR kit (Qiagen-Italia, Milan, Italy) in a final volume of 25 μl. Fifteen microliters of PCR product was used for quantitative DNA methylation by pyrosequencing on a Pyromark Q48 Autoprep system with the PyroMark Q48 Adv. CpG Reagents Kit (Qiagen-Italia, Milan, Italy) and PyroMark Q48 Magnetic Beads. Results were analyzed by using the Pyromark Q48 Autoprep software. The primers, used for PCR amplification and sequencing, were designed with Pyromark Assay Design SW 2.0 and reported in Additional file 9: Table S5.

Genome-wide methylation array profiling

Genomic DNA was extracted from PBL of probands, siblings, mothers, and 12

controls (whose sex, gender, and age are listed in Additional file 10: Table S6) and subjected to bisulfite conversion methylation array processing as described previously [22]. Data was analyzed using R version 3.6.1. Beta values were extracted from "idat" files by using the "champ.load" module of the "Champ" R package (v. 2.16.2), with quality control options set as default and array type as "EPIC." The quality control step retained 736,048 probes, which were used for further analysis. Probes overlapped with SNP (96, 368) and with a detection *p* value < 0.01 (3456), bead count < 3 in at least 5% of samples in one or more samples (10,453) were eliminated. X or Y chromosome probes (16,664), non-cg probes (2918), and probes aligning to multiple locations (11) and unannotated probes were also removed from analysis. No background correction was carried out. BMIQ normalization was ap- plied, with the default options for array type as "EPIC." BMIQ normalized samples were assigned with respective genome coordinates using B4 annotation of Illumina manifest file (hg19) (Illumina Inc., USA). Since six out of twelve controls included in our study were from our pre- vious study (GSE133774), we have performed batch correction using Combat function from sva package (sva version). PCA plots (scale = TRUE) were generated to understand clustering with CpGs underlying whole genome and human ICRs. Hierarchical clustering was per- formed and visualized as heatmap using pheatmap package with "Euclidean" distance and "Ward.D" clustering method. Violin plots were created using ggplot2 (3.2.1). Heatmap was plotted using pheatmap (v.1.0.12). The raw and processed files are deposited in GEO under the accession GSE153211. Methylation level of imprinted DMRs was calculated as average of the methylation levels of their respective CpGs. Hypo- and hypermethylation of the DMRs were indicated if methylation level exceeded ± 2 standard deviation from average of 12 controls.

DNA sequencing

Whole-exome sequencing analysis

Whole-exome sequencing was performed on genomic DNA of the probands and their parents (II-1, II-2, III-1, III-2) of family 1, I-4 of family 2, and I-2 of family 3. The DNA samples were sequenced 150 bp pair-end at BIODI- VERSA srl Service (Milan, Italy) using the Agilent SureSelect V6 + UTR (~ 89 Mbp target) library and

the Illumina NovaSeq6000 platform. The bioinformatic analysis was performed as previously reported [22]. In brief: reads were aligned to the human genome reference assembly (Genome Reference Consortium Human GRCh37) using the BWAmem software package v0.7.15 [33]. PCR duplicates were filtered out by Picard v2.9 (http://picard.source forge. net), and the GATK v3.7 suite was used to locally realign around inferred Insertion/Deletions (InDels) and recalibrate base quality scores. Single-nucleotide variants and InDels were called using GATK HaplotypeCaller and GenotypeGVCFs [34] and recalibrated with VariantRecalibrator. Recalibrated variants were annotated using wAN- NOVAR [35]. Genome variants with low coverage (< 15) or low quality (< 20) or in VQSRTran- cheSNP99.00to99 or frequently occurring in general population (MAF > 0.01 in 1000 Genomes Project [36] or Exac (http://exac.broadinstitute.org/) [37] or gnomAD [38]) were filtered out.

The effect of the variants was predicted using the sequence-based tool PolyPhen-2 (http://genetics.bwh. harvard.edu/pph2) [39], whose scores correlate with the residual activity of the protein affected by the mutation [40], as well as Site Directed Mutator (SDM) that analyzes if the variants occurring at specific structural environment are tolerated within the family of homologous [41]. Template-based modeling was used to obtain a 3D structural model of human PADI6, due to the lack of crystal structure for this protein. The sequence was retrieved from the UniProt database (UniProt ID: Q6TGC4). Residues 10–694 of PADI6 could be aligned to protein-arginine deiminase type-4 (PADI4). A 3D model of a PADI6 dimer was built with SWISS-MODEL [42] in the automated mode. The program generated a list of the best possible templates. We chose 4dkt [43], which is the X-ray structure of PADI4, and built the model. Amino acids from 10 to the carboxy-terminus 694 could be modeled. The accessible surface area (ASA) was calculated with the protein interfaces, surfaces, and assemblies service PISA at the European Bio- informatics Institute [44].

Sanger sequencing

About 100 ng of genomic DNA was amplified by PCR and then sequenced by Sanger sequencing (Eurofins Genomics) to obtain the genotypes. PCR primers are

listed in Additional file 11: Table S7.

Web resources

The following web resources were used:

PolyPhen-2, http://genetics.bwh.harvard.edu/pph2 GnomAD, http://gnomad.broadinstitute.org/ SDM, http://marid.bioc.cam.ac.uk/sdm2 PISA, http://www.ebi.ac.uk/pdbe/prot_int/pistart.html

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13148-020-00925-2.

Additional file 1: Table S1. Pathogenic variants of *PADI6* and corresponding clinical phenotype.

Additional file 2: Figure S1. Pyrosequencing analysis of seven imprinted DMRs.

Additional file 3 Figure S2. Batch effect adjustment of array datasets.

Additional file 4: Figure S3. Violin plots showing whole-genome DNA methylation profiles of probands, their siblings and mothers, and 12 controls.

Additional file 5: Table S2. Methylation defects of imprinted DMRs in patients affected by BWS-MLID whose mothers are carriers of loss of function mutations in *PADI6*.

Additional file 6: Table S3. Maternal genetic variants in homozigosity or compound heterozigosity identified by WES.

Additional file 7: Figure S4. Sanger sequencing validating the *PADI6*

variants identified by WES.

Additional file 8: Table S4 . List of the maternal-effect gene variants associated with MLID in the offspring identified so far and corresponding clinical phenotype.

Abbreviations

BWS: Beckwith-Wiedemann syndrome; DMRs: Differentially methylated regions; gDMR: Germline-derived differentially methylated region; GOM: Gain of methylation; ICs: Imprinting centers; IDs: Imprinting disorders; LOM: Loss of methylation; MLID: Multi-locus imprinting disturbance; PBL: Peripheral blood leukocyte; PCA: Principal component analysis; SCMC: Subcortical maternal complex; SRS: Silver-Russell syndrome; SNV: Single-nucleotide variant; TNDM: Transient neonatal diabetes mellitus; UPD: Mosaic segmental paternal unidisomy;

WES: Whole-exome sequencing

Acknowledgements

The authors are grateful to the patients and their family members and the Italian BWS patient support group AIBWS for their collaboration.

Authors' contributions

LP, AS, SR, LC, and RDP carried out most of the molecular analyses of the patients. MVC and MM carried out the exome-seq analysis and identified the PADI6 variants. AV performed the bioinformatic analysis of the methylome array. VA, RT, and DM identified the families and made the clinical diagnosis. LP, AS, FC, and AR carried out the study design, data analysis, interpretation, manuscript writing, and revision. All authors read and approved the final manuscript.

Funding

The work was supported by the followings grants: INCIPIT H2020-MSCA-COFUND grant agreement - N. 665403 project (AR and AV), Associazione Italiana Ricerca sul Cancro IG 2016 N.18671 (AR), Telethon-Italia GGP15131 (AR), "Progetti per la ricerca oncologica della Regione Campania" Grant: I-Cure (AR and FC), "Progetti competitivi intraAteneo" Programma V:ALERE (VAnviteLli pEr la RicErca) 2019 – grant MIRIAM from Università degli Studi della Campa- nia "Luigi Vanvitelli" (AR, FC and AS).

Availability of data and materials

Methylation array data that support the findings of this study have been deposited under accession code GSE153211 in the Gene Expression Omnibus repository.

Ethics approval and consent to participate

Approval to conduct this study was obtained from the Ethical Committee of the University of Campania "Luigi Vanvitelli" together with the signed informed consent from parents of the individuals involved in this study.

Consent for publication

Consent for publication of clinical and molecular data was obtained from parents of the individuals involved in this study.

Competing interests

The authors declare that they have no potential competing interests to disclose. The authors declare that they have no financial relationships relevant to this article to disclose.

Author details

[1]Department of Biology, Università degli Studi di Napoli "Federico II", Napoli, Italy.

[2]Department of Environmental Biological and Pharmaceutical Sciences and Technologies (DiSTABiF), Università degli Studi della Campania "Luigi Vanvitelli", Caserta, Italy.

[3]Institute of Genetics and Biophysics (IGB) "Adriano Buzzati-Traverso", Consiglio Nazionale delle Ricerche (CNR), Naples, Italy.

[4]Medical Cytogenetics and Molecular Genetics Laboratory, Centro di Ricerche e Tecnologie Biomediche IRCCS, Istituto Auxologico Italiano, Milan, Italy.

[5]Department of Sciences for Health Promotion and Mother and Child Care "G. D'Alessandro", University of Palermo, Palermo, Italy.

[6]Medical, Surgical, and Dental Department, Università degli Studi di Salerno, Salerno, Italy.

[7]Department of Pediatrics, Clinical Genetics, Università di Padova, Padova, Italy.

Received: 23 June 2020 Accepted: 25 August 2020

References

1. Monk D, Mackay DJG, Eggermann T, Maher ER, Riccio A. Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. Nat Rev Genet. 2019;20(4):235–48.

2. Carli D, Riberi E, Ferrero GB, Mussa AJ. Syndromic disorders caused by

disturbed human imprinting. Clin Res Pediatr Endocrinol. 2020;12(1):1–16.

3. Mussa A, Russo S, De Crescenzo A, Chiesa N, Molinatto C, Selicorni A, et al. Prevalence of Beckwith-Wiedemann syndrome in North West of Italy. Am J Med Genet A. 2013;161A(10):2481–6.

4. Brioude F, Kalish JM, Mussa A, Foster AC, Bliek J, Ferrero GB, et al. Expert consensus document: clinical and molecular diagnosis, screening and management of Beckwith-Wiedemann syndrome: an international consensus statement. Nat Rev Endocrinol. 2018;14(4):229–49.

5. Sanchez-Delgado M, Riccio A, Eggermann T, Maher ER, Lapunzina P, Mackay D, et al. Causes and consequences of multi-locus imprinting disturbances in humans. Trends Genet. 2016;32(7):444–55.

6. Cerrato F, Sparago A, Ariani F, Brugnoletti F, Calzari L, Coppedè F, et al. DNA methylation in the diagnosis of monogenic diseases. Disturbed genomic imprinting and its relevance for human reproduction: causes and clinical consequences. Genes. 2020;11(4):355.

7. Elbracht M, Mackay D, Begemann M, Kagan KO, Eggermann T. Disturbed genomic imprinting and its relevance for human reproduction: causes and clinical consequences. Hum Reprod Update. 2020;26(2):197–213.

8. Mackay DJ, Callaway JL, Marks SM, White HE, Acerini CL, Boonen SE, et al. Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. Nat Genet. 2008; 40(8):949–51.

9. Lu X, Gao Z, Qin D, Li L. a maternal functional module in the mammalian oocyte-to-embryo transition. Trends Mol Med. 2017;23(11):1014–23.

10. Witalison EE, Thompson PR, Hofseth LJ. Protein arginine deiminases and associated citrullination: physiological functions and diseases associated with dysregulation. Curr Drug Targets. 2015;16(7):700–10.

11. Raijmakers R, Zendman AJ, Egberts WV, Vossenaar ER, Raats J, Soede-Huijbregts C, et al. Methylation of arginine residues interferes with citrullination by peptidylarginine deiminases in vitro. J Mol Biol. 2007;367(4): 1118–29.

12. Li L, Baibakov B, Dean J. A subcortical maternal complex essential for preimplantation mouse embryogenesis. Dev Cell. 2008;15(3):416–25.

13. Qian J, Nguyen NMP, Rezaei M, Huang B, Tao Y, Zhang X, et al. Biallelic PADI6 variants linking infertility, miscarriages, and hydatidiform moles. Eur J Hum Genet. 2018;26(7):1007–13.

14. Yurttas P, Vitale AM, Fitzhenry RJ, Cohen-Gould L, Wu W, Gossen JA, et al. Role for PADI6 and the cytoplasmic lattices in ribosomal storage in oocytes and translational control in the early mouse embryo. Development. 2008; 135(15):2627–36.

15. Esposito G, Vitale AM, Leijten FP, Strik AM, Koonen-Reemst AM, Yurttas P, et al. Peptidylarginine deiminase (PAD) 6 is essential for oocyte cytoskeletal sheet formation and female fertility. Mol Cell Endocrinol. 2007;273(1-2):25–31.

16. Xu Y, Shi Y, Fu J, Yu M, Feng R, Sang Q, et al. Mutations in PADI6 cause female infertility characterized by early embryonic arrest. Am J Hum Genet. 2016;99(3):744–52.

17. Maddirevula S, Coskun S, Awartani K, Alsaif H, Abdulwahab FM, Alkuraya FS. The human knockout phenotype of PADI6 is female sterility
caused by cleavage failure of their fertilized eggs. Clin Genet. 2017; 91(2):344–5.

18. Wang X, Song D, Mykytenko D, Kuang Y, Lv Q, Li B, et al. Novel mutations in genes encoding subcortical maternal complex proteins may cause human embryonic developmental arrest. Reprod BioMed Online. 2018;36(6): 698–704.

19. Zheng W, Chen L, Dai J, Dai C, Guo J, Lu C, et al. New biallelic mutations in PADI6 cause recurrent preimplantation embryonic arrest characterized by direct cleavage. Assist Reprod Genet. 2020;37(1):205–12.

20. Begemann M, Rezwan FI, Beygo J, Docherty LE, Kolarova J, Schroeder C, et al. Maternal variants in NLRP and other maternal effect proteins are associated with multilocus imprinting disturbance in offspring. J Med Genet. 2018;55(7):497–504.

21. Meyer E, Lim D, Pasha S, Tee LJ, Rahman F, Yates JR, et al. Germline mutation in NLRP2 (NALP2) in a familial imprinting disorder (Beckwith-Wiedemann syndrome). PLoS Genet. 2009;5(3):e1000423.

22. Sparago A, Verma A, Patricelli MG, Pignata L, Russo S, Calzari L, et al. The phenotypic variations of multi-locus imprinting disturbances associated with maternal-effect variants of NLRP5 range from overt imprinting disorder to

apparently healthy phenotype. Clin Epigenetics. 2019;11(1):190.

23. Docherty LE, Rezwan FI, Poole RL, Turner CL, Kivuva E, Maher ER, et al. Mutations in NLRP5 are associated with reproductive wastage and multilocus imprinting disorders in humans. Nat Commun. 2015;6:8086.

24. Zink F, Magnusdottir DN, Magnusson OT, Walker NJ, Morris TJ, Sigurdsson A, et al. Insights into imprinting from parent-of-origin phased methylomes and transcriptomes. Nat Genet. 2018;50(11):1542–52.

25. Monk D, Morales J, den Dunnen JT, Russo S, Court F, Prawitt D, et al. Recommendations for a nomenclature system for reporting methylation aberrations in imprinted domains. Epigenetics. 2018;13(2):117–21.

26. Liu YL, Chiang YH, Liu GY, Hung HC. Functional role of dimerization of human peptidylarginine deiminase 4 (PAD4). PLoS One. 2011;6(6):e21314.

27. Bakker B, Sonneveld LJ, Woltering MC, Bikker H, Kant SGJ. A girl with Beckwith-Wiedemann syndrome and pseudohypoparathyroidism type 1b due to multiple imprinting defects. Clin Endocrinol Metab. 2015;100(11): 3963–6.

28. Sano S, Matsubara K, Nagasaki K, Kikuchi T, Nakabayashi K, Hata K, et al. Beckwith-Wiedemann syndrome and pseudohypoparathyroidism type Ib in a patient with multilocus imprinting disturbance: a female-dominant phenomenon? Hum Genet. 2016;61(8):765–9.

29. Bliek J, Verde G, Callaway J, Maas SM, De Crescenzo A, Sparago A, et al. Hypomethylation at multiple maternally methylated imprinted regions including PLAGL1 and GNAS loci in Beckwith-Wiedemann syndrome. Eur J Hum Genet. 2009;17(5):611–9.

30. Sharma S, Schiller MR. The carboxy-terminus, a key regulator of protein function. Crit Rev Biochem Mol Biol. 2019;54(2):85–102.

31. Demond H, Anvar Z, Jahromi BN, Sparago A, Verma A, Davari M, et al. A KHDC3L mutation resulting in recurrent hydatidiform mole causes genome- wide DNA methylation loss in oocytes and persistent imprinting defects post-fertilisation. Genome Med. 2019;11(1):84.

32. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3): 1215. https://doi.org/10.1093/nar/16.3.1215.

33. Li H, Durbin R. Fast and accurate short read alignment with BurrowsWheeler transform. Bioinformatics. 2009;25(14):1754–60.

34. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491–8.

35. Chang X, Wang K. wANNOVAR: annotating genetic variants for personal genomes via the web. J Med Genet. 2012;49(7):433–6.

36. Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56–65.

37. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016; 536(7616):285–91.

38. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv. 2019:531210 https://doi.org/10.1101/531210.

39. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248–9.

40. Cimmaruta C, Citro V, Andreotti G, Liguori L, Cubellis MV, Mele BH. Challenging popular tools for the annotation of genetic variations with a real case, pathogenic mutations of lysosomal alpha-galactosidase. BMC Bioinformatics. 2018;19:433.

41. Pandurangan AP, Ochoa-Montaño B, Ascher DB, Blundell TL. SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Res. 2017;45:W229–35.

42. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. Swiss-model: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46:W296–303.

43. Jones JE, Slack JL, Fang P, Zhang X, Subramanian V, Causey CP. Synthesis and screening of a haloacetamidine containing library to identify PAD4 selective

inhibitors. ACS Chem Biol. 2012;7:160–5.

44. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372:774–97.

7. PHARMACOLOGICAL CHAPERONES: A THERAPEUTIC APPROACH FOR DISEASES CAUSED BY DESTABILIZING MISSENSE MUTATIONS

International Journal of
Molecular Sciences

MDPI

Review

Pharmacological Chaperones: A Therapeutic Approach for Diseases Caused by Destabilizing Missense Mutations

Ludovica Liguori [1,2], Maria Monticelli [3], Mariateresa Allocca [1,2], Bruno Hay Mele [4], Jan Lukas [5,6], Maria Vittoria Cubellis [2,3,*] and Giuseppina Andreotti [2]

[1] Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università degli Studi dellaCampania "Luigi Vanvitelli", 81100 Caserta, Italy; lud.liguori@gmail.com (L.L.); mariateresa.allocca@gmail.com (M.A.)

[2] Istituto di Chimica Biomolecolare–CNR, 80078 Pozzuoli, Italy; gandreotti@icb.cnr.it

[3] Dipartimento di Biologia, Università Federico II, 80126 Napoli, Italy; maria.monticelli@yahoo.com

[4] Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy; bruno.haymele@unina.it

[5] Translational Neurodegeneration Section "Albrecht-Kossel", Department of Neurology, University Medical Center Rostock, University of Rostock, 18147 Rostock, Germany; jan.lukas@med.uni-rostock.de

[6] Center for Transdisciplinary Neurosciences Rostock (CTNR), University Medical Center Rostock,University of Rostock, 18147 Rostock, Germany

***** Correspondence: cubellis@unina.it; Tel.: +39-081-679118; Fax: +39-081-679233

Received: 12 December 2019; Accepted: 9 January 2020; Published: 13 January 2020

Abstract: The term "pharmacological chaperone" was introduced 20 years ago. Since then the approach with this type of drug has been proposed for several diseases, lysosomal storage disorders representing the most popular targets. The hallmark of a pharmacological chaperone is its ability to bind a protein specifically and stabilize it. This property can be beneficial for curing diseases that are associated with protein mutants that are intrinsically active but unstable. The total activity of the affected proteins in the cell is lower than normal because they are cleared by the quality control system. Although most pharmacological chaperones are reversible

competitive inhibitors or antagonists of their target proteins, the inhibitory activity is neither required nor desirable. This issue is well documented by specific examples among which those concerning Fabry disease. Direct specific binding is not the only mechanism by which small molecules can rescue mutant proteins in the cell.These drugs and the properly defined pharmacological chaperones can work together with different and possibly synergistic modes of action to revert a disease phenotype caused by an unstable protein.

Keywords: disease; pharmacological chaperones; low molecular weight drugs; protein stability; drug repositioning; lysosomal storage disorders

Introduction

Pharmacological chaperones have ceased to be a niche category and have entered the clinical practice for some rare diseases caused primarily by protein instability. Many reviews exist to cover this subject with different points of view (a few examples [1–15]). We wanted to contribute by providing the readers with a list of research papers organized per disease that covers the years from 2000 to 2018. Different protein targets have been addressed, although enzymes associated with lysosomal storage disorders represent altogether half of the cases and most pharmacological chaperones that have been proposed are their reversible specific inhibitors. It is difficult to write a comprehensive report hence we chose to use the example of Fabry disease for which a drug has recently been approved, as a *leitmotiv* to put in evidence certain concepts that we believe require some clarification. We discussed a few papers to emphasize successes, stress limits, and indicate possible solutions.

Even in the most favorable cases, such as that of the drug approved for Fabry disease, inhibitors are not the ideal drugs as they can be able to stabilize their target proteins but might not be able to fullyrevert the disease phenotype. It has been proposed to modify first generation pharmacological chaperone to enhance their therapeutic effects.

Specifically, binding and stabilization of a protein target define a pharmacological chaperone, but a certain degree of confusion exists about the assignment of this term.

Other small molecules thatrescue mutant proteins in the cell without direct binding. Beyond the correctness of the definition, the possibility of employing different types of small molecules in synergy with pharmacological chaperones can potentiate their therapeutic effect.

Pharmacological Chaperones: The Time-Lapse

The term "pharmacological chaperone" was introduced by Morello and co-workers in 2000 [16] to define the action of a specific antagonist of the receptor of vasopressin. Missense mutations of the receptor cause Nephrogenic Diabetes Insipidus. If the antagonist is administered to cells carrying a mutant form of the receptor that is unable to reach the cell surface, it favors the accumulation of the mature protein. The authors wrote that molecules like the antagonist of vasopressin "would act as pharmacological chaperones that promote receptor processing through their specific binding activity" [16].

Specific binding is the hallmark that serves to distinguish pharmacological chaperones from other small molecules that can be useful in the treatment of diseases caused by unstable proteins. Since the pioneering work of Morello, the fortune of the term raised and the number of research articles mentioning "pharmacological chaperone" per year increased significantly. The approach with pharmacological chaperones was extended to other diseases.

We looked in UniProt [17] for the human proteins associated with MalaCards [18], the human disease database. We extracted the names of the diseases associated with each entry from the annotation in UniProt. The list (column 2 in Table S1) is partially redundant because we wanted to take into account that a given disease can be known with different names. For instance mutations affecting the Uniprot entry P10253 cause a disease known with different names, acid maltase deficiency, alpha-1,4-glucosidase deficiency, and, cardiomegalia glycogenica, gaa deficiency, glycogen storage disease, glycogenosis ii or Pompe disease. We queried Scopus to search for articles that contained the term "pharmacological chaperone" and one of the disease names extracted from Uniprot in either the title, the abstract or the keywords, respectively. In Figure 7.1 we provide a histogram describing the distribution of the researchpapers per year.

The association of the papers with specific target proteins required manual curation of data. Those papers for which the association with the protein target was not evident in the abstract were excluded from the analysis. For the others, we added the annotations of the protein target, i.e., the functional type, the localization (Table 7.1).

Disease	**Gene**	**Uniprot Code**	**Protein Type**	**Subcellular Location**	**N. Of Articles**
Gaucher	*GBA*	P04062	enzyme	lysosome	64
Fabry	*GLA*	P06280	enzyme	lysosome	42
GM-1, Morqio B	*GLB1*	P16278	enzyme	lysosome	16
Pompe	*GAA*	P10253	enzyme	lysosome	14
Cystic fibrosis	*CFTR*	P13569	transporter	plasma membrane	14
Retinitis pigmentosa	*RHO*	P08100	receptor	plasma membrane	12
Phenylketonuria	*PAH*	P00439	enzyme	cytosol	9
Krabbe disease	*GALC*	P54803	enzyme	lysosome	9
Nephrogenic diabetes insipidus	*V2R*	P30518	receptor	plasma membrane	8
Long QT Syndrome	*KCNH2*	Q12809	transporter	plasma membrane	7
Parkinson	*PARK7*	Q99497	enzyme	plasma membrane, nucleus,mitochondrion	5
Niemann-Pick	*NPC1*	O15118	receptor	lysosome	5
Hyperoxaluria	*AGXT*	Q86XE5	enzyme	mitochondrion	5
Obesity	*MC4R*	P32245	receptor	plasma membrane	4
GM-2, Sanfilippo syndrome	*GNRHR*	P07686	enzyme	lysosome	4
GM-2, Tay-Sachs syndrome	*HEXB*	P06865	enzyme	lysosome	4
Galactosemia	*HEXA*	P07902	enzyme	cytosol	4
Hypoparathyroidism	*PTH*	P01270	hormone	extracellular or secreted	3
Parkinson	*GALT*	P04062	enzyme	lysosome	3
Hypogonadotropic hypogonadism	*ATP7B*	P30968	receptor	plasma membrane	2
Wilson	*PMM2*	P35670	transporter	Golgi apparatus	2

PMM2-CDG	*SLC26A4*	O15305	enzyme	cytosol	2
Pendred	*MMAB*	O43511	transporter	plasma membrane	2
Methylmalonic aciduria	*ABCB4*	Q96EY8	enzyme	mitochondrion	2
Intrahepatic cholestasis	*DRD4*	P21439	transporter	plasma membrane	2
Hyperactivity disorder	*ABCC8*	P21917	receptor	plasma membrane	2
Diabetes	*GPR56*	Q09428	receptor	plasma membrane	2
Polymicrogyria	*PGK1*	Q9Y653	receptor	plasma membrane, extracellular or secreted	1
Phosphoglycerate kinase 1 deficiency	*SNCA*	P00558	enzyme	cytosol	1
Parkinson	*SUMF1*	P37840	regulator	presynaptic vesicle	1
Multiple sulfatase deficiency	*NPM*	Q8NBK3	enzyme	E.R.	1
Leukemia	*PKR2*	P06748	regulator	nucleus, cytoskeleton	1
Intrahepatic cholestasis	*ABCB11*	O95342	transporter	plasma membrane	1
Nocturnal frontal lobe epilepsy	*CHRNB2*/*CHRNA4*	P17787/P43681	transporter	plasma membrane	1
Hypomagnesemia	*CLDN16*	Q9Y5I7	transporter	plasma membrane	1
Creutzfeld-Jacob, Kuru	*PRNP*	P04156	unclear/prion	plasma membrane	1
Homocystinuria	*CBS*	P35520	enzyme	nucleus	1
Fibrodysplasia ossificans	*ACVR1*	Q04771	enzyme	plasma membrane	1
Epilepsy, Migraine	*SCN1A*	P35498	transporter	plasma membrane	1
Dystonia	*SLC2A1*	P11166	transporter	plasma membrane	1
Diarrhea (cholera toxin)	*NHE3*	P48764	transporter	plasma membrane	1
Diabetes	*KCNJ11*	Q14654	transporter	plasma membrane	1
Intrahepatic cholestasis	*ATP8B1*	O43520	transporter	plasma membrane, Golgiapparatus, E.R.	1
Breast cancer	*NBS*	O60934	regulator	nucleus	1
Amyotrophic lateral sclerosis	*SOD1*	P00441	enzyme	nucleus, mitochondrion	1
Amyloidosis	*VPS29*/*VPS26*	Q9UBQ0/O75436	transporter	endosome	1
Allan-Herndon-Dudley	*SLC16A2*	P36021	transporter	plasma membrane	1
Alkaptonuria	*HGD*	Q93099	enzyme	cytosol	1
Aspartylglucosaminuria	*AGA*	P20933	enzyme	lysosome	1

Ceroid lipofuscinosis	*PPT1*	P50897	enzyme	lysosome	1
Schindler disease	*NAGA*	P17050	enzyme	lysosome	1
Diabetes mellitus	*IAPP*	P10997	hormone	extracellular or secreted	1
GM-1	*IDS*	P22304	enzyme	lysosome	1
Morquio A, Hunter disease	*GALNS*	P34059	enzyme	lysosome	1

Table 7.1. Distribution of research papers per disease and per protein target. The table summarizes how many articles use the term "pharmacological chaperone" andthe corresponding disease context (we chose arbitrarily what we considered the most representative name of the disease). The UniProt entry of the affected protein, the name of the gene, the protein type, and the protein localization are also shown.

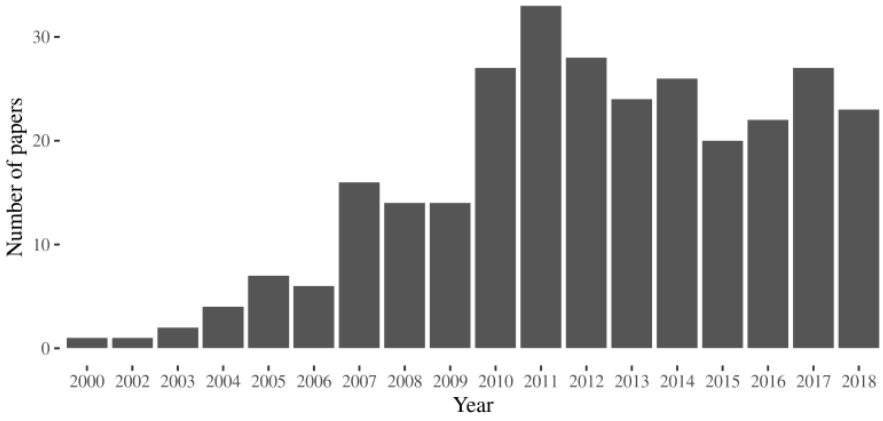

Figure 7.1. The number of the research articles indexed in Scopus that cite the term "pharmacological chaperone" in the title, in the abstract or the keywords and a specific disease.

Pharmacological Chaperones: Rational and Application

Pharmacological chaperones have been tested for many diseases but the distribution of papers is not even. Approximately one-fourth of them concern Gaucher disease, and more than half the lysosomal storage disorders (LSD: Gaucher, Fabry, GM-1, Pompe) all together (Figure 7.2). Besides LSDs, the diseases most frequently associated with pharmacological chaperones are cystic fibrosis, retinitis pigmentosa,phenylketonuria, Parkinson.

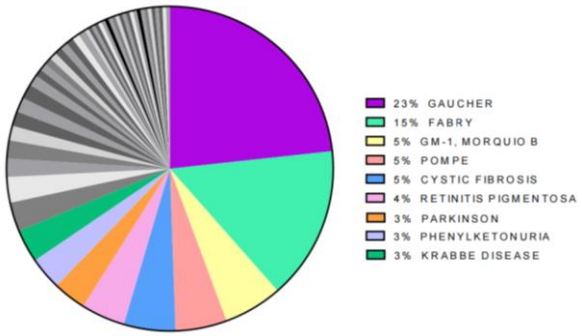

Figure 7.2. Distribution of papers citing "pharmacological chaperone" per disease. The less representeddiseases are in gray.

Protein targets of pharmacological chaperone can be diverse. Enzymes, in particular transferases, represent the most populated group (45%) followed by transporters (28%) and receptors (15%) (Figure 7.3).

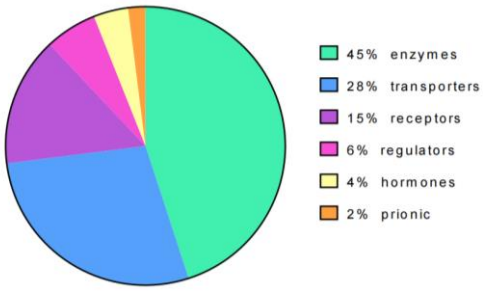

Figure 7.3. Protein targets of pharmacological chaperones based on their function.

The proteins targeted by pharmacological chaperones can have different subcellular localizationand, in the majority of the cases, are translocated across the endoplasmic reticulum (Figure 7.4).

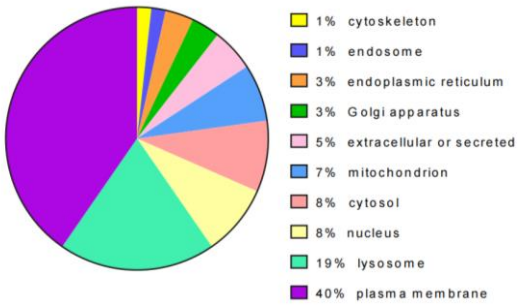

Figure 7.4. Distribution of protein targets of pharmacological chaperones per cell localization.

Pharmacological Chaperones: Direct Specific Binding to Folded or Partially Folded TargetProteins as a Mechanistic Hallmark

Pharmacological chaperones are low molecular weight chemical molecules and should not be confused with molecular chaperones that are proteins [19,20]. According to the original definition by Morello and co-workers [16], pharmacological chaperones exert their action by binding specifically their target proteins. Direct specific binding differentiates pharmacological chaperones from other low molecular weight chemical molecules, such as chemical chaperones and regulators of protein homeostasis, that promote correct processing of a pathological mutant by different mechanisms.

In principle, direct binding and stabilization should be tested on purified mutants. This is unpractical for mutants that have disulfide bridges and post-translational modifications such as thosethat are secreted, that are resident in membranes (e.g., plasma membrane) or the endomembrane system (e.g., lysosomes, Golgi, endoplasmic reticulum or endosomes) because they cannot be produced and purified with a high yield. They represent the majority of the cases because proteins that transit in the endoplasmic reticulum are the most frequent targets of pharmacological chaperones (Figure 7.4). However, although the effect of pharmacological chaperones is more evident on responsive mutants, it can be observed on the corresponding wild type proteins too (an example is providedin [21] and is shown in Figure 7.5). This means that a preliminary screening can be run on the wild type protein.

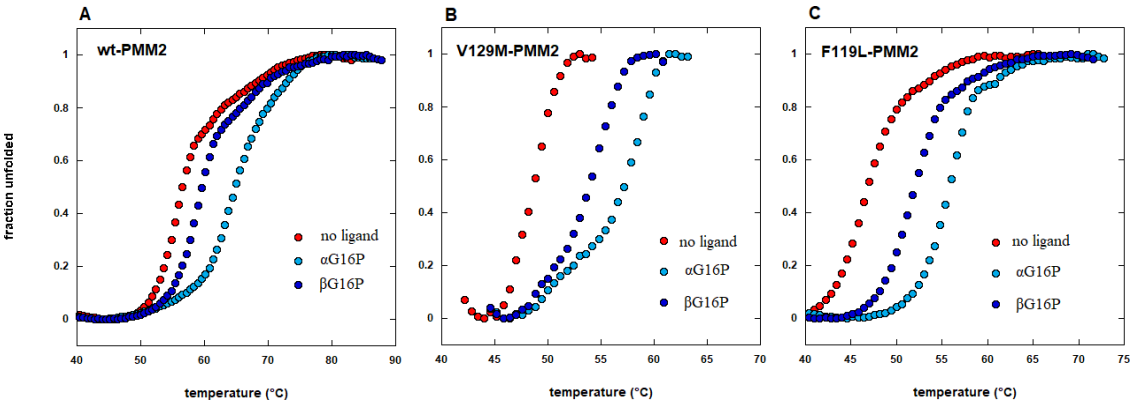

Figure 7.5. The effect of pharmacological chaperones is observed on responsive mutants as well as on wild type proteins. Thermal shift analysis of wild type phosphomannomutase2 (wt-PMM2 in panel A) and two pathological mutants (V129M-PMM2 in panel B and F119L-PMM2 in panel C) in the absence or in the presence of two different pharmacological chaperones, 1,6-alpha glucose-bisphosphate (αG16P)or 1,6-beta glucose-bisphosphate (βG16P) (extracted from reference [21]).

Direct protein stabilization can be proved by chemical-physical techniques such as differential scanning calorimetry [22], by circular dichroism, *i.e.*, measures of ellipticity as a function of temperature [23], or as a function of denaturant concentration [24]. Recently thermal shift assay has gained great popularity because it can process many samples at the same time; it requires small amounts of protein and equipment that is available in the majority of biomedical laboratories [25–28]. Incubation at a relatively high temperature of purified wild type protein targets in the presence or the absence of drugs followed by the measurement of residual enzymatic activity was used for the screening of relatively large libraries [29]. In some cases, the stabilizing effect of chaperones was tested *in vitro* on mutant proteins that were present in cellular extracts. Often the extracts were incubated in the absence or the presence of the chaperone at different temperatures and then assayed under standard conditions to measure the residual activity [30]. Alternatively, the extracts were incubated atdifferent concentrations of urea, denatured proteins were selectively degraded by limited proteolysis and folded undigested proteins were detected by Western blot [24].

Pharmacological Chaperones: The Quick Path to Success is not Always the Best One

Most pharmacological chaperones have been discovered among reversible competitive inhibitors and have been tested for their stabilizing effect with the techniques mentioned beforehand. In this paper, we would like to stress on the prevalence of inhibitors among pharmacological chaperones occurred for practical reasons since many targets are enzymes (Figure 7.3). Competitive inhibitors often resemble chemically known substrates or products and an activity-based screening is straightforward. Indeed, competitive inhibitors are ligands that bind preferentially the folded state of an enzyme, and for this reason, they stabilize the protein. However, the stabilizing activity is required for chaperones, whereas the inhibitory activity is neither required nor desirable. Hence the concept of chaperone andthat of inhibitor should not be assimilated.

The development of drugs for Fabry disease well illustrates this issue.

Fabry disease is an X-linked lysosomal storage disorder due to mutations in the gene *GLA* that encodes acid alpha-galactosidase A (AGAL). The enzyme cleaves globotriaosylceramide, generatinggalactose that is a product and a reversible inhibitor of the enzyme. There exist more than 400 pathological missense mutations, the majority of which reduce the stability of the protein but do not affect the active site [31,32]. In 1995, some years before Morello et al. had coined the term "pharmacological chaperone," Okumya et al. [33] showed that galactose stabilizes missense mutants of alpha-galactosidase. They showed that administering 100 mM galactose to cells expressing certain missense mutants, a greater amount of the protein could be detected by Western blot and consequently a higher activity could be measured in the extracts. The authors explained the effect suggesting that the ligand stabilized responsive mutant AGAL. A few years later, Frustaci et al. [34] reported that infusions of galactose (1 g per kg body weight), improved the cardiac function of a Fabry patient. Interestingly the therapywas administered every other day. Although galactose cannot be considered a practicable therapeutic agent because of the high dosage needed to measure the effect *in vitro* and *in vivo*, the road was marked toward the usage of pharmacological chaperones for the treatment of Fabry patients carryingmissense destabilizing mutations. In 1986 1,5-dideoxy-

1,5-iminogalactitol, which is an iminosugar analog of galactose where the endocyclic oxygen is substituted by nitrogen, was synthesized [35,36]. The IUPAC name of that molecule is (2R,3S,4R,5S)-2-(Hydroxymethyl)-3,4,5-piperidinetriol, the usual name is 1-deoxygalactonojirimicin abbreviated in DGJ, but other names have also been employed, 1,5-dideoxy-1,5-imino-D-galactitol,1-deoxygalactonojirimycin, 1-deoxygalactostatin, Amigal, Migalastat, Galafold (trademark). DGJ was considered a good candidate as a practicable therapeutic agent because its K_i is at least two or three orders of magnitude lower than that of galactose [37]. In 1999 it was demonstrated that DGJ is more potent than galactose as a stabilizer of responsive AGAL mutants [30]. When DGJ was administered at 20 micromolar for four days to the R301Q or Q279E lymphoblasts from Fabry patients, it raised the residual activity in the cell lysates by seven–eight folds. Although Fan and co-workers showed that DGJ binds R301Q *in vitro* and correctly interpreted the mode of action of the drug, they proposed "the concept of 'chemical chaperon' to describe a small molecule whose function is to assist a protein to fold properly and enter normal processing pathway smoothly." The term "chemical chaperon" for specific stabilizing ligands would have been soon overtaken by "pharmacological chaperone."

Since the pioneering work by Fan and co-workers, the effect of DGJ has been tested on hundreds of missense mutations either on cells derived from patients or on transfected cells. To overcome the difficulty of obtaining lymphoblasts or fibroblasts from patients, Benjamin and co-workers [38] have developed a cell-based assay in HEK-293 cells that consents to test any mutations. After having been transiently transfected with expression vectors encoding *GLA* mutants, the HEK-293 cells are exposed to the drug, washed, and lysed. The amount of AGAL is quantified in the extracts by Western blot and standard enzymatic assays with a fluorogenic substrate. Almost 50% of the missense mutations tested so far are responsive to DGJ [39]. However, activity in cell extracts should not be confused withintracellular activity. AGAL dosage on cell extracts, in transfected cells as well as in cells derived from patients, represents a simple and effective way to demonstrate the stabilization of responsive mutantsbut does not prove that AGAL activity is increased in the presence of the drug *in vivo*.

DGJ is a potent inhibitor of AGAL, it binds and stabilizes the enzyme at the neutral pH as well as at acidic pH, ($K_{DpH7} = 3.7 \times 10^{-9}$ mol/L, $K_{DpH5} = 6.0 \times 10^{-9}$ mol/L) [37]. It is unlikely that DGJ stabilizes AGAL in ER without inhibiting it in the lysosomes. A proof of the double intracellular activity of the drug, the beneficial AGAL stabilization on one side and the deleterious AGAL inhibition on the other side, comes from the experiment carried out by Valenzano and co-workers [40]. They treated the Fabry fibroblasts carrying the responsive mutations R301Q or L300P with DGJ either with a discontinuous regimen, seven days followed by a three-day washout, or for a continuous regimen, ten days without washout. They observed augmented intracellular AGAL activity, namely the clearance of the AGAL substrate that accumulates in the cells of the patients, only when the treatment with the drug was discontinuous. The result was confirmed in a mouse model where a "4 on/3 off" regimen was more effective than a daily administration of DGJ to reduce the substrate of AGAL [41]. The benefits of an intermittent administration of DGJ can be explained by the fact that the half-life of most AGAL mutants is longer than the half-life of DGJ *in vivo*. The therapy would have two alternating phases. In the first one, *i.e.*, in the presence of the drug, the enzyme would be stabilized but inhibited; in the second one, *i.e.*, in the absence of the drug, the enzyme would be active but slowly degraded. DGJ was approved by FDA for the use of responsive mutations [42]. An intermittent regimen is adopted in patients, and 150 mg of the drug is administered orally every other day.

The first reports on real clinical cases suggest that the beneficial long terms effects of DGJ should be evaluated carefully. In fact, a raise in AGAL residual activity has been observed in the leucocytes in patients of a Swiss cohort but regrettably, it did not correlate with a decrease in lyso-Gb3 concentration measured in dried blood spots [43].

Pharmacological Chaperones: Promising Drugs with a Restriction

Pharmacological chaperones cannot be used by all patients affected by a given genetic disease. They act by binding specifically unstable folded (or partially folded) mutant proteins. Hence they are not useful in all the cases, e.g., in which

the protein is absent because the gene is affected by a deletion, a stop gain mutation, a splicing mutation, or a mutation occurring in the regulatory regions. Frameshift or stop loss could be treated if the mutation caused the production of a protein with some few extra amino acids at the carboxy-terminus that is able to fold. Pharmacological chaperones are of no use if the active site is affected. Hence, some mutations can be excluded *ab initio*, the other ones, which are potentially responsive, must be tested individually. The situation is even more complex in cases where patients are composite heterozygous and hetero-oligomeric proteins can be formed [44]. For some diseases the percentage of responsive mutations is low (for example Pompe disease [45]), for others, it is relatively high (for example Fabry disease [39]). For some diseases, there are prevalent mutations that can be prioritized in the analysis (for example cystic fibrosis [46,47] and Gaucher disease [48,49]). For other diseases, this not the case and all potentially responsive disease mutations must be systematically tested. In the last case, i.e., when there are a large number of private mutations, it is very important to have a cellular model for the test because it is unpractical to obtain cells from patients. Moreover, it is important that the test is carried out by independent labs in order to assure the robustness of the results. Preliminary tests are carried out assessing the ability of a pharmacological chaperone (and more in general of small molecules) to increase the intracellular concentration of the target protein. In the case of Fabry disease, an enzymatic assay is carried out on the extracts of transiently transfected cells [38]. This test can be carried out on any mutation; it is relatively simple and practical protocols are available [50]. Successive experiments should be carried out to assess the ability of the drug to revert the phenotype of the disease, in cell models or, more ambitiously, in animal models. DGJ has been approved for the therapy of responsive Fabry patients. Hence it is not surprising that for this disease a large number of potentially responsive genotypes have been tested by the company that developed the drug into a commercial product, galafold (trademark) [38,40,51] or by independent laboratories [50,52–54]. A database that keeps track of all results, FABRY_CEP, helps clinicians Choose Eligible Patients [55].

Pharmacological Chaperones: Improvement of a Drug

DGJ is able to stabilize a large proportion of missense mutations associated with Fabry disease and has been approved for the therapy and its long-term effects on the phenotype are still under verification. It might not be the ideal drug. However, certainly, it represents a good starting point. In the meantime, moving from the structure of DGJ, similar glycomimetic molecules have been proposed.

Its enantiomer L-DGJ is 1000-fold less active than D-DGJ and its mode of action is different because it is a non-competitive inhibitor of AGAL. Interestingly the two sugars act synergistically [56].

Iminosugars like DGJ are charged molecules with a negative octanol-water partition coefficient (Log Po/w) that indicates a low capacity to cross membranes [57]. The most obvious modification to enhance lipophilicity was alkylation. Unfortunately, N-alkylation, which was useful for molecules such as 1-Deoxynojirimycin, lowered DGJ efficacy [58].

Druglikeness could be improved synthesizing DGJ-thioureas derivatives. DGJ is an iminosugar analog of galactose. Galactose, in turn, is the terminal part of the natural substrate globotriaosylceramide. Hence it could be expected that extending the structure of DGJ it would be possible to derive a more selective and potent pharmacological chaperone. Although DGJ-thioureas derivatives inhibited AGAL at neutral and at acidic pH similarly to DGJ, the authors [59] stated that *p*-methoxyphenyl and the fluorophenyl thiourea derivatives had a chaperoning effect superior to that of DGJ in cells. DGJ-thioureas can be described as compounds in which thiourea bridges two moieties, a sugar-like part, which in chemical terms is known the glycone to a non-sugar-like part, which in chemical termsis defined as a non-glycon. Ortiz Mellet and co-workers [60] proposed synthesizing a pH-sensitive thiourea. The non-glycone moiety is an orthoester group, which is stable at neutral pH but it is easily degraded at acidic pH. Hence the extended contacts with the target enzyme are possible inthe endoplasmic reticulum but not in the lysosomes. When the DGJ-thiourea-orthoester PC was administered to fibroblasts derived from Fabry patients, it produced a reduction of Gb3 far more significant than that obtained with the parental drug DGJ. Remarkably, the approach can be extendedin principle to all lysosomal diseases and a proof of concept was offered by the same

authors who synthesized 1-deoxynojirimycin (DNJ)-thiourea-orthoesters for the cure of Gaucher disease.

pH-sensitive inhibitors can be found among molecules that do not look like substrates or products, screening large collections of small molecules. This was the case of ambroxol that is a mixed-type inhibitor and a stabilizer of lysosomal acid glucosylceramidase (Gaucher disease) and was found screening a library of 1040 FDA-approved compounds. It acts in the endoplasmic reticulum better than in the lysosome because its K_i is 5 times lower at pH 7 than at pH 5.6 and reduces the storage of the glucosylceramide in cells harboring a responsive mutation [29]. A similar reduction of glucosylceramide was not obtained with another inhibitor, the iminosugar isofagomine, which binds the enzyme at neutral pH as well as at acidic pH. Ambroxol can rescue some mutants of lysosomal alpha-galactosidase and alpha-glucosidase in cells harboring responsive mutations of Fabry or Pompe disease too, but inthese cases, it is effective only if it is co-administered with the respective iminosugar pharmacological chaperones 1-deoxygalactonojirimycin (DGJ) and 1-deoxynojirimycin (DNJ) [61]. How comes that ambroxol is active on different enzymes? It might be a promiscuous ligand because lysosomal glycosidases, such as glucosylceramidase, alpha-galactosidase, and alpha-glucosidase share the sameTIM-barrel architecture and have very similar structures [22]. On the other hand, ambroxol could stabilize glucosylceramidase not only by binding the protein but also by other mechanisms and theseother mechanisms could be effective for other lysosomal storage disorders too.

Indeed most efforts to develop a pharmacological chaperone useful in clinical practice were concentrated on Gaucher disease (Figure 7.2). Isofagomine, an unmodified iminosugar, appeared to be very promising in stabilizing mutant glucosylceramidase [62] but did not reduce the accumulation oflipids significantly [14,63]. Besides the approaches already mentioned with alkylated or pH-sensitive molecules, another interesting option was evaluated. It consisted of the synthesis of compounds where nitrogen with high sp^2-hybridation character takes the place of the amine-type endocyclic nitrogen found in iminosugars [64].

Pharmacological Chaperones: Besides the Inhibitors

Besides inhibitors, other ligands can act as pharmacological chaperones. It has been proposed for Fabry disease to look for allosteric ligands that stabilize missense unstable mutants but do not bind the active site and do not inhibit the activity. These molecules would be the ideal pharmacological chaperones. In a pilot study one molecule, 2,6-dithiopurine, was identified by *in silico* docking and tested by thermal shift assay before passing to tests on transfected cells [65]. Another example is offered by an elegant approach by Millet and co-workers [66] who were looking for drugs to cure congenital erythropoietic porphyria, which is caused by deficiencies of uroporphyrinogen III synthase. They used *in silico* docking to identify an allosteric binding site on the surface of the enzyme, they looked for ligands among 2500 diverse chemical fragments, they tested the binding of the best hits by thermal shift assay and NMR. The most promising chemical fragments were used to search a library of approved drugs and look for similar compounds. Millet and co-workers found that ciclopirox, an anti-fungal, could stabilize uroporphyrinogen III synthase in cells. This paper shows that drug repositioning should be always considered since this would cut the costs and the time needed to offer actual therapy to the patients [67]. Besides soluble enzyme multi-pass, G-coupled receptors have been targeted with allosteric modulators. Three molecules, cinacalcet, NPS-2143, and NPS-568 respectively, an FDA approved and two investigational drugs, were able to recover membrane expression and impaired function of some pathological mutants of the human calcium-sensing receptor [68,69]. The thienopyrimdine, Org 42599, a derivative of an allosteric modulator of the luteinising hormone receptor, was able to rescue several pathogenic mutants that are retained intracellularly and promote their localization on the plasma membrane and maturation. However, only some of the rescued mutants regained activity [70]. Besides allosteric non-inhibitory ligands, other types of molecules should be considered such as activators compounds. In this case, the example comes from PMM2-CDG a disorder caused by deficiencies in the gene encoding phosphomanno-mutase2. The enzyme catalyzes the conversion of mannose 6-phosphate into mannose 1-phosphate and uses as activator glucose 1,6 bisphosphate. The activator binds the active site but is not consumed. It stabilizes

several pathological mutants [44,71,72]. Other non-inhibitory ligands, which bind the active site and stabilize some mutants in the cells, have been described, 11-cis-retinal and tetrahydrobiopterin, which are the natural cofactors of rhodopsin [73] and phenylalanine-hydroxylase respectively, for retinitis pigmentosa and phenylketonuria [74–76].

Pharmacological and Chemical Chaperones: Two Types of Drugs Often Confused

So far we have described pharmacological chaperones. We now move to other small molecules, the so-called "chemical chaperones." They have been proposed for the treatment of genetic diseases that are caused by unstable proteins too, but "chemical chaperones," such as osmolites and hydrophobic compounds, do not bind their targets specifically. It has been proposed that osmolites such as betaine,sarcosine, glycerol, glycine, trimethylamine N-oxide increase the energy of the unfolded state of proteins. Hence pharmacological chaperones would raise the difference between the free energy of the folded and unfolded states by lowering the first one whereas osmolites would obtain the same effectmainly by raising the second one [2] (Figure 7.6).

Osmolytes are relatively aspecific and were proved active on different proteins such as cystathionine β-synthase [77], aspartylglucosaminidase [78], adenosine triphosphate (ATP)-binding cassette subfamily A member 3 [79], alpha-1 anti-trypsin [80] in cellular models. Particularly interesting is the example offered by another chemical chaperone, 4-phenyl-butyrate (4-PBA). 4-PBA is a drug approved for the treatment of disorders of the urea cycle. Besides being an ammonia scavenger, it can stabilize a wide range of proteins. In fact, 4-PBA contains an aromatic moiety and is able to bind hydrophobic patches of unfolded or partially folded proteins. For this reason, it acts as a "hydrophobic chaperone." It can rescue DeltaF508 cystic fibrosis transmembrane conductance regulator (CFTR) [81], cystathionine β-synthase [77], alpha-1 anti-trypsin [80], ABCC6 protein [82], ATP-binding cassette transporter A1 [83], parathyroid hormone [84], and Wilson protein (Copper-transporting P-type ATPase ATP7B) [85] and the list might not be exhaustive. However, the interaction with hydrophobic patches of unstable proteins

can only in part account for the effects of 4-PBA. In fact, it down-regulates the expression of Hsc70 [86], transiently increases Hsp70 expression [87], reduces the upregulation of ER stress markers CHOP and GRP78 [88], and inhibits histone deacetylases activity [89,90] and signaling [87]. Disentangling the different roles of 4-PBA can be difficult [91].

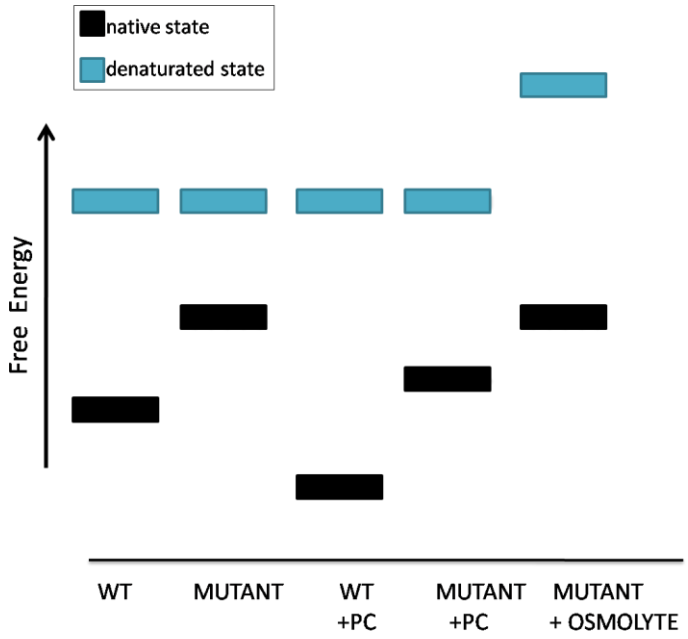

Figure 7.6. Schematic representation of the effects of pharmacological chaperones or osmolyte chemicalchaperones on protein stability.

Other "Small Molecules": Alternative or Synergistic Approaches to Treat Diseases Caused byDestabilizing Missense Mutations

Small molecules can stabilize missense mutants acting indirectly. This is the case of heat-shock protein inducers. The first example is provided by 4-PBA that could act as a hydrophobic chaperone as well as an inducer of hsp70. Other heat-shock response co-inducers have been described. Arimoclomol is an aromatic heteromonocyclic compound that underwent clinical trial phase 3 for Amyotrophic Lateral Sclerosis [92]. It rescues glucocerebrosidase in fibroblasts of Gaucher patients both in terms of quantity and in terms of maturation [93] and prevents aggregation of a missense mutant of rhodopsin, P23H, which is frequently

encountered in retinitis pigmentosa [94]. Kelly et al. [95] suggested that co-administration of a proteostasis regulator and of pharmacological chaperone could have a synergistic effect on rescuing missense pathogenic mutants. They proved their concept utilizing Celastrol, a triterpene derived from the traditional Chinese herb known as Thunder of God Vine and an alkylated deoxynojirimycin. The first molecule activates the heat shock protein response in various cell types [96] and inhibits proteasome [97], while the latter binds and stabilizes specific mutants of lysosomal acid glucosylceramidase which are responsible for Gaucher disease [98]. A similar approach with proteostasis regulators was tested with Fabry disease [99].

Conclusions

Most disease mutations destabilize the affected protein but do not affect the active site and mutant pathogenic proteins can be rescued. This goal can be achieved using small molecular weightmolecules that can be administered orally and reach difficult districts such as the central nervous system. The approach is not limited to enzymopathies and has been transversally proposed for manydiseases. Admittedly the pathway from tests *in vitro* to a clinical trial is long but the steps to be followed are clear since they are well established in traditional medicinal chemistry. On the contrary,other very promising approaches such as gene therapy are still at their infancy. Different classes of small molecules (Figure 7.7) can be considered.

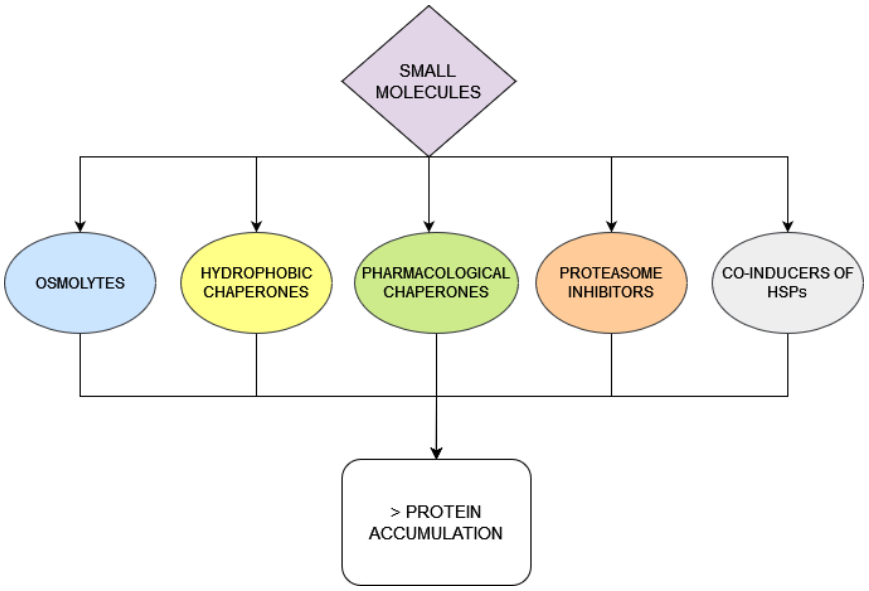

Figure 7.7. Different classes of small molecules can rescue proteins destabilized by mutations and increase their intracellular concentration.

Pharmacological chaperones play a special role among them. In fact, they are by far more specific than the others for a given protein. It is important to realize that most pharmacological chaperones tested so far are inhibitors or antagonists of their targets and that this occurred only for practical reasons. Special regimens must be envisaged to maximize their stabilizing activity and minimize their inhibitory activity. We might consider these molecules as first-generation drugs. They played an important role to prove the value of the approach but more effective molecules should be sought. If this was not feasible, we should consider combined therapies, using a specific pharmacological chaperone at a low dosage but potentiated by other types of small molecules. Unspecific drugs that promote protein folding and restore proteostasis can be useful for diverse diseases and can be foundamong approved or experimental drugs in the first place.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/21/2/489/s1. Research papers citing "pharmacological chaperone" per disease Table S1.

Author Contributions: L.L., M.M., M.A. carried out the analysis, B.H.M. searched Scopus, J.L., M.V.C. and G.A.devised the study, M.V.C. and L.L. wrote the paper, J.L., M.V.C. and G.A. carefully read the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant from MIUR PRIN 2015 2015JHLY35 (to MVC). The funding body had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Acknowledgments: Technical support was provided by the NMR Service of the Institute of Biomolecular, Chemistry–National Research Council, Pozzuoli (NA)

Italy. We are grateful to Emilio, Castelluccio for technical assistance. L.L., program at the DISTABIF, Università degli Studi della Campania "Luigi Vanvitelli", fellowship PON RI 2014/2020 "Dottorati innovativi con caratterizzazione industriale". M.A., program at the DISTABIF, Università degli Studi della Campania "Luigi Vanvitelli", fellowship POR Campania FSE 2014/2020 "Dottorati di Ricerca Con Caratterizzazione Industriale". This work is dedicated to our friend and colleague Maria Malanga.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hou, Z.S.; Ulloa-Aguirre, A.; Tao, Y.X. Pharmacoperone drugs: Targeting misfolded proteins causing lysosomal storage- ion channels-, and g protein-coupled receptors-associated conformational disorders. *Expert Rev. Clin. Pharmacol.* 2018, *11*, 611–624. [CrossRef]
2. Pereira, D.M.; Valentão, P.; Andrade, P.B. Tuning protein folding in lysosomal storage diseases: The chemistry behind pharmacological chaperones. *Chem. Sci.* 2018, *9*, 1740–1752. [CrossRef] [PubMed]
3. Tao, Y.X.; Conn, P.M. Pharmacoperones as novel therapeutics for diverse protein conformational diseases.
Physiol. Rev. 2018, *98*, 697–725. [CrossRef] [PubMed]
4. Gámez, A.; Yuste-Checa, P.; Brasil, S.; Briso-Montiano, Á.; Desviat, L.R.; Ugarte, M.; Pérez-Cerdá, C.; Pérez, B. Protein misfolding diseases: Prospects of pharmacological treatment. *Clin. Genet.* 2018, *93*, 450–458. [CrossRef] [PubMed]
5. Sanchez-Fernandez, E.M.; Garcia Fernandez, J.M.; Mellet, C.O. Glycomimetic-based pharmacological chaperones for lysosomal storage disorders: Lessons from gaucher, gm1-gangliosidosis and fabry diseases. *Chem. Commun.* 2016, *52*, 5497–5515. [CrossRef] [PubMed]
6. Matalonga, L.; Gort, L.; Ribes, A. Small molecules as therapeutic agents for inborn errors of metabolism.
J. Inherit. Metab. Dis. 2017, *40*, 177–193. [CrossRef] [PubMed]
7. Leidenheimer, N.J. Pharmacological chaperones: Beyond conformational

disorders. In *Targeting Traffic*king *in Drug Development*; Springer: Cham, Switzerland, 2017; pp. 135–153.

8. Shin, M.H.; Lim, H.-S. Screening methods for identifying pharmacological chaperones. *Mol. Biosyst.* 2017,
13, 638–647. [CrossRef]

9. Betancor-Fernández, I.; Timson, D.J.; Salido, E.; Pey, A.L. Natural (and unnatural) small molecules as pharmacological chaperones and inhibitors in cancer. In *Targeting Traffic*king *in Drug Development*; Springer: Cham, Switzerland, 2017; pp. 155–190.

10. Convertino, M.; Das, J.; Dokholyan, N.V. Pharmacological chaperones: Design and development of new therapeutic strategies for the treatment of conformational diseases. *ACS Chem. Biol.* 2016, *11*, 1471–1489. [CrossRef]

11. Tao, Y.X.; Conn, P.M. Chaperoning g protein-coupled receptors: From cell biology to therapeutics. *Endocr. Rev.* 2014, *35*, 602–647. [CrossRef]

12. Aymami, J.; Barril, X.; Rodríguez-Pascau, L.; Martinell, M. Pharmacological chaperones for enzyme enhancement therapy in genetic diseases. *Pharm. Pat. Anal.* 2013, *2*, 109–124. [CrossRef]

13. Boyd, R.E.; Lee, G.; Rybczynski, P.; Benjamin, E.R.; Khanna, R.; Wustman, B.A.; Valenzano, K.J. Pharmacological chaperones as therapeutics for lysosomal storage diseases. *J. Med. Chem.* 2013, *56*, 2705–2725. [CrossRef] [PubMed]

14. Benito, J.M.; García Fernández, J.M.; Mellet, C.O. Pharmacological chaperone therapy for gaucher disease: A patent review. *Expert Opin. on Ther. Pat.* 2011, *21*, 885–903. [CrossRef] [PubMed]

15. Parenti, G. Treating lysosomal storage diseases with pharmacological chaperones: From concept to clinics.
EMBO Mol. Med. 2009, *1*, 268–279. [CrossRef] [PubMed]

16. Morello, J.P.; Salahpour, A.; Laperriere, A.; Bernier, V.; Arthus, M.F.; Lonergan, M.; Petaja-Repo, U.; Angers, S.; Morin, D.; Bichet, D.G.; et al. Pharmacological chaperones rescue cell-surface expression and function of misfolded v2 vasopressin receptor mutants. *J. Clin. Investig.* 2000, *105*, 887–895. [CrossRef]

17. UniProt Consortium, T. Uniprot: The universal protein knowledgebase.

Nucleic Acids Res. 2018, *46*, 2699. [CrossRef]

18. Rappaport, N.; Twik, M.; Plaschkes, I.; Nudel, R.; Iny Stein, T.; Levitt, J.; Gershoni, M.; Morrey, C.P.; Safran, M.; Lancet, D. Malacards: An amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. *Nucleic Acids Res.* 2017, *45*, D877–D887. [CrossRef]

19. Ringe, D.; Petsko, G.A. Q&a: What are pharmacological chaperones and why are they interesting? *J. Biol.* 2009, *8*, 80.

20. Balchin, D.; Hayer-Hartl, M.; Hartl, F.U. In vivo aspects of protein folding and quality control. *Science* 2016, *353*, aac4354. [CrossRef]

21. Monticelli, M.; Liguori, L.; Allocca, M.; Andreotti, G.; Cubellis, M.V. Beta-glucose-1,6-bisphosphate stabilizes pathological phophomannomutase2 mutants in vitro and represents a lead compound to develop pharmacological chaperones for the most common disorder of glycosylation, pmm2-cdg. *Int. J. Mol. Sci.* 2019, *20*, 4164. [CrossRef]

22. Lieberman, R.L.; D´Aquino, J.A.; Ringe, D.; Petsko, G.A. Effects of ph and iminosugar pharmacological chaperones on lysosomal glycosidase structure and stability. *Biochemistry* 2009, *48*, 4816–4827. [CrossRef]

23. Guce, A.I.; Clark, N.E.; Rogich, J.J.; Garman, S.C. The molecular basis of pharmacological chaperoning in human alpha-galactosidase. *Chem. Biol.* 2011, *18*, 1521–1526. [CrossRef] [PubMed]

24. Andreotti, G.; Citro, V.; Correra, A.; Cubellis, M.V. A thermodynamic assay to test pharmacological chaperones for fabry disease. *Biochim. Biophys. Acta* 2014, *1840*, 1214–1224. [CrossRef] [PubMed]

25. Lo, M.C.; Aulabaugh, A.; Jin, G.; Cowling, R.; Bard, J.; Malamas, M.; Ellestad, G. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.* 2004, *332*, 153–159. [CrossRef] [PubMed]

26. Pantoliano, M.W.; Petrella, E.C.; Kwasnoski, J.D.; Lobanov, V.S.; Myslik, J.; Graf, E.; Carver, T.; Asel, E.; Springer, B.A.; Lane, P.; et al. High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screen* 2001, *6*, 429–440. [CrossRef] [PubMed]

27. Huynh, K.; Partch, C.L. Analysis of protein stability and ligand interactions

by thermal shift assay. *Curr. Protoc. Protein. Sci.* 2015, *79*, 28.9.1–28.9.14. [CrossRef] [PubMed]

28. Andreotti, G.; Monticelli, M.; Cubellis, M.V. Looking for protein stabilizing drugs with thermal shift assay.
Drug Test. Anal. 2015, *7*, 831–834. [CrossRef]

29. Maegawa, G.H.; Tropak, M.B.; Buttner, J.D.; Rigat, B.A.; Fuller, M.; Pandit, D.; Tang, L.; Kornhaber, G.J.; Hamuro, Y.; Clarke, J.T.; et al. Identification and characterization of ambroxol as an enzyme enhancement agent for gaucher disease. *J. Biol. Chem.* 2009, *284*, 23502–23516. [CrossRef]

30. Fan, J.Q.; Ishii, S.; Asano, N.; Suzuki, Y. Accelerated transport and maturation of lysosomal alpha-galactosidase a in fabry lymphoblasts by an enzyme inhibitor. *Nat. Med.* 1999, *5*, 112–115. [CrossRef]

31. Sugawara, K.; Ohno, K.; Saito, S.; Sakuraba, H. Structural characterization of mutant α-galactosidases causing fabry disease. *J. Hum. Genet.* 2008, *53*, 812. [CrossRef]

32. Citro, V.; Cammisa, M.; Liguori, L.; Cimmaruta, C.; Lukas, J.; Cubellis, M.V.; Andreotti, G. The large phenotypic spectrum of fabry disease requires graduated diagnosis and personalized therapy: A meta-analysis can help to differentiate missense mutations. *Int. J. Mol. Sci.* 2016, *17*, 2010. [CrossRef]

33. Okumiya, T.; Ishii, S.; Takenaka, T.; Kase, R.; Kamei, S.; Sakuraba, H.; Suzuki, Y. Galactose stabilizes various missense mutants of alpha-galactosidase in fabry disease. *Biochem. Biophys. Res. Commun.* 1995, *214*, 1219–1224. [CrossRef] [PubMed]

34. Frustaci, A.; Chimenti, C.; Ricci, R.; Natale, L.; Russo, M.A.; Pieroni, M.; Eng, C.M.; Desnick, R.J. Improvement in cardiac function in the cardiac variant of fabry's disease with galactose-infusion therapy. *N. Engl. J. Med.* 2001, *345*, 25–32. [CrossRef] [PubMed]

35. Legler, G.; Pohl, S. Synthesis of 5-amino-5-deoxy-d-galactopyranose and 1,5-dideoxy-1,5-imino-d-galactitol, and their inhibition of alpha- and beta-d-galactosidases. *Carbohydr. Res.* 1986, *155*, 119–129. [CrossRef]

36. Bernotas, R.C.; Pezzone, M.A.; Ganem, B. Synthesis of (+)-1,5-dideoxy-1,5-imino-d-galactitol, a potent alpha-d-galactosidase inhibitor. *Carbohydr. Res.* 1987,

167, 305–311. [CrossRef]

37. Tsukimura, T.; Chiba, Y.; Ohno, K.; Saito, S.; Tajima, Y.; Sakuraba, H. Molecular mechanism for stabilization of a mutant alpha-galactosidase a involving m51i amino acid substitution by imino sugars. Mol. Genet. Metab. 2011, 103, 26–32. [CrossRef]

38. Wu, X.; Katz, E.; Della Valle, M.C.; Mascioli, K.; Flanagan, J.J.; Castelli, J.P.; Schiffmann, R.; Boudes, P.; Lockhart, D.J.; Valenzano, K.J.; et al. A pharmacogenetic approach to identify mutant forms of alpha-galactosidase a that respond to a pharmacological chaperone for fabry disease. *Hum. Mutat.* 2011, *32*, 965–977. [CrossRef]

39. Galafold. Available online: https://www.ema.europa.eu/en/medicines/human/EPAR/galafold (accessed on 9 June 2016).

40. Benjamin, E.R.; Flanagan, J.J.; Schilling, A.; Chang, H.H.; Agarwal, L.; Katz, E.; Wu, X.; Pine, C.; Wustman, B.; Desnick, R.J.; et al. The pharmacological chaperone 1-deoxygalactonojirimycin increases alpha-galactosidase a levels in fabry patient cell lines. *J. Inherit. Metab. Dis.* 2009, *32*, 424–440. [CrossRef]

41. Khanna, R.; Benjamin, E.R.; Pellegrino, L.; Schilling, A.; Rigat, B.A.; Soska, R.; Nafar, H.; Ranes, B.E.; Feng, J.; Lun, Y.; et al. The pharmacological chaperone isofagomine increases the activity of the gaucher disease l444p mutant form of beta-glucosidase. *Febs. J.* 2010, *277*, 1618–1638. [CrossRef]

42. Moran, N. Fda approves galafold, a triumph for amicus. *Nat. Biotechnol.* 2018, *36*, 913. [CrossRef]

43. Nowak, A.; Uyen, H.D.; Krayenbuehl, P.A.; Beuschlein, F.; Schiffmann, R.; Barbey, F. Fabry disease genotype, phenotype and migalastat amenability: Insights from a national cohort. *J. Inherit. Metab. Dis.* 2019, 1–14. [CrossRef]

44. Andreotti, G.; Monti, M.C.; Citro, V.; Cubellis, M.V. Heterodimerization of two pathological mutants enhances the activity of human phosphomannomutase2. *PLoS ONE* 2015, *10*, e0139882. [CrossRef] [PubMed]

45. Flanagan, J.J.; Rossi, B.; Tang, K.; Wu, X.; Mascioli, K.; Donaudy, F.; Tuzzi, M.R.; Fontana, F.; Cubellis, M.V.;Porto, C.; et al. The pharmacological chaperone 1-deoxynojirimycin increases the activity and lysosomal trafficking of multiple

mutant forms of acid alpha-glucosidase. *Hum. Mutat.* 2009, *30*, 1683–1692. [CrossRef][PubMed]

46. Estivill, X.; Casals, T.; Morral, N.; Chillon, M.; Bosch, A.; Nunes, V.; Gasparini, P.; Seia, A.; Pignatti, P.; Novelli, G. Δf508 gene deletion in cystic fibrosis in southern europe. *Lancet* 1989, *334*, 1404. [CrossRef]

47. Bobadilla, J.L.; Macek, M., Jr.; Fine, J.P.; Farrell, P.M. Cystic fibrosis: A worldwide analysis of cftr mutations—correlation with incidence data and application to screening. *Hum. Mutat.* 2002, *19*, 575–606. [CrossRef] [PubMed]

48. Zimran, A.; Gelbart, T.; Westwood, B.; Grabowski, G.; Beutler, E. High frequency of the gaucher disease mutation at nucleotide 1226 among ashkenazi jews. *Am. J. Hum. Genet.* 1991, *49*, 855.

49. Slatkin, M. A population-genetic test of founder effects and implications for ashkenazi jewish diseases.
Am. J. Hum. Genet. 2004, *75*, 282–293. [CrossRef]

50. Lukas, J.; Knospe, A.M.; Seemann, S.; Citro, V.; Cubellis, M.V.; Rolfs, A. In vitro enzyme measurement to test pharmacological chaperone responsiveness in fabry and pompe disease. *J. Vis. Exp.* 2017. [CrossRef]

51. Benjamin, E.R.; Della Valle, M.C.; Wu, X.; Katz, E.; Pruthi, F.; Bond, S.; Bronfin, B.; Williams, H.; Yu, J.; Bichet, D.G.; et al. The validation of pharmacogenetics for the identification of fabry patients to be treated with migalastat. *Genet. Med.* 2017, *19*, 430–438. [CrossRef]

52. Lukas, J.; Giese, A.K.; Markoff, A.; Grittner, U.; Kolodny, E.; Mascher, H.; Lackner, K.J.; Meyer, W.; Wree, P.; Saviouk, V.; et al. Functional characterisation of alpha-galactosidase a mutations as a basis for a new classification system in fabry disease. *PLoS Genet* 2013, *9*, e1003632. [CrossRef]

53. Andreotti, G.; Citro, V.; De Crescenzo, A.; Orlando, P.; Cammisa, M.; Correra, A.; Cubellis, M.V. Therapy of fabry disease with pharmacological chaperones: From in silico predictions to in vitro tests. *Orphanet J. Rare Dis.* 2011, *6*, 66. [CrossRef]

54. Giugliani, R.; Waldek, S.; Germain, D.P.; Nicholls, K.; Bichet, D.G.; Simosky, J.K.; Bragat, A.C.; Castelli, J.P.; Benjamin, E.R.; Boudes, P.F. A phase 2 study of migalastat hydrochloride in females with fabry disease: Selection of

population, safety and pharmacodynamic effects. *Mol. Genet. Metab.* 2013, *109*, 86–92. [CrossRef] [PubMed]

55. Cammisa, M.; Correra, A.; Andreotti, G.; Cubellis, M.V. Fabry_cep: A tool to identify fabry mutations responsive to pharmacological chaperones. *Orphanet J. Rare Dis.* 2013, *8*, 111. [CrossRef] [PubMed]

56. Jenkinson, S.F.; Fleet, G.W.; Nash, R.J.; Koike, Y.; Adachi, I.; Yoshihara, A.; Morimoto, K.; Izumori, K.; Kato, A. Looking-glass synergistic pharmacological chaperones: Dgj and l-dgj from the enantiomers of tagatose. *Org. Lett.* 2011, *13*, 4064–4067. [CrossRef] [PubMed]

57. Katritzky, A.R.; Kuanar, M.; Slavov, S.; Hall, C.D.; Karelson, M.; Kahn, I.; Dobchev, D.A. Quantitative correlation of physical and chemical properties with chemical structure: Utility for prediction. *Chem. Rev.* 2010, *110*, 5714–5789. [CrossRef]

58. Asano, N.; Ishii, S.; Kizu, H.; Ikeda, K.; Yasuda, K.; Kato, A.; Martin, O.R.; Fan, J.Q. In vitro inhibition and intracellular enhancement of lysosomal alpha-galactosidase a activity in fabry lymphoblasts by1-deoxygalactonojirimycin and its derivatives. *Eur. J. Biochem.* 2000, *267*, 4179–4186. [CrossRef]

59. Yu, Y.; Mena-Barragan, T.; Higaki, K.; Johnson, J.L.; Drury, J.E.; Lieberman, R.L.; Nakasone, N.; Ninomiya, H.; Tsukimura, T.; Sakuraba, H.; et al. Molecular basis of 1-deoxygalactonojirimycin arylthiourea binding to human alpha-galactosidase a: Pharmacological chaperoning efficacy on fabry disease mutants. *ACS Chem. Biol.*2014, *9*, 1460–1469. [CrossRef]

60. Mena-Barragan, T.; Narita, A.; Matias, D.; Tiscornia, G.; Nanba, E.; Ohno, K.; Suzuki, Y.; Higaki, K.; Garcia Fernandez, J.M.; Ortiz Mellet, C. Ph-responsive pharmacological chaperones for rescuing mutant glycosidases. *Angew. Chem. Int. Ed. Engl.* 2015, *54*, 11696–11700. [CrossRef]

61. Lukas, J.; Pockrandt, A.M.; Seemann, S.; Sharif, M.; Runge, F.; Pohlers, S.; Zheng, C.; Glaser, A.; Beller, M.; Rolfs, A.; et al. Enzyme enhancers for the treatment of fabry and pompe disease. *Mol. Ther.* 2015, *23*, 456–464. [CrossRef]

62. Sun, Y.; Liou, B.; Xu, Y.H.; Quinn, B.; Zhang, W.; Hamler, R.; Setchell, K.D.; Grabowski, G.A. Ex vivo and in vivo effects of isofagomine on acid beta-glucosidase variants and substrate levels in gaucher disease.*J. Biol. Chem.* 2012,

287, 4275–4287. [CrossRef]

63. Dasgupta, N.; Xu, Y.H.; Li, R.; Peng, Y.; Pandey, M.K.; Tinch, S.L.; Liou, B.; Inskeep, V.; Zhang, W.; Setchell, K.D.; et al. Neuronopathic gaucher disease: Dysregulated mrnas and mirnas in brain pathogenesis and effects of pharmacologic chaperone treatment in a mouse model. *Hum. Mol. Genet.* 2015, *24*, 7031–7048. [CrossRef]

64. Mena-Barragan, T.; Garcia-Moreno, M.I.; Sevsek, A.; Okazaki, T.; Nanba, E.; Higaki, K.; Martin, N.I.; Pieters, R.J.; Fernandez, J.M.G.; Mellet, C.O. Probing the inhibitor versus chaperone properties of sp(2)-iminosugars towards human beta-glucocerebrosidase: A picomolar chaperone for gaucher disease. *Molecules* 2018, *23*, 927. [CrossRef] [PubMed]

65. Citro, V.; Pena-Garcia, J.; den-Haan, H.; Perez-Sanchez, H.; Del Prete, R.; Liguori, L.; Cimmaruta, C.; Lukas, J.; Cubellis, M.V.; Andreotti, G. Identification of an allosteric binding site on human lysosomal alpha-galactosidase opens the way to new pharmacological chaperones for fabry disease. *PLoS ONE* 2016, *11*, e0165463. [CrossRef] [PubMed]

66. Urquiza, P.; Lain, A.; Sanz-Parra, A.; Moreno, J.; Bernardo-Seisdedos, G.; Dubus, P.; Gonzalez, E.; Gutierrez-de-Juan, V.; Garcia, S.; Erana, H.; et al. Repurposing ciclopirox as a pharmacological chaperone in a model of congenital erythropoietic porphyria. *Sci. Transl. Med.* 2018, *10*, eaat7467. [CrossRef] [PubMed]

67. Hay Mele, B.; Citro, V.; Andreotti, G.; Cubellis, M.V. Drug repositioning can accelerate discovery of pharmacological chaperones. *Orphanet J. Rare Dis.* 2015, *10*, 55. [CrossRef] [PubMed]

68. Leach, K.; Wen, A.; Cook, A.E.; Sexton, P.M.; Conigrave, A.D.; Christopoulos, A. Impact of clinically relevant mutations on the pharmacoregulation and signaling bias of the calcium-sensing receptor by positive and negative allosteric modulators. *Endocrinology* 2013, *154*, 1105–1116. [CrossRef]

69. Nakamura, A.; Hotsubo, T.; Kobayashi, K.; Mochizuki, H.; Ishizu, K.; Tajima, T. Loss-of-function and gain-of-function mutations of calcium-sensing receptor: Functional analysis and the effect of allosteric modulators nps r-568 and

nps 2143. *J. Clin. Endocrinol. Metab.* 2013, *98*, E1692–E1701. [CrossRef]

70. Newton, C.L.; Anderson, R.C. Pharmacoperones for misfolded gonadotropin receptors. *Handb. Exp. Pharmacol.* 2018, *245*, 111–134.

71. Andreotti, G.; Pedone, E.; Giordano, A.; Cubellis, M.V. Biochemical phenotype of a common disease-causing mutation and a possible therapeutic approach for the phosphomannomutase 2-associated disorder of glycosylation. *Mol. Genet. Genomic. Med.* 2013, *1*, 32–44. [CrossRef]

72. Andreotti, G.; Cabeza de Vaca, I.; Poziello, A.; Monti, M.C.; Guallar, V.; Cubellis, M.V. Conformational response to ligand binding in phosphomannomutase2: Insights into inborn glycosylation disorder. *J. Biol. Chem.* 2014, *289*, 34900–34910. [CrossRef]

73. Noorwez, S.M.; Malhotra, R.; McDowell, J.H.; Smith, K.A.; Krebs, M.P.; Kaushal, S. Retinoids assist the cellular folding of the autosomal dominant retinitis pigmentosa opsin mutant p23h. *J. Biol. Chem.* 2004, *279*,16278–16284. [CrossRef]

74. Kure, S.; Hou, D.C.; Ohura, T.; Iwamoto, H.; Suzuki, S.; Sugiyama, N.; Sakamoto, O.; Fujii, K.; Matsubara, Y.; Narisawa, K. Tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. *J. Pediatr.* 1999, *135*, 375–378. [CrossRef]

75. Muntau, A.C.; Roschinger, W.; Habich, M.; Demmelmair, H.; Hoffmann, B.; Sommerhoff, C.P.; Roscher, A.A. Tetrahydrobiopterin as an alternative treatment for mild phenylketonuria. *N. Engl. J. Med.* 2002, *347*, 2122–2132. [CrossRef]

76. Bernegger, C.; Blau, N. High frequency of tetrahydrobiopterin-responsiveness among hyperphenylalaninemias: A study of 1,919 patients observed from 1988 to 2002. *Mol. Genet. Metab.* 2002, *77*, 304–313. [CrossRef]

77. Ismail, H.M.; Krishnamoorthy, N.; Al-Dewik, N.; Zayed, H.; Mohamed, N.A.; Giacomo, V.D.; Gupta, S.; Haberle, J.; Thony, B.; Blom, H.J.; et al. In silico and in vivo models for qatari-specific classical homocystinuria as basis for development of novel therapies. *Hum. Mutat.* 2019, *40*, 230–240. [CrossRef]

78. Banning, A.; Gulec, C.; Rouvinen, J.; Gray, S.J.; Tikkanen, R. Identification of small molecule compounds for pharmacological chaperone therapy of aspartylglucosaminuria. *Sci. Rep.* 2016, *6*, 37583. [CrossRef] [PubMed]

79. Kinting, S.; Hoppner, S.; Schindlbeck, U.; Forstner, M.E.; Harfst, J.; Wittmann, T.; Griese, M. Functional rescue of misfolding abca3 mutations by small

molecular correctors. *Hum. Mol. Genet.* 2018, *27*, 943–953. [CrossRef] [PubMed]

80. Burrows, J.A.; Willis, L.K.; Perlmutter, D.H. Chemical chaperones mediate increased secretion of mutant alpha 1-antitrypsin (alpha 1-at) z: A potential pharmacological strategy for prevention of liver injury and emphysema in alpha 1-at deficiency. *Proc. Natl. Acad. Sci. USA* 2000, *97*, 1796–1801. [CrossRef]

81. Choo-Kang, L.R.; Zeitlin, P.L. Induction of hsp70 promotes deltaf508 cftr trafficking. *Am. J. Physiol. Lung. Cell. Mol. Physiol.* 2001, *281*, L58–L68. [CrossRef]

82. Le Saux, O.; Fulop, K.; Yamaguchi, Y.; Ilias, A.; Szabo, Z.; Brampton, C.N.; Pomozi, V.; Huszar, K.; Aranyi, T.; Varadi, A. Expression and in vivo rescue of human abcc6 disease-causing mutants in mouse liver. *PLoS ONE* 2011, *6*, e24738. [CrossRef]

83. Sorrenson, B.; Suetani, R.J.; Williams, M.J.; Bickley, V.M.; George, P.M.; Jones, G.T.; McCormick, S.P. Functional rescue of mutant abca1 proteins by sodium 4-phenylbutyrate. *J. Lipid Res.* 2013, *54*, 55–62. [CrossRef]

84. Cinque, L.; Sparaneo, A.; Penta, L.; Mencarelli, A.; Rogaia, D.; Esposito, S.; Fabrizio, F.P.; Baorda, F.; Verrotti, A.; Falorni, A.; et al. Autosomal dominant pth gene signal sequence mutation in a family with familial isolated hypoparathyroidism. *J. Clin. Endocrinol. Metab.* 2017, *102*, 3961–3969. [CrossRef] [PubMed]

85. Van den Berghe, P.V.; Stapelbroek, J.M.; Krieger, E.; de Bie, P.; van de Graaf, S.F.; de Groot, R.E.; van Beurden, E.; Spijker, E.; Houwen, R.H.; Berger, R. Reduced expression of atp7b affected by wilson disease–causing mutations is rescued by pharmacological folding chaperones 4-phenylbutyrate and curcumin. *Hepatology* 2009, *50*, 1783–1795. [CrossRef] [PubMed]

86. Rubenstein, R.C.; Lyons, B.M. Sodium 4-phenylbutyrate downregulates hsc70 expression by facilitating mrna degradation. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 2001, *281*, L43–L51. [CrossRef] [PubMed]

87. Suaud, L.; Miller, K.; Panichelli, A.E.; Randell, R.L.; Marando, C.M.; Rubenstein, R.C. 4-phenylbutyrate stimulates hsp70 expression through the elp2 component of elongator and stat-3 in cystic fibrosis epithelial cells. *J. Biol. Chem.* 2011, *286*, 45083–45092. [CrossRef]

88. Yue, Z.S.; Zeng, L.R.; Quan, R.F.; Tang, Y.H.; Zheng, W.J.; Qu, G.; Xu, C.D.; Zhu, F.B.; Huang, Z.M. 4-phenylbutyrate protects rat skin flaps against ischemia-reperfusion injury and apoptosis by inhibiting endoplasmic reticulum stress. *Mol. Med. Rep.* 2016, *13*, 1227–1233. [CrossRef]

89. Konsoula, Z.; Barile, F.A. Epigenetic histone acetylation and deacetylation mechanisms in experimental models of neurodegenerative disorders. *J. Pharmacol. Toxicol. Methods* 2012, *66*, 215–220. [CrossRef]

90. Cousens, L.S.; Gallwitz, D.; Alberts, B.M. Different accessibilities in chromatin to histone acetylase. *J. Biol. Chem.* 1979, *254*, 1716–1723.

91. Mimori, S.; Ohtaka, H.; Koshikawa, Y.; Kawada, K.; Kaneko, M.; Okuma, Y.; Nomura, Y.; Murakami, Y.; Hamana, H. 4-phenylbutyric acid protects against neuronal cell death by primarily acting as a chemical chaperone rather than histone deacetylase inhibitor. *Bioorganic Med. Chem. Lett.* 2013, *23*, 6015–6018. [CrossRef]

92. McDermott, C.J. Clinical trials in amyotrophic lateral sclerosis. *Curr. Opin. Neurol.* 2019, *32*, 758–763. [CrossRef]

93. Fog, C.K.; Zago, P.; Malini, E.; Solanko, L.M.; Peruzzo, P.; Bornaes, C.; Magnoni, R.; Mehmedbasic, A.; Petersen, N.H.; Bembi, B. The heat shock protein amplifier arimoclomol improves refolding, maturation and lysosomal activity of glucocerebrosidase. *EBioMedicine* 2018, *38*, 142–153. [CrossRef]

94. Parfitt, D.; Aguila, M.; McCulley, C.; Bevilacqua, D.; Mendes, H.; Athanasiou, D.; Novoselov, S.; Kanuga, N.; Munro, P.; Coffey, P. The heat-shock response co-inducer arimoclomol protects against retinal degeneration in rhodopsin retinitis pigmentosa. *Cell Death Dis.* 2014, *5*, e1236. [CrossRef]

95. Mu, T.-W.; Ong, D.S.T.; Wang, Y.-J.; Balch, W.E.; Yates, J.R.; Segatori, L.; Kelly, J.W. Proteostasis regulators and pharmacologic chaperones synergize to correct protein misfolding diseases. *Cell* 2008, *134*, 769. [CrossRef][PubMed]

96. Chow, A.M.; Brown, I.R. Induction of heat shock proteins in differentiated human and rodent neurons by celastrol. Cell Stress Chaperones 2007, 12, 237. [CrossRef] [PubMed]

97. Yang, H.; Chen, D.; Cui, Q.C.; Yuan, X.; Dou, Q.P. Celastrol, a triterpene extracted from the chinese “thunder of god vine,” is a potent proteasome inhibitor

and suppresses human prostate cancer growth in nude mice. Cancer Res. 2006, 66, 4758–4765. [CrossRef] [PubMed]

98. Sawkar, A.R.; Cheng, W.C.; Beutler, E.; Wong, C.H.; Balch, W.E.; Kelly, J.W. Chemical chaperones increase the cellular activity of n370s β-glucosidase: A therapeutic strategy for gaucher disease. Proc. Natl. Acad. Sci. USA 2002, 99, 15428–15433. [CrossRef] [PubMed]

99. Seemann, S.; Ernst, M.; Cimmaruta, C.; Struckmann, S.; Cozma, C.; Koczan, D.; Knospe, A.M.; Haake, L.R.; Citro, V.; Brauer, A.U.; et al. Proteostasis regulators modulate proteasomal activity and gene expression to attenuate multiple phenotypes in fabry disease. Biochem. J. 2020, 477. [CrossRef]

8. β-GLUCOSE-1,6-BISPHOSPHATE STABILIZES PATHOLOGICAL PHOPHOMANNOMUTASE2 MUTANTS *IN VITRO* AND REPRESENTS A LEAD COMPOUND TO DEVELOP PHARMACOLOGICAL CHAPERONES FOR THE MOST COMMON DISORDER OF GLYCOSYLATION, PMM2-CDG

International Journal of
Molecular Sciences

MDPI

Article

β-Glucose-1,6-Bisphosphate Stabilizes Pathological Phophomannomutase2 Mutants In Vitro and Represents a Lead Compound to Develop Pharmacological Chaperones for the Most Common Disorder of Glycosylation, PMM2-CDG

Maria Monticelli [1], Ludovica Liguori [2,3], Mariateresa Allocca [2,3], Giuseppina Andreotti [3,*] and Maria Vittoria Cubellis [1,3]

[1]Dipartimento di Biologia, Università Federico II, 80126 Napoli, Italy
[2]Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Università degli Studi della Campania "Luigi Vanvitelli", 81100 Caserta, Italy
[3]Istituto di Chimica Biomolecolare—CNR, 80078 Pozzuoli, Italy
[*]Correspondence: gandreotti@icb.cnr.it; Tel.: +39-081-867-5241

Received: 15 July 2019; Accepted: 22 August 2019; Published: 26 August 2019

Abstract

A large number of mutations causing PMM2-CDG, which is the most frequent disorder of glycosylation, destabilize phosphomannomutase2. We looked for a pharmacological chaperone to cure PMM2-CDG, starting from the structure of a natural ligand of phosphomannomutase2, α-glucose-1,6-bisphosphate. The compound, β-glucose-1,6-bisphosphate, was synthesized and characterized via ^{31}P-NMR. β-glucose-1,6-bisphosphate binds its target enzyme *in silico*. The binding induces a large conformational change that was predicted by the program PELE and validated *in vitro* by limited proteolysis. The ability of the compound to stabilize wild type phosphomannomutase2, as well as frequently encountered pathogenic mutants, was measured using thermal shift assay. β-glucose-1,6-bisphosphate is relatively resistant to the enzyme that specifically hydrolyses natural esose-bisphosphates.

Keywords: PMM2-CDG; pharmacological chaperone; glucose-1,6-bisphosphate

Introduction

In humans there exist two phosphomannomutases, PMM1 and PMM2, that interconvert mannose-6-phosphate (M6P) into α-mannose-1-phosphate (αM1P) and to a minor extent, glucose-6-phosphate into α-glucose-1-phosphate (αG1P) [1]. Both enzymes require a sugar bisphosphate, α-mannose-1,6-bisphosphate or α-glucose-1,6-bisphospate, for their mutase activity. Besides being a mutase, PMM1 is phosphatase and plays a prevalent role in the degradation of esose bisphosphate in particular in the presence of inosine monophosphate (IMP) [2,3].
The following scheme summarizes the activities of PMM1 and PMM2.

α-hexose-1,6-P2 + PMM : hexose-6P (or α-hexose-1P) + PMM-P (1)
α-hexose-1P (or hexose-6P) + PMM-P : α-hexose-1,6-P2 + PMM (2)
PMM-P + H2O : PMM + Pi (3)

hexose = mannose, (1) + (2) = phosphomannomutase activity
hexose = glucose, (1) + (2) = phosphoglucomutase activity
hexose = mannose or glucose, (1) + (3) = phosphatase activity

Only PMM2 has been associated with human diseases and is responsible for the most common congenital disorder of glycosylation (CDG) [4]. PMM2-CDG, also known as CDG-1A or Jaeken syndrome, is a rare autosomic recessive disease without a cure. The disease is particularly frequent in Scandinavian countries but is pan-ethnically distributed. More than 110 pathological mutations have been associated with the disease [5,6]. A complete absence of PMM2 activity is not compatible with life [7]. Usually, patients are composite heterozygous with one inactivating mutation and one hypomorphic mutation, more rarely they carry two different hypomorphic mutations in heterozygosis or one in homozygosity [8,9]. R141H, which is the most common allele, is inactive because the mutation affects the active site [10], but has never been observed in homozygosity [6].
Pharmacological chaperones (PCs) are low molecular weight molecules that are able to rescue the activity of hypomorphic mutants if the active site is integer [11]. These mutants are intrinsically functional and the low residual activity in the cell is

due to their instability. Usually [12], but not necessarily, pharmacological chaperones are non-covalent competitive inhibitors of their target enzymes [13–15]. They bind and stabilize mutant enzymes, but natural substrates can compete with them for the active site. In fact, the ultimate test for PC is proving that after the administration to cells, the activity of their target enzyme increases. The genotypes associated with PMM2-CDG retain residual activity and are in principle amenable of treatment with PCs since they all retain residual activity. Even a slight increase in PMM2 activity produced by PCs could alleviate the symptoms of the patients since heterozygous carriers with just 50% normal phosphomannomutase activity in fibroblasts are asymptomatic, patients with approximately 25% normal enzymatic activity exhibit a moderate clinical picture whereas patients with less than 10% normal enzymatic activity present a severe phenotype [16,17]. Common hypomorphic PMM2 mutants are thermo-sensitive, but can be stabilized by low molecular weight molecules [5,10,18–21]. In general, this has been observed on single mutants, but in some cases, it has been proved in heterozygosity [22].

Some attempts have been carried out to find PC for PMM2-CDG either starting from large collections of commercially available products [19,23] or from known ligands of PMM2 [10,18,22], yet no drug is available.

Perez and coworkers [19] screened 10,000 small molecules from the Myria screen diversity collection and found four hits that did not resemble chemically the natural ligands of PMM2. These molecules were able to stabilize PMM2 mutants *in vitro* and rescued the activity in human cells. One molecule, 1-(3-chlorophenyl)-3,3-bis(pyridin-2-yl) urea, appeared very appealing for drug development since it did not inhibit PMM2 and did not possess predictable adverse features. Perlestein and co-workers [23] undertook drug repositioning testing 2560 compounds consisting of FDA approved drugs, bioactive tool compounds, and natural products, on a yeast strain whose phophomannomutase gene is defective.

Drug repositioning is a very useful approach to look for medicines for rare diseases because, in case of success, it decreases the gap from bench to bedside [24]. Three compounds, α-cyano-4-hydroxycinnamic acid, suramin hexasodium, and 2J,2J-bisepigallocatechin digallate, suppressed growth deficiencies

of the defective yeast strain. Their mechanism of action is not known because they were not tested on purified mutant proteins. None of these molecules resembles natural ligands of PMM2. α -cyano-4-hydroxycinnamic acid enhanced PMM2 enzymatic activity in fibroblasts derived from patients and in a nematode model of PMM2-CDG [25]. Since α -cyano-4-hydroxycinnamic acid shares the carboxylic acid-containing pharmacophore of aldose reductase inhibitors, Perlestein and co-workers [25] tested other commercially available inhibitors of the same enzyme in the nematode model and fibroblasts. They found that epalrestat, which is a safe, orally bioavailable, and brain penetrant aldose reductase inhibitor used to treat diabetic peripheral neuropathy, rescued PMM2 enzymatic activity in both species. Epalrest is a monocarboxylic acid and contains a phenyl and a rhodanine group and does not resemble chemically PMM2 ligand. Direct binding and stabilization to mutant PMM2 proteins is not among the modes of actions proposed by Perlestein and co-workers for epalrest [25].

α-D-Glucose 1,6-bisphosphate (αG16P) could be considered an effective chaperone to rescue unstable pathological variants because it binds PMM2, induces the closure of the enzyme, raises the melting temperature, and activates catalysis [10,18,22]. However, it is rapidly hydrolyzed by PMM1 in particular in the presence of IMP [2,3].

In this paper, we describe a novel approach. We synthesized a molecule, β-glucose-1,6-bisphosphate (βG16P) that is an analog of a natural ligand of PMM2, αG16P. βG16P and αG16P were fully characterized by ^{31}P-NMR. This technique can be very useful to measure the activity of PMM1 and PMM2 directly without the aid of ancillary enzymes. *In silico* observations using the program PELE [26] show that βG16P induces a large conformational change in the structure of PMM2 and the closure of the active site. βG16P can bind PMM2 and stabilize the wild type and hypomorphic PMM2. As it is the case for the majority of PCs used so far for other diseases, βG16P is an inhibitor of its target enzyme [12]. Compared to αG16P, the β anomer is more resistant to the hydrolysis by PMM1.

Results

Synthesis and ^{31}P-NMR Characterization of βG16P

βG1P was synthesized from maltose with bacterial phosphorylase and phosphorylated with phosphofructokinase to generate the bisphosphate sugar (about 40 mg were obtained starting from 650 mg of maltose). βG16P was analyzed by ^{31}P-NMR spectroscopy and compared to commercial αG16P. ^{1}H-decoupled one-dimensional ^{31}P spectrum shows two signals, one at 5.30 pmm and one at 7.45 ppm, for the P nucleus in the position 1, P(1), and in position 6, P(6) respectively (Figure 8.1A). Both signals are deshielded compared to those of the α anomer (5.06 and 7.36 ppm). 2D HSQC ^{1}H-^{31}P spectra were recorded for both anomers. The signals of βG16P are clearly resolved and distinguishable from those of αG16P (Figure 8.1 B,C).

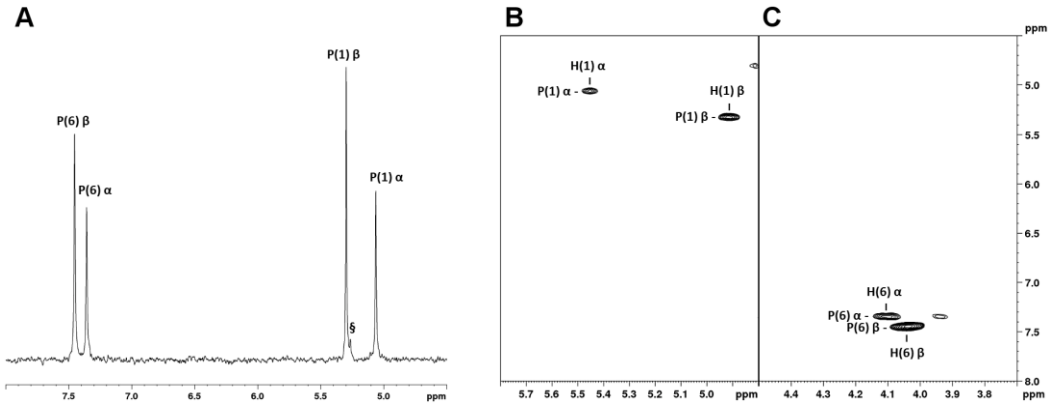

Figure 8.1. ^{31}P-NMR characterization of α-D-Glucose 1,6-bisphosphate (αG16P) and β-glucose-1,6-bisphosphate (βG16P). (A) ^{1}H-decoupled one-dimensional ^{31}P spectrum. (B,C) Selected regions of the ^{1}H-^{31}P HSQC spectrum. Both the spectra were acquired in H_2O + D_2O 10% in the presence of EDTA 50 mM; ppm were referred to creatine phosphate (0 ppm, ^{31}P scale) and to trimethylsilylpropanoic acid (0 ppm, ^{1}H scale). P(6) and P(1) indicate the positions of phosphorous nuclei in the molecules; §: inorganic phosphate.

βG16P Binds PMM2 and Induces a Conformational Change

In silico docking was carried out to test the binding of βG16P onto PMM2 with the program PELE [26]. PMM2 is a homodimer and each subunit is made up by a core

(res 1–81; 189–247) and a cap (res 86–185) domain, connected by hinge peptides. The structure of the enzyme is deposited in the PDB with the code 2AMY. Due to the absence of a few atoms and to some disordered regions in 2AMY, the structure cannot be used as such for *in silico* docking. In a previous paper of ours, we described how 2AMY can be fixed and a native-like model can be generated [10]. The native-like model, which is in open conformation as 2AMY, was used to carry out *in silico* docking of βG16P. The ligand was initially placed far from the active site (~30 Å) and was free to explore the protein surface with no bias. After a global search, a refinement was carried out. The binding of βG16P binding into the active site induces a large backbone motion and the closing up of PMM2. Actually, we observed two binding modes, in one case the phosphate at position 1, P(1), interacts with the catalytic site (Asp12, Asp14, and Asp 217 that form an acidic triad coordinated with Mg^{2+}), in the other case, it is the phosphate at position 6, P(6) to interact with the catalytic site. We will refer to the first case as P1_Mg mode and the second case as P6_Mg mode. In Figure 8.2 we show the initial unbound PMM2 structure and the two bound structures. It can be observed that the closure is tighter in the P6_Mg mode than in P1_Mg mode. In Figure 8.3A,B, we show the residues that interact with βG16P by hydrogen bonds or salt bridges in P1_Mg mode and P6_Mg mode respectively.

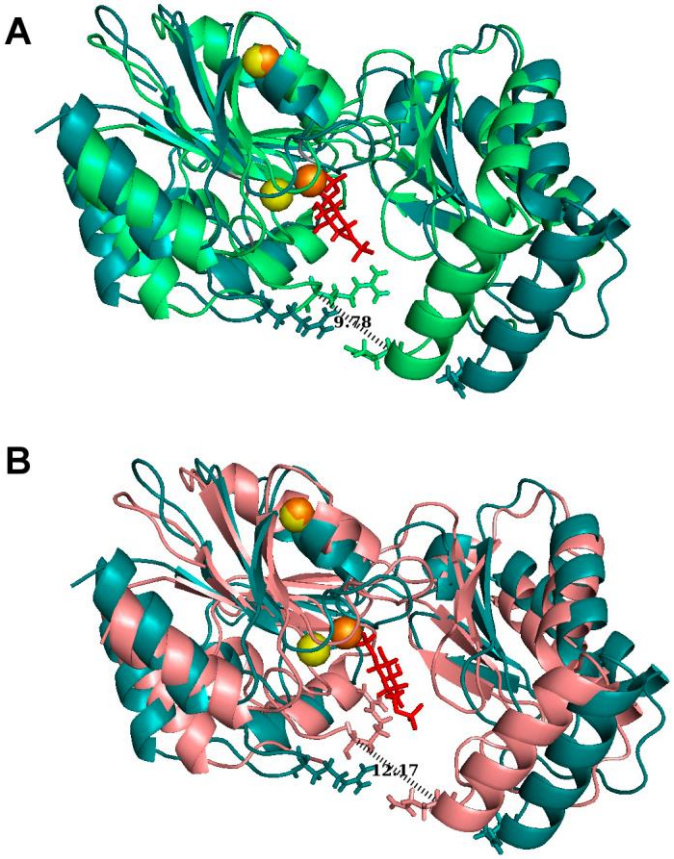

Figure 8.2. *In silico* binding models for PMM2 and βG16P: either the phosphate proximal to C6, P6_Mg mode (A), or the phosphate proximal to C1, P1_Mg mode (B), can interact with the catalytic site. PMM2 in the initial state is in cyan (A,B); the two closed conformations are in green (A) and pink (B); βG16P is red; Mg^{2+} are shown as spheres.

For comparison, we show the interactions that were observed when αG16P was docked onto the same PMM2 native-like structure (Figure 8.3C,D) with the same protocol [10]. PMM2 forms fewer contacts with βG16P than with αG16P, in particular in P1_Mg mode, and Asp14 of acidic triad does not come in contact with the phosphate. Upon βG16P binding, the core and the cap domain close up and the distance between Arg21 and Gln138 alpha carbons experiences the largest change. The number of contacts correlates with the tightness of the closure of PMM2. In fact, we measured that the distance between Arg21 and Gln138 alpha carbons is 12.2 and 9.8 Å for βG16P respectively in P1_Mg mode and P6_Mg mode and 7.7 and 7.4 for αG16P in P1_Mg mode and P6_Mg mode respectively.

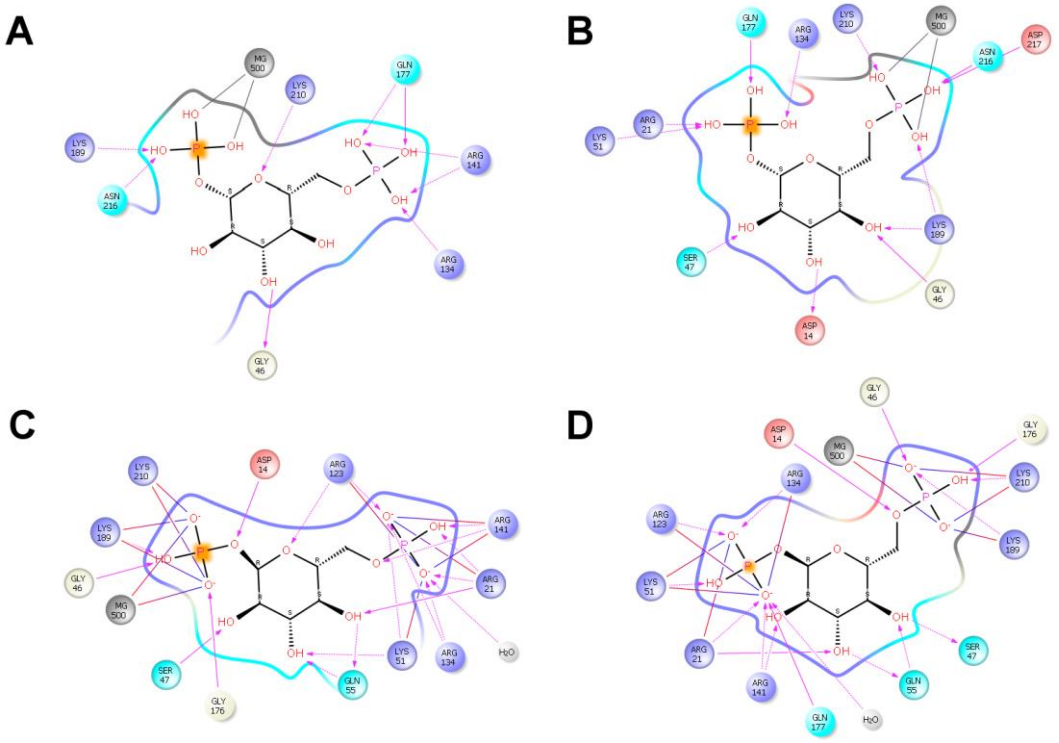

Figure 8.3. Interactions between PMM2 and ligands: residues that interact with βG16P by hydrogen bonds or salt bridges in P1_Mg mode (A) and P6_Mg mode (B) respectively; residues that interact with αG16P by hydrogen bonds or salt bridges in P1_Mg (C) mode and P6_Mg mode (D) respectively

When the core and cap domains close up, the solvent accessibility of some residues changes. In particular, the exposure of Arg21 to solvent decreases dramatically passing from 98% in the open model to 15% in the closed model. The conformational change that is observed *in silico* can be tested by limited proteolysis with trypsin. Under the conditions used in the experiment, neutral pH NaCl 150 mM, it had been demonstrated that wt-PMM2 is a dimer [18]. In Figure 8.4 we show the results obtained incubating wt-PMM2 with the protease in the absence or in the presence of ligands and demonstrate that βG16P binding protects PMM2 from the protease as well as αG16P possibly by rendering the enzyme more compact. The results obtained by SDS-PAGE are qualitative but were observed in three independent experiments. The conformational change induced by αG16P has already been discussed in a previous paper of ours [10].

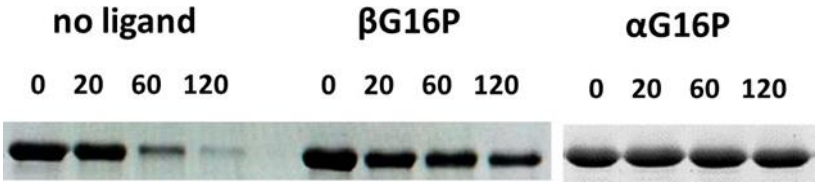

Figure 8.4. Limited proteolysis of wild type PMM2 by trypsin. Wild type PMM2 was incubated at 37 ◦C with trypsin in a 50:1 ratio in the presence or the absence of βG16P or αG16P 0.5 mM. Aliquots were withdrawn at specified times (0, 20, 60, 120 min) and analyzed by SDS-PAGE and Coomassie staining.

βG16P Inhibits PMM2

PMM2 has two enzymatic activities since it acts as a phosphomannomutase (interconversion of αM1P and M6P) or as phosphoglucomutase (interconversion of αG1P and G6P) and in both cases a bisphosphate sugar activator is needed [1,18]. Standard enzymatic tests for PMM2 require the use of ancillary enzymes. Dosing phosphoglucomutase activity is simpler and requires only glucose6P dehydrogenase to monitor the formation of G6P measuring the production of NADPH spectrophotometrically or fluorimetrically. We carried out the fluorimetric assay of phophoglucomutase activity using different pairs of substrates and activators. In Figure 8.5A, it can be observed that βG1P is not a substrate (βG1P + αG16P) and that βG16P is not an activator (αG1P + βG16P). The controls were carried out employing only beta (βG1P + βG16P) or only alpha anomers (αG1P + αG16P).

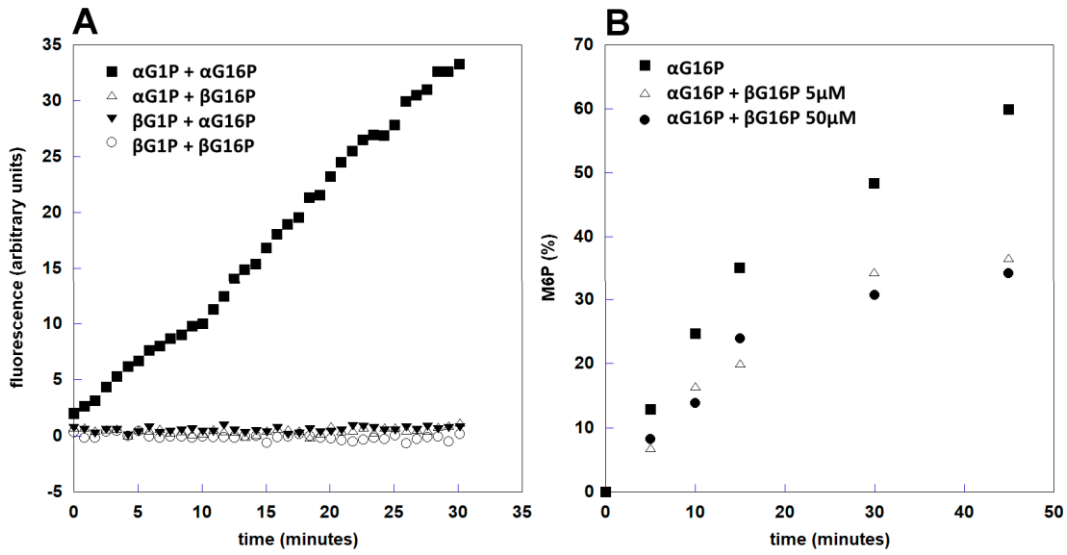

Figure 8.5. PMM2 phosphoglucomutase activity monitored by fluorescence spectroscopy (A) and phosphomannomutase activity monitored by ^{31}P-NMR (B). (A) A total of 0.08 μg PMM2 were incubated at room temperature with different combinations of α or βG1P as substrates and α or βG16P as activators, in the presence of G6PDH and $NADP^+$. (B) A total of 0.12 μg PMM2 were incubated at 32°C with 1 mM α-mannose-1-phosphate (αM1P); αG16P 5 μM was used as an activator, in the absence or the presence of βG16P 5 or 50 μM.

Standard spectrophotometric tests of mannomutase activity are more complex and require three ancillary enzymes to generate NADPH. Alternatively, the formation of M6P and the consumption of αM1P can be monitored directly by recording ^{31}P-NMR spectra.

These experiments require the accumulation of several scans for each NMR spectrum to have a good S/N ratio. They cannot be used to measure initial velocities, but show clearly all the phosphorylated species present in the solution and permit to avoid the usage of ancillary enzymes. These experiments confirm that βG16P is not an activator and show that it inhibits PMM2 (Figure 8.5B).

βG16P Stabilizes PMM2

We carried out a thermal shift assay to test the ability of βG16P to work as a chaperone, *i.e.*, to stabilize PMM2. The effects of αG16P and βG16P were compared (Figure 8.6A). It can be observed that both anomers stabilize PMM2

although the extent of the stabilization is higher for αG16P: the increase of the melting temperature was about 4°C in the presence of βG16P (from 54.8 ± 0.3 to 58.5 ±0.5°C) and about 9°C in the presence of αG16P (from 54.8 ± 0.3 to 63.4 ± 0.4°C).

αG1P is a stabilizer of PMM2 too. Its effect is strengthened by vandate, an inhibitor of PMM2 which mimics phosphate and recreates a non-covalent complex with αG1P similar to sugar 1,6-bisphosphate in the active site. βG1P does not stabilize PMM2 either in the absence or in the presence of vanadate (Figure 8.6B). This observation and the lack of activity (Figure 8.4) suggests that βG1P does not bind PMM2.

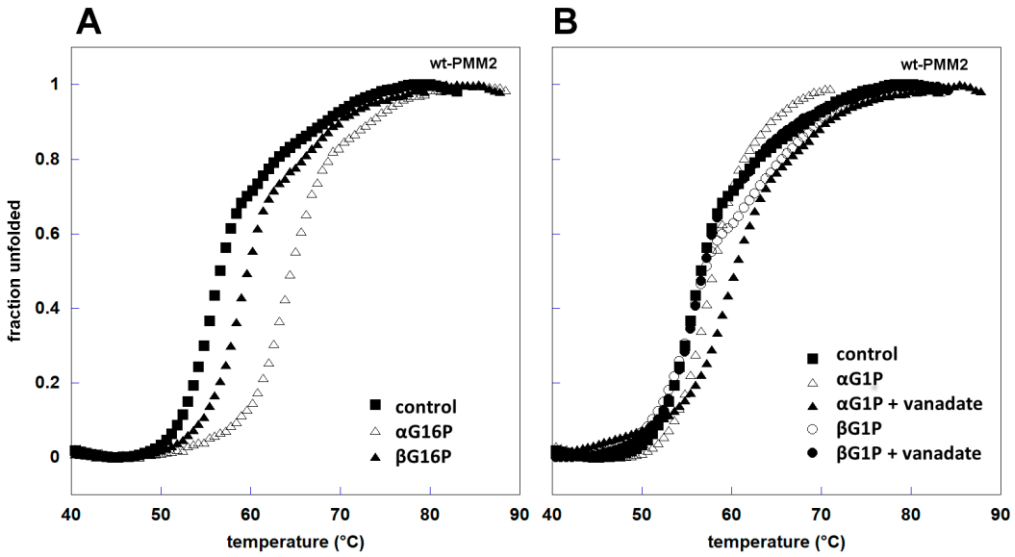

Figure 8.6. Thermal stability in the presence of ligands. wt-PMM2 0.3 mg/mL was incubated with different ligands 0.5 mM ((A) αG16P and βG16P; (B) αG1P and βG1P, with and without vanadate); the melting curves were measured in the presence of dithiothreitol (DTT) 1 mM and Sypro Orange 2.4x, from 20 to 90°C with increments of 1°C/min.

We also tested the effect βG16P on two pathological mutants, F119L and V129M. Overall, βG16P 0.5 mM caused an increase of about 4°C, from 46.1 ± 0.4 to 50.3 ± 0.4°C, while αG16P at the same concentration produces an increase of about 8°C, up to 54.2 ± 0.8°C. As far as V129M-PMM2 is concerned the melting temperature in the presence of βG16P increases of about 7°C, from 47.3 ± 0.4 to

53.9 ± 0.4°C, and of about 10°C in the presence of αG16P, up to 58.1 ± 1.3°C (Figure 8.7B).

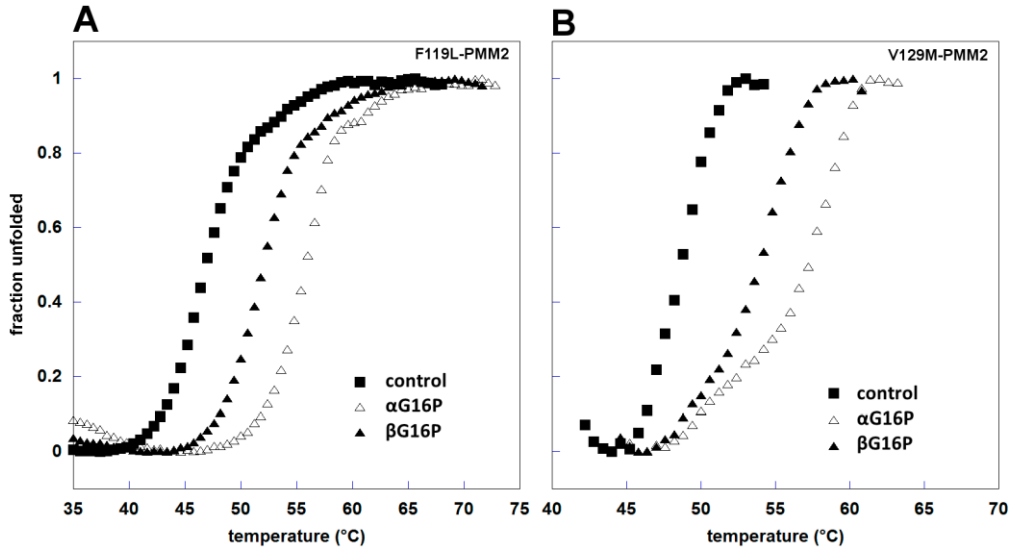

Figure 8.7. Thermal stability of pathological mutants in the presence of ligands. F119L-PMM2 (A) and V129M-PMM2 (B) 0.3 mg/mL were incubated with different ligands, αG16P (0.5 mM), and βG16P (0.5 mM). The melting curves were measured in the presence of DTT 1 mM and Sypro Orange 2.4x, from 20 to 90°C with increments of 1°C/min.

Cell viability tests showed that βG16P is well tolerated by cells when administered at 0.25 or 0.5 mM.

βG16P is a Poor Substrate of PMM1

αG16P is degraded rapidly by PMM1 in particular in the presence of IMP [2,3]. We tested the stability of βG16P to the hydrolysis by PMM1 both in the absence and in the presence of IMP using ^{31}P-NMR. It is very convenient that the signals of αG16P and βG16P can be clearly distinguished in the spectra. Hence, the experiment can be carried out by incubating both bisphosphates with PMM1 and examining whether or not only the alpha anomer is hydrolyzed.

The results are summarized in Table 8.1. PMM1 hydrolyses αG16P preferentially and the different susceptibilities of the two anomers are larger in the presence of IMP.

Table 8.1. Phosphatase activity monitored by ^{31}P-NMR. A total of 40 µg PMM1 were incubated for 90 min at 32°C with 145 µM αG16P or βG16P or a combination of both, in the absence (−) or the presence (+) of inosine monophosphate (IMP) 170 µM.

Bisphosphate (145 µM)	**IMP (170 µM)**	**Residual αG16P (%)**	**Residual βG16P (%)**
αG16P	-	54.2 ± 1.6	-
βG16P	-	-	62.5 ± 7.0
αG16P + βG16P	-	53.3 ± 1.3	93.6 ± 1.2
αG16P	+	0	-
βG16P	+	-	65.7 ± 11.4
αG16P + βG16P	+	0	74.6 ± 3.1

It was also noticed that the hydrolysis of αG16P produces Pi, αG1P, or G6P conversely the hydrolysis of βG16P produces Pi and βG1P only, suggesting that βG1P is not a substrate of PMM1 (data not shown).

Although several therapeutic approaches for PMM2-CDG [27,28] have been proposed including mannose supplementation [29,30], membrane-permeable α-mannose-1-P [31,32], and the increase mannose-6-P flux into glycosylation pathways [33], metformin [34] and acetazolamide [35], a drug addressing unstable PMM2 mutants is not yet available.

In many cases, the reduced activity of PMM2 hypomorphic mutants is a consequence of their reduced intracellular concentration that in turn is the consequence of the clearance of unstable proteins by the quality control systems of the cell. Using the program SDM [36] we could predict that 70% of pathological mutations are destabilizing. In such cases, PCs represent a good choice for patients [19].

There are two approaches to find pharmacological chaperones. The first one requires the screening of large libraries of molecules. The effort, which is big both in terms of time and budget, does not guarantee finding a drug ready for clinical trials but may produce a lead compound that requires further maturation. The other approach requires an educated guess. One can start from the structures of known

ligands of the enzyme and try to modify them to improve their potency, their safety, and their half-life. Although the second approach may seem reductive, it has produced drugs for other pathologies that have been approved and introduced into clinical practice [37,38]. In this paper, we applied the second approach. We synthesized an analog of αG16P, which is a natural ligand of PMM2. αG16P itself could be considered an effective chaperone to rescue unstable pathological variants because it binds PMM2, induces the closure of the enzyme, raises the melting temperature, and activates catalysis. Regrettably, αG16P is hydrolyzed by PMM1 in particular in the presence of IMP [2,3].

Although data concerning the half-life of αG16P *in vivo* are not available, it can be predicted that its hydrolysis by PMM1 can hinder its utility as a drug. For this reason, we explored the possibility of employing an analogous bisphosphate sugar that would not be degraded as easily as αG16P by PMM1.

We synthesized the anomer βG16P and characterized it by NMR. ^{31}P-NMR proved to be a useful technique because it consents to monitor most phosphorylated metabolites and to distinguish αG16P from βG16P anomers. We proved that βG16P binds PMM2 *in silico*. Incidentally, our experiments confirm the usefulness of PELE [26] as a program for *in silico* docking. In fact, the number of contacts and the extent of closure predicted by PELE for the two anomers are fully in line with the experimental results.

βG16P is a "classical" PC since it is a mild non-covalent inhibitor of PMM2. Usually, the chaperoning effect of drugs is tested *in vitro* on wild type enzymes. In the case of PMM2 some mutants can be expressed in *E. coli* and purified. Hence the chaperoning effect of βG16P can be tested on pathological mutants as well as on the wild type enzyme. We have chosen two mutants, F119L and V129M.

Particularly, F119L is a very frequent variant and it has been observed both in association with R141H and in homozygosis [8,9,39]. Since F119 does not occur in the active site, the mutant is active although less stable than the wild type. V129M is relatively common in Italy where it has been observed in association with R141H [40]. Position 129 is not located in the active site and the mutant retains more than 50% activity but is less stable than wt-PMM2 [41]. F119L and V129M are amenable for the cure with pharmacological chaperones in principle. We could

prove that βG16P stabilizes pathogenic mutants. βG16P is a poor substrate of PMM1 and for this reason, it can be preferred to αG16P as a PC. We are confident that the pharmacokinetics of βG16P can be superior to those of αG16P. The next step of the process that starts from the known ligand, would be enhancing its bioavailability reducing the charges by chemical modification with nontoxic acetoxymethyl groups as proposed for mannose-1 phosphate [31] and using the hydrophobic derivative βG16P as a prodrug. Once inside the cell, the modifying groups would be hydrolyzed generating βG16P. Another possibility to facilitate the entrance of βG16P is offered by liposomes as proposed by Glycomine [42]. The stabilizing effect of βG16P (through the usage of hydrophobic derivatives or with liposomes) could be tested on fibroblasts derived from patients either measuring the increase of PMM2 activity or monitoring the normalization of N-glycosylated biomarkers [28].

Materials and Methods

Materials

α-D-Glucose 1,6-bisphosphate potassium salt hydrate, α-D-Glucose 1-phosphate disodium salt hydrate, α-D(+)Mannose 1-phosphate sodium salt hydrate, Inosine 5'-monophosphate disodium salt hydrate, Adenosine 5'-triphosphate disodium salt hydrate, Maltose phosphorylase from *Enterococcus sp.*, Fructose-6-phosphate Kinase from *Bacillus stearothermophilus*, Trimethylamine, 3-(4,5-dimethylthazol-2-yl)-2,5-diphenyltetrazolium bromide were purchased from Sigma-Aldrich (Sigma-Aldrich, Milan, Italy). Trypsin was purchased from ICN Pharmaceuticals (MP Biomedicals Germany GmbH).

AG1x8 Resin 200-400 Mesh Hydroxide Form was purchased from Bio-Rad (Bio-Rad Laboratories Srl, Milan, Italy).

Sypro Orange was from Invitrogen Molecular Probes (Invitrogen Molecular Probes, Monza, Italy), StepOne^{TM} Real Time PCR System from Applied Biosystems (Applied Biosystems, Foster City, CA, USA), strips from Sarstedt (Multiply–μStrip Pro 8-strip low profile) (Sarstedt Srl, Milan, Italy).

D(+)Maltose monohydrate, β-Nicotinamide adenine dinucleotide phosphate sodium salt, Glucose-6-phosphate Dehydrogenase, Creatine phosphate disodium salt thetrahydrate were from Alfa Aesar (Alfa Aesar, Thermo Scientific, Kandel, Germany).

Synthesis of α-Glucose-1-Phosphate and β-Glucose-1,6-Bisphosphate

β-glucose-1-phosphate and β-glucose-1,6-bisphosphate were prepared as described by [43] with minor changes. Briefly, for β-glucose-1-phosphate a solution containing maltose 1 M, potassium phosphate 100 mM pH 7.5, $MgCl_2$ 2 mM was incubated with maltose phosphorylase 5 U/mL for 2 h at room temperature. Synthesis of βG1P was verified by ^{31}P-NMR and the concentration of βG1P was estimated by comparison with an internal standard. β-glucose-1,6-bisphosphate was produced incubating βG1P and ATP-Mg^{2+} (1:1 ratio) with phosphofructokinase 15 U/mL overnight at room temperature in a buffer containing potassium phosphate 25 mM, Hepes 14 mM pH 7.5.

βG1P and βG16P were purified on an AG1x8 Hydroxide Form column equilibrated with 0.01 M triethylammonium bicarbonate pH 7.5. A step gradient of triethylammonium bicarbonate from 0.01 M to 1 M was used to elute the sugars; fractions (lyophilized and dissolved in water) were analyzed by ^{31}P-NMR. βG1P was eluted at 0.3–0.4 M, βG16 was eluted at 0.8–0.95 M. Triethylammonium bicarbonate 1 M was prepared by bubbling CO_2 gas into a triethylammine 1 M solution until reaching the desired pH.

^{31}P-NMR Spectroscopy

A Bruker AVANCETMIII HD spectrometer 400 MHz, equipped with a BBO BB-H&F-D CryoProbe^{TM} Prodigy fitted with a gradient along the Z-axis, was used for the NMR analysis (Bruker Italia Srl, Milan, Italy).

The ^{1}H-decoupled one-dimensional ^{31}P (zgpg) spectra were recorded at 161.976 MHz; spectral width 120 ppm, delay time 1.2 s, pulse width of 12.0 µs were applied. The phase-sensitive 2D HSQC using Echo/Antiecho-TPPI gradient selection, with decoupling during acquisition, (hsqcetgp) ^{1}H-^{31}P spectra were recorded at 400.13/161.97 MHz; spectral width 32 ppm, delay time 1.2 sec, pulse width of 12.0

μs, frequency offset of 2nd nucleus −4.0 ppm were applied. Coupling constant was 120 Hz; function type was non-uniform sampling, with a NUS amount of 70%.

Docking and Structure Analysis

In silico docking of βG16P was carried out as described for αG16P [10].

Briefly two different exploration runs were performed: (a) a global free search and (b) a local refinement.

(a) The global search was performed by combining a long (6 Å) and a short (1.5 Å) ligand perturbation steps, with a 75%/25% probability, respectively. Rotations were kept in the [0◦–90◦] range. A randomly chosen search direction was kept for two Monte Carlo steps, allowing a more complete exploration of the entire protein surface. No information about the bound structure was used to drive the search. Anisotropic normale mode perturbation included the lowest six modes, with maximum displacements of the alpha carbon of 1 Å. Within the lowest six modes, a randomly chosen mode was kept for six steps to facilitate large conformational exploration.

(b) The local search used translations of only 0.5 A and rotations in the [0°–180°] range. Furthermore, to keep the ligand in the active site, random search direction was maintained to only one iteration.

Residue percent accessibility was calculated with PISA (PISA v1.48 European Bioinformatics Institute, Hinxton, UK) [44]. Active site residues were identified with DrosteP [45]. The figure of superimposed proteins was prepared with PyMOL (PyMOL 2.3.0 Schrödinger, LLC, New York, NY, USA) [46]. Ligand protein interactions were drawn with Maestro (Maestro, Schrödinger Release 2015-2: Maestro, Schrödinger, LLC, New York, NY, USA, 2015) [47].

Protein Expression and Purification

wt-PMM1 [3,48], wt-PMM2 and its mutants (F119L-PMM2 and V129M-PMM2) [18,41] were expressed and purified as already described.

All the PMMs were expressed (using the vector Pet22b^{+}) in *E. coli* BL21(DE3) strain grown at 37°C in LB broth containing ampicillin 0.2 mg/mL. The expression of wt-PMM1 was performed by adding IPTG 0.4 mM when the optical density was

0.5 and prolonging the incubation for 4 h after induction. Bacteria were then harvested, washed with PBS, suspended in Hepes 50 mM pH 7.5 (containing 2-mercaptoethanol 0.1 mM, EDTA 1 mM, phenylmethylsulfonyl fluoride 0.1 mM), and enzymatically lysed with lysozyme 1 mg/mL, treated with Deoxyribonuclease I 0.005 mg/mL after adding $MgCl_2$ 10 mM, and centrifuged. Ammonium sulphate was added to the clear homogenate up to 50% saturation. The precipitate was recovered, dissolved, dialyzed (in Hepes 50 mM pH 7.1, $MgCl_2$ 5 mM, 2-mercaptoethanol 1 mM), and fractionated on a DEAE-Sepharose ff with a salt gradient (0–0.7 M NaCl in Hepes 50 mM pH 7.1, $MgCl_2$ 5 mM, 2-mercaptoethanol 1 mM). A subsequent fractionation step was conducted on a Butyl-Sepharose ff column (equilibrated in Tris 50 mM pH 7.1 containing $MgCl_2$ 5 mM, 2-mercaptoethanol 1 mM, ammonium sulphate 20%). Ammonium sulphate 20% was added to the sample and then loaded onto the column, the proteins were eluted with a gradient 20%–0% ammonium sulphate. The active fractions, judged pure by SDS-PAGE, were dialyzed (in Hepes 20 mM pH 7.5, $MgCl_2$ 1 mM, NaCl 150 mM), concentrated, and stored at −20 ◦C.

The expression of PMM2s (wt and mutants) was performed by adding IPTG 0.4 mM when the optical density was 0.5 (V129M) or 0.8 (F119L and wt) and prolonging the incubation for 4 h after induction. Bacteria were harvested, washed with PBS, suspended in Tris 50 mM, pH 7.5 (containing 2-mercaptoethanol 1 mM, EDTA 5 mM, and phenylmethylsulfonyl fluoride 1 mM), and enzymatically lysed with lysozyme 1 mg/mL, treated with Deoxyribonuclease I 0.005 mg/mL after adding $MgCl_2$ 10 mM, then centrifuged. Ammonium sulphate was added to the clear homogenate up to 60% saturation. The precipitate was recovered, dissolved, and dialyzed (in Hepes 50 mM pH 7.1 containing 5 mM $MgCl_2$ and 2-mercaptoethanol 1 mM), then loaded on a DEAE-Sepharose ff column. The pass-through was collected, concentrated, and subsequently fractionated on a Superdex 75 column (in Hepes 20 mM, MgCl2 1 mM, NaCl 150 mM, pH 7.5). The active fractions, judged pure by SDS-PAGE, were pooled, concentrated, and stored at −20°C.

Limited Proteolysis

Limited proteolysis is a useful assay to test PCs [49]. wt-PMM2 to a final concentration 0.3 mg/mL was incubated with trypsin in a 50:1 ratio, in Hepes 20 mM pH 7.5, $MgCl_2$ 1 mM, NaCl 150 mM, in the presence or the absence of βG16P or αG16P 0.5 mM. The incubation was carried out at 37°C for 120 min, collecting samples at specified time intervals. The reaction was stopped by addition of sample buffer up to Tris 50 mM pH 6.8, 10% glycerol, 2% SDS, dithiothreitol (DTT) 100 mM, 0.1 bromophenol blue, then boiling for 5 min and quick cooling. Samples were analyzed by SDS-PAGE.

Enzyme Assay by Fluorescence Spectroscopy

Phosphoglucomutase activity was measured recording the reduction of $NADP^+$ to NADPH. 0.08 µg PMM2 were incubated at 25°C in a solution containing Hepes 20 mM pH 7.5, $MgCl_2$ 5 mM, G6PDH 2.6 U/mL, $NADP^+$ 0.25 mM, BSA 0.1 mg/mL. αG1P or βG1P 40 µM were used as the substrate; αG16P or βG16P 27 µM were used as the activator. Fluorescence at 340/445 nm (ex/em) was recorded for 30 min using a Varian Cary Eclipse Fluorescence Spectrophotometer equipped with a Microplate Reader (Agilent Technologies Italia SpA, Milan, Italy).

Enzyme Assay by ^{31}P-NMR

Enzymatic activities of phosphomannomutases (phosphomannomutase and phosphatase activities) can be analyzed recording ^{31}P-NMR spectra [3]. Both these activities were measured in Hepes 20 mM pH 7.3, $MgCl_2$ 1 mM, creatine phosphate (CP) 0.5 mM, D_2O 10%. The reaction was stopped by addition of EDTA 11 mM and heating for 5 min at 60°C. The samples were then quickly cooled and stored at 20°C until NMR analysis.

^{31}P-NMR spectra were acquired as described in the appropriate paragraph and the content of the substrate and/or the product of the enzyme reaction was measured by integrating the area of the signal. CP 0.5 mM was used as an internal standard. Quantitative results are the average of two independent experiments.

Phosphomannomutase activity was assayed by incubating 0.12 µg PMM2 at 32°C with 1 mM αM1P as substrate and αG16P or βG16P (5 or 50 µM) or a combination of both, as the activator. Samples were collected over a period of 30 min.
Phosphatase activity was assayed by incubating 42 µg PMM1 at 32°C for 90 min with αG16P or βG16P 145 µM or a combination of both; at t0 (0J) and at t1 (90J) samples were collected. Each condition was also tested in the presence of IMP 170 µM.

Thermal Shift Assay

Thermal shift assay was carried out as described [50]. In particular wt-PMM2, F119L-PMM2 or V129M-PMM2 to a final concentration of 0.3 mg/mL were equilibrated in Hepes 20 mM pH 7.5, $MgCl_2$ 6.25 mM, NaCl 150 mM, DTT 1.25 mM, Sypro Orange 3x, then distributed in 0.2 mL PCR-strip (20 µL each). The appropriate ligand solution (5 µL) was added and the strips were sealed and heated from 20 to 90°C with temperature increments of 1°C/min using a StepOne Real Time PCR System. Ligands αG1P, βG1P, vanadate, αG16P, and βG16P were used to a final concentration of 0.5 mM. Four replicas were run for each condition.

Miscellaneous

Cell viability assay was performed as described in [51] on Hek-293 and Caco2 cells using βG16P 0.25 and 0.5 mM for 18 h.
Proteins were quantified with the Quick Start Bradford (Bio-Rad Laboratories Srl, Milan, Italy) using BSA as the standard [52]. SDS-PAGE was performed as in the standard procedures [53].

Author Contributions: Conceptualization: M.V.C. and G.A.; Writing-Original Draft Preparation: M.V.C.; Writing-Review & Editing: G.A., M.M., L.L., M.A.; Investigation: M.M, L.L, M.A.; Validation: M.M, L.L, M.A.; Supervision, M.V.C. and G.A.

Funding: Publication costs for this manuscript were sponsored by a grant from MIUR PRIN 2015 2015JHLY35 (to MVC). The funding body had no role in the

design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Acknowledgments: We are grateful to Victor Guallar, Joint BSC-CRG-IRB Research Program in Computational Biology, Barcelona Supercomputing Center, for his help with PELE. Technical support was provided by the NMR Service of the Institute of Biomolecular Chemistry—National Research Council, Pozzuoli (NA) Italy. We are grateful to Emilio P. Castelluccio for technical assistance. L.L., PhD program at the DISTABIF, Università degli Studi della Campania "Luigi Vanvitelli", PhD fellowship PON RI 2014/2020 "Dottorati innovativi con caratterizzazione industriale". M.A., PhD program at the DISTABIF, Università degli Studi della Campania "Luigi Vanvitelli", PhD fellowship POR Campania FSE 2014/2020 "Dottorati di Ricerca Con Caratterizzazione Industriale". G.A. acknowledges Short Term Mobility program 2018 (CNR). This work is dedicated to our friend and colleague Maria Malanga.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

CP	creatine phosphate
DTT	dithiothreitol
αG1P	α-glucose-1-phosphate
αG16P	α-glucose-1,6-bisphosphate
βG16P	β-glucose-1,6-bisphosphate
IMP	inosine monophosphate
αM1P	α-mannose-1-phosphate
M6P	mannose-6-phosphate
PC	pharmacological chaperone

References

1. Pirard, M.; Achouri, Y.; Collet, J.F.; Schollen, E.; Matthijs, G.; Van Schaftingen, E. Kinetic properties and tissular distribution of mammalian

phosphomannomutase isozymes. Biochem. J. 1999, 339, 201–207. [CrossRef] [PubMed]

2. Veiga-da-Cunha, M.; Vleugels, W.; Maliekal, P.; Matthijs, G.; Van Schaftingen, E. Mammalian phosphomannomutase PMM1 is the brain IMP-sensitive glucose-1,6-bisphosphatase. J. Biol. Chem. 2008, 283, 33988–33993. [CrossRef] [PubMed]

3. Citro, V.; Cimmaruta, C.; Liguori, L.; Viscido, G.; Cubellis, M.V.; Andreotti, G. A mutant of phosphomannomutase1 retains full enzymatic activity, but is not activated by IMP: Possible implications for the disease PMM2-CDG. PLoS ONE 2017, 12, e0189629. [CrossRef] [PubMed]

4. Van Schaftingen, E.; Jaeken, J. Phosphomannomutase deficiency is a cause of carbohydrate-deficient glycoprotein syndrome type I. FEBS Lett. 1995, 377, 318–320. [PubMed]

5. Kjaergaard, S.; Skovby, F.; Schwartz, M. Carbohydrate-deficient glycoprotein syndrome type 1A: Expression and characterisation of wild type and mutant PMM2 in E. coli. Eur. J. Hum. Genet. 1999, 7, 884–888. [CrossRef] [PubMed]

6. Matthijs, G.; Schollen, E.; Bjursell, C.; Erlandson, A.; Freeze, H.; Imtiaz, F.; Kjaergaard, S.; Martinsson, T.; Schwartz, M.; Seta, N.; et al. Mutations in PMM2 that cause congenital disorders of glycosylation, type Ia (CDG-Ia). Hum. Mutat. 2000, 16, 386–394. [CrossRef]

7. Pirard, M.; Matthijs, G.; Heykants, L.; Schollen, E.; Grunewald, S.; Jaeken, J.; van Schaftingen, E. Effect of mutations found in carbohydrate-deficient glycoprotein syndrome type IA on the activity of phosphomannomutase 2. FEBS Lett. 1999, 452, 319–322. [CrossRef]

8. Kjaergaard, S.; Skovby, F.; Schwartz, M. Absence of homozygosity for predominant mutations in PMM2 in Danish patients with carbohydrate-deficient glycoprotein syndrome type 1. Eur. J. Hum. Genet. 1998, 6, 331–336. [CrossRef]

9. Matthijs, G.; Schollen, E.; Van Schaftingen, E.; Cassiman, J.J.; Jaeken, J. Lack of homozygotes for the most frequent disease allele in carbohydrate-deficient glycoprotein syndrome type 1A. Am. J. Hum. Genet. 1998, 62, 542–550. [CrossRef]

10. Andreotti, G.; Cabeza de Vaca, I.; Poziello, A.; Monti, M.C.; Guallar, V.; Cubellis, M.V. Conformational response to ligand binding in phosphomannomutase2: Insights into inborn glycosylation disorder. J. Biol. Chem. 2014, 289, 34900–34910. [CrossRef]

11. Wang, Y.-J.; Di, X.-J.; Mu, T.-W. Using pharmacological chaperones to restore proteostasis. Pharmacol. Res.
2014, 83, 3–9. [CrossRef] [PubMed]

12. Fan, J.Q. A counterintuitive approach to treat enzyme deficiencies: Use of enzyme inhibitors for restoring mutant enzyme activity. Biol. Chem. 2008, 389, 1–11. [CrossRef] [PubMed]

13. Jorge-Finnigan, A.; Brasil, S.; Underhaug, J.; Ruiz-Sala, P.; Merinero, B.; Banerjee, R.; Desviat, L.R.; Ugarte, M.; Martinez, A.; Pérez, B. Pharmacological chaperones as a potential therapeutic option in methylmalonic aciduria cblB type. Hum. Mol. Genet. 2013, 22, 3680–3689. [CrossRef] [PubMed]

14. Leidenheimer, N.J.; Ryder, K.G. Pharmacological chaperoning: A primer on mechanism and pharmacology.
Pharmacol. Res. 2014, 83, 10–19. [CrossRef] [PubMed]

15. Citro, V.; Pena-Garcia, J.; den-Haan, H.; Perez-Sanchez, H.; Del Prete, R.; Liguori, L.; Cimmaruta, C.; Lukas, J.; Cubellis, M.V.; Andreotti, G. Identification of an Allosteric Binding Site on Human Lysosomal Alpha-Galactosidase Opens the Way to New Pharmacological Chaperones for Fabry Disease. PLoS ONE 2016, 11, e0165463. [CrossRef] [PubMed]

16. Jaeken, J.; Matthijs, G. Congenital disorders of glycosylation. Annu Rev. Genom. Hum. Genet. 2001, 2, 129–151.
[CrossRef] [PubMed]

17. Grunewald, S.; Schollen, E.; Van Schaftingen, E.; Jaeken, J.; Matthijs, G. High residual activity of PMM2 in patients' fibroblasts: Possible pitfall in the diagnosis of CDG-Ia (phosphomannomutase deficiency). Am. J. Hum. Genet. 2001, 68, 347–354. [CrossRef]

18. Andreotti, G.; Pedone, E.; Giordano, A.; Cubellis, M.V. Biochemical phenotype of a common disease-causing mutation and a possible therapeutic

approach for the phosphomannomutase 2-associated disorder of glycosylation. Mol. Genet. Genom. Med. 2013, 1, 32–44. [CrossRef]

19. Yuste-Checa, P.; Brasil, S.; Gamez, A.; Underhaug, J.; Desviat, L.R.; Ugarte, M.; Pérez-Cerdá, C.; Martinez, A.; Pérez, B. Pharmacological Chaperoning: A Potential Treatment for PMM2-CDG. Hum. Mutat. 2017, 38, 160–168. [CrossRef]

20. Yuste-Checa, P.; Gamez, A.; Brasil, S.; Desviat, L.R.; Ugarte, M.; Perez-Cerda, C.; Pérez, B. The Effects of PMM2-CDG-Causing Mutations on the Folding, Activity, and Stability of the PMM2 Protein. Hum. Mutat. 2015, 36, 851–860. [CrossRef]

21. Vega, A.I.; Perez-Cerda, C.; Abia, D.; Gamez, A.; Briones, P.; Artuch, R.; Desviat, L.R.; Ugarte, M.; Pérez, B. Expression analysis revealing destabilizing mutations in phosphomannomutase 2 deficiency (PMM2-CDG): Expression analysis of PMM2-CDG mutations. J. Inherit. Metab. Dis. 2011, 34, 929–939. [CrossRef] [PubMed]

22. Andreotti, G.; Monti, M.C.; Citro, V.; Cubellis, M.V. Heterodimerization of Two Pathological Mutants Enhances the Activity of Human Phosphomannomutase2. PLoS ONE 2015, 10, e0139882. [CrossRef]

23. Lao, J.P.; DiPrimio, N.; Prangley, M.; Sam, F.S.; Mast, J.D.; Perlstein, E.O. Yeast Models of Phosphomannomutase 2 Deficiency, A Congenital Disorder of Glycosylation. G3 (Bethesda) 2019, 9, 413–423. [CrossRef]

24. Hay Mele, B.; Citro, V.; Andreotti, G.; Cubellis, M.V. Drug repositioning can accelerate discovery of pharmacological chaperones. Orphanet J. Rare Dis. 2015, 10, 55. [CrossRef]

25. Iyer, S.; Murthy, K.; Parton, Z.; Tsang, H.; Sam, F.S.; DiPrimio, N.; Lao, J.; Perlstein, E.O. Repurposing the aldose reductase inhibitor and diabetic neuropathy drug epalrestat for the congenital disorder of glycosylation PMM2-CDG. bioRxiv 2019. [CrossRef]

26. Borrelli, K.W.; Vitalis, A.; Alcantara, R.; Guallar, V. PELE: Protein energy landscape exploration. A novel Monte Carlo based technique. J. Chem. Theory Comput. 2005, 1, 1304–1311. [CrossRef]

27. Freeze, H.H. Towards a therapy for phosphomannomutase 2 deficiency, the defect in CDG-Ia patients.
Biochim. Biophys. Acta 2009, 1792, 835–840. [CrossRef]
28. Brasil, S.; Pascoal, C.; Francisco, R.; Marques-da-Silva, D.; Andreotti, G.; Videira, P.A.; Morava, E.; Jaeken, J.; Dos Reis Ferreira, V. CDG Therapies: From Bench to Bedside. Int. J. Mol. Sci. 2018, 19, 1304. [CrossRef]
29. Kjaergaard, S.; Kristiansson, B.; Stibler, H.; Freeze, H.H.; Schwartz, M.; Martinsson, T.; Skovby, F. Failure of short-term mannose therapy of patients with carbohydrate-deficient glycoprotein syndrome type 1A. Acta Paediatr. 1998, 87, 884–888. [CrossRef]
30. Mayatepek, E.; Kohlmuller, D. Mannose supplementation in carbohydrate-deficient glycoprotein syndrome type I and phosphomannomutase deficiency. Eur. J. Pediatr. 1998, 157, 605–606. [CrossRef]
31. Eklund, E.A.; Merbouh, N.; Ichikawa, M.; Nishikawa, A.; Clima, J.M.; Dorman, J.A.; Norberg, T.; Freeze, H.H. Hydrophobic Man-1-P derivatives correct abnormal glycosylation in Type I congenital disorder of glycosylation fibroblasts. Glycobiology 2005, 15, 1084–1093. [CrossRef]
32. Rutschow, S.; Thiem, J.; Kranz, C.; Marquardt, T. Membrane-permeant derivatives of mannose-1-phosphate.
Bioorg. Med. Chem. 2002, 10, 4043–4049. [CrossRef]
33. Sharma, V.; Ichikawa, M.; He, P.; Scott, D.A.; Bravo, Y.; Dahl, R.; Ng, B.G.; Cosford, N.D.; Freeze, H.H. Phosphomannose isomerase inhibitors improve N-glycosylation in selected phosphomannomutase-deficient fibroblasts. J. Biol. Chem. 2011, 286, 39431–39438. [CrossRef]
34. Shang, J.; Lehrman, M.A. Metformin-stimulated mannose transport in dermal fibroblasts. J. Biol. Chem. 2004,
279, 9703–9712. [CrossRef]
35. Martinez-Monseny, A.F.; Bolasell, M.; Callejon-Poo, L.; Cuadras, D.; Freniche, V.; Itzep, D.C.; Gassiot, S.; Arango, P.; Casas-Alba, D.; de la Morena, E.; et al. AZATAX: Acetazolamide safety and efficacy in cerebellar syndrome in PMM2 congenital disorder of glycosylation (PMM2-CDG). Ann. Neurol. 2019, 85, 740–751. [CrossRef]

36. Pandurangan, A.P.; Ochoa-Montano, B.; Ascher, D.B.; Blundell, T.L. SDM: A server for predicting effects of mutations on protein stability. Nucleic Acids Res. 2017, 45, W229–W235. [CrossRef]

37. Levy, H.L.; Milanowski, A.; Chakrapani, A.; Cleary, M.; Lee, P.; Trefz, F.K.; Whitley, C.B.; Feillet, F.; Feigenbaum, A.S.; Bebchuk, J.D.; et al. Efficacy of sapropterin dihydrochloride (tetrahydrobiopterin, 6R-BH4) for reduction of phenylalanine concentration in patients with phenylketonuria: A phase III randomised placebo-controlled study. Lancet 2007, 370, 504–510. [CrossRef]

38. Markham, A. Migalastat: First Global Approval. Drugs 2016, 76, 1147–1152. [CrossRef]

39. Kjaergaard, S.; Schwartz, M.; Skovby, F. Congenital disorder of glycosylation type Ia (CDG-Ia): Phenotypic spectrum of the R141H/F119L genotype. Arch. Dis. Child. 2001, 85, 236–239. [CrossRef]

40. Barone, R.; Carrozzi, M.; Parini, R.; Battini, R.; Martinelli, D.; Elia, M.; Spada, M.; Lilliu, F.; Ciana, G.; Burlina, A.; et al. A nationwide survey of PMM2-CDG in Italy: High frequency of a mild neurological variant associated with the L32R mutation. J. Neurol. 2015, 262, 154–164. [CrossRef]

41. Citro, V.; Cimmaruta, C.; Monticelli, M.; Riccio, G.; Hay Mele, B.; Cubellis, M.V.; Andreotti, G. The Analysis of Variants in the General Population Reveals That PMM2 Is Extremely Tolerant to Missense Mutations and That Diagnosis of PMM2-CDG Can Benefit from the Identification of Modifiers. Int. J. Mol. Sci. 2018, 19, 2218. [CrossRef]

42. Glycomine. Available online: http://glycomine.com (accessed on 12 August 2019).

43. Dai, J.; Wang, L.; Allen, K.N.; Radstrom, P.; Dunaway-Mariano, D. Conformational cycling in beta-phosphoglucomutase catalysis: Reorientation of the beta-d-glucose 1,6-(Bis)phosphate intermediate. Biochemistry 2006, 45, 7818–7824. [CrossRef]

44. PISA. Available online: https://www.ebi.ac.uk/pdbe/pisa/pi_link.html (accessed on 10 July 2019).

45. Cammisa, M.; Correra, A.; Andreotti, G.; Cubellis, M.V. Identification and analysis of conserved pockets on protein surfaces. BMC Bioinform. 2013, 14 (Suppl. 7), S9. [CrossRef]
46. PyMOL. Available online: https://pymol.org/2/ (accessed on 10 July 2019).
47. Maestro. Schrödinger Release 2015–2: Maestro, Schrödinger; Maestro Molecular Modeling Interface, Tool "Ligand Interaction Diagram"; LLC: New York, NY, USA, 2015.
48. Silvaggi, N.R.; Zhang, C.; Lu, Z.; Dai, J.; Dunaway-Mariano, D.; Allen, K.N. The X-ray crystal structures of human alpha-phosphomannomutase 1 reveal the structural basis of congenital disorder of glycosylation type 1a. J. Biol. Chem. 2006, 281, 14918–14926. [CrossRef]
49. Andreotti, G.; Citro, V.; Correra, A.; Cubellis, M.V. A thermodynamic assay to test pharmacological chaperones for Fabry disease. Biochim. Biophys. Acta. 2014, 1840, 1214–1224. [CrossRef]
50. Andreotti, G.; Monticelli, M.; Cubellis, M.V. Looking for protein stabilizing drugs with thermal shift assay.
Drug Test. Anal. 2015, 7, 831–834. [CrossRef]
51. D'Aniello, E.; Fellous, T.; Iannotti, F.A.; Gentile, A.; Allara, M.; Balestrieri, F.; Gray, R.; Amodeo, P.; Vitale, R.M.; Di Marzo, V.; et al. Identification and characterization of phytocannabinoids as novel dual PPARalpha/gamma agonists by a computational and in vitro experimental approach. Biochim. Biophys. Acta Gen. Subj. 2019, 1863, 586–597. [CrossRef]
52. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 1976, 72, 248–254. [CrossRef]
53. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature
1970, 227, 680–685. [CrossRef]

9. DRUG REPOSITIONING FOR FABRY DISEASE: ASPIRIN POTENTIATES THE STABILIZATION OF LYSOSOMAL ALPHA-GALACTOSIDASE BY PHARMACOLOGICAL CHAPERONES

The content of this chapter is currently under revision.

International Journal of
Molecular Sciences

MDPI

Article

Drug repositioning for Fabry disease: aspirin potentiates the stabilization of lysosomal alpha-galactosidase by pharmacological chaperones

Maria Monticelli[1], Ludovica Liguori[2,3], Mariateresa Allocca[2,3], Andrea Bosso[1,4], Giuseppina Andreotti[3], Jan Lukas[5,6], Maria Vittoria Cubellis[1,3]* and Bruno Hay Mele[1]

[1] Dept. Biology, University of Napoli "Federico II", complesso Universitario Monte Sant'Angelo, via Cinthia 80126, Napoli, Italy; maria.monticelli@unina.it, andrea.bosso@unina.it, bruno.haymele@unina.it

[2] Dept. Environmental, Biological and Pharmaceutical Sciences and Technologies (DiSTABiF), University of Campania "Luigi Vanvitelli", Via Vivaldi, 43-81100 Caserta, Italy; lud.liguori@gmail.com, mariate-resa.allocca@unicampania.it

[3] Institute of Biomolecular Chemistry ICB, CNR, Via Campi Flegrei 34, 80078 Pozzuoli, Italy; gandreot-ti@icb.cnr.it

[4] Biochemistry and Cellular Biology, National Research Council, via Pietro Castellino 111, 80131, Napoli, Italy

[5] Translational Neurodegeneration Section "Albrecht-Kossel", Department of Neurology, University Medical Center Rostock, 18147 Rostock, Germany; jan.lukas@med.uni-rostock.de

[6] Center for Transdisciplinary Neurosciences Rostock (CTNR), University Medical Center Rostock, 18147 Rostock, Germany

* Correspondence: cubellis@unina.it; Tel.: +39 081 679152

Abstract: Fabry disease is caused by a deficiency of lysosomal alpha galactosidase and has a very large genotypic and phenotypic spectrum. Some patients who carry hypomorphic mutations can benefit from an oral therapy with a pharmacological chaperone. The drug requires a very precise regimen because it is a reversible inhibitor of alpha-galactosidase. We looked for molecules that can potentiate the pharmacological chaperone specific for Fabry disease among drugs that have already been approved for the other diseases. We tested the candidates on fibroblasts derived from a patient carrying a large deletion in the gene *GLA* and stably transfected with plasmid expressing hypomorphic mutants. On our cell

model three drugs were able to potentiate the action of the pharmacological chaperone. We focused our attention on one of them, acetylsalicylic acid (aspirin). We trust that aspirin can be used in synergy with the pharmacological chaperone approved for Fabry disease and prolong its stabilizing effect on alpha-galactosidase.

Keywords: drug repositioning; pharmacological chaperones; aspirin; AGAL; Fabry disease; lysosomal storage diseases

Introduction

Fabry disease (FD) is caused by a deficiency of the enzyme lysosomal alpha-galactosidase (AGAL) that removes a terminal galactose residue from globotriaosylceramides like Gb3. It is encoded by the gene GLA on the X chromosome [1–5].

FD exhibits a large spectrum of phenotypes, ranging from severe, early-onset forms to atypical or mild, late-onset forms, and is associated with more than 1000 genotypes, none of which is prevalent [6–9].

Contrary to what was observed in other X-linked diseases, FD affects heterozygous females as well as hemizygous males [10–12]. Patients who do not produce AGAL require enzyme replacement therapy (ERT). ERT consists of repeated intravenous in-fusions of a recombinant AGAL [13]. Patients who carry a missense mutation in *GLA* may benefit from oral therapy with a pharmacological chaperone (PC) [14], but, un-fortunately, not all of them are eligible. Since hundreds of different genotypes are known, a big effort was put in place to predict and test experimentally which missense variants are treatable with PCs [15,16]. Only in a few cases, the experiments have been carried out *ex vivo*, on fibroblasts or leucocytes, due to the difficulty of obtaining cells from patients [17–20]. In most cases, experiments have been carried out with vectors containing the *GLA* gene variants for heterologous expression in COS1, COS7, or HEK293 cells [9,21–28]. PCs bind and stabilize variants mildly destabilized by missense mutations mainly occurring in flexible exposed regions of the protein [29] but cannot rescue severely destabilizing mutations as well as those affecting the active site or di-sulfide bridges [30]. 1-Deoxygalactonojirimycin (DGJ), also known as Migalastat, is an imino sugar

analog of galactose, that stabilizes wild-type AGAL as well as responsive variants [31]. It binds the active site and acts as a competitive inhibitor specifically. Although a few attempts have been carried out to find a molecule that stabilizes AGAL but does not inhibit it in the lysosomes [32,33], DGJ is the only approved PC for FD so far and it is merchandised under the name of Galafold® (Amicus Therapeutics, Philadelphia, PA, USA) [34]. PCs and ERT are not a cure for FD but chronic treatments. DGJ requires a precise dosage and an intermitted regimen where the stabilizing and inhibitory effects are carefully balanced [17,35–37]. Combining DGJ with other drugs that raise the activity of AGAL variants acting by different mechanisms would be a possibility to ameliorate the therapy [38].

AGAL is synthesized as a high molecular weight precursor, imported into the endoplasmic reticulum, transferred to the lysosome, and partly released extracellularly [39]. The enzyme undergoes maturation which includes proteolysis, glycosylation, and phosphorylation on mannose residues during this process [40,41]. Unstable variants are cleared by the quality control system and their concentration, and consequently, the total activity in the cell is lower than normal. Several pathways contribute to proteostasis and could represent the source of novel targets for the treatment of FD as it was proposed for other lysosomal storage disorders [42,43]. A paper by Seemann et al. described the screening of proteostasis regulators as potentiators of DGJ. The authors found that Clasto-Lactacystin β-lactone, MG132, and Bortezomib enhance the effect of DGJ in two lines of fibroblasts derived from Fabry patients [44]. On the same cellular model, they could not confirm the positive effect of ambroxol (ABX) that had been previously observed in transiently transfected cells [45]. An effector of heat-shock proteins, 4-phenylbutyrate (4-PBA), was tested on fibroblasts derived from FD patients in monotherapy. The drug raised the amount of intracellular AGAL precursor but did not enhance the enzymatic activity [46]. The mode of action of proteostasis and heat-shock regulators is not as straightforward as one could predict. For example, Bortezomib affects proteostasis as well as *GLA* expression [45] and 4-PBA acts as a chemical chaperone binding hydrophobic patches of misfolded proteins, [47] controls ER stress. Disentangling its precise role can be difficult [48].

We propose a practical approach based on repositioning in which the effects of drugs on AGAL total activity, is tested in synergy with DGJ on a suitable cellular mod-el. Transient transfection of a plasmid expressing AGAL mutants in COS or HEK cells is very useful for PCs because any missense mutation can be tested but may cause un-foreseen artificial results due to overexpression, e.g. an unphysiological overload of the ER protein folding apparatus not observed in cells even in pathological conditions. We propose a cellular model where the advantages of FD-derived cells and the versatility of transfection are combined. We produced stably transfected fibroblasts derived from an FD patient carrying a large deletion to test FDA-approved drugs in synergy with DGJ. We started from 4-PBA and ABX and moved on to other DGJ potentiators looking for safe and cheap drugs that can be used for life-long chronic treatment.

We found that aspirin can raise the total amount and activity of AGAL in synergy with DGJ on responsive mutations.

Results

Establishment of versatile cell models that do not overexpress AGAL

DGJ is routinely tested on different missense variants in transiently transfected HEK293 cells [9,15]. In the genome of these cells wt-*GLA* is present, nonetheless the effect of the drug can be evaluated since the missense mutant to be tested is overexpressed. We felt that testing proteostasis regulators under conditions of super-expression is not appropriate. Nonetheless the versatility of transfected cells is needed to test any possible missense variant. It is for this reason that we stably transfected a line of fibroblasts derived from a male FD patient carrying a large deletion of exons 3 and 4 in *GLA*. The cells were immortalized (IF cells) as described in methods. Starting from a single clone, IF were stably transfected with a plasmid encoding wt-*GLA*, obtaining IF-GLA cells, or encoding hypomorphic *GLA* mutants, obtaining IF-GLA-MUTs, or with an empty vector, obtaining IF-NULL cells (Figure 9.1, panel A). Transcriptional levels of *GLA* in the cell lines were comparable or lower than healthy fibroblasts (Figure 9.1, panel B).

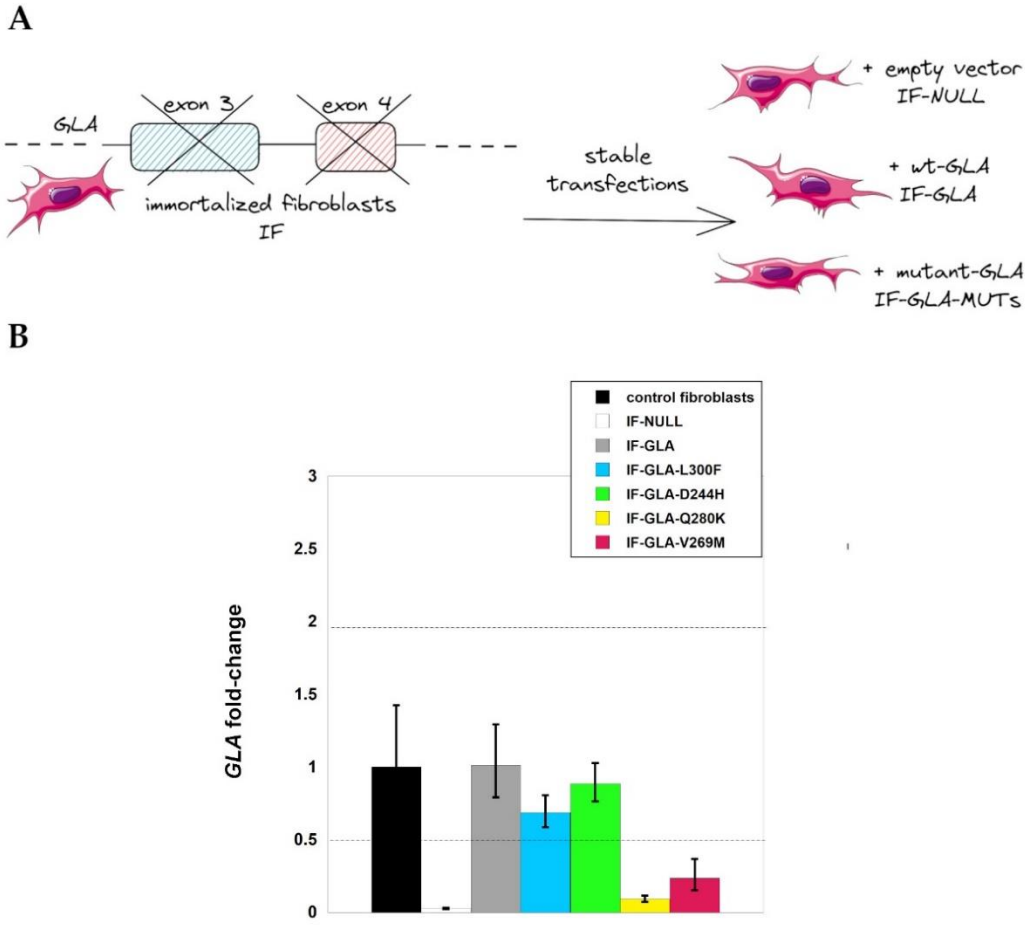

Figure 9.1. Cell lines establishment. Fabry patient's fibroblasts carrying a deletion of exons 3 and 4 in the *GLA* gene were obtained from Telethon biobank, immortalized (IF) then stably transfected with an empty vector (IF-NULL), the wt-*GLA* (IF-GLA) or different *GLA* pathogenic mutants (IF-GLA-MUTs, panel A, workflow). *GLA* expression was measured by RT-qPCR; the selected cell lines did not over-express *GLA* if compared to healthy fibroblasts (panel B).

Intracellular stabilization of AGAL by DGJ is enhanced by ambroxol and 4-phenylbutyrate

Ambroxol (Figure 9.2 A and B) and Phenylbutyrate (4-PBA) (Figure 9.3 A and B), were tested in IF-GLA-MUTs. Both drugs enhanced the DGJ stabilizing effect. A different mechanism of action with respect to DGJ was revealed by immunoblots. The presence of the PC-enhancer alone increases AGAL precursor levels (higher molecular weight band) while the chaperone stabilizes the active form (lower

molecular weight band). The combined treatment results in a strong increase in the active AGAL. Figure 9.2 highlights the different molecular weights of the AGAL precursor and active form as an example.

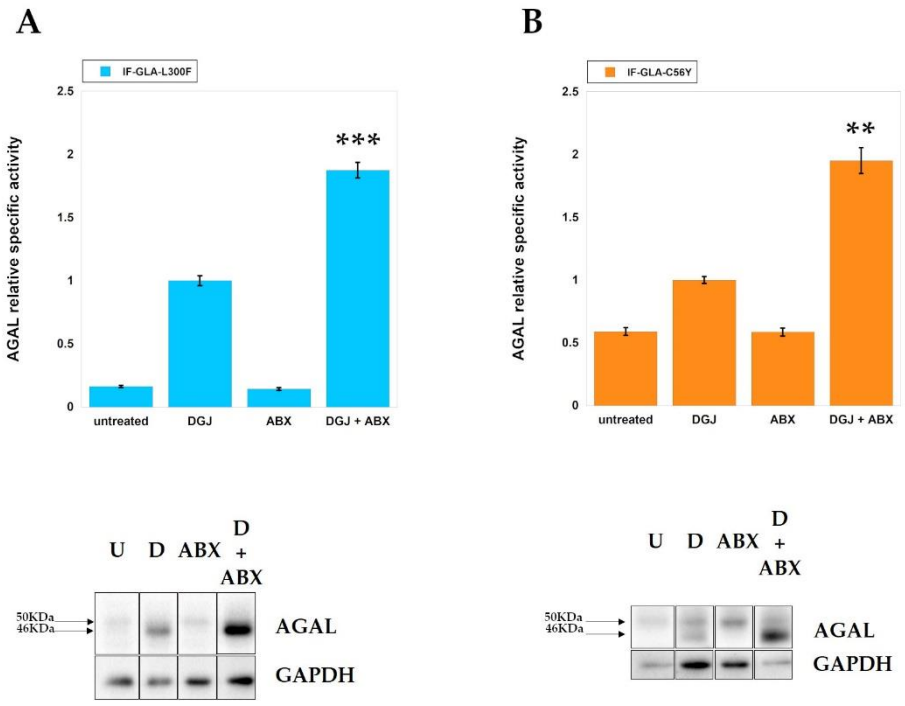

Figure 9.2. Ambroxol is a PC-enhancer. IF-GLA-L300F (panel A) and IF-GLA-C56Y (panel B) were treated for 72 hours with 10 μM DGJ in the presence or the absence of 40 μM ambroxol (ABX). AGAL relative specific activity measured on protein extracts is shown and values are referred to DGJ monotherapy. The effects of combined treatments are significantly higher than those of monotherapy (**: *p*-value < 0.01; ***: *p*-value < 0.001). Immunoblots confirmed the results (U = untreated; D = DGJ 10 μM; ABX = ambroxol 40 μM; D+ABX = DGJ 10 μM + ambroxol 40 μM). Arrows on the immunoblots highlight the higher molecular weight band (precursor) and the lower molecular weight band (active form) of AGAL.

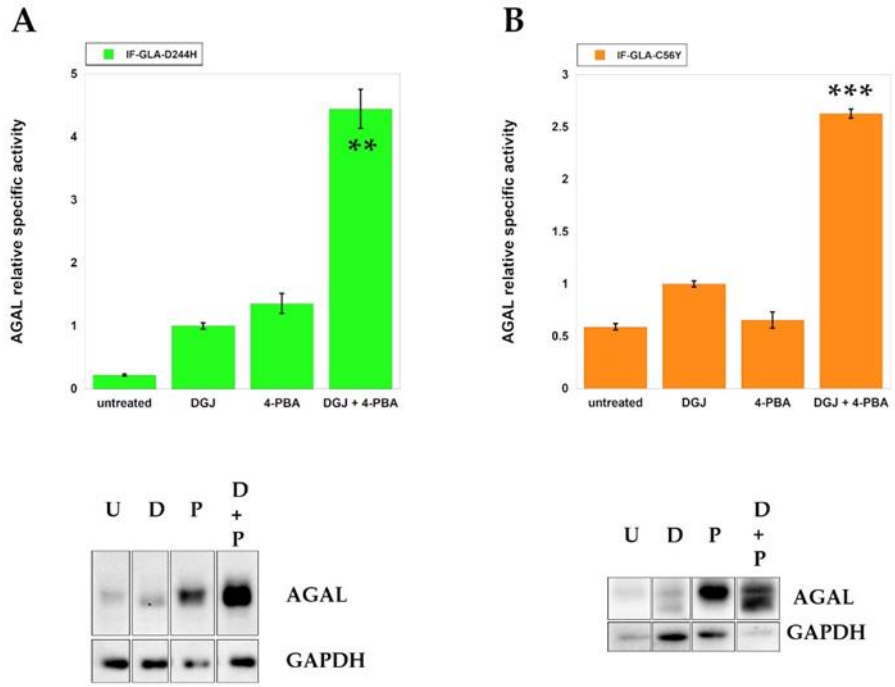

Figure 9.3. 4-phenylbutirrate is a PC-enhancer. IF-GLA-D244H (panel A) and IF-GLA-C56Y (panel B) were treated for 72 hours with 10 μM DGJ in the presence or the absence of 4 mM 4-phenylbutyrate (4-PBA). AGAL relative specific activity measured on protein extracts is shown and values are referred to DGJ monotherapy. The effects of combined treatments are significantly higher than those of monotherapy (**: *p*-value < 0.01; ***: *p*-value < 0.001). Immunoblots confirmed the results (U = untreated; D = DGJ 10 μM; P = 4-PBA 4 mM; D+P = DGJ 10 μM + 4-PBA 4 mM).

Intracellular stabilization of AGAL by DGJ is enhanced by Acetylsalicilic acid

To find useful PC enhancers among drugs that are used for chronic treatment of patients since long time, we tested aspirin. Acetylsalicylic acid (ASA) enhanced the DGJ stabilizing effect. The presence of the ASA increases AGAL precursor levels (higher molecular weight band) while the chaperone stabilizes the active form (lower molecular weight). The combined treatment results in a strong increase in the active AGAL.

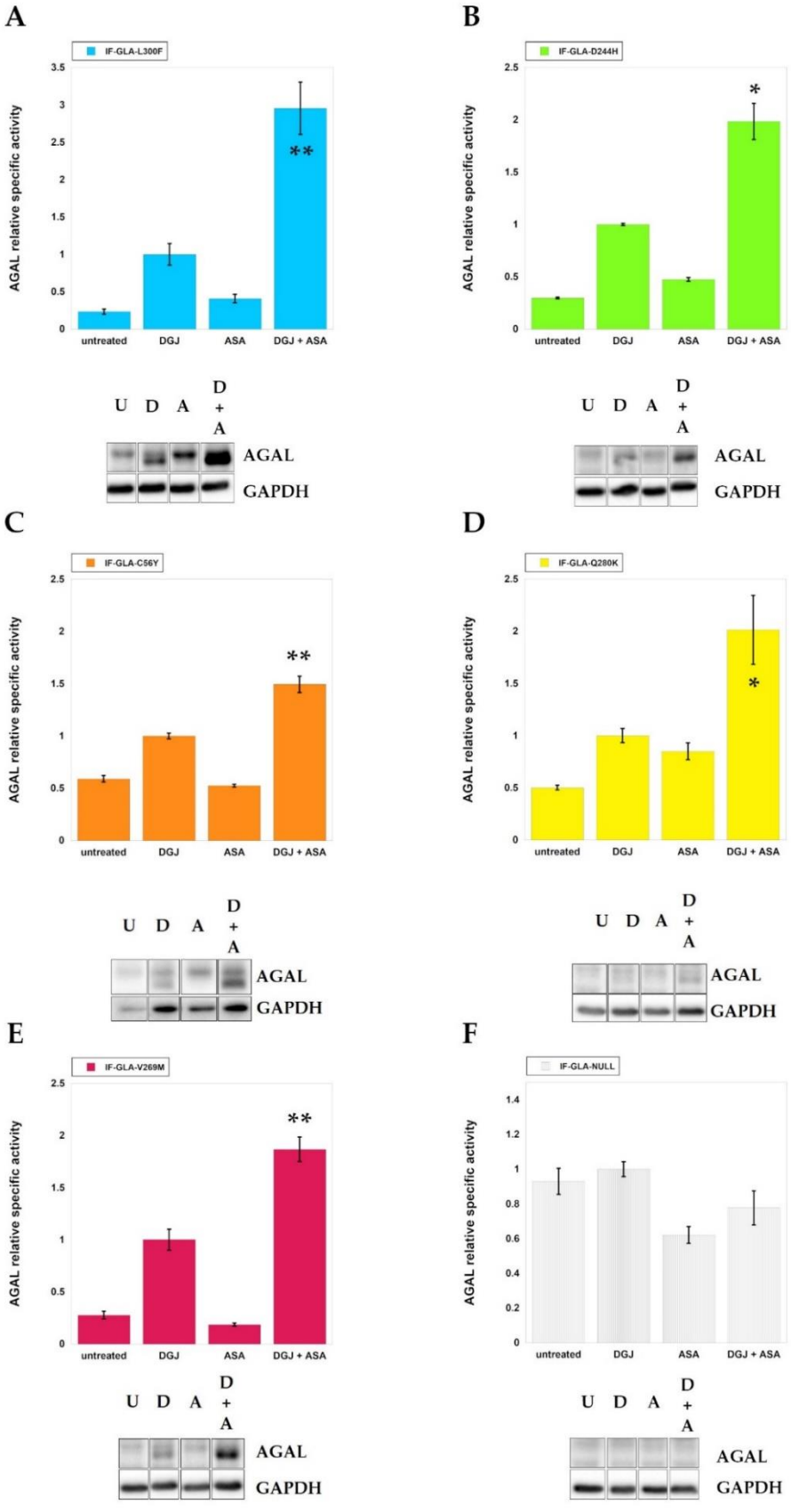
A
IF-GLA-L300F
AGAL relative specific activity
3.5
3
2.5
2
1.5
1
0.5
0
**
untreated
DGJ
ASA
DGJ + ASA
U D A D + A
AGAL
GAPDH
B
IF-GLA-D244H
AGAL relative specific activity
2.5
2
1.5
1
0.5
0
*
untreated
DGJ
ASA
DGJ + ASA
U D A D + A
AGAL
GAPDH
C
IF-GLA-C56Y
AGAL relative specific activity
2.5
2
1.5
1
0.5
0
**
untreated
DGJ
ASA
DGJ + ASA
U D A D + A
AGAL
GAPDH
D
IF-GLA-Q280K
AGAL relative specific activity
2.5
2
1.5
1
0.5
0
*
untreated
DGJ
ASA
DGJ + ASA
U D A D + A
AGAL
GAPDH
E
IF-GLA-V269M
AGAL relative specific activity
2.5
2
1.5
1
0.5
0
**
untreated
DGJ
ASA
DGJ + ASA
U D A D + A
AGAL
GAPDH
F
IF-GLA-NULL
AGAL relative specific activity
1.4
1.2
1
0.8
0.6
0.4
0.2
0
untreated
DGJ
ASA
DGJ + ASA
U D A D + A
AGAL
GAPDH

Figure 9.4. Acetylsalicylic acid is a PC-enhancer. IF-GLA-MUTs, specifically L300F (panel A), D244H (panel B), C56Y (panel C), Q280K (panel D), V269M (panel E), were treated for 72 hours with 10 μM DGJ in the presence or the absence of 4 mM acetylsalicylic acid (ASA). IF-NULL was used as a control (Panel F). AGAL relative specific activity measured on protein extracts is shown and values are referred to DGJ monotherapy. The effects of combined treatments are significantly higher than those of monotherapy (*: *p*-value < 0.05; **: *p*-value < 0.01). Immunoblots confirmed the results (U = untreated; D = DGJ 10 μM; A = ASA 4 mM; D+A = DGJ 10 μM + ASA 4 mM).

The effect of aspirin was tested at different concentrations (Figure 9.5).

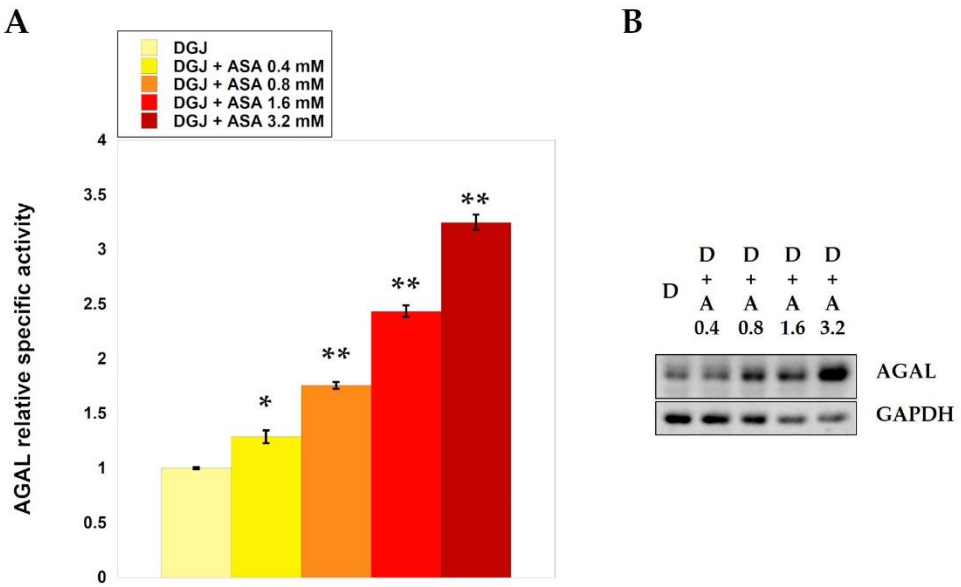

Figure 9.5. – Acetylsalicylic acid effect is dose-dependant. IF-GLA-L300F was treated with different doses of acetylsalicylic acid (range 0 – 3.2 mM) in the presence of 10 μM DGJ. AGAL relative specific activity measured on protein extracts is shown and values are referred to DGJ monotherapy. Immunoblots confirmed the results (D = DGJ 10 μM; D+A = DGJ 10 μM + ASA).

The effect of ASA is dose-dependent, and each dose improves the activity significantly respect to the preceding (*: *p*-value < 0.05; **: *p*-value < 0.01).
Although ASA stabilizes certain mutants in a statistically significant manner in monotherapy, in general it acts as a potentiator of DGJ. We tested whether ASA prolongs the effects of DGJ, thus suggesting that a less frequent administration of the drug could be considered in patients. The experimental design is sketched in figure 9.6.

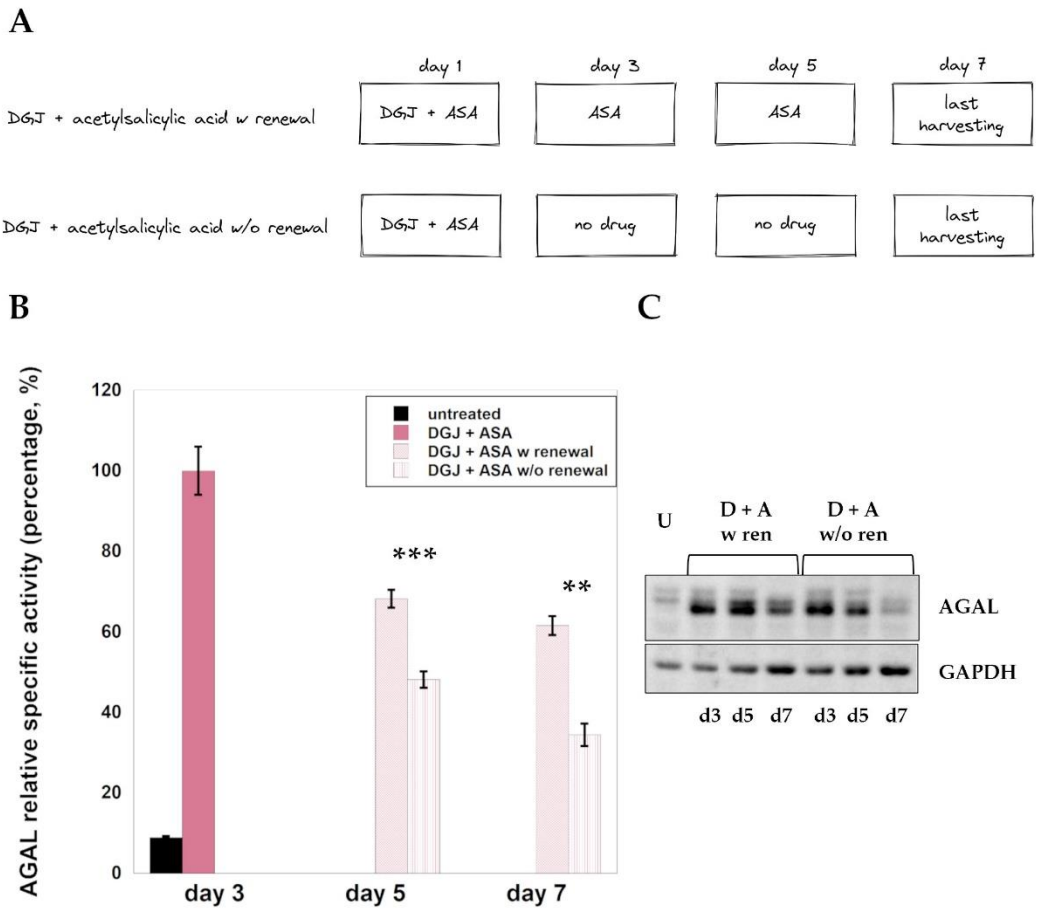

Figure 9.6. - Acetylsalicylic acid prolongs the effects of DGJ. IF-GLA-L300F cells were treated with 10 μM DGJ in the presence of 4 mM acetylsalicylic acid (ASA). Acetylsalicylic acid was renewed every other day or not for 7 days (panel A, workflow). AGAL content was evaluated through the 7 days. AGAL relative specific activity measured on protein extracts is reported in panel B, and values are referred to DGJ + ASA day 3. Panel C shows the immunoblot analysis (U = untreated; D+A w ren = DGJ 10 μM + ASA 4 mM with ASA renewal; D+A w/o ren = DGJ 10 μM + ASA 4 mM without ASA renewal).

Acetylsalicylic acid prolongs DGJ effect over the seven days (**: p-value < 0.01; ***: p-value < 0.001).

Mode of action of aspirin

Aspirin can acetylate a large range of cellular proteins [49] and in so doing it prevents protein aggregation in certain cases [50]. This does not appear to be the case

since AGAL stabilization in synergy with DGJ can be obtained using salicylate (Figure 9.7). Aspirin does not inhibit AGAL (*data not shown*). This finding does not exclude the possibility that aspirin binds allosterically without interfering with the enzymatic activity. However, it may point towards an indirect effect on stabilization.

Eighteen interactions of aspirin with human proteins are reported in the Biological General Repository for Interaction Datasets (BioGRID) [51]. As expected, most of the proteins are involved in inflammation. Interestingly, besides these, we found a protein involved in protein folding and quality control, Heat shock 70kDa protein 5, also known as BiP. This interaction has been analysed by DENG et al. [52].

A regulated expression of several genes by aspirin was reported in the Expression Omnibus (GEO) repository. We analysed the GSE58162 series because the data had been obtained under experimental conditions (2.0 mM aspirin and 72 hours of treatment) similar to those employed by us. Differential expression analysis of this dataset highlighted 1192 DEGs ($p <= 0.05$) (Supplementary file 1).

Interestingly the genes *PSMA1*, *PSMB10*, *PSMC5*, *PSMD1*, *PSMD12*, *PSME1* and *PSME4*, which encode proteasome-associated proteins, are downregulated (0.33 to 0.60 foldchange range).

Some genes that encode molecular chaperones are affected too. *TSC1*, *FKBP10* and *DNAJB2* are hyper-expressed (respectively 2.27, 2.11, 2.00 fold-change), while *HSPB3*, *DNAJC15*, *FKBP5*, *HSPA14* and *DNAJC7* are hypo-expressed. The list of proteasome-related and molecular chaperones genes obtained from the analysis is available as Supplementary file 2.

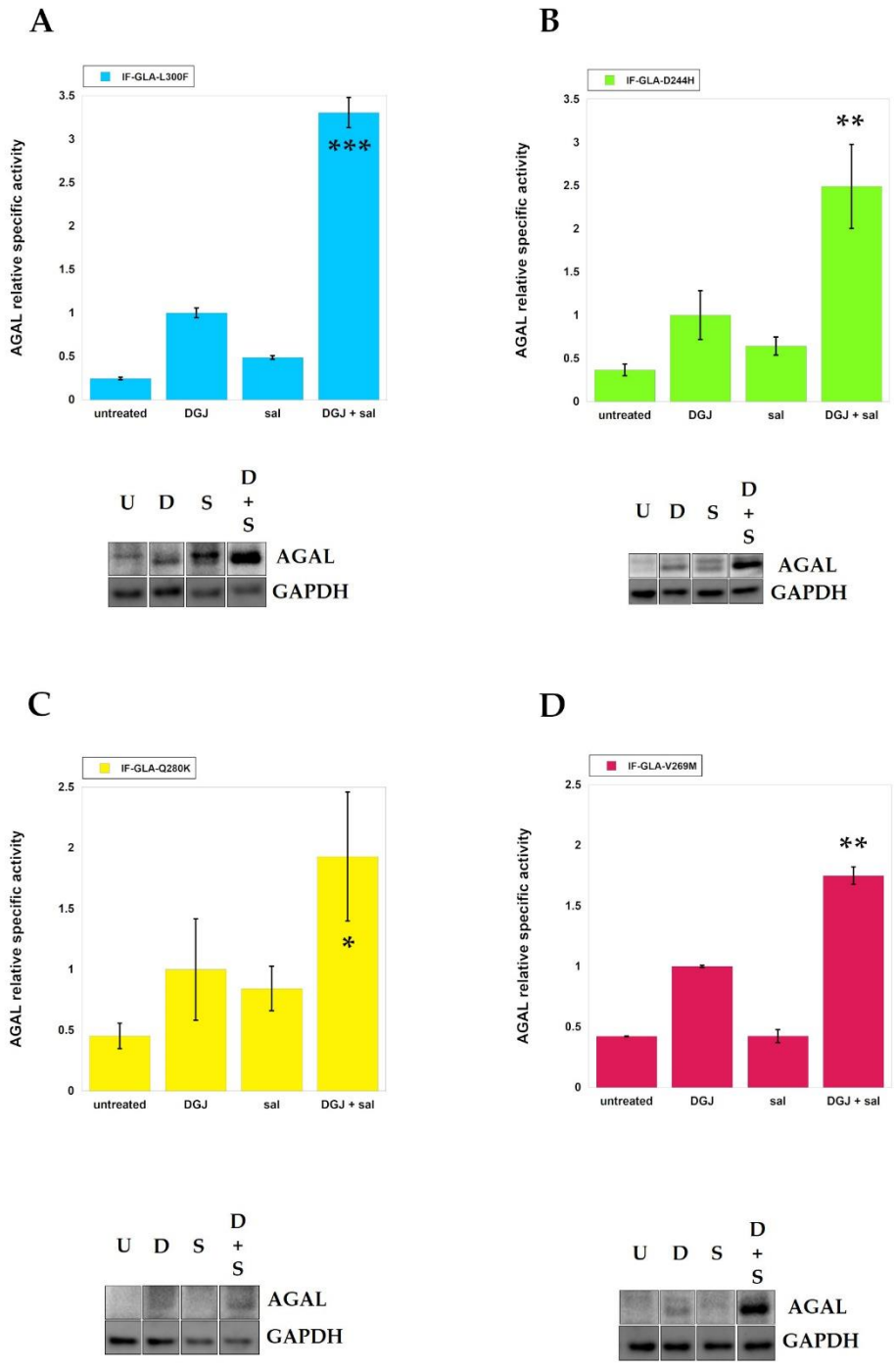

Figure 9.7. – Salicylic acid acts as acetylsalicylic acid. IF-GLA-MUTs, specifically L300F (panel A), D244H (panel B), Q280K (panel C), V269M (panel C), were treated for 72 hours with 10 μM DGJ in the presence or the absence of 4 mM salicylic acid (sal). AGAL relative specific activity measured on protein extracts is shown and values are referred to DGJ monotherapy. The effects of

combined treatments are significantly higher than those of monotherapy (*: *p*-value < 0.05; **: *p*-value < 0.01, ***: *p*-value < 0.001). Immunoblots confirmed the results (U = untreated; D = DGJ 10 μM; S = sal 4 mM; D+S = DGJ 10 μM + sal 4 mM).

Discussion

In this paper we present a case of drug repositioning for the treatment of rare diseases. This approach has been proposed by many authors [53–55]. Indeed iminosugar, which have been proposed for the treatment of lysosomal storage disorders, are them-selves a successful example of drug repositioning since they had been proposed as antiviral agents for the cure of HIV [56].

Finding potentiators of DGJ was our aim. We employed two drugs already tested for FD, namely ambroxol and 4-PBA, and proved that they work in co-administration with DGJ in cells that do not overexpress AGAL.

We moved to aspirin, a molecule widely employed in the chronical management of different pathologies. In the case of FD patients, it is often associated to specific treatments such as ERT, for stroke prophylaxis [57,58]. Our results highlighted a novel role of aspirin, that is its ability to enhance the stabilizing effect of DGJ on responsive *GLA* mutants. We are aware of the fact aspirin can have adverse effects such as bleeding and gastric mucosal damage that cannot be considered by tests on cells, nonetheless medical doctors could test whether co-administration of aspirin affects AGAL residual activity in patients who are treated with Galafold®.

Often a new usage for an old drug is found serendipitously and the mechanism by which a certain molecule, that has been designed for a specific target, works on a different one, is not clear. We can put forward some hypothesis to explain how aspirin accomplish its role. We are aware of the fact that further work will be needed to disentangle the complex matter.

Acetylsalicylic acid has been described as a proteasome inhibitor [59]. This can be confirmed *in silico*. The expression of the genes *PSMA1 PSMB2*, *PSMB8* and *PSMB10* encoding components of the 20S subunit (α6, β24, β5i and β2i), which are components of the 20S subunit as well as *PSMC5*, *PSMD1*, and *PSMD12* encoding components of the 19S subunit (Rpt6, Rpn2 and Rpn5), are down-regulated. The

same applies *PSME1* and *PMSE4* that encodes the proteasomal activators DA28α and PA200 [58].

The expression of molecular chaperones is regulated too by aspirin. Three genes are up-regulated. *TSC1* encodes hamartin, a co-chaperone and facilitator of Hsp90 chaperoning of kinase and non-kinase clients. *FKBP10* encodes peptidyl-prolyl cis-trans isomerase FKBP10, also known as FKBP-65, that forms a type I protocollagen chaperone complex with BiP and Hsp47 in the endoplasmic reticulum. *DNAJB2* encodes DnaJ homolog subfamily B member 2, also known as HSJ1. HSJ1 is a co-chaperone of Hsp70 [61] and regulates substrate binding and ATPase activity stimulating the binding of ubiquitylated proteins and their delivery to the proteasome [60]. The other chaperones are downregulated. *FKBP5* encodes a peptidyl-prolyl cis-trans isomerase FKBP51, that is a co-chaperone for Hsp90. [63]. HSPB3 protein belongs to the small heat shock protein family and interacts with MKBP/HSPB2 in muscle cells [62]. DNAJC15 belongs to the HSP40/DNAJ family of co-chaperones and acts as a Hsp70 activator by ATPase activity [63]. *HSPA14* encodes a member of the Hsp70s family, the heat shock 70 kDa protein 14, smaller than the other members and lacking a strong sequence homology [66]. DNAJC7 is a co-chaperone that has been shown to interact with the protein UFD2b [65], belonging to the U-box ubiquitin ligase [66].

Conclusions

DGJ was approved by FDA for the use on responsive mutations [69]. Several positive reports have been accumulated since its approval [70–72]. Nonetheless some concerns have been raised regarding its effectiveness [14,73]. It should not be underestimated the fact that DGJ inhibits AGAL at neutral as well as at acidic pH, and it is un-likely that it stabilizes the enzyme in the ER and does not inhibit it in the lysosome [31]. For this reason, continuous administration of the drug as well as over dosage are counterproductive. Several regimens were tested in a mouse model and in cell models. It was demonstrated that an intermitting administration of DGJ is more effective of daily administration in terms of substrate reduction [17,37]. Presently an intermittent regimen is adopted in patients, and 150 mg of the drug is administered orally every other day. In this paper we showed that it is possible to

combine drugs to potentiate the effects of DGJ. This finding opens the possibility of prolonging the stabilizing effect of DGJ and lowering its inhibitory effect. In our opinion, the fact the potentiators can be found among drugs, such as aspirin, that have been used since long time for the chronicle treatment of patients, even if their mode of action has not been elucidated, facilitates off-label usage for FD eligible patients.

Materials and Methods

Materials

Cell culture media and reagents were purchased from Gibco; Fetal Bovine Serum (South America) and trypsin from Euroclone. TRIzol™ Reagent was from ThermoFisher Scientific; QuantiTect Rev.Transcription Kit from Qiagen; SYBR Green from Biorad; cell transfection kit from InVitrogen.

pCMV6-AC vector encoding Galactosidase alpha (*GLA*) (NM_000169) Human Untagged Clone was purchased from Origene and the vectors carrying individual *GLA* mutants were then obtained as described in [25].

Enzyme assay substrate and inhibitor (4-methylumbelliferylgalactopyranoside and N-acetylgalactosamine) and lysis-M reagent for protein extraction were purchased from Sigma-Aldrich (Merck); SYBR Green and Bradford reagent from Bio-Rad.

AGAL Polyclonal Antibody (PA5-27349) and GAPDH Monoclonal Antibody (MA5-15738) were purchased from ThermoFisher Scientific. Anti-mouse secondary antibody (115-035-003) was from Jackson ImmuoResearch Laboratories; anti-rabbit secondary antibody (170-6515) from Bio-Rad.

Specific primers for RT-qPCR were purchased from either InVitrogen (*GLA*) or Sigma-Aldrich (*RPLP0*).

Fluorescence was detected using a Synergy HT Microplate Reader. qRT-PCR were performed with a StepOnePlus™ Real-Time PCR System.

Statistical analysis and graphs drawing were performed with KaleidaGraph™ 4.5.

Cell cultures and stable transfections

Cells were cultured in RPMI 1640 medium, supplemented with fetal bovine serum 10%, glutamine 2 mM, penicillin 0.5 mg/mL, streptomycin 0.5 mg/ml and non-essential amino acids at 37°C in humidified 5% CO_2.

Patient-derived fibroblasts carrying a large deletion in GLA exons 3 and 4 were obtained from Telethon Biobank and were immortalized as described by Miceli et al [72].

Briefly the cells were co-infected with HPV16 E6/E7 and hTERT lentiviral vectors (infection number 1). After a week, the cells were split and infected again only with hTERT (infection number 2) and cultured until stabilization.

Immortalized fibroblasts (IF) were transfected with individual pCMV6-AC plasmids carrying *GLA* mutants (IF-GLA-MUTs) or with the empty vector (IF-GLA-NULL) by electroporation. 5 10^6 cells from a 150 cm^2 plate were transfected with 20 μg plasmid following the manufacturer's instructions, then plated in a 60 cm^2 plate with an antibiotic free medium. 48 hours after the transfection, geneticin 0.1 mg/mL was added to the medium for the selection of transfected cells. Geneticin concentration was slowly raised up to 0.4 mg/mL then brought back to 0.1 mg/mL for maintenance. Treatments with drugs were performed in the absence of geneticin.

Quantitative real-time PCR

7 10^5 cells were plated in 60 cm^2 plate and let grown until 80-90% confluency. Cells were harvested in Trizol reagent and stored at -20°C until usage. RNA was extracted according to manufacter's instructions. RNA integrity was verified by electrophoresis on agarose gel, then 1 μg was reverse transcribed and 0.01 μg cDNA was analysed. Primer sequences were 5'-TTCAAAAGCCCAATTATACAGAAA-3' (forward) and 5'-CTGGTCCAGCAACATCAACA-3' (reverse) for *GLA* and 5'-GACGGATTACACCTTCCCACTT-3' (forward) 5'-GGCAGATGGATCAGCCAAGA-3' (reverse) for *RPLP0*. The $2^{-\Delta\Delta Ct}$ method [75] was used to calculate the relative mRNA expression.

AGAL enzymatic activity assay

Cells from a 90% confluent 20 cm^2 plate were harvested in 100 uL Roche M cOmplete lysis buffer and centrifuged at 14,000 x g for 10 minutes. The enzymatic

activity assay was performed as described in [25] with minor changes. 40 μg protein extract were incubated at 37°C for 60' in McIlvaine buffer pH 4.4, 0.4 mM 4-methylumbelliferylgalactopyranoside and 8.7 mM N-acetylgalactosamine in a total volume of 55 μL using a 96 multiwell. Reaction was stopped by addition of 140 μL GlyNaOH 1 M pH 10.5 and fluorescence at 340/445 nm ex/em was read. 4-methylumbelliferone was used for the calibration curve. Then, nmol/min of product per μg of proteins were calculated and each set of experiments was normalized setting the DGJ monotherapy activity as reference.

Bioinformatics analysis

Candidates for protein-aspirin interactions were mined from the Biological General Repository for Interaction Datasets (BioGRID, https://thebiogrid.org/chemical/935/acetylsalicylic-acid.html) using "acetylsalicylic acid" as the query.

We explored the Gene Expression Omnibus (GEO), looking for studies where aspirin administration to human cell cultures resulted in differential gene expression. We selected the GSE58162 Series because the experimental conditions (72 hours treatment, 2.0 mM aspirin) are similar to those employed in our tests. We used the interactive GEO2R tool (https://www.ncbi.nlm.nih.gov/geo/geo2r/), based on the geoquery (v2.60.0) and limma (v 3.48.3) R packages, to evaluate the effect of aspirin over gene expression.

The data in GSE58162 Series derive from 3 samples "Treated with 2.0 mM aspirin" and 3 controls, "Untreated with aspirin". Visual inspection of the mean-variance trends convinced us to reject the constant variance approximation and use precision weights. Both the trends and precision weights were calculated and visualised with the vooma()function. Logarithm of Fold Change and its significance were calculated using the eBayes() function, and the significantly differentially expressed genes (DEG) were extracted with the topTable() function. The DEG list was then intersected with two gene sets: chaperones [76] and proteasome [60]. A slightly edited and reduced version of the R script used by GEO2R for DEG list generation and subsetting is available as Supplementary File 3.

Miscellaneous

Protein concentration was determined using the Bradford method and BSA as the standard [77].

Immunoblot was performed as in [32], using 20 μg protein extract.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, File S1: Differential expression analysis of GSE58162 GEO dataset; File S2: Proteasome-related and molecular chaperones genes list derived from the GSE58162 GEO dataset analysis; S3: R script used for the analysis.

Author Contributions: Conceptualization, MM, MA, LL, AB; validation, MVC, JL, GA; formal analysis, MM, BHM; data curation, BHM; writing—original draft preparation, MVC, JL; visuali-zation, MVC, GA. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi – accession number: GSE58162].

Acknowledgments: L.L.: program at the DISTABIF, Università degli Studi della Campania “Luigi Vanvitelli”, fellowship PON RI 2014/2020 “Dottorati innovativi con caratterizzazione industriale”. M.A., program at the DISTABIF, Università degli Studi della Campania “Luigi Vanvitelli”, fellowship POR Campania FSE 2014/2020 “Dottorati di Ricerca Con Caratterizzazione Industriale”

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations:

ABX: ambroxol

AGAL: lysosomal alpha-galactosidase

ASA: acetylsalicylic acid (aspirin)

DGJ: 1-deoxygalactonojirimycin (Migalastat)

FD: Fabry Disease

IF-GLA-MUTs: immortalized fibroblasts transfected with individual pCMV6-AC plasmids carrying *GLA* mutants

IF-GLA-NULL: immortalized fibroblasts transfected with the empty vector

4-PBA: 4-phenylbutyrate

References

1. Tuttolomondo, A.; Simonetta, I.; Riolo, R.; Todaro, F.; di Chiara, T.; Miceli, S.; Pinto, A. Pathogenesis and Molecular Mechanisms of Anderson–Fabry Disease and Possible New Molecular Addressed Therapeutic Strategies. International Journal of Molecular Sciences 2021, 22.

2. Bernardes, T.P.; Foresto, R.D.; Kirsztajn, G.M. Fabry Disease: Genetics, Pathology, and Treatment. Revista da Associacao Medica Brasileira 2020, 66.

3. Chan, B.; Adam, D.N. A Review of Fabry Disease. Skin therapy letter 2018, 23.

4. Mehta, A.; Hughes, D.A. Fabry Disease; Adam, M., Ardinger, H., Pagon, R., Wallace, S., Bean, L., Mirzaa, G., Amemiya, A., Eds.; 1993;

5. Germain, D.P. Fabry Disease. Orphanet Journal of Rare Diseases 2010, 5.

6. Michaud, M.; Mauhin, W.; Belmatoug, N.; Garnotel, R.; Bedreddine, N.; Catros, F.; Ancellin, S.; Lidove, O.; Gaches, F. When and How to Diagnose Fabry Disease in Clinical Pratice. American Journal of the Medical Sciences 2020, 360.

7. Cairns, T.; Müntze, J.; Gernert, J.; Spingler, L.; Nordbeck, P.; Wanner, C. Hot Topics in Fabry Disease. Postgraduate Medical Journal 2018, 94.

8. Citro, V.; Cammisa, M.; Liguori, L.; Cimmaruta, C.; Lukas, J.; Vittoria, M.; Andreotti, G. The Large Phenotypic Spectrum of Fabry Disease Requires Graduated Diagnosis and Personalized Therapy: A Meta-Analysis Can Help to

Differentiate Missense Mutations. International Journal of Molecular Sciences 2016, 17.

9. Lukas, J.; Cimmaruta, C.; Liguori, L.; Pantoom, S.; Iwanov, K.; Petters, J.; Hund, C.; Bunschkowski, M.; Hermann, A.; Cubellis, M.V.; et al. Assessment of Gene Variant Amenability for Pharmacological Chaperone Therapy with 1-Deoxygalactonojirimycin in Fabry Disease. International Journal of Molecular Sciences 2020, 21, doi:10.3390/ijms21030956.

10. di Risi, T.; Vinciguerra, R.; Cuomo, M.; della Monica, R.; Riccio, E.; Cocozza, S.; Imbriaco, M.; Duro, G.; Pisani, A.; Chiariotti, L. DNA Methylation Impact on Fabry Disease. Clinical Epigenetics 2021, 13.

11. Echevarria, L.; Benistan, K.; Toussaint, A.; Dubourg, O.; Hagege, A.A.; Eladari, D.; Jabbour, F.; Beldjord, C.; de Mazancourt, P.; Germain, D.P. X-Chromosome Inactivation in Female Patients with Fabry Disease. Clinical Genetics 2016, 89, doi:10.1111/cge.12613.

12. Germain, D.P.; Arad, M.; Burlina, A.; Elliott, P.M.; Falissard, B.; Feldt-Rasmussen, U.; Hilz, M.J.; Hughes, D.A.; Ortiz, A.; Wanner, C.; et al. The Effect of Enzyme Replacement Therapy on Clinical Outcomes in Female Patients with Fabry Disease – A Systematic Literature Review by a European Panel of Experts. Molecular Genetics and Metabolism 2019, 126.

13. Germain, D.P.; Elliott, P.M.; Falissard, B.; Fomin, V. v.; Hilz, M.J.; Jovanovic, A.; Kantola, I.; Linhart, A.; Mignani, R.; Namdar, M.; et al. The Effect of Enzyme Replacement Therapy on Clinical Outcomes in Male Patients with Fabry Disease: A Systematic Literature Review by a European Panel of Experts. Molecular Genetics and Metabolism Reports 2019, 19.

14. Nowak, A.; Huynh-Do, U.; Krayenbuehl, P.A.; Beuschlein, F.; Schiffmann, R.; Barbey, F. Fabry Disease Genotype, Phenotype, and Migalastat Amenability: Insights from a National Cohort. Journal of Inherited Metabolic Disease 2020, 43, doi:10.1002/jimd.12167.

15. Benjamin, E.R.; della Valle, M.C.; Wu, X.; Katz, E.; Pruthi, F.; Bond, S.; Bronfin, B.; Williams, H.; Yu, J.; Bichet, D.G.; et al. The Validation of Pharmacogenetics for the Identification of Fabry Patients to Be Treated with Migalastat. Genetics in Medicine 2017, 19, doi:10.1038/gim.2016.122.

16. Cammisa, M.; Correra, A.; Andreotti, G.; Cubellis, M.V. Fabry-CEP: A Tool to Identify Fabry Mutations Responsive to Pharmacological Chaperones. Orphanet Journal of Rare Diseases 2013, 8, doi:10.1186/1750-1172-8-111.

17. Benjamin, E.R.; Flanagan, J.J.; Schilling, A.; Chang, H.H.; Agarwal, L.; Katz, E.; Wu, X.; Pine, C.; Wustman, B.; Desnick, R.J.; et al. The Pharmacological Chaperone 1-Deoxygalactonojirimycin Increases α-Galactosidase A Levels in Fabry Patient Cell Lines. In Proceedings of the Journal of Inherited Metabolic Disease; 2009; Vol. 32.

18. Shin, S.H.; Kluepfel-Stahl, S.; Cooney, A.M.; Kaneski, C.R.; Quirk, J.M.; Schiffmann, R.; Brady, R.O.; Murray, G.J. Prediction of Response of Mutated Alpha-Galactosidase A to a Pharmacological Chaperone. Pharmacogenetics and Genomics 2008, 18, doi:10.1097/FPC.0b013e32830500f4.

19. Shin, S.H.; Murray, G.J.; Kluepfel-Stahl, S.; Cooney, A.M.; Quirk, J.M.; Schiffmann, R.; Brady, R.O.; Kaneski, C.R. Screening for Pharmacological Chaperones in Fabry Disease. Biochemical and Biophysical Research Communications 2007, 359, doi:10.1016/j.bbrc.2007.05.082.

20. Ishii, S.; Chang, H.H.; Kawasaki, K.; Yasuda, K.; Wu, H.L.; Garman, S.C.; Fan, J.Q. Mutant α-Galactosidase A Enzymes Identified in Fabry Disease Patients with Residual Enzyme Activity: Biochemical Characterization and Restoration of Normal Intracellular Processing by 1-Deoxygalactonojirimycin. Biochemical Journal 2007, 406, doi:10.1042/BJ20070479.

21. Spada, M.; Pagliardini, S.; Yasuda, M.; Tukel, T.; Thiagarajan, G.; Sakuraba, H.; Ponzone, A.; Desnick, R.J. High Incidence of Later-Onset Fabry Disease Revealed by Newborn Screening. American Journal of Human Genetics 2006, 79, doi:10.1086/504601.

22. Lukas, J.; Scalia, S.; Eichler, S.; Pockrandt, A.M.; Dehn, N.; Cozma, C.; Giese, A.K.; Rolfs, A. Functional and Clinical Consequences of Novel α-Galactosidase A Mutations in Fabry Disease. Human Mutation 2016, 37, doi:10.1002/humu.22910.

23. Ishii, S.; Suzuki, Y.; Fan, J.Q. Role of Ser-65 in the Activity of α-Galactosidase A: Characterization of a Point Mutation (S65T) Detected in a Patient

with Fabry Disease. Archives of Biochemistry and Biophysics 2000, 377, doi:10.1006/abbi.2000.1743.

24. Wu, X.; Katz, E.; Valle, M.C. della; Mascioli, K.; Flanagan, J.J.; Castelli, J.P.; Schiffmann, R.; Boudes, P.; Lockhart, D.J.; Valenzano, K.J.; et al. A Pharmacogenetic Approach to Identify Mutant Forms of α-Galactosidase a That Respond to a Pharmacological Chaperone for Fabry Disease. Human Mutation 2011, 32, doi:10.1002/humu.21530.

25. Andreotti, G.; Citro, V.; de Crescenzo, A.; Orlando, P.; Cammisa, M.; Correra, A.; Cubellis, M.V. Therapy of Fabry Disease with Pharmacological Chaperones: From in Silico Predictions to in Vitro Tests. Orphanet Journal of Rare Diseases 2011, 6, doi:10.1186/1750-1172-6-66.

26. Park, J.Y.; Kim, G.H.; Kim, S.S.; Ko, J.M.; Lee, J.J.; Yoo, H.W. Effects of a Chemical Chaperone on Genetic Mutations in α-Galactosidase A in Korean Patients with Fabry Disease. Experimental and Molecular Medicine 2009, 41, doi:10.3858/emm.2009.41.1.001.

27. Giugliani, R.; Waldek, S.; Germain, D.P.; Nicholls, K.; Bichet, D.G.; Simosky, J.K.; Bragat, A.C.; Castelli, J.P.; Benjamin, E.R.; Boudes, P.F. A Phase 2 Study of Migalastat Hydrochloride in Females with Fabry Disease: Selection of Population, Safety and Pharmacodynamic Effects. Molecular Genetics and Metabolism 2013, 109, doi:10.1016/j.ymgme.2013.01.009.

28. Lukas, J.; Giese, A.K.; Markoff, A.; Grittner, U.; Kolodny, E.; Mascher, H.; Lackner, K.J.; Meyer, W.; Wree, P.; Saviouk, V.; et al. Functional Characterisation of Alpha-Galactosidase A Mutations as a Basis for a New Classification System in Fabry Disease. PLoS Genetics 2013, 9, doi:10.1371/journal.pgen.1003632.

29. Cubellis, M.V.; Baaden, M.; Andreotti, G. Taming Molecular Flexibility to Tackle Rare Diseases. Biochimie 2015, 113, doi:10.1016/j.biochi.2015.03.018.

30. Andreotti, G.; Guarracino, M.R.; Cammisa, M.; Correra, A.; Cubellis, M.V. Prediction of the Responsiveness to Pharmacological Chaperones: Lysosomal Human Alpha-Galactosidase, a Case of Study. Orphanet Journal of Rare Diseases 2010, 5, doi:10.1186/1750-1172-5-36.

31. Tsukimura, T.; Chiba, Y.; Ohno, K.; Saito, S.; Tajima, Y.; Sakuraba, H. Molecular Mechanism for Stabilization of a Mutant α-Galactosidase A Involving

M51I Amino Acid Substitution by Imino Sugars. Molecular Genetics and Metabolism 2011, 103, doi:10.1016/j.ymgme.2011.01.013.

32. Citro, V.; Peña-García, J.; Den-Haan, H.; Pérez-Sánchez, H.; del Prete, R.; Liguori, L.; Cimmaruta, C.; Lukas, J.; Cubellis, M.V.; Andreotti, G. Identification of an Allosteric Binding Site on Human Lysosomal Alpha-Galactosidase Opens the Way to New Pharmacological Chaperones for Fabry Disease. PLoS ONE 2016, 11, doi:10.1371/journal.pone.0165463.

33. Urquiza, P.; Laín, A.; Sanz-Parra, A.; Moreno, J.; Bernardo-Seisdedos, G.; Dubus, P.; González, E.; Gutiérrez-De-Juan, V.; García, S.; Eraña, H.; et al. Repurposing Ciclopirox as a Pharmacological Chaperone in a Model of Congenital Erythropoietic Porphyria. Science Translational Medicine 2018, 10, doi:10.1126/scitranslmed.aat7467.

34. Markham, A. Migalastat: First Global Approval. Drugs 2016, 76, doi:10.1007/s40265-016-0607-y.

35. Pharmacoeconomic Review Report: Migalastat (Galafold): (Amicus Therapeutics): Indication: Fabry Disease; 2018;

36. McCafferty, E.H.; Scott, L.J. Migalastat: A Review in Fabry Disease. Drugs 2019, 79, doi:10.1007/s40265-019-01090-4.

37. Khanna, R.; Soska, R.; Lun, Y.; Feng, J.; Frascella, M.; Young, B.; Brignol, N.; Pellegrino, L.; Sitaraman, S.A.; Desnick, R.J.; et al. The Pharmacological Chaperone 1-Deoxygalactonojirimycin Reduces Tissue Globotriaosylceramide Levels in a Mouse Model of Fabry Disease. Molecular Therapy 2010, 18, doi:10.1038/mt.2009.220.

38. Liguori, L.; Monticelli, M.; Allocca, M.; Mele, B.H.; Lukas, J.; Cubellis, M.V.; Andreotti, G. Pharmacological Chaperones: A Therapeutic Approach for Diseases Caused by Destabilizing Missense Mutations. International Journal of Molecular Sciences 2020, 21.

39. Lemansky, P.; Bishop, D.F.; Desnick, R.J.; Hasilik, A.; von Figura, K. Synthesis and Processing of α-Galactosidase A in Human Fibroblasts. Evidence for Different Mutations in Fabry Disease. Journal of Biological Chemistry 1987, 262, doi:10.1016/s0021-9258(18)61618-7.

40. Matsuura, F.; Ohta, M.; Ioannou, Y.A.; Desnick, R.J. Human α-Galactosidase A: Characterization of the N-Linked Oligosaccharides on the Intracellular and Secreted Glycoforms Overexpressed by Chinese Hamster Ovary Cells. Glycobiology 1998, 8, doi:10.1093/glycob/8.4.329.

41. Garman, S.C.; Garboczi, D.N. The Molecular Defect Leading to Fabry Disease: Structure of Human α-Galactosidase. Journal of Molecular Biology 2004, 337, doi:10.1016/j.jmb.2004.01.035.

42. Mohamed, F.E.; Al-Gazali, L.; Al-Jasmi, F.; Ali, B.R. Pharmaceutical Chaperones and Proteostasis Regulators in the Therapy of Lysosomal Storage Disorders: Current Perspective and Future Promises. Frontiers in Pharmacology 2017, 8.

43. Wang, F.; Segatori, L. Remodeling the Proteostasis Network to Rescue Glucocerebrosidase Variants by Inhibiting ER-Associated Degradation and Enhancing ER Folding. PLoS ONE 2013, 8, doi:10.1371/journal.pone.0061418.

44. Seemann, S.; Ernst, M.; Cimmaruta, C.; Struckmann, S.; Cozma, C.; Koczan, D.; Knospe, A.M.; Haake, L.R.; Citro, V.; Bräuer, A.U.; et al. Proteostasis Regulators Modulate Proteasomal Activity and Gene Expression to Attenuate Multiple Phenotypes in Fabry Disease. Biochemical Journal 2020, 477, doi:10.1042/BCJ20190513.

45. Lukas, J.; Pockrandt, A.M.; Seemann, S.; Sharif, M.; Runge, F.; Pohlers, S.; Zheng, C.; Gläser, A.; Beller, M.; Rolfs, A.; et al. Enzyme Enhancers for the Treatment of Fabry and Pompe Disease. Molecular Therapy 2015, 23, doi:10.1038/mt.2014.224.

46. Yam, G.H.F.; Roth, J.; Zuber, C. 4-Phenylbutyrate Rescues Trafficking Incompetent Mutant α-Galactosidase A without Restoring Its Functionality. Biochemical and Biophysical Research Communications 2007, 360, doi:10.1016/j.bbrc.2007.06.048.

47. He, B.; Moreau, R. Lipid-Regulating Properties of Butyric Acid and 4-Phenylbutyric Acid: Molecular Mechanisms and Therapeutic Applications. Pharmacological Research 2019, 144.

48. Mimori, S.; Ohtaka, H.; Koshikawa, Y.; Kawada, K.; Kaneko, M.; Okuma, Y.; Nomura, Y.; Murakami, Y.; Hamana, H. 4-Phenylbutyric Acid Protects against

Neuronal Cell Death by Primarily Acting as a Chemical Chaperone Rather than Histone Deacetylase Inhibitor. Bioorganic and Medicinal Chemistry Letters 2013, 23, doi:10.1016/j.bmcl.2013.08.001.

49. Alfonso, L.F.; Srivenugopal, K.S.; Bhat, G.J. Does Aspirin Acetylate Multiple Cellular Proteins? (Review). Molecular Medicine Reports 2009, 2.

50. Ayyadevara, S.; Balasubramaniam, M.; Kakraba, S.; Alla, R.; Mehta, J.L.; Shmookler Reis, R.J. Aspirin-Mediated Acetylation Protects Against Multiple Neurodegenerative Pathologies by Impeding Protein Aggregation. Antioxidants and Redox Signaling 2017, 27, doi:10.1089/ars.2016.6978.

51. Oughtred, R.; Rust, J.; Chang, C.; Breitkreutz, B.J.; Stark, C.; Willems, A.; Boucher, L.; Leung, G.; Kolas, N.; Zhang, F.; et al. The BioGRID Database: A Comprehensive Biomedical Resource of Curated Protein, Genetic, and Chemical Interactions. Protein Science 2021, 30, doi:10.1002/pro.3978.

52. DENG, W.-G.; RUAN, K.-H.; DU, M.; SAUNDERS, M.A.; WU, K.K. Aspirin and Salicylate Bind to Immunoglobulin Heavy Chain Binding Protein (BiP) and Inhibit Its ATPase Activity in Human Fibroblasts. The FASEB Journal 2001, 15, doi:10.1096/fj.01-0259com.

53. Hay Mele, B.; Citro, V.; Andreotti, G.; Cubellis, M.V. Drug Repositioning Can Accelerate Discovery of Pharmacological Chaperones. Orphanet Journal of Rare Diseases 2015, 10, doi:10.1186/s13023-015-0273-2.

54. Sun, W.; Zheng, W.; Simeonov, A. Drug Discovery and Development for Rare Genetic Disorders. American Journal of Medical Genetics, Part A 2017, 173.

55. Sardana, D.; Zhu, C.; Zhang, M.; Gudivada, R.C.; Yang, L.; Jegga, A.G. Drug Repositioning for Orphan Diseases. Briefings in Bioinformatics 2011, 12, doi:10.1093/bib/bbr021.

56. Tierney, M.; Pottage, J.; Kessler, H.; Fischl, M.; Richman, D.; Merigan, T.; Powderly, W.; Smith, S.; Karim, A.; Sherman, J.; et al. The Tolerability and Pharmacokinetics of N-Butyl-Deoxynojirimycin in Patients with Advanced HIV Disease (ACTG 100). Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology 1995, 10, doi:10.1097/00042560-199510050-00008.

57. Hughes, D.A.; Mehta, A.B. Vascular Complications of Fabry Disease: Enzyme Replacement and Other Therapies. In Proceedings of the Acta Paediatrica, International Journal of Paediatrics, Supplement; 2005; Vol. 94.

58. Sirrs, S.M.; Bichet, D.G.; Casey, R.; Clarke, J.T.R.; Lemoine, K.; Doucette, S.; West, M.L. Outcomes of Patients Treated through the Canadian Fabry Disease Initiative. Molecular Genetics and Metabolism 2014, 111, doi:10.1016/j.ymgme.2014.01.014.

59. Chiaviello, A.; Paciello, I.; Postiglione, I.; Crescenzi, E.; Palumbo, G. Combination of Photodynamic Therapy with Aspirin in Human-Derived Lung Adenocarcinoma Cells Affects Proteasome Activity and Induces Apoptosis. Cell Proliferation 2010, 43, doi:10.1111/j.1365-2184.2010.00698.x.

60. Gomes, A. v. Genetics of Proteasome Diseases. Scientifica 2013, 2013, doi:10.1155/2013/637629.

61. Cheetham, M.E.; Jackson, A.P.; Anderton, B.H. Regulation of 70-KDa Heat-Shock-Protein ATPase Activity and Substrate Binding by Human DnaJ-Like Proteins, HSJ1a and HSJ1b. European Journal of Biochemistry 1994, 226, doi:10.1111/j.1432-1033.1994.tb20030.x.

62. Westhoff, B.; Chapple, J.P.; van der Spuy, J.; Höhfeld, J.; Cheetham, M.E. HSJ1 Is a Neuronal Shuttling Factor for the Sorting of Chaperone Clients to the Proteasome. Current Biology 2005, 15, doi:10.1016/j.cub.2005.04.058.

63. Sabbagh, J.J.; Cordova, R.A.; Zheng, D.; Criado-Marrero, M.; Lemus, A.; Li, P.; Baker, J.D.; Nordhues, B.A.; Darling, A.L.; Martinez-Licha, C.; et al. Targeting the FKBP51/GR/Hsp90 Complex to Identify Functionally Relevant Treatments for Depression and PTSD. ACS Chemical Biology 2018, 13, doi:10.1021/acschembio.8b00454.

64. Sugiyama, Y.; Suzuki, A.; Kishikawa, M.; Akutsu, R.; Hirose, T.; Waye, M.M.Y.; Tsui, S.K.W.; Yoshida, S.; Ohno, S. Muscle Develops a Specific Form of Small Heat Shock Protein Complex Composed of MKBP/HSPB2 and HSPB3 during Myogenic Differentiation. Journal of Biological Chemistry 2000, 275, doi:10.1074/jbc.275.2.1095.

65. Alessandrini, F.; Pezzè, L.; Menendez, D.; Resnick, M.A.; Ciribilli, Y. ETV7-Mediated DNAJC15 Repression Leads to Doxorubicin Resistance in Breast Cancer Cells. Neoplasia (United States) 2018, 20, doi:10.1016/j.neo.2018.06.008.
66. Wan, T.; Zhou, X.; Chen, G.; An, H.; Chen, T.; Zhang, W.; Liu, S.; Jiang, Y.; Yang, F.; Wu, Y.; et al. Novel Heat Shock Protein Hsp70L1 Activates Dendritic Cells and Acts as a Th1 Polarizing Adjuvant. Blood 2004, 103, doi:10.1182/blood-2003-08-2828.
67. Hatakeyama, S.; Matsumoto, M.; Yada, M.; Nakayama, K.I. Interaction of U-Box-Type Ubiquitin-Protein Ligases (E3s) with Molecular Chaperones. Genes to Cells 2004, 9, doi:10.1111/j.1356-9597.2004.00742.x.
68. Hatakeyama, S.; Nakayama, K.I.I. U-Box Proteins as a New Family of Ubiquitin Ligases. Biochemical and Biophysical Research Communications 2003, 302.
69. Moran, N. FDA Approves Galafold, a Triumph for Amicus. Nature Biotechnology 2018, 36.
70. Germain, D.P.; Nicholls, K.; Giugliani, R.; Bichet, D.G.; Hughes, D.A.; Barisoni, L.M.; Colvin, R.B.; Jennette, J.C.; Skuban, N.; Castelli, J.P.; et al. Efficacy of the Pharmacologic Chaperone Migalastat in a Subset of Male Patients with the Classic Phenotype of Fabry Disease and Migalastat-Amenable Variants: Data from the Phase 3 Randomized, Multicenter, Double-Blind Clinical Trial and Extension Study. Genetics in Medicine 2019, 21, doi:10.1038/s41436-019-0451-z.
71. Müntze, J.; Gensler, D.; Maniuc, O.; Liu, D.; Cairns, T.; Oder, D.; Hu, K.; Lorenz, K.; Frantz, S.; Wanner, C.; et al. Oral Chaperone Therapy Migalastat for Treating Fabry Disease: Enzymatic Response and Serum Biomarker Changes After 1 Year. Clinical Pharmacology and Therapeutics 2019, 105, doi:10.1002/cpt.1321.
72. Mauer, M.; Sokolovskiy, A.; Barth, J.A.; Castelli, J.P.; Williams, H.N.; Benjamin, E.R.; Najafian, B. Reduction of Podocyte Globotriaosylceramide Content in Adult Male Patients with Fabry Disease with Amenable GLA Mutations Following 6 Months of Migalastat Treatment. Journal of Medical Genetics 2017, 54, doi:10.1136/jmedgenet-2017-104826.
73. Körver, S.; Feldt-Rasmussen, U.; Svarstad, E.; Kantola, I.; Langeveld, M. Oral Chaperone Therapy Migalastat for the Treatment of Fabry Disease: Potentials

and Pitfalls of Real-World Data. Clinical Pharmacology and Therapeutics 2019, 106.

74. Miceli, M.; Franci, G.; Dell'Aversana, C.; Ricciardiello, F.; Petraglia, F.; Carissimo, A.; Perone, L.; Maruotti, G.M.; Savarese, M.; Martinelli, P.; et al. MePR: A Novel Human Mesenchymal Progenitor Model with Characteristics of Pluripotency. Stem Cells and Development 2013, 22, doi:10.1089/scd.2012.0498.

75. Kenneth J. Livak; Thomas D. Schmittgen Analysis of Relative Gene Expression Data Using Real- Time Quantitative PCR and the 22DDCT Method. METHODS 2001, 25.

76. Shemesh, N.; Jubran, J.; Dror, S.; Simonovsky, E.; Basha, O.; Argov, C.; Hekselman, I.; Abu-Qarn, M.; Vinogradov, E.; Mauer, O.; et al. The Landscape of Molecular Chaperones across Human Tissues Reveals a Layered Architecture of Core and Variable Chaperones. Nature Communications 2021, 12, doi:10.1038/s41467-021-22369-9.

77. Bradford, M.M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. Analytical Biochemistry 1976, 72, doi:10.1016/0003-2697(76)90527-3.

10. DISCUSSION

The relationship between missense mutations and pathological conditions is complex and can be metaphorically described as “two sides of a coin”.

The inheritance of damaging mutations is the core of Mendelian diseases, which accounts for almost all of the rare diseases currently known. On the contrary, variants that are frequent in the population and are not pathogenic *per se* but have an impact on the protein, might influence the course of different diseases.

The influence of genetic factors has had a great relevance in public opinion about the pandemic emergence of COVID-19. The great variability of the phenotype, ranging from asymptomatic individuals to the most severe manifestations that required intubation, suggested the possibility of an individual predisposition. Since the beginning of the emergence, much evidence have been published (Benetti et al., 2020; Freitas et al., 2021; Latini et al., 2020; Saponi-Cortes et al., 2021).

The work presented in this thesis contributed to addressing this topic. In particular, we identified five common variants in the general population which have different allele frequencies in different populations and can be related to SARS-CoV* in literature. Three of these variants were in the gene *MBL2*, encoding a mannose-binding lectin with a fundamental role in innate immunity; one was in *TMPRSS2*, the gene that encodes the serin-protease described to cleave the Spike protein on SARS-CoV-2, before the virus internalization; the last one was in *CD27*, encoding a receptor that connects the CD4 and CD8 branches of the immune system. The different frequencies of these mutations in the different populations were correlated with the incidence and/or mortality rates in the different associated countries.

Many factors contribute to the difficulty of establishing such a kind of correlation: socio-behavioural habits, politics, the possibility of co-existence with other viruses, and so on. Thus, the suggestion that the host genetic background is of utmost importance in the COVID-19 course, found strong support in the analysis of whole-exome sequencing data from patients. The analysis of a large cohort of Italian patients (1177) showed the previously identified *TMPRSS2* variant p.Val197Met as an important predictive factor. In particular, a relevant protective effect of this variant was revealed in two sub-cohorts of patients: young men (age $\leq$ 60 years) and elderly women (age $\geq$ 58 years). *TMPRSS2* is regulated by sexual hormones; its expression is induced by androgens, while estrogens (especially estradiol) have

the opposite effect. Thus, it is not surprising that two identified categories show the greatest sensitivity to the mutation, for the opposite reasons: 1) young men would be at higher risk because of the highest androgens amount; 2) elderly women would be at higher risk because of the lowest estrogens amount. Both the categories would have the higher expression of *TMPRSS2*, thus they would benefit the most from its destabilizing mutation.

Missense mutations play a central role in rare diseases.

Almost 80% of rare diseases are genetic diseases, and nearly all of them are Mendelian diseases, which means they are caused by mutations in a single gene (Rahit & Tarailo-Graovac, 2020).

"Rarity" has not been defined in the scientific community, since different countries refers to different definitions. Nevertheless, the characteristic of being rare implies many difficulties that are common to diseases very different from each other. For example, the diagnosis is a major challenge in these cases. Many different rare diseases share common symptoms and phenotypical aspects, thus the expertise and the experience required from the medical doctors are not commonly found. Moreover, the differential diagnosis needs to be supported by genetic and biochemical tests. In particular, the advent of next-generation sequencing and the possibility to perform whole-exome sequencing represented a game-changer in the field (Alkuraya, 2021).

WES allows an overview of the most interesting mutations in the genome, *i.e.* the ones occurring in coding regions. Typically, by WES analysis, mutations of interest can be identified, then a validation based on bioinformatics and biochemistry is required. The effect of a mutation on the protein is variable, and needs to be tested *in silico*, *in vitro* and, if possible, *in vivo*.

In the present work, we investigated four cases of Beckwith-Wiedemann syndrome associated with Multi-Locus Imprinting Disturbance, confirming the phenotype-based diagnosis. WES analysis was followed by *in silico* prediction of the mutations' effects. Our results supported the role of *PADI6* as a maternal effect gene and strongly established its role in imprinting disorders, identifying five novel mutations carried by the patients' mothers. The most interesting case was

represented by our family 1, where the healthy mother of two girls with BWS was carriers of a truncating mutation and a deleterious missense one. The presence of the truncating variant in the grandmother highlighted the insufficiency of monoallelic inactivation to cause BWS in the progeny. The pedigree of family 1, with a more severe maternal *PADI6* genotype associated with two affected siblings, also suggested that the severity of the variants may influence recurrence risk.

Among the characteristics associated with rarity, the lack of a cure is a big issue. Some patients benefit from symptomatic treatments, that ameliorate the phenotype but do not represent a cure.

Also, drug discovery in the field of rare diseases faces many difficulties. In particular, the duration and costs associated to *de novo* drug discovery render it less attractive for pharmaceutical industries and investments. In this work, two different approaches were presented to search for drugs to cure rare diseases.

On the one side, the canonical approach was used to search for a pharmacological chaperone; on the other side, a drug repositioning strategy was applied to identify approved drugs that potentiate pharmacological chaperones.

Phosphomannomutase-2 deficiency is a rare disease currently lacking any therapeutic option. It is caused by mutations in the gene *PMM2*, whose complete absence is non-compatible with life (Van Schaftingen & Jaeken, 1995). Thus, each patient carries at least one missense destabilizing mutation and retains some enzymatic activity, thus it would be in principle treatable with a pharmacological chaperone approach. PCs can be found using two different approaches: 1) through the screening of large library molecules; 2) starting from structures of known protein ligands.

First-generation pharmacological chaperones are usually substrate analogous, that binds the active site. We performed our discovery using this second procedure and investigated the suitability as a PC of the β anomer of glucose-1,6-bisphosphate, an enzyme activator. *In silico* analysis revealed the binding of the molecule, which was then confirmed *in vitro*. β-Glc-1,6-P was synthesized, purified, and characterized by ^{31}P-NMR. Its chemical characteristics allow a good separation in

terms of ppm of the ^{31}P-NMR signals with respect to those of the α-Glc-1,6-P. This feature allowed the possibility of analyzing the different behaviors of the sugars.
β-Glc-1,6-P acts as a classical pharmacological chaperone: it binds the active site, stabilizes the wt and the pathogenic mutants, but inhibits the enzymatic activity. The greatest advantage represented by the β anomer is its resistance to hydrolysis from a specific phosphatase, namely PMM1.
PMM1 is the paralogous enzyme of PMM2, that has a strong phosphatase activity on bisphosphate sugars. In particular, its phosphatase activity is enhanced by the presence of inosine monophosphate (IMP). We could prove that β-Glc-1,6-P is more resistant to PMM1 hydrolysis. The differences with the α-Glc-1,6-P become more evident when the two anomers are mixed, indicating that the α anomer is a preferential substrate, but most importantly, in the presence of IMP. In fact, in these conditions ,α-Glc-1,6-P is completely hydrolysed, while the percentage of residual β-Glc-1,6-P is around 70%. For these reasons, we believe that the pharmacokinetics of β-Glc-1,6-P could be superior to those of α-Glc-1,6-P, making this sugar a lead compound to develop PCs for phosphomannomutase-2 deficiency. The main future perspective associated to this drug development is related to the enhancement of its bioavailability. In fact, the entrance of such a polar molecule inside the cells requires a strategy. Two options are suitable: 1) chemical modification with hydrophobic derivatives that would be cleaved once the pro-drug is inside the cell; 2) use of liposomes.

Beside the ease of their discovery, first-generation PCs developed starting by the substrate chemical structure have a disadvantage, since they also act as enzyme inhibitors. This is the case, for example, of the approved PC for the Fabry-Anderson disease, namely 1-deoxygalactonojirimycin (DGJ). DGJ is currently used as a therapeutic option with an intermittent regimen.
Recently, the FDA approval of combined therapy for cystic fibrosis revealed the efficacy of potentiating PCs effects by combination with drugs influencing other pathways, such as proteostasis (Ridley & Condren, 2020).
Our aim was to apply this approach to Fabry disease, to reduce the side effects caused by the use of DGJ. To this scope, we established versatile cell models,

starting from immortalized FD fibroblasts carrying a large deletion of two exons in the gene *GLA*, the gene that encodes α-galactosidase A (AGAL). These cells were stably transfected with different AGAL mutants. Since testing proteostasis regulators under conditions of super-expression is not appropriate, thus we selected the cell lines with *GLA* expression comparable to that of healthy fibroblasts.

Drug repositioning, *i.e.* the application of an approved drug for a new therapeutic purpose, represents a convenient approach to reduce the drug discovery timing. We employed this methodology to test the usefulness of acetylsalicylic acid (aspirin) as DGJ enhancer. The combination of drugs was tested on different pathogenic mutant cell lines, and the enzyme activity was compared to the monotherapy with DGJ. Immunoblots were then used to validate the results and appreciate the increase of different AGAL forms, namely the precursor and the active form.

Our results supported the hypothesis and showed that the continuous treatment with aspirin prolongs DGJ stabilizing effect and enhances AGAL quantity in cells.

The paper presented in chapter nine is currently under revision and it is here presented in its submitted version. It is noteworthy to mention that the revision process pointed out new results, that will be included in the future published version. In particular, a long-term treatment performed with the combined therapy indicated a reduction of the AGAL substrate (Gb3), which typically accumulates in FD patients.

The mechanism of action of aspirin in increasing AGAL is not fully understood. Bioinformatics analysis of transcriptomics data available in literature leads us to hypothesize a general under-regulation of the proteasome associated genes and a differential expression of molecular chaperones, that would act synergically.

Proteomic analysis on aspirin-treated vs untreated cells confirmed a general influence of this drug on proteostasis.

The benefits of drug repositioning for rare diseases are huge.

The cell lines and the methodology that we established allow an easy screening of approved drugs and nutraceuticals, that we are currently testing.

11. BIBLIOGRAPHY

Aebi, M. (2013). N-linked protein glycosylation in the ER. In *Biochimica et Biophysica Acta - Molecular Cell Research* (Vol. 1833, Issue 11). https://doi.org/10.1016/j.bbamcr.2013.04.001

Alkuraya, F. S. (2021). How the human genome transformed the study of rare diseases. In *Nature* (Vol. 590, Issue 7845). https://doi.org/10.1038/d41586-021-00294-7

Andreotti, G., Monticelli, M., & Cubellis, M. V. (2015). Looking for protein stabilizing drugs with thermal shift assay. In *Drug Testing and Analysis* (Vol. 7, Issue 9). https://doi.org/10.1002/dta.1798

Ashburn, T. T., & Thor, K. B. (2004). Drug repositioning: Identifying and developing new uses for existing drugs. In *Nature Reviews Drug Discovery* (Vol. 3, Issue 8). https://doi.org/10.1038/nrd1468

Baker, D. (2000). A surprising simplicity to protein folding. *Nature*, *405*(6782). https://doi.org/10.1038/35011000

Bebbere, D., Masala, L., Albertini, D. F., & Ledda, S. (2016). The subcortical maternal complex: multiple functions for one biological structure? In *Journal of Assisted Reproduction and Genetics* (Vol. 33, Issue 11). https://doi.org/10.1007/s10815-016-0788-z

Benetti, E., Tita, R., Spiga, O., Ciolfi, A., Birolo, G., Bruselles, A., Doddato, G., Giliberti, A., Marconi, C., Musacchia, F., Pippucci, T., Torella, A., Trezza, A., Valentino, F., Baldassarri, M., Brusco, A., Asselta, R., Bruttini, M., Furini, S., … Pinto, A. M. (2020). ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *European Journal of Human Genetics*, *28*(11). https://doi.org/10.1038/s41431-020-0691-z

Bjursell, C., Wahlström, J., Berg, K., Stibler, H., Kristiansson, B., Matthijs, G., & Martinsson, T. (1998). Detailed mapping of the phosphomannomutase 2 (PMM2) gene and mutation detection enable improved analysis for Scandinavian CDG type I families. *European Journal of Human Genetics*, *6*(6). https://doi.org/10.1038/sj.ejhg.5200234

Cammisa, M., Correra, A., Andreotti, G., & Cubellis, M. V. (2013). Fabry-CEP: A tool to identify Fabry mutations responsive to pharmacological

chaperones. *Orphanet Journal of Rare Diseases*, *8*(1). https://doi.org/10.1186/1750-1172-8-111

Cha, Y., Erez, T., Reynolds, I. J., Kumar, D., Ross, J., Koytiger, G., Kusko, R., Zeskind, B., Risso, S., Kagan, E., Papapetropoulos, S., Grossman, I., & Laifenfeld, D. (2018). Drug repurposing from the perspective of pharmaceutical companies. In *British Journal of Pharmacology* (Vol. 175, Issue 2). https://doi.org/10.1111/bph.13798

Cimmaruta, C., Liguori, L., Monticelli, M., Andreotti, G., & Citro, V. (2017). E-learning for rare diseases: An example using Fabry disease. *International Journal of Molecular Sciences*, *18*(10). https://doi.org/10.3390/ijms18102049

Citro, V., Cammisa, M., Liguori, L., Cimmaruta, C., Lukas, J., Vittoria, M., & Andreotti, G. (2016). The large phenotypic spectrum of Fabry disease requires graduated diagnosis and personalized therapy: A Meta-Analysis can help to differentiate missense mutations. In *International Journal of Molecular Sciences* (Vol. 17, Issue 12). https://doi.org/10.3390/ijms17122010

Citro, V., Peña-García, J., Den-Haan, H., Pérez-Sánchez, H., Del Prete, R., Liguori, L., Cimmaruta, C., Lukas, J., Cubellis, M. V., & Andreotti, G. (2016). Identification of an allosteric binding site on human lysosomal alpha-galactosidase opens the way to new pharmacological chaperones for Fabry disease. *PLoS ONE*, *11*(10). https://doi.org/10.1371/journal.pone.0165463

Cubellis, M. V., Pignata, L., Verma, A., Sparago, A., Del Prete, R., Monticelli, M., Calzari, L., Antona, V., Melis, D., Tenconi, R., Russo, S., Cerrato, F., & Riccio, A. (2020). Loss-of-function maternal-effect mutations of PADI6 are associated with familial and sporadic Beckwith-Wiedemann syndrome with multi-locus imprinting disturbance. *Clinical Epigenetics*, *12*(1). https://doi.org/10.1186/s13148-020-00925-2

Demond, H., Anvar, Z., Jahromi, B. N., Sparago, A., Verma, A., Davari, M., Calzari, L., Russo, S., Jahromi, M. A., Monk, D., Andrews, S., Riccio, A., & Kelsey, G. (2019). A KHDC3L mutation resulting in recurrent hydatidiform mole causes genome-wide DNA methylation loss in oocytes and persistent

imprinting defects post-fertilization. *Genome Medicine*, *11*(1). https://doi.org/10.1186/s13073-019-0694-y

DePristo, M. A., Weinreich, D. M., & Hartl, D. L. (2005). Missense meanderings in sequence space: A biophysical view of protein evolution. In *Nature Reviews Genetics* (Vol. 6, Issue 9). https://doi.org/10.1038/nrg1672

European Commission. (n.d.). *No Title*. https://ec.europa.eu/health/non_communicable_diseases/rare_diseases_en

Fang, L., Karakiulakis, G., & Roth, M. (2020). Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? In *The Lancet Respiratory Medicine* (Vol. 8, Issue 4). https://doi.org/10.1016/S2213-2600(20)30116-8

Freitas, A. T., Calhau, C., Antunes, G., Araújo, B., Bandeira, M., Barreira, S., Bazenga, F., Braz, S., Caldeira, D., Santos, S. C. R., Faria, A., Faria, D., Fraga, M., Nogueira-Garcia, B., Gonçalves, L., Kovalchuk, P., Lacerda, L., Lopes, H., Luís, D., … Pinto, F. J. (2021). Vitamin D-related polymorphisms and vitamin D levels as risk biomarkers of COVID-19 disease severity. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-99952-z

Grünewald, S. (2009). The clinical spectrum of phosphomannomutase 2 deficiency (CDG-Ia). In *Biochimica et Biophysica Acta - Molecular Basis of Disease* (Vol. 1792, Issue 9). https://doi.org/10.1016/j.bbadis.2009.01.003

Haendel, M., Vasilevsky, N., Unni, D., Bologa, C., Harris, N., Rehm, H., Hamosh, A., Baynam, G., Groza, T., McMurry, J., Dawkins, H., Rath, A., Thaxon, C., Bocci, G., Joachimiak, M. P., Köhler, S., Robinson, P. N., Mungall, C., & Oprea, T. I. (2020). How many rare diseases are there? In *Nature Reviews Drug Discovery* (Vol. 19, Issue 2). https://doi.org/10.1038/d41573-019-00180-y

Jaeken, J., Artigas, J., Barone, R., Fiumara, A., De Koning, T. J., Poll-The, B. T., De Rijk-Van Andel, J. F., Hoffmann, G. F., Assmann, B., Mayatepek, E., Pineda, M., Vilaseca, M. A., Saudubray, J. M., Schlüter, B., Wevers, R., & Van Schaftingen, E. (1997). Phosphomannomutase deficiency is the main cause of carbohydrate-deficient glycoprotein syndrome with type I isoelectrofocusing pattern of serum sialotransferrins. *Journal of Inherited*

Metabolic Disease, *20*(3). https://doi.org/10.1023/A:1005331523477

Kane, M. S., Davids, M., Adams, C., Wolfe, L. A., Cheung, H. W., Gropman, A., Huang, Y., Ng, B. G., Freeze, H. H., Adams, D. R., Gahl, W. A., & Boerkoel, C. F. (2016). Mitotic Intragenic Recombination: A Mechanism of Survival for Several Congenital Disorders of Glycosylation. *American Journal of Human Genetics*, *98*(2). https://doi.org/10.1016/j.ajhg.2015.12.007

Kint, J. A. (1970). Fabry's disease: Alpha-galactosidase deficiency. *Science*, *167*(3922). https://doi.org/10.1126/science.167.3922.1268

Latini, A., Agolini, E., Novelli, A., Borgiani, P., Giannini, R., Gravina, P., Smarrazzo, A., Dauri, M., Andreoni, M., Rogliani, P., Bernardini, S., Helmer-citterich, M., Biancolella, M., & Novelli, G. (2020). COVID-19 and genetic variants of protein involved in the SARS-CoV-2 entry into the host cells. *Genes*, *11*(9). https://doi.org/10.3390/genes11091010

Lemansky, P., Bishop, D. F., Desnick, R. J., Hasilik, A., & von Figura, K. (1987). Synthesis and processing of α-galactosidase A in human fibroblasts. Evidence for different mutations in Fabry disease. *Journal of Biological Chemistry*, *262*(5). https://doi.org/10.1016/s0021-9258(18)61618-7

Liguori, L., Monticelli, M., Allocca, M., Mele, B. H., Lukas, J., Cubellis, M. V., & Andreotti, G. (2020). Pharmacological chaperones: A therapeutic approach for diseases caused by destabilizing missense mutations. In *International Journal of Molecular Sciences* (Vol. 21, Issue 2). https://doi.org/10.3390/ijms21020489

Lukas, J., Cimmaruta, C., Liguori, L., Pantoom, S., Iwanov, K., Petters, J., Hund, C., Bunschkowski, M., Hermann, A., Cubellis, M. V., & Rolfs, A. (2020). Assessment of gene variant amenability for pharmacological chaperone therapy with 1-deoxygalactonojirimycin in fabry disease. *International Journal of Molecular Sciences*, *21*(3). https://doi.org/10.3390/ijms21030956

Maeda, Y., & Kinoshita, T. (2008). Dolichol-phosphate mannose synthase: Structure, function and regulation. In *Biochimica et Biophysica Acta - General Subjects* (Vol. 1780, Issue 6). https://doi.org/10.1016/j.bbagen.2008.03.005

Matthijs, G., Schollen, E., Pirard, M., Budarf, M. L., Van Schaftingen, E., & Cassiman, J. J. (1997). PMM (PMM1), the human homologue of SEC53 or yeast phosphomannomutase, is localized on chromosome 22q13. *Genomics*, *40*(1). https://doi.org/10.1006/geno.1996.4536

Molinari, M. (2007). N-glycan structure dictates extension of protein folding or onset of disposal. In *Nature Chemical Biology* (Vol. 3, Issue 6). https://doi.org/10.1038/nchembio880

Monin, M. L., Mignot, C., De Lonlay, P., Héron, B., Masurel, A., Mathieu-Dramard, M., Lenaerts, C., Thauvin, C., Gérard, M., Roze, E., Jacquette, A., Charles, P., de Baracé, C., Drouin-Garraud, V., Khau Van Kien, P., Cormier-Daire, V., Mayer, M., Ogier, H., Brice, A., … Héron, D. (2014). 29 French adult patients with PMM2-congenital disorder of glycosylation: outcome of the classical pediatric phenotype and depiction of a late-onset phenotype. *Orphanet Journal of Rare Diseases*, *9*. https://doi.org/10.1186/s13023-014-0207-4

Monk, D., Mackay, D. J. G., Eggermann, T., Maher, E. R., & Riccio, A. (2019). Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. In *Nature Reviews Genetics* (Vol. 20, Issue 4). https://doi.org/10.1038/s41576-018-0092-0

Monticelli, M., Liguori, L., Allocca, M., Andreotti, G., & Cubellis, M. V. (2019). β-Glucose-1,6-Bisphosphate stabilizes pathological phophomannomutase2 mutants in vitro and represents a lead compound to develop pharmacological chaperones for the most common disorder of glycosylation, PMM2-CDG. *International Journal of Molecular Sciences*, *20*(17). https://doi.org/10.3390/ijms20174164

Monticelli, M., Mele, B. H., Andreotti, G., Cubellis, M. V., & Riccio, G. (2021). Why does SARS-CoV-2 hit in different ways? Host genetic factors can influence the acquisition or the course of COVID-19. In *European Journal of Medical Genetics* (Vol. 64, Issue 6). https://doi.org/10.1016/j.ejmg.2021.104227

Monticelli, M., Mele, B. H., Benetti, E., Fallerini, C., Baldassarri, M., Furini, S., Frullanti, E., Mari, F., Andreotti, G., Cubellis, M. V., & Renieri, A. (2021).

Protective role of a tmprss2 variant on severe covid-19 outcome in young males and elderly women. *Genes*, *12*(4). https://doi.org/10.3390/genes12040596

Morello, J. P., Salahpour, A., Laperrière, A., Bernier, V., Arthus, M. F., Lonergan, M., Petäjä-Repo, U., Angers, S., Morin, D., Bichet, D. G., & Bouvier, M. (2000). Pharmacological chaperones rescue cell-surface expression and function of misfolded V2 vasopressin receptor mutants. *Journal of Clinical Investigation*, *105*(7). https://doi.org/10.1172/JCI8688

Ng, S. B., Buckingham, K. J., Lee, C., Bigham, A. W., Tabor, H. K., Dent, K. M., Huff, C. D., Shannon, P. T., Jabs, E. W., Nickerson, D. A., Shendure, J., & Bamshad, M. J. (2010). Exome sequencing identifies the cause of a mendelian disorder. *Nature Genetics*, *42*(1). https://doi.org/10.1038/ng.499

NIH. (n.d.). *No Title*. https://www.genome.gov/genetics-glossary/Missense-Mutation

Nosengo, N. (2016). Can you teach old drugs new tricks? *Nature*, *534*(7607). https://doi.org/10.1038/534314a

Nowak, A., Mechtler, T. P., Desnick, R. J., & Kasper, D. C. (2017). Plasma LysoGb3: A useful biomarker for the diagnosis and treatment of Fabry disease heterozygotes. *Molecular Genetics and Metabolism*, *120*(1–2). https://doi.org/10.1016/j.ymgme.2016.10.006

Orphan Drug Act, (1983).

Quental, R., Moleirinho, A., Azevedo, L., & Amorim, A. (2010). Evolutionary history and functional diversification of phosphomannomutase genes. *Journal of Molecular Evolution*, *71*(2). https://doi.org/10.1007/s00239-010-9368-5

Rahit, K. M. T. H., & Tarailo-Graovac, M. (2020). Genetic Modifiers and Rare Mendelian Disease. In *Genes* (Vol. 11, Issue 3). https://doi.org/10.3390/genes11030239

Richter, T., Nestler-Parr, S., Babela, R., Khan, Z. M., Tesoro, T., Molsen, E., & Hughes, D. A. (2015). Rare Disease Terminology and Definitions-A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group. *Value in Health*, *18*(6).

https://doi.org/10.1016/j.jval.2015.05.008

Ridley, K., & Condren, M. (2020). Elexacaftor-tezacaftor-ivacaftor: The first triple-combination cystic fibrosis transmembrane conductance regulator modulating therapy. *Journal of Pediatric Pharmacology and Therapeutics, 25*(3). https://doi.org/10.5863/1551-6776-25.3.192

Rod, J. E., Oviedo-Trespalacios, O., & Cortes-Ramirez, J. (2020). A brief-review of the risk factors for covid-19 severity. *Revista de Saude Publica, 54*. https://doi.org/10.11606/S1518-8787.2020054002481

Salvatore, M., Polizzi, A., De Stefano, M. C., Floridia, G., Baldovino, S., Roccatello, D., Sciascia, S., Menegatti, E., Remuzzi, G., Daina, E., Iatropoulos, P., Bembi, B., Da Riol, R. M., Ferlini, A., Neri, M., Novelli, G., Sangiuolo, F., Brancati, F., & Taruscio, D. (2020). Improving diagnosis for rare diseases: The experience of the Italian undiagnosed Rare diseases network. *Italian Journal of Pediatrics*, *46*(1). https://doi.org/10.1186/s13052-020-00883-8

Saponi-Cortes, J. M. R., Rivas, M. D., Calle-Alonso, F., Sanchez, J. F., Costo, A., Martin, C., & Zamorano, J. (2021). IFNL4 genetic variant can predispose to COVID-19. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-00747-z

Seemann, S., Ernst, M., Cimmaruta, C., Struckmann, S., Cozma, C., Koczan, D., Knospe, A. M., Haake, L. R., Citro, V., Bräuer, A. U., Andreotti, G., Cubellis, M. V., Fuellen, G., Hermann, A., Giese, A. K., Rolfs, A., & Lukas, J. (2020). Proteostasis regulators modulate proteasomal activity and gene expression to attenuate multiple phenotypes in Fabry disease. *Biochemical Journal*, *477*(2). https://doi.org/10.1042/BCJ20190513

Serrano, M., De Diego, V., Muchart, J., Cuadras, D., Felipe, A., Macaya, A., Velázquez, R., Poo, M. P., Fons, C., O'Callaghan, M. M., García-Cazorla, A., Boix, C., Robles, B., Carratalá, F., Girós, M., Briones, P., Gort, L., Artuch, R., Pérez-Cerdá, C., … Pérez-Dueñas, B. (2015). Phosphomannomutase deficiency (PMM2-CDG): Ataxia and cerebellar assessment. *Orphanet Journal of Rare Diseases*, *10*(1). https://doi.org/10.1186/s13023-015-0358-y

Taruscio, D., Groft, S. C., Cederroth, H., Melegh, B., Lasko, P., Kosaki, K., Baynam, G., McCray, A., & Gahl, W. A. (2015). Undiagnosed Diseases Network International (UDNI): White paper for global actions to meet patient needs. In *Molecular Genetics and Metabolism* (Vol. 116, Issue 4). https://doi.org/10.1016/j.ymgme.2015.11.003

Van Schaftingen, E., & Jaeken, J. (1995). Phosphomannomutase deficiency is a cause of carbohydrate-deficient glycoprotein syndrome type I. *FEBS Letters*, *377*(3). https://doi.org/10.1016/0014-5793(95)01357-1

Wouters, O. J., McKee, M., & Luyten, J. (2020). Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. In *JAMA - Journal of the American Medical Association* (Vol. 323, Issue 9). https://doi.org/10.1001/jama.2020.1166

Yamamoto, N., & Bauer, G. (2020). Apparent difference in fatalities between Central Europe and East Asia due to SARS-COV-2 and COVID-19: Four hypotheses for possible explanation. *Medical Hypotheses, 144*. https://doi.org/10.1016/j.mehy.2020.110160

Zhou, Y., Chi, J., Lv, W., & Wang, Y. (2021). Obesity and diabetes as high-risk factors for severe coronavirus disease 2019 (Covid-19). In *Diabetes/Metabolism Research and Reviews* (Vol. 37, Issue 2). https://doi.org/10.1002/dmrr.3377

APPENDIX A – PUBLICATIONS

The following list is comprehensive of all the papers produced during the three years of Ph.D., either discussed in the present thesis or not.

1. **Monticelli, M.**; Liguori, L.; Allocca, M.; Andreotti, G.; Cubellis, M.V. "β-Glucose-1,6-Bisphosphate Stabilizes Pathological Phophomannomutase2 Mutants In Vitro and Represents a Lead Compound to Develop Pharmacological Chaperones for the Most Common Disorder of Glycosylation, PMM2-CDG". Int. J. Mol. Sci. 2019, 20, 4164.
2. L. Liguori, **M. Monticelli**, M. Allocca, M.V. Cubellis and B. Hay Mele, "Bioinformatics tools for marine biotechnology: a practical tutorial with a metagenomic approach", BMC Bioinformatics. 2020 Aug 21;21(Suppl 10):348.
3. Cubellis M. V., Pignata L., Verma A., Sparago A., Del Prete R., **Monticelli M.**, Calzari L., Antona V., Melis D., Tenconi R., Russo S., Cerrato F., & Riccio A. (2020). "Loss-of-function maternal-effect mutations of PADI6 are associated with familial and sporadic Beckwith-Wiedemann syndrome with multi-locus imprinting disturbance." Clinical epigenetics, 12(1), 139.
4. Liguori L, **Monticelli M**, Allocca M, et al. "Pharmacological Chaperones: A Therapeutic Approach for Diseases Caused by Destabilizing Missense Mutations." Int J Mol Sci. 2020;21(2):489.
5. **M. Monticelli**, B. Hay Mele, G. Andreotti, M.V. Cubellis, G. Riccio, "Why does SARS-CoV-2 hit in different ways? Host genetic factors can influence the acquisition or the course of COVID-19", Eur J Med Genet 2021 Jun;64(6):104227.
6. **M. Monticelli**, B. Hay Mele, E. Benetti, C. Fallerini, M. Baldassarri, S. Furini, E. Frullanti, F. Mari, GEN-COVID Multicenter Study, G. Andreotti, M.V. Cubellis and A. Renieri, "Protective Role of a TMPRSS2 Variant on Severe COVID-19 Outcome in Young Males and Elderly Women", Genes 2021, 12(4), 596.
7. Bosso, R. Gaglione, R. Di Girolamo, E. JA Veldhuizen, P. García-Vello, S. Fusco, V. Cafaro, **M. Monticelli**, R. Culurciello, E. Notomista, A. Arciello, E. Pizzo, "Human cryptic host defence peptide GVF27 exhibits anti-infective properties against biofilm forming members of the Burkholderia cepacia complex", Pharmaceuticals 2022, 15, 260.

8. **M. Monticelli**, L. Liguori, M.Allocca, A. Bosso, G. Andreotti, J. Lukas, M.V. Cubellis and B. Hay Mele, "Drug repositioning for Fabry disease: aspirin potentiates the stabilization of lysosomal alpha-galactosidase by pharmacological chaperones", *under revision*

9. L. Pignata, F. Cecere, A. Verma, B. Hay Mele, **M. Monticelli,** B. Acurzio, C. Giaccari, A. Sparago, J. R. Hernandez Mora, A. Monteagudo-Sánchez, M. Esteller, A. Pereda, J. Tenorio-Castano, O. Palumbo, M. Carella, P. Prontera, C. Piscopo, M. Accadia, P. Lapunzina, M. V. Cubellis, G. Perez de Nanclares, D. Monk, A. Riccio, F. Cerrato, "Novel genetic variants of KHDC3L and other members of the subcortical maternal complex associated with Beckwith-Wiedemann syndrome or Pseudohypoparathyroidism 1B and multi-locus imprinting disturbances", *submitted*

10. **M. Monticelli**, R. Francisco, S. Brasil, D. Marques-da-Silva, T. Rijoff, C. Pascoal, J. Jaeken, P. A. Videira, V. dos Reis Ferreira, "Stakeholders views on drug development: the Congenital Disorders of Glycosylation Community perspective", *manuscript in preparation*

RINGRAZIAMENTI

Se è vero che "chi parte sa da che cosa fugge, ma non sa che cosa cerca", come direbbe Lello – o con una più elegante forma, Michel de Montaigne –, mi viene da pensare a quanto imprevedibile sia stato il viaggio intrapreso poco più di tre anni fa, del quale credevo di conoscere una rotta che si è invece rivelata inaspettata, giorno dopo giorno. Le poche parole che seguono assumono l'onere di rappresentare i passi finali di questo cammino, e pertanto si caricano di un peso simbolico che mi spinge a farne l'uso migliore possibile. E non c'è altra via per onorare un percorso, che menzionarne i compagni di viaggio, in un lungo elenco dall'ordine improbabile, per quanto è arduo il compito di ribaltare su un foglio di carta una platea di persone che vivono l'una accanto all'altra.

Coloro che hanno consentito la partenza, suggerendo via via che direzione prendere, con i propri consigli e i preziosi rimproveri: la Professoressa Cubellis e la Dottoressa Andreotti, e con loro il Professore Riccio.

Coloro che hanno condiviso pezzi di ordinaria vita di laboratorio, esultanze, gioie, rabbia e delusioni: Bruno, Mariateresa, Ludovica, Valentina, Evelina, Tania. I colleghi del Dipartimento di Biologia e quelli dell'Istituto di Chimica Biomolecolare, che negli anni mi hanno accolta e supportata, con un pensiero rivolto soprattutto ad Elio, e poi Simone, Andrea, Debora, Marco, Fabio, Luigi, Emilio, Alessandro. E ancora, i colleghi del trentaquattresimo ciclo, con menzione speciale per Rosanna e Fabiano, "vicini di casa". E con loro, Andrea, Ilaria, Francesca, Giovanna, e le meravigliose persone che hanno vissuto più o meno a lungo la pigna del nostro cuore. Il "Glycoimmunology Group" dell'Università NOVA di Lisbona, guidato da Paula e Vanessa, travolgenti e appassionate ad ogni passo.

Chi ha vissuto la quotidianità dei miei giorni, ad ogni cambio di programma, imprevisto, euforia, amarezza. Mamma e Papà, sempre lì, nonostante me. Antonella e Antonio, l'essenziale. Gianni, Mena e Federica, oltre le distanze. Raffaele, il paracadute dei miei pensieri. Lisa e Lorena, e la meraviglia di riscoprirsi ogni volta come se non fosse passato un giorno. I miei Compagni, con cui è un onore condividere ideali e passione, politica e amicizia, in un crescendo di legami che si

intrecciano e si fortificano. Ilaria, come farei senza di lei a riconoscere le porte giuste?
E siccome un percorso così intenso merita una chiusura speciale, in volo verso il mio viaggio più bello: Francesco, il mio ritorno a casa.

Grazie.