1506
UNIVERSITÀ
DEGLI STUDI
DI URBINO
CARLO BO

Department of Biomolecular Sciences
Ph.D. Course in Life Sciences, Health and Biotechnologies
Curriculum Cell and Organism Biology
XXXIII cycle

N-glycosylation as a regulatory process in the IGF-1 system: from mechanisms to clinical implications

SSD: BIO/13

Supervisor

Prof. Elena Barbieri

Ph.D. Student

Dr. Laura Di Patria

Co-Advisor

Dr. Giosuè Annibalini

ACADEMIC YEAR 2019-2020

Alla mia famiglia, porto sicuro da cui salpare e approdare,

in ogni momento, con una nave carica di speranze.

A mia sorella,

mio raggio di sole.

Grazie infinite.

À Paris, la Ville Lumière, mon cœur et mes rêves sont toujours là-bas.

Et surtout, «Impossible n'est pas français».

Merci.

To all the people I have met during these years of hard work and smiles.

Thank you.

A cada viaje realizado, único instrumento para conocerme a mí misma y al mundo.

Gracias.

ABSTRACT

The introductory chapter of the Thesis provides a detailed description of glycosylation, as one of the most common and complex protein post-translation modifications critical for physiological and pathological cellular functions. Particular attention is focused on the importance of the N-glycosylation in the insulin-like growth factor-1 (IGF-1) system. IGF-1 is a polypeptide growth factor with essential roles in normal tissue development and differentiation. More than 90% of IGF-1 in the bloodstream is bound to IGF binding protein-3 (IGFBP-3) and acid labile subunit (ALS). This ternary complex of 150 kDa protects IGF-1 from proteolytic degradation, and regulates its interaction with the IGF-1 receptor (IGF-1R).

In the first chapter, we provide novel evidence for structural features of IGF-1 prohormone domains and their role as regulators of IGF-1 production. Our study highlights that IGF-1 prohormones are composed of both protein structural domain, i.e. the mature IGF-1, and intrinsically disordered regions, i.e. the C-terminal E-domains. We document that disordered E-domains have distinct regulatory functions on IGF-1 prohormones production. In particular, N-glycosylation status of Ea-domain regulates the stability and secretion of IGF-1Ea prohormone and mature IGF-1. Accordingly, the interference with IGF-1Ea prohormone N-glycosylation by tunicamycin (TUN), glucose starvation or 2-deoxyglucose directly affects protein IGF-1 level. Furthermore, we demonstrate that the alternative IGF-1Eb and IGF-1Ec prohormones are devoid of N-glycosylation sites, and hence are insensitive to modulation of glycosylation. Notably, the Eb- and Ec- disordered tails promote the nuclear accumulation of IGF-1Eb and IGF-1Ec prohormones. Thus, disordered E-domains are regulatory elements that control the structure, production, regulation, functioning and secretion of IGF-1.

The second chapter describes the effects of N-glycosylation inhibition on C2C12 myoblast differentiation. TUN treatment or genetic knockdown of *PMM2* gene, which encodes a key enzyme in N-glycosylation, inhibits myoblast fusion and interferes with the early steps of the myogenic program. In addition, we find that N-glycosylation inhibition by TUN or *PMM2* downregulation also affects the IGF-1R signaling pathway reducing the IGF-1R expression, inhibiting the AKT activation and deregulating *IGF-1* mRNA expression and protein secretion.

In the third chapter of the Thesis, we evaluate the IGF-1 system components in

fibroblasts obtained from patients with different types of Congenital Disorders of Glycosylation (CDG). We observe that primary fibroblasts from CDG patients have reduced levels of IGF-1R, associated with an impairment in the IGF-1R signaling pathway activation. Moreover, the IGF-1Ea prohormone is underglycosylated in CDG fibroblasts and it is also associated with lower IGF-1 secretion in the cell culture media.

In conclusion, our results provide new insights into the importance of N-glycosylation for the regulation of the IGF-1 system. These findings may open new diagnostic options and therapeutic strategies for disorders characterized by altered N-glycosylation.

CONTENT

ORIGINAL PAPERS

This Thesis is based on following original research articles, which will be referred to their Roman numerals.

I. L. Di Patria*, G. Annibalini*, R. Saltarelli et al. *N-glycosylation as a regulatory process in the IGF-1 system: from mechanisms to clinical implications.* – In preparation. * Both authors contributed equally.

II. G. Annibalini, S. Contarelli, M. De Santi, R. Saltarelli, L. Di Patria, M. Guescini, A. Villarini, G. Brandi, V. Stocchi & E. Barbieri, The intrinsically disordered E-domains regulate the IGF-1 prohormones stability, subcellular localisation and secretion. *Scientific Reports,* vol. 8:(1), 8919 2018. DOI:10.1038/s41598-018-27233-3.

III. G. Annibalini, L. Di Patria, G. Valli, et al. *Inhibition of N-glycosylation impaired myoblast differentiation and IGF-1 receptor signalling pathways activation.* - In preparation.

IV. L. Di Patria, R. Saltarelli, E. Barbieri et al. *Defective insulin-like growth factor 1 (IGF-1) pro-hormone N-glycosylation and reduced IGF-1 receptor signalling activation in congenital disorders of glycosylation.* - In preparation.

INTRODUCTION

Glycans

Composed of monosaccharide units in various combinations and linkages, and featuring diverse and asymmetric types of branching, glycans are among the most complex biological molecules found in nature [1].

All cells and numerous macromolecules are characterized by the presence of covalently attached sugars, monosaccharides or sugar chains, called oligosaccharides, commonly defined as glycans. Glycans have a broad range of biological functions that are essential to the development, growth, maintenance and survival of the organism that synthesizes them. Among the many structural and modulatory processes involving glycans is the formation of cell surface glycocalyx, which plays a role in cellular adhesion, migration, growth, differentiation and nutrient storage and sequestration. Host-pathogen relations are mediated by interactions with complementary glycan binding proteins (GBPs or lectins). In humans, there are more than 80 GBPs involved in both self and pathogen recognition. GBPs give rise to inflammation, direct ongoing inflammatory responses and recognize and destroy pathogens, yet pathogens can evade the immune system. Glycan recognition is the basis for the interaction between microbes and pathogens, the molecular mimicry of host glycans and the glycan gimmickry for modulating host immunity towards increased tolerance [2-4]. Glycans produced in the secretory pathway play a role in multiple mechanisms of cellular regulation. They influence protein-protein interactions, including a number of cellular functions that span from nascent protein folding and intracellular trafficking to roles in extracellular compartments where cell-cell communication is modulated by adhesion, molecular and cellular homeostasis, receptor activation and signal transduction, and endocytosis [5]. However, the function of many glycans is not yet known. The same glycan may also have different functions depending on the type of linkage it has with the attached aglycone. The aglycone, a chemical compound formed from a glycoside when a hydrogen atom replaces a glycosyl group, can be a protein or a lipid [6]. Glycans may give rise to macromolecules defined as glycoconjugates, such as glycoproteins and glycolipids [3]. A glycoprotein is a glycoconjugate in which a protein carries one or more glycans that are covalently attached to a polypeptide backbone, usually via N- or O-linkages. A glycosylphosphatidylinositol anchor (GPI) is a glycan bridge between phosphatidylinositol and a phosphoethanolamine that is in amide linkage to the

carboxyl terminus of a protein [6]. This structure typically constitutes the membrane anchors of many eukaryotic cell surface proteins [7]. A glycosphingolipid (or glycolipid) belongs to a class of membrane lipids consisting of complex glycan moieties, usually attached via glucose or galactose to the terminal primary hydroxyl group of the lipid ceramide, which is composed of a long chain base (sphingosine) and a fatty acid [6,8]. A glycosaminoglycan (also called a mucopolysaccharide or GAG) is a linear polysaccharide present in every mammalian tissue, and it is composed of disaccharide building blocks, an amino sugar and uronic acid or galactose. There are four primary groups classified according to their core disaccharide units, including heparin/heparan sulfate, chondroitin sulfate/dermatan sulfate, keratan sulfate, and hyaluronic acid. They have different functions, including regulation of cell growth and proliferation, promotion of cell adhesion, anticoagulation, and wound repair [9,10]. There are several other less studied glycan types found in animal cells and others in plants, algae, and prokaryotes. Unlike nucleic acid and protein synthesis, which are primary gene products, glycan structures are not encoded directly in the genome and are secondary gene products. The biosynthesis of glycans, namely glycosylation, is a non-template driven process carried out by the action of different enzymes such as glycosyltransferases and glycosidases. There are ~200 glycosyltransferases, enzymes that extend acceptor glycan structures using nucleotide or lipid-linked sugars as activated donor substrates. Each glycosyltransferase is more or less specific for the sugar donor, anomeric linkage and hydroxyl on the acceptor sugar [2]. On the other hand, glycosidases are a large family of enzymes with 135 members that play a role in the degradation of glycan structures, hydrolizing specific glycan linkages, for the uptake and metabolism of sugars and for the turnover of glycoconjugates in various cellular processes. Glycosidases are also involved in the formation of intermediates used as substrates for glycosyltransferases during the glycosylation. Almost 700 proteins are necessary to create the ≥7,000 mammalian glycan structures found in nature [11]. Glycans are assembled from only ten monosaccharide precursors: glucose (Glc), fucose (Fuc), galactose (Gal), mannose (Man), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc), glucuronic acid (GlcA), iduronic acid (IdoA), sialic acid (SA) and xylose (Xyl), all derivable from glucose in every cell (Fig. 1) [12]. Glucose is the central monosaccharide in carbohydrate metabolism, and it can be converted into all other sugars. It is first converted to glucose-6-P by the

hexokinase enzyme. In the glycolytic pathway, glucose-6-P is transformed to fructose-6-P by phosphoglucose isomerase or into glucose-1-P by phosphoglucomutase. The reaction of glucose-1-P with UTP forms the high-energy donor UDP-Glc. The UDP-Glc pool is used to synthesize glycogen and other glucose-containing molecules, such as glucosylceramide and dolichol-P-glucose. The pool is used in the N-linked glycan biosynthetic pathway [6]. Fucose is involved in a common form of glycosylation called fucosylation. This process consists in the incorporation of monosaccharide L-fucose (also known as 6-deoxy-L-galactose) into many glycans, and it requires both the generation of the donor substrate guanosine diphosphate L-fucose (GDP-fucose) and the presence of a set of fucosyltransferases that transfer L-fucose to the specific glycan [13]. Activated uridine diphosphate galactose (UDP-Gal) is derived by direct phosphorylation at C-1 to give galactose-1-P, which reacts with UTP to form UDP-Gal. Alternatively, galactose-1-P can be converted to UDP-Gal via a uridyl transferase exchange reaction with UDP-Glc that displaces glucose-1-P. Galactose-1-P is involved mainly in O-glycosylation. The deficiency of galactose-1-phosphate uridyltransferase (GALT) represents the most common form of galactosemia, a serious human disease which leads to intellectual disability, liver damage, and even death if galactose intake is not controlled. Mannose is used for the production of several types of glycans. Guanosine diphosphate mannose (GDP-Man) is the primary activated donor for lipid-linked oligosaccharides and a precursor for dolichol-P-mannose, necessary for the glycosylation pathway. First, mannose-6-P is synthesized and then converted to mannose-1-P. The production of mannose-6-P can occur in two ways: by direct phosphorylation via hexokinase or conversion of fructose-6-P to mannose-6-P by the phosphomannose isomerase enzyme [6]. The synthesis of the amino sugar uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) from glucose occurs via the hexosamine biosynthetic pathway. After glucose phosphorilation by hexokinase, glutamine is required to produce glucosamine-6-phosphate from fructose-6-phosphate by transamination. Glucosamine-6-P is then N-acetylated via an acetyl-CoA-mediated reaction to form N-acetylglucosamine-6-P, which is then isomerized to N-acetylglucosamine-1-P via a 1,6-bis-phosphate intermediate. Similar to the other activation reactions, N-acetylglucosamine-1-P reacts with UTP to form UDP-GlcNAc and pyrophosphate. UDP-GlcNAc is an abundant building block for N- and O-linked glycosylation, with

structural and cell signalling functions [14,15]. N-Acetylgalactosamine is another amino sugar that, in its active donor form, uridine diphosphate N-acetylgalactosamine (UDP-GalNAc), is the first monosaccharide that gives rise to O-glycosylation. It can arise by the reaction of N-acetylgalactosamine-1-P with UTP or by epimerization of UDP-GlcNAc with the same epimerase that converts UDP-glucose to UDP-galactose [16]. Glucuronic acid (GlcA) is synthesized directly from uridin-diphosphate glucose (UDP-Glc). It is used primarily for glycosaminoglycan biosynthesis, but it is also contained in some N- and O-linked glycans and glycosphingolipids. Iduronic acid is an epimer of glucuronic acid, and it is present in glycosaminoglycans dermatan sulfate, heparan sulfate, and heparin. Unlike all the other monosaccharides that constitute glycans, it is not directly obtained from a nucleotide sugar donor, but by the epimerization reaction of GlcA and its incorporation into the growing glycosaminoglycan chain. The initiation of sialic acid (known as N-acetylneuraminic acid in vertebrates) biosynthesis occurs in the cytoplasm with the main substrate UDP-N-acetylglucosamine. UDP-acetylglucosamine is converted into N-acetylmannosamine (ManNAc) by UDP- N-acetylglucosamine 2-epimerase. N-acetylmannosamine is then phosphorylated by N-acetylmannosamine kinase. Mutations in this enzyme cause two completely distinct metabolic disorders: sialuria and inclusion body myopathy type 2. Thus, sialic acid becomes “activated” by cytidine-5′-monophospho (CMP)-sialic acid synthetase in the nucleus. Cytidine-5′-monophospho (CMP)-sialic acid acts as a sialic acid donor to sialylate glycans on nascent glycoproteins and glycolipids in the Golgi apparatus during sialylation modification. Indeed, sialic acids are typically located at the terminal end of glycan structures. Xylose derives from decarboxylation of UDP-GlcA, used to initiate glycosaminoglycan synthesis in vertebrates. It is also present on proteins that have O-glucose modifications in epidermal growth factor (EGF) modules and on O-mannose-based glycans on α-dystroglycan, as well as on plant N-glycans [6]. Thus, monosaccharide residues are linked to one another enzymatically to create linear or branched oligosaccharides and to form macromolecules, named glycoconjugates to which proteins or lipids can subsequently attach. The correct assembling of glycan is regulated by a wide range of factors, including the presence of the appropriate enzymes and their localization, the availability of glycoprotein substrates and activated sugar donors in different secretory compartments in the pathway [11].

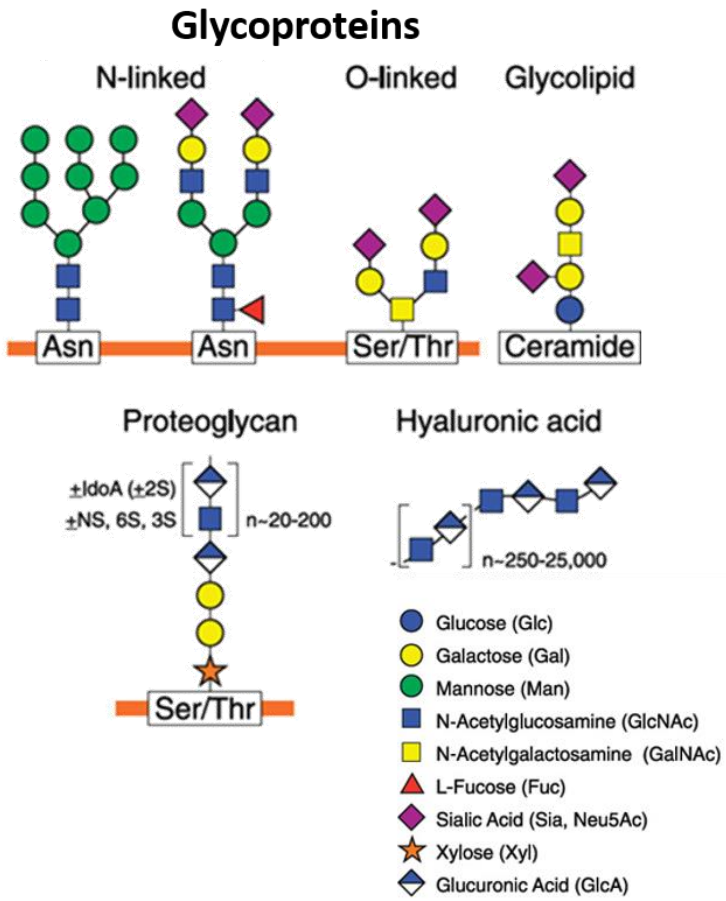


Figure 1. Major human glycans. Glycans are assembled from different monosaccharide precursors through various linkages. (*Image modified from* [2]).

Glycosylation

Specifically, glycosylation is the attachment and subsequent processing of polysaccharide chains – glycans – to proteins, to form glycoproteins. It is one of the most common and important post translational modifications that occur during or after protein synthesis [5]. Protein glycosylation is a ubiquitous phenomenon found in all domains of life [11]. Most of the blood plasma, tissue fluids, mucus, and membrane proteins are glycosylated. There are different patterns of protein glycosylation, mainly due to the expression of a variety of competing and sequentially acting enzymes within the Endoplasmic Reticulum (ER) and Golgi apparatus of each cell. This gives cells a great dynamism, allowing them to create glycoproteins with specific characteristics related to their function from a limited number of monosaccharides [17]. Glycosylation is involved in multiple cellular mechanisms that contribute to health and disease. The carbohydrate structures alone are able to determine many functions of glycoproteins of which they are part, including bioactivity, cellular/molecular trafficking and localization, binding specificity, folding, secretion and endocytosis, solubility, cell recognition, receptor activation and signalling, adhesion, antigenicity, thermodynamic stability and cellular phenotypes, such as growth, development, and disease [18-20]. Indeed, anomalous glycosylation represents a hallmark for several pathologies, including, immune deficiencies, neurodegenerative diseases, different types of cancer and microbial pathogenesis [21]. A fascinating and problematic feature of protein glycosylation is its microheterogeneity. Indeed, a given polypeptide encoded by a single gene can exist in numerous “glycoforms”, with different possibilities for branching and anomeric linkage. These glycoforms are highly complex and are much more elaborate than linear nucleic acid or polypeptide structures. Mammalian glycans are synthesized in intricate biosynthetic pathways, with a strong competition among multiple enzymes for the same glycan substrates, thus increasing their variability and complexity [6]. Another interesting property of glycoproteins is their complex pleiotropy, which means that specific changes on one glycosylation site may affect the function or recognition within a specific cellular context, yet may induce other effects or be functionally silent in other contexts [11]. In addition, the hydroxyl groups of different monosaccharides can be subject to further subsequent modifications, such as phosphorylation, sulfation, methylation, O-acetylation, or

fatty acylation, thus further increasing glycan diversity in nature and allowing glycans to mediate specific biological functions [6]. The glycosylation properties of microheterogeneity and pleiotropy make the study and structural analysis of glycoproteins difficult and complex. There are two major types of protein glycosylation in mammalian cells, namely N-linked glycosylation and O-linked glycosylation, and both types often coexist in the same protein. The carbohydrate component (N- or O-glycan) of glycoproteins constitutes at least 20% and, in some cases, can even account for up to 80% of the molecular weight. There are more than ~10^{12} different branched glycan structures in nature [4]. Correct glycosylation is essential to the normal biological activity of proteins, and its impairment leads to the synthesis of glycoproteins with altered, reduced or lost function [22]. The structures of N- and O-linked oligosaccharides are very different from one another and each of them uses distinct sugar residues for its pathways [23]. A study on potential glycosylation sites showed that three quarters of proteins may be glycosylated, but only 10% of these are O-glycosylated [24].

The biosynthesis of oligosaccharides starts with the action of sugar donors.

All the sugar nucleotides used in the synthesis of glycoproteins originate from nucleoside triphosphates and sugar phosphates in the cytosol. Many of them derive from the primary metabolism, others are salvaged following glycan catabolism or they are created from other sugars within cells. Subsequently, specific antiport proteins present in the rough ER, and Golgi cisternae import the sugar nucleotides into the lumina of these organelles and export free nucleotides (uridine monophosphate or triphosphate, cytidine-5′-monophosphate and guanosine monophosphate or triphosphate) produced within the organelles. The presence of these antiporters regulates the concentration of sugar nucleotides at a constant level, a fundamental prerequisite for oligosaccharide synthesis [6]. Monosaccharides must enter the pathway as nucleotide activated sugars to be used in glycosylation. Their activation is carried out by the formation of a high-energy ester bond between the phosphate residue and the carbon atom in the sugar. This allows the transfer of the sugar residue to an acceptor hydroxyl group, on a serine or threonine residue or on another sugar residue [23]. Glycans are subsequently formed by the sequential transfer of monosaccharides from nucleotide-sugar donors to the glycoprotein acceptor by glycosyltransferases that generate different saccharide linkages. Glycosyltransferases that act on secretory proteins are integral

membrane proteins located on the luminal side of the ER and Golgi apparatus. Each glycosyltransferase is specific to both the nucleotide sugar donor and acceptor substrate [24]. There is also evidence of competition among glycosyltransferases *in vivo* for substrates in the secretory pathway, which can modify glycan structures. Some glycosyltransferases that generate different saccharide linkages have distinct specificities for nucleotide sugar donors but the same acceptor substrate specificity, whereas others bear identical donor specificity but act on different acceptor substrates [5].

N-Glycosylation

N-glycans can be relatively large with 12 to 25 monosaccharide residues [17]. N-glycosylation sites are highly conserved in the primary sequence of acceptor proteins: the well-known NxS/T sequon, where x can be any amino acid except proline, defines a potential N-glycosylation site of proteins. The same sequon is also present as a key determinant of N-glycosylation sites in the homologous prokaryotic N-glycosylation systems [25].

In all *N*-linked oligosaccharides, *N*-acetylglucosamine (GlcNAc) is linked to the nitrogen atom of the asparagine amide group via the N-glycosid bond (GlcNAcb-Asn). This glycan type represents the most common form of N-linked protein glycosylation. In addition, typical *N*-linked oligosaccharides always contain mannose as well as *N*-acetylglucosamine and usually have several branches, each terminating with a negatively charged sialic acid residue. N-glycosylation begins in the RER with the addition of a preformed oligosaccharide precursor, linked by a pyrophosphoryl residue to the polyisoprenoid lipid dolichol. Dolichol is a long chain unsaturated compound made up of numerous isoprene units (18 – 21 carbon atoms in mammalian cells) present in the ER membrane. Here dolichol performs the function of oligosaccharide carrier (Fig. 2) [23].

A series of glycosyltransferases, encoded by the ALG defined asparagine linked glycosylation genes, perform the biosynthesis of lipid-linked oligosaccharides.

Firstly, ALG7 N-acetylglucosamine-phosphate transferase adds N-acetylglucosamine–phosphate (GlcNAc-P) to dolichol-phosphate (Dol-P) forming anhydride dolichylpyrophosphate- N-acetylglucosamine (Dol-PP-GlcNAc).

During the second step, a protein complex, composed of ALG13 and ALG14 transferases, attaches the second N-acetylglucosamine (GlcNAc) residue.

Five mannose residues are then inserted with guanosine diphosphate mannose (GDP-Man) as a substrate. The first mannose is added by the ALG1 (β-1,4 mannosyltransferase) and the formation of the two branching mannose residues is catalysed by the Alg2p enzyme.

The ALG11 enzyme acts by elongating the $Man_3GlcNAc_2$ pentasaccharide (by adding three mannose and two N-acetylglucosamine residues), yielding the $Man_5GlcNAc_2$ oligosaccharide, the final product of cytoplasmic oligosaccharide biosynthesis.

Subsequently, linked-oligosaccharide biosynthesis continues on the luminal side of the ER. Here, four mannose and three glucose residues are included in the oligosaccharide structure by specific transmembrane glycosyltransferases that catalyze α-glycosidic linkages: α-1,3 mannoslytransferase (ALG3) and ALG9 mannosyltransferase with the addition of an α-1,2 linked mannose to create the b-antenna of the oligosaccharide. The c-antenna is realized by the inclusion of two mannose residues by ALG12 α-1,6 and ALG9p α-1,2 mannosyltransferases. In parallel, Alg6p starts the glucosylation of the a-antenna of the oligosaccharide (with the addition of a glucose residue) and subsequently, Alg8p attaches the second α-1,3 linked glucose. The final step in lipid-linked oligosaccharide synthesis consists in the incorporation of the third and last glucose residue, with an α-1,2 linkage, catalysed by the Alg10p enzyme, to form the “capping” α-1,2-linked glucose [25].

This final product is a branched oligosaccharide composed of three glucose (Glc), nine mannose (Man), and two *N*-acetylglucosamine (GlcNAc): [$Glc_3Man_9(GlcNAc)_2$] [23].

The capping α-1,2-linked glucose is the key-determinant for substrate recognition by oligosaccharyltransferase and also represents the ordered assembly of the lipid-linked oligosaccharide. This allows oligosaccharyltransferase to transfer completed oligosaccharides to protein to form N-glycoproteins. Oligosaccharyltransferase (OST) is the central enzyme in the pathway of N-linked protein glycosylation. It acts by transferring the $Glc_3Man_9(GlcNAc)_2$ oligosaccharide from the lipid carrier dolichyl-pyrophosphate to the amide group of selected asparagine residues of polypeptide chains.

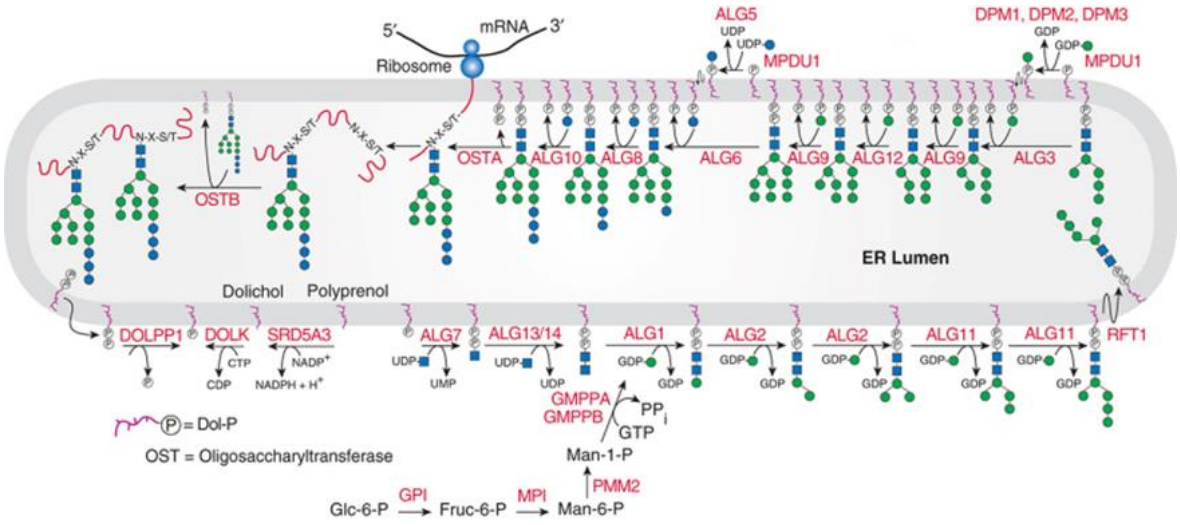


Figure 2. Steps of the dolichol-P-P-$GlcNAc_2Man_9Glc_3$ synthesis (*Image modified from* [6]).

Following the translocation of the oligosaccharide to a nascent polypeptide, all three glucose units and one mannose residue are removed by three different enzymes. A glucosyltransferase enzyme subsequently adds back one glucose to a protein-oligosaccharide. This enzyme glucosylates unfolded and misfolded, but not native, folded glycoproteins. In addition, two lectins (or carbohydrate-binding proteins – CBPs), calnexin and calreticulin, selectively bind re-glucosylated oligosaccharides and prevent folding of the adjacent amino acid segments. Rarely, proteins spontaneously dissociate from calnexin or calreticulin and are immediately deglucosylated; if they are folded correctly, they are not re-glucosylated or re-bound to a lectin, and can then pass to the Golgi. Newly formed proteins move to the Golgi complex to undergo several changes. In the *cis, medial,* and *trans* cisternae of the Golgi, there are several enzymes that introduce additional modifications to secretory and membrane proteins forming N-linked complex glycoproteins in vertebrate cells [6, 27]. After the removal of three mannose residues in the *cis*-Golgi, the protein moves to the medial-Golgi, where two more mannose residues are removed, three N-acetylglucosamine (GlcNAc) are added, and a fucose is attached. Processing is completed in the trans-Golgi with the addition of three galactose by galactosyltransferase and a final linkage of an N-acetylneuraminic acid residue to each of the galactose residues by sialyltransferase.

Defects in the N-linked glycosylation biosynthesis lead to an accumulation of intermediates and a hypoglycosylation of proteins with the consequent onset of the

serious diseases known as congenital disorders of glycosylation (CDG). Glycosylation affects all organs and accordingly, CDG show a wide spectrum of phenotypes with multisystem manifestations [26]. The effect of glycosylation inhibition on muscle differentiation will be the topic examined in chapter 2, while the impact of impaired N-glycosylation on the IGF-1 system will be described in chapter 3.

O-Glycosylation

There are different categories of O-glycan biosynthesis: N-acetylgalactosamine–to–serine/threonine type or *mucin-type*; *non-mucin* O-glycan glycosylation, such as α-linked O-mannose, α-linked O-fucose, β-linked O-xylose, β-linked O- N-acetylglucosamine, α-/β-linked O-galactose, and α-/β-linked O-glucose glycans.

In humans, O-glycosylation pathways make use of N-acetylgalactosamine (GalNAc), galactose (Gal), N-acetylglucosamine (GlcNAc), fucose (Fuc), and sialic acid (Sia) as precursors, while mannose (Man), glucose (Glc), or xylose (Xyl) residues are not required [6]. Unlike the biosynthesis of *N*-linked oligosaccharides that begins with the insertion of a large preformed oligosaccharide composed of 14 sugar residues, in O-linked glycosylation, certain sugars are removed and others attached, one at a time, in a specific order with each reaction catalyzed by a different enzyme [23, 28, 29]. O-linked glycan biosynthesis is simpler in many ways than N-linked glycan biosynthesis. The frequency of O-glycosylation in many glycoproteins is high, especially among secreted or membrane-bound mucins, which are rich in serine and threonine. Mucins are a class of glycoproteins present throughout the body and are characterized by the greatest number of O-N-acetylgalactosamine glycans [6]. Mucin type O-glycosylation, also called O-linked N-acetylgalactosamine – O-GalNAc, is the most common type of O-glycosylation in mammals and it is highly conserved through the evolution of many species [30]. Mammalian O-N-acetylgalactosamine glycans may vary from a single N-acetylgalactosamine to more than 20 sugars attached to a protein via monosaccharide N-acetylgalactosamine or (in collagens) to the hydroxyl group of hydroxylysine via galactose [6,17]. The biosynthesis of O-liked oligosaccharide consists in a sequential transfer of *N*-acetylgalactosamine (GalNAc) from Uridin diphosphate – *N*-acetylgalactosamine (UDP-GalNac) to the hydroxyl group of a serine or threonine residue in the protein via an O-glycosidic bond [23]. Unlike N-glycans, O-glycans do not have a linear consensus sequence from which their location in a protein can be determined [17]. In addition, unlike the initial step of N-glycosylation, lipid-linked intermediates and glycosidases are not required in O-*N*-acetylgalactosamine glycan biosynthesis. Donor nucleotide sugars are transported into the Golgi from the cytoplasm and, subsequently, reaction products are translocated out of the Golgi complex. O-glycosylation starts with a large polypeptide N-acetylgalactosamine-transferase

family that catalyzes the transfer of an N-acetyl-α-D-galactosamine from uridine diphosphate-N-acetylgalactosamine (UDP-GalNAc) to a hydroxyl group of serine or threonine in a nascent polypeptide. These enzymes are localized in the RER or the *cis*-Golgi network [23]. There are at least 20 distinct N-acetylgalactosamine transferases in humans, with different substrate specificity and 17 isoforms characterized to date, namely GALNT1 to 20 [24]. Protein subsequently passes to the *trans*-Golgi vesicles and undergoes the addition of a galactose residue to *N*-acetylgalactosamine by a specific *trans*-Golgi galactosyltransferase [23]. The biosynthesis continues with the elongation of the O-glycan chain to form four common core structures which will be further extended to form a mature linear or branched O-GalNAc glycan. These building blocks are created by the attachment of galactose, N-acetylgalactosamine or N-acetylglucosamine residues to the 3- and 6-linked positions of the principal N-acetylgalactosamine precursor. Core 1 is derived from the attachment of a β-1,3-linked galactose residue by core 1 β-1,3-galactosyltransferase (C1GALT1). The formation of core 2 requires the action of the β1-6 N-acetylglucosaminyltransferases 1, 3, and 4 (C2GnT-1, GCNT1s 1, 3, 4), which add an N-acetylglucosamine to the N-acetylgalactosamine of core 1. In parallel, β1-3 N-acetylglucosaminyltransferase 6 (C3GnT6; B3GNT6) catalyzes the insertion of-N-acetylglucosamine linked to α-N-acetylgalactosamine-Ser/Thr in core 3. Core 3 is also branched by the inclusion of β1-6 N-acetylglucosamine to form the core 4. These core structures represent the majority of O-glycan types found in nature. Core structures 5–8 have an extremely restricted occurrence, and core 7 has not been found in humans [29]. Subsequently, core structures may be elongated and subject to many other modifications to realize complex O-GalNAc glycans. The termini of O-GalNAc glycans may also present fucose and sialic acid in α-linkages, galactose, N-acetylgalactosamine and N-acetylglucosamine in both α- and β-linkages, and sulfate [6]. A thoroughly studied example of an O-glycosylation product is glycophorin, a heavily glycosylated sialoglycoprotein of the red blood cells. During the second step of glycophorin synthesis, a specific glycosyltransferase catalyzes the addition of a galactose residue from uridine diphosphate-galactose (UDP-Gal) to the carbon atom 3 of N-acetylgalactosamine attached to the protein forming a β1 → 3 linkage. Two different enzymes then add galactose to the 4 carbon of *N*-acetylglucosamine and a new galactose residue to the third carbon of galactose.

In vertebrate cells, this type of biosynthesis is completed by the insertion of two negatively charged *N*-acetylneuraminic acid (sialic acid) residues from a CMP precursor (Cytidine-5′-monophospho-N-acetylneuraminic acid) in the *trans*-Golgi [23].

The IGF-1 System

IGF-1

Insulin-like growth factor 1 (IGF-1) is a mitogenic polypeptide produced in most tissues and cell lines [31]. IGF-1 plays an essential role in the first decades of life in normal development and growth, cellular proliferation and differentiation and in apoptosis and necrosis inhibition. IGF-1 levels vary over the course of a lifetime. They are relatively low at birth, increase during childhood, reaching their peak in adolescence, and then begin to decline in the third decade of life [32]. The liver is the main site of production of circulating IGF-1, but every tissue is able to secrete it for autocrine, paracrine and endocrine actions [33]. The physiology of IGF-1 is complex because it behaves like a circulating hormone and also acts as a local growth factor. The hepatic synthesis of systemic IGF-1 occurs primarily in the liver, under the control of growth hormone (GH). It is also produced in multiple extrahepatic tissues, where it plays the role of a local growth factor [34,35]. The human *igf-1* gene is located within a region of over 90 kb of DNA at the long arm of chromosome 12 (12q22-24.1) and consists of six exons separated by five introns. The *igf-1* gene is subjected to numerous transcriptional and post-transcriptional modifications that give rise to several isoforms, six of which are known in the literature [43]. The existence of these different isoforms contributes to the diversity of IGF-1 actions [38,43]. The molecular mechanism generating IGF-1 isoforms and the role of N-glycosylation on the control of IGF-1 production will be the topic examined in chapter 1.

The IGF-1 Receptor and IGF-1 Signaling Pathway

Another important component of the IGF-1 system is the IGF-1 receptor. The cellular actions of IGF-1 are mediated by the IGF-1 receptor (IGF-1R), a transmembrane receptor tyrosine kinase that is expressed in IGF target tissues. IGF-1R belongs to the insulin receptor family, together with IGF-2R and insulin receptor (IR). IGF-1R is a glycoprotein containing several putative N-glycosylation sites [67]. Specifically, the α-subunit presents 11 possible glycosylation sites, whereas the β-subunit presents only 5. This indicates that the whole α-β-β-α oligotetramer may encompass 32 glycosylation positions [68]. It binds IGF-1 and IGF-2 with high affinity, mediating their biological functions by activating a complex intracellular signaling cascade leading to the transcription of IGF target genes. IGF-1R is located on chromosome 15q26.3 and spans 315 kb [69]. The IGF-1R presents a membrane spanning tetrameric structure. In its mature functional form, it consists of two identical extra cellular α-subunits responsible for hormone binding and two identical transmembrane β-subunits containing the tyrosine kinase domain activity with a cluster characterized by the presence of three tyrosine residues at positions 1131, 1135 and 1136 (Fig. 3) [70, 71].

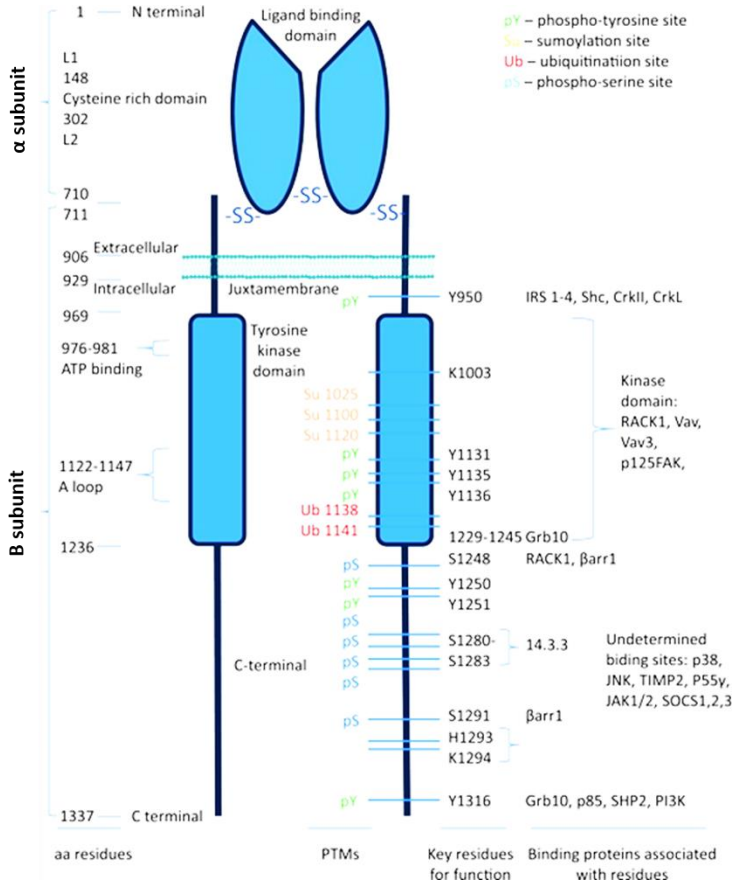


Figure 3. IGF-1 receptor structure. It presents two extra cellular α-subunits for hormone binding and two transmembrane β-subunits for the tyrosine kinase activity *(Image modified from* [70]*)*.

Ligand-circulating IGF-1 binding to the cysteine-rich domain, contained in the α-subunits of the receptor, causes a structural rearrangement in the transmembrane β-subunits. Tyrosine kinase activity is then activated and the IGF-1R dimerizes and undergoes the autophosphorylation of its cytoplasmic tyrosine kinase domain, as the kinase domains phosphorylate one another. These autophosphorylation and conformation changes permit unlimited access to a wide range of protein substrates such as members of the insulin receptor substrate (IRS) proteins, which trigger a complex signal transduction network. This transduction cascade involves the phosphatidylinositol-3 kinase (PI3K) and AKT pathways, leading to protein synthesis, cell survival and inhibition of apoptosis, as well as the pathway of MAP kinases, which stimulates cell proliferation and differentiation. Following phosphorylation, the IRS interacts with the SH2 domains of the PI3K cytoplasmic protein, which actively catalyzes the phosphorylation of PIP2 (phosphatidylinositol 4,5-bisphosphate) with the synthesis of PIP3 (phosphatidylinositol 3,4,5-triphosphate). The accumulation of high concentrations of PIP3 allows the

recruitment and activation of PDK-1 (phosphoinositide-dependent kinases). PDK-1 then phosphorylates another protein kinase: AKT at the Thr308 residue [72]. The activated AKT increases protein synthesis, promoting the stimulation of mTOR, which phosphorylates other protein substrates: p70SK and 4E-BP (eukaryotic initiation factor 4E-binding protein) in bone and muscle tissues (Fig. 4) [73].

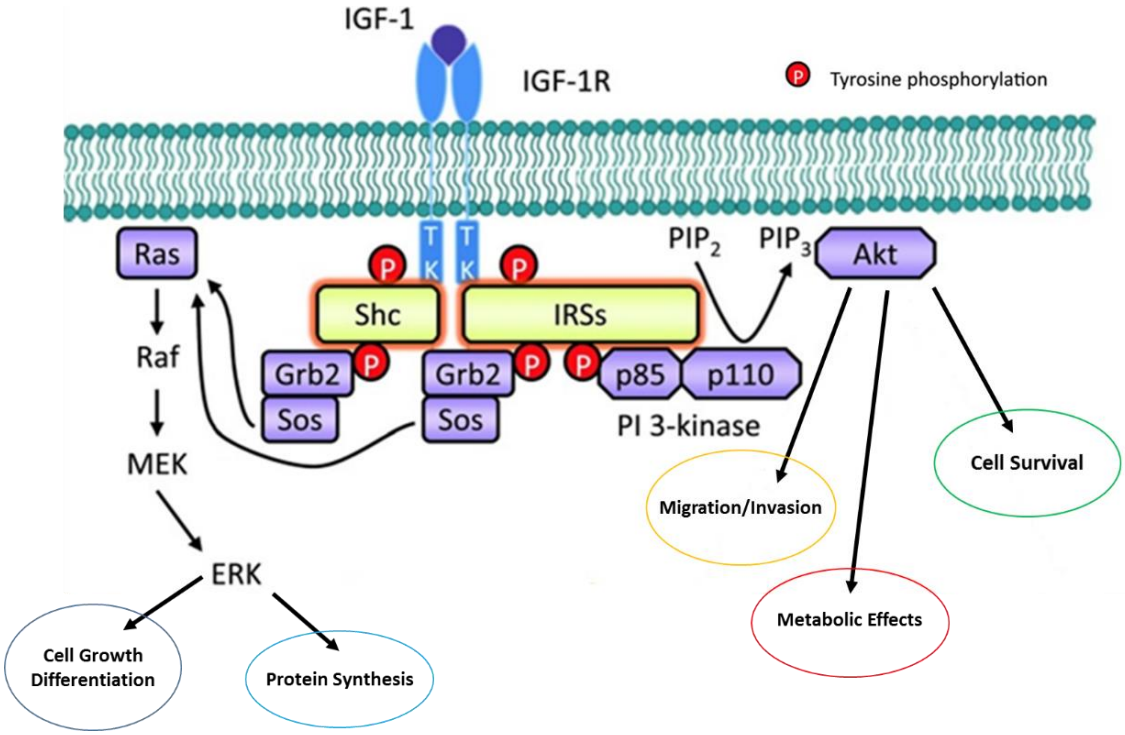


Figure 4. IGF-1 Receptor kinase-dependent signaling pathway *(Image modified from* [70]*)*.

Furthermore, AKT is implicated in another pathway, inhibiting GSK3 (glycogen synthase kinase-3), which results from its phosphorylation in an N-terminal serine residue. In response to IGF-1, GSK3 inhibition promotes dephosphorylation and activation of glycogen synthase, contributing to the stimulation of glycogen synthesis. Activated AKT also phosphorylates PKC (kinase-C protein). PKC, together with AKT, increases the cell's glucose intake by facilitating the translocation of GLUT4, glucose transporters, from the intracellular vesicle to the membrane. AKT is also involved in apoptosis regulation by inhibiting BAD, a B-cell lymphoma 2 (BCL2) associated agonist of cell death, a protein involved in the apoptosis process. When BAD proteins are not phosphorylated, they remain on the mitochondrial membrane and interact with BCL2, preventing its anti-apoptotic action. When, on the contrary, BAD is phosphorylated by AKT, it is associated with a cytoplasmic protein, and is unable to counteract the action of BCL2. In addition, AKT also

phosphorylates FOXO family members, promoting their export from the nucleus to the cytoplasm, thus reducing the induction of genes such as atrogin-1 and MuRF1, which are ubiquitin ligases involved in the protein degradation process [74]. Another action of AKT is to phosphorylate different pro-apoptotic members of the forkhead family (transcription factor), FKHRL1, FKHR and to impair their activity. AKT diminishes the expression of Fas ligand (FasL), thus decreasing Fas-mediated apoptosis. On the other hand, AKT increases levels of anti-apoptotic proteins, including BCL2 and BCL-X and various adhesion molecules of the extracellular matrix. AKT also regulates the expression of the anti-apoptotic transcription factor NF-kB, which in turn can move into the nucleus and initiate the transcription of anti-apoptotic genes. Another pathway activated by IGF-1R/IRS1 is the MAP kinase pathway, implicated in cell proliferation and migration. Here, the activated IRS interacts with the Shc that binds to the SH2 domain of Grb2, which in turn forms a complex with Sos, a guanine nucleotide exchange factor. This results in the activation of the small G-protein Ras and initiates a chain reaction of phosphorylation. Ras then stimulates the protein serine kinase Raf, which phosphorylates and activates MEK, leading to the phosphorylation and subsequent activation of ERK1/2 (MAPK). This results in the translocation of ERK1/2 to the nucleus, which phosphorylates and recruits transcriptional factors including elk-1 and c-jun, which in turn stimulate cell proliferation and cell cycle progression. These nuclear transcription factors enhance cyclin D1 and reduce p21 and p27 expression, with cell cycle progression from G1 to S phase and proliferation.

The IGF-Binding Proteins

The IGF-1 binding proteins (IGFBPs) are a family of six proteins that function as transport proteins for IGF-1. In the serum, > 90% of IGF-1 is found in a 150-kDa protein complex formed by IGF-1, the glycoprotein acid labile subunit (ALS) and the IGFBP-3, the predominant circulating binding protein, or IGFBP-5 (also detected in the ternary complex). Only less than 1% of total serum IGF-1 circulates as a free hormone [75]. The hepatic production of all three components of the 150-kDa complex is regulated by GH and their serum levels are impaired and reduced in growth hormone deficiency and resistance and elevated in growth hormone excess. The 150-kDa ternary complex plays an important role in determining the endocrine function of IGF-1 [34]. IGFBPs are a family of specific proteins with several functions: extension of the half-life of IGF-1 in the bloodstream; regulation of IGF-1-induced hypoglycaemia; control of the passage of IGF-1 from the intravascular to the extravascular space; limitation of the bioavailability of free IGF-1 to interact with the IGF-1R; improvement of IGF-1 roles by the formation of a pool of slow-release IGF-1; and direct cellular functions mediated through their own receptors, acting independently of IGF-1 [76]. These proteins present the same cysteine pattern (a highly conserved set of at least 16 cysteines) in the amino-terminal and carboxy-terminal, which is indispensable to the binding affinity of IGF-1. IGFBP-1 and IGFBP-2 are nonglycosylated, while IGFBP-3 and IGFBP-4 are N-glycosylated and IGFBP-5 and IGFBP-6 are O-glycosylated proteins [75]. IGFBP-3 is located on chromosome 7 and it is involved primarily in growth. It has three glycosylation sites and it is found in the bloodstream in the glycosylated state with a molecular weight between 40 and 44 kDa. It is produced by the hepatic endothelia and Kupffer cells. IGFBP-3 is the most abundant IGFBP in the blood. The formation of the ternary complex protects and increases the half-life of IGFBP-3 and IGFs. In fact, the half-life of unbound IGFBP-3 is between 30 and 90 minutes and the half-life of free IGF-1 is less than 15 minutes, whereas the half-life of the 150-kDa complex is approximately 16-20 hours [76]. In addition, the major IGF transport function can be attributed to IGFBP-3: it carries 75% or more of serum IGF-1 and IGF-2 in heterotrimeric complexes. IGFBP-5, present in about 10% of the molar concentration of IGFBP-3, can form similar ternary complexes. Approximately 90% of IGFBP-3 and 55% of IGFBP-5 circulate in these complexes in healthy adults. All

six IGFBPs are also found in the bloodstream in the free form or in binary complexes with IGF-1. IGFBP-3 can increase and at the same time inhibit IGF-1 actions. It has been observed that cotreatment of cells with IGFBP-3 and IGF-1 induces IGFBP-3 to inhibit IGF-1 mediated effects via high affinity sequestration of the ligand, leading to prevention of IGF-1-induced IGF-1R autophosphorylation and signaling. By contrast, preincubation of cells with IGFBP-3 before IGF-1 treatment leads to the accumulation of cell-bound forms of IGFBP-3 with lowered affinity for IGF-1, which may enhance the presentation of IGF-1 to IGF-1R. However, this mechanism is still unclear [77]. IGFBP-1, -2, -4 and -6 can also bind IGF-1 in the bloodstream and peripheral tissues but do not form part of the ternary complex. IGFBPs are present in greater concentrations than IGF-1.

The acid-labile subunit (ALS) is a soluble glycoprotein of 85 kDa with approximately 20 kDa composed of N-linked oligosaccharides, members of the leucine-rich repeat superfamily. This family is characterized by its ability to participate in protein-protein interactions. It is synthesized in the liver under the control of GH, expressed by hepatocytes and secreted into the bloodstream [69]. This protein is encoded by the IGFALS – Insulin like growth factor binding protein acid labile subunit gene, located on chromosome 16p 13.3 and spanning 3.3 kb. Structural features of ALS conserved across species include the presence of 12-13 cysteine residues, six or seven asparagine-linked glycosylation sites, and 18-20 repeating leucine-rich domains of 24 amino acids [78]. ALS is mostly stimulated by GH, as are both IGF-1 and IGFBP-3. Presence of ALS after birth is concomitant with improved sensitivity to GH as a result of an increase in GH secretion and hepatic GH receptors, and is an important key driver for the formation of the 150 kDa complex. As its name suggests, the ability of ALS to form ternary complexes is irreversibly destroyed under acidic conditions (pH < 4-5) [79]. In addition, few studies have shown that the removal of negatively charged sialic acid from the glycan chains of ALS decreases the affinity of ALS for the IGF-1 and -2 binary complexes, but does not prevent complex formation [78]. After the initial increase in ALS after puberty, ALS concentrations largely remain unchanged throughout adulthood, thereby ensuring that IGF-1 are constantly sequestered as 150 kDa complexes. ALS concentrations in the bloodstream are greater than the other components of the complex, showing that ALS has a crucial role in the storage and release of IGF-1. ALS can be found free or bound to IGF-1 or -2 and IGFBP-3 or -5 to form a ternary complex, which

prevents IGFs, free or bound to IGFBPs, from leaving the bloodstream, thus increasing their half-lives and reducing their availability at a tissue level [80]. ALS is essential for the accumulation and maintenance of serum IGF-1 and IGFBP-3. Indeed, marked decreases in serum IGF-1 and IGFBP-3 have been found in *als null* mice because the ternary complex cannot be formed. Despite this reduction in serum IGF-1 level, growth of *als null* mice is only slightly impaired, but is consistent with the poor growth observed in mice with deletions of igf-1 in the liver and highlights that systemic IGF-1 is necessary for postnatal linear growth. Besides its role as a systemic hormone, IGF-1 plays a function in the control of the autocrine and paracrine regulation of cell metabolism in multiple tissues, including cartilage and bone. Locally, the availability and activity of IGF-1 is also regulated by IGFBPs, and *in vitro* studies have showed that at the tissue level, most of the IGF is bound to IGFBPs, with a small fraction present in the unbound free form. However, *in vivo* binding interactions between IGF-1 and IGFBPs at the tissue level have not been investigated [34].

References

[1] A. Varki *et al.*, “Symbol nomenclature for graphical representations of glycans,” *Glycobiology*, vol. 25, no. 12, pp. 1323–1324, 2015.

[2] R. L. Schnaar, “Glycobiology simplified: diverse roles of glycan recognition in inflammation,” *J. Leukoc. Biol.*, vol. 99, no. 6, pp. 825–838, 2016.

[3] A. Varki, “Biological roles of glycans,” *Glycobiology*, vol. 27, no. 1, pp. 3–49, 2017.

[4] K. T. Shade, M. E. Conroy, and R. M. Anthony, “IgE Glycosylation in Health and Disease,” *Curr. Top. Microbiol. Immunol.*, vol. 423, no. 6, pp. 77–93, 2019.

[5] K. Ohtsubo and J. D. Marth, “Glycosylation in Cellular Mechanisms of Health and Disease,” *Cell*, vol. 126, no. 5, pp. 855–867, 2006.

[6] A.Varki et al., “Essentials of Glycobiology, 3rd edition,” *Ed. Cold Spring Harb. Lab. Press*. 2015-2017.

[7] T. Kinoshita and M. Fujita, “Biosynthesis of GPI-anchored proteins: Special emphasis on GPI lipid remodeling,” *J. Lipid Res.*, vol. 57, no. 1, pp. 6–24, 2016.

[8] D. Russo, S. Parashuraman, and G. D’Angelo, “Glycosphingolipid–protein interaction in signal transduction,” *Int. J. Mol. Sci.*, vol. 17, no. 10, pp. 1–21, 2016.

[9] J. Casale; J. S. Crane., Biochemistry, Glycosaminoglycans., *StatPearls Publ. LLC.* 2020.

[10] S. Morla, “Glycosaminoglycans and glycosaminoglycan mimetics in cancer and inflammation,” *Int. J. Mol. Sci.*, vol. 20, no. 8, 2019.

[11] K. W. Moremen, M. Tiemeyer, and A. V Nairn, “Vertebrate protein glycosylation: diversity, synthesis and function”, *Nat Rev Mol Cell Biol*, vol. 13, no. 7, pp. 448–462, 2014.

[12] H. Wang, A. Ramakrishnan, S. Fletcher, E. V Prochownik, and M. Genetics, “HHS Public Protein glycosylation in cancer.,” *Annu Rev Pathol*, vol. 2, no. 2, pp. 473–510, 2015.

[13] B. G. Ng *et al.*, “Pathogenic Variants in Fucokinase Cause a Congenital Disorder of Glycosylation,” *Am. J. Hum. Genet.*, vol. 103, no. 6, pp. 1030–1037, 2018.

[14] P. Santoyo-Ramos, M. Cristina Castañeda-Patlán and Martha Robles-Flores, “The Role of O-Linked β-N-Acetylglucosamine (GlcNAc) Modification in Cell

Signaling", *Petrescu, Book IntechOpen*, 2012.

[15] J. B. Konopka, "N-Acetylglucosamine Functions in Cell Signaling," *Hindawi Publishing Corporation Scientifica*, vol. 2012, pp. 1–15, 2012.

[16] V. Lorenz *et al.*, "Extrinsic functions of lectin domains in O-N-acetylgalactosamine glycan biosynthesis," *J. Biol. Chem.*, vol. 291, no. 49, pp. 25339–25350, 2016.

[17] R. Easton, "Glycosylation of Proteins - Structure, Function and Analysis," *Life Sci. - Tech. Bull.*, vol. 1, no. 48, pp. 1–5, 2011.

[18] V. Restelli and M. Butler, "The Effect of Cell Culture Parameters on Protein Glycosylation". *Cell Engineering,* 61-92, 2002.

[19] R. J. Solá, J. A. Rodríguez-Martínez, and K. Griebenow, "Modulation of protein biophysical properties by chemical glycosylation: Biochemical insights and biomedical implications," *Cell. Mol. Life Sci.*, vol. 64, no. 16, pp. 2133–2152, 2007.

[20] R. J. Solá and K. A. I. Griebenow, "Effects of Glycosylation on the Stability of Protein Pharmaceuticals," *Biochemistry*, vol. 98, no. 4, pp. 1223–1245, 2010.

[21] K. K. Palaniappan and C. R. Bertozzi, "Chemical Glycoproteomics," *Chem. Rev.*, vol. 116, no. 23, pp. 14277–14306, 2016.

[22] B. Cylwik, M. Naklicki, L. Chrostek, and E. Gruszewska, "Congenital disorders of glycosylation. part I. defects of protein N-glycosylation," *Acta Biochim. Pol.*, vol. 60, no. 2, pp. 151–161, 2013.

[23] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell, "Molecular Cell Biology". 4th edition. Section 17.7. New York: W. H. Freeman; 2000.

[24] A. G. McDonald, K. F. Tipton, and G. P. Davey, "A Knowledge-Based System for Display and Prediction of O-Glycosylation Network Behaviour in Response to Enzyme Knockouts," *PLoS Comput. Biol.*, vol. 12, no. 4, pp. 1–24, 2016.

[25] M. Aebi, "N-linked protein glycosylation in the ER," *Biochim. Biophys. Acta - Mol. Cell Res.*, vol. 1833, no. 11, pp. 2430–2437, 2013.

[26] E. Aronica *et al.*, "Congenital disorder of glycosylation type Ia: A clinicopathological report of a newborn infant with cerebellar pathology," *Acta Neuropathol.*, vol. 109, no. 4, pp. 433–442, 2005.

[27] X. Zhang and Y. Wang., "Glycosylation quality control by the Golgi structure.," *J Mol Biol.*, vol. 428(16): 3, 2016.

[28] R. Kornfeld and S. Kornfeld, "Rosalind Kornfeld and Stuart Kornfeld," *Annu. Rev. Biochem.*, vol. 54, no. 1, pp. 631–664, 1985.

[29] W. L. Ho, W. M. Hsu, M. C. Huang, K. Kadomatsu, and A. Nakagawara, “Protein glycosylation in cancers and its potential therapeutic applications in neuroblastoma,” *J. Hematol. Oncol.*, vol. 9, no. 1, pp. 1–15, 2016.

[30] W. van Tol, H. Wessels, and D. J. Lefeber, “O-glycosylation disorders pave the road for understanding the complex human O-glycosylation machinery,” *Curr. Opin. Struct. Biol.*, vol. 56, pp. 107–118, 2019.

[31] S. J. Duguay , J. Lai-Zhang , D. Steiner, “Mutational analysis of the insulin-like growth factor I prohormone processing site.,” *J. Biol. Chem.*, vol. 270, no. 29, pp. 17566–17574, 1995.

[32] C. T. Roberts, Jr. S. R. Lasky, Jr. W. L. Lowe, W. T. Seaman, and D. LeRoith, “Molecular cloning of rat insulin-like growth factor I complementary deoxyribonucleic acids: differential messenger ribonucleic acid processing and regulation by growth hormone in extrahepatic tissues.,” *Mol Endocrinol*, vol. 1(3, pp. 243–8, 1987.

[33] J. E. Puche and I. Castilla-Cortázar, “Human conditions of insulin-like growth factor-I (IGF-I) deficiency,” *J. Transl. Med.*, vol. 10, no. 1, pp. 1–29, 2012.

[34] A. Giustina, G. Mazziotti, and E. Canalis, “Growth hormone, insulin-like growth factors, and the skeleton,” *Endocr. Rev.*, vol. 29, no. 5, pp. 535–559, 2008.

[35] A. M. Oberbauer, “The regulation of IGF-1 gene transcription and splicing during development and aging,” *Front. Endocrinol. (Lausanne).*, vol. 4, no. MAR, pp. 1–9, 2013.

[36] L. Temmerman, E. Slonimsky, N. Rosenthal, “Class 2 IGF-1 isoforms are dispensable for viability, growth and maintenance of IGF-1 serum levels.,” *Growth Horm IGF Res.*, vol. 20, no. 3, pp. 255–63, 2010.

[37] R. W. Matheny, B. C. Nindl, and M. L. Adamo, “Minireview: Mechano-Growth Factor: A Putative Product of IGF-I Gene Expression Involved in Tissue Repair and Regeneration.,” *Endocrinology*, vol. 151, no. 3, pp. 865–875., 2010.

[38] M. S. Hede *et al.*, “E-Peptides Control Bioavailability of IGF-1,” *PLoS One*, vol. 7, no. 12, 2012.

[39] Y. Kajimoto and P. Rotwein, “Structure of the chicken insulin-like growth factor I gene reveals conserved promoter elements,” *J. Biol. Chem.*, vol. 266, no. 15, pp. 9724–9731, 1991.

[40] A. M. Sparkman *et al.*, “Rates of molecular evolution vary in vertebrates for insulin-like growth factor-1 (IGF-1), a pleiotropic locus that regulates life history traits,” *Gen. Comp. Endocrinol.*, vol. 178, no. 1, pp. 164–173, 2012.

[41] D. M. Tiago, V. Laizé, and M. L. Cancela, “Alternatively spliced transcripts of Sparus aurata insulin-like growth factor 1 are differentially expressed in adult tissues and during early development,” *Gen. Comp. Endocrinol.*, vol. 157, no. 2, pp. 107–115, 2008.

[42] M. Reinecke and C. Collet, “The phylogeny of the insulin-like growth factors,” *Int. Rev Cytol.*, vol. 183, pp. 1-94., 1998.

[43] G. Annibalini, P. Bielli, M. de Santi, D. Agostini, M. Guescini, D. Sisti, S. Contarelli, G. Brandi, A. Villarini, V. Stocchi, C. Sette, E. Barbieri, “MIR retroposon exonization promotes evolutionary variability and generates species-specific expression of IGF-1 splice variants.,” *Biochim. Biophys. Acta - Gene Regul. Mech.*, vol. 1859, no. 5, pp. 757–768, 2016.

[44] M. Wallis, “New insulin-like growth factor (IGF)-precursor sequences from mammalian genomes: the molecular evolution of IGFs and associated peptides in primates,” *Growth Horm. IGF Res.*, vol. 19, no. 1, pp. 12–23, 2009.

[45] A. Armakolas *et al.*, “Oncogenic role of the Ec peptide of the IGF-1Ec isoform in prostate cancer,” *Mol. Med.*, vol. 21, pp. 167–179, 2015.

[46] Y. Xing and C. Lee, “Evidence of functional selection pressure for alternative splicing events that accelerate evolution of protein subsequences,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 38, pp. 13526–13531, 2005.

[47] Y. Xing and C. Lee, “Can RNA selection pressure distort the measurement of Ka / Ks?,” *Gene*, vol. 370, no. 1–2, pp. 1–5, 2006.

[48] P. Rotwein,“Editorial: The Fall of Mechanogrowth Factor?,” *Mol Endocrinol.*, vol. 28, no. 2., pp. 155–156., 2014.

[49] J. V.Chamary, J. L. Parmley, L. D. Hurst, “Hearing silence: non-neutral evolution at synonymous sites in mammals,” *Nat Rev Genet*, vol. 7, no. 2, pp. 98–108, 2006.

[50] M. F. Lin, P. Kheradpour, S. Washietl, B. J. Parker, J. S. Pedersen, and M. Kellis, “Locating protein-coding sequences under selection for additional, overlapping functions in 29 mammalian genomes,” *Genome Res.*, vol. 21, no. 11, pp. 1916–1928, 2011.

[51] M. Blanchette *et al.*, “Aligning multiple genomic sequences with the threaded blockset aligner,” *Genome Res.*, vol. 14, no. 4, pp. 708–715, 2004.

[52] R. C. Edgar, “MUSCLE: Multiple sequence alignment with high accuracy and high throughput,” *Nucleic Acids Res.*, vol. 32, no. 5, pp. 1792–1797, 2004.

[53] J. D. Thompson, T. J. Gibson, D. G. Higgins, “Multiple sequence alignment using ClustalW and ClustalX,” *Curr Protoc Bioinforma.*, 2002.

[54] J. Jurka, “Repbase update: a database and an electronic journal of repetitive elements,” *Trends Genet*, vol. 16, no. 9, pp. 418–20, 2000.

[55] T. J. Wheeler *et al.*, “Dfam: A database of repetitive DNA based on profile hidden Markov models,” *Nucleic Acids Res.*, vol. 41, no. D1, pp. 70–82, 2013.

[56] P. Bielli *et al.*, “The transcription factor FBI-1 inhibits SAM68-mediated BCL-X alternative splicing and apoptosis,” *EMBO Rep.*, vol. 15, no. 4, pp. 419–427, 2014.

[57] E. R. Barton, J. DeMeo, and H.Lei. “The insulin-like growth factor (IGF)-I E-peptides are required for isoform-specific gene expression and muscle hypertrophy after local IGF-I production,” *J Appl Physiol*, vol. 108, no. 5, pp. 1069–1076, 2010.

[58] Z. Dosztányi, V. Csizmok, P. Tompa, and I. Simon, “IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content,” *Bioinformatics*, vol. 21, no. 16, pp. 3433–3434, 2005.

[59] R. Sorek, “The birth of new exons: Mechanisms and evolutionary consequences,” *Rna*, vol. 13, no. 10, pp. 1603–1608, 2007.

[60] I. Vorechovsky, “Transposable elements in disease-associated cryptic exons,” *Hum. Genet.*, vol. 127, no. 2, pp. 135–154, 2010.

[61] M. J. Oberbauer AM, Belanger JM, Rincon G, Cánovas A, Islas-Trejo A, Gularte-Mérida R, Thomas MG, “Bovine and murine tissue expression of insulin like growth factor-I.,” *Gene*, vol. 535, no. 2, pp. 101–105, 2013.

[62] S. Naftelberg, I. E. Schor, G. Ast, and A. R. Kornblihtt, “Regulation of alternative splicing through coupling with transcription and chromatin structure,” *Annu. Rev. Biochem.*, vol. 84, pp. 165–198, 2015.

[63] N. N. Singh, M. N. Lawler, E. W. Ottesen, D. Upreti, J. R. Kaczynski, and R. N. Singh, “An intronic structure enabled by a long-distance interaction serves as a novel target for splicing correction in spinal muscular atrophy,” *Nucleic Acids Res.*, vol. 41, no. 17, pp. 8144–8165, 2013.

[64] E. Kovacs, P. Tompa, K. Liliom, and L. Kalmar, “Dual coding in alternative reading frames correlates with intrinsic protein disorder,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 12, pp. 5429–5434, 2010.

[65] M. Macossay-Castillo, S. Kosol, P. Tompa, and R. Pancsa, “Synonymous Constraint Elements Show a Tendency to Encode Intrinsically Disordered Protein Segments,” *PLoS Comput. Biol.*, vol. 10, no. 5, 2014.

[66] J. Brosius and S. J. Gould, “On ‘genomenclature’: A comprehensive (and respectful) taxonomy for pseudogenes and other ‘junk DNA,’” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 89, no. 22, pp. 10706–10710, 1992.

[67] A. Dricu, M. Carlberg, M. Wang, and O. Larsson, “Inhibition of N-linked glycosylation using tunicamycin causes cell death in malignant cells: Role of down-regulation of the insulin-like growth factor 1 receptor in induction of apoptosis,” *Cancer Res.*, vol. 57, no. 3, pp. 543–548, 1997.

[68] M. Carlberg *et al.*, “Mevalonic acid is limiting for N-linked glycosylation and translocation of the insulin-like growth factor-1 receptor to the cell surface. Evidence for a new link between 3-hydroxy-3-methylglutaryl-coenzyme A reductase and cell growth,” *J. Biol. Chem.*, vol. 271, no. 29, pp. 17453–17462, 1996.

[69] M. O. Savage, V. Hwa, A. David, R. G. Rosenfeld, and L. A. Metherell, “Genetic defects in the growth hormone-IGF-I axis causing growth hormone insensitivity and impaired linear growth,” *Front. Endocrinol. (Lausanne).*, vol. 2, no. DEC, pp. 1–12, 2011.

[70] L. Girnita, C. Worrall, S. I. Takahashi, S. Seregard, and A. Girnita, “Something old, something new and something borrowed: Emerging paradigm of insulin-like growth factor type 1 receptor (IGF-1R) signaling regulation,” *Cell. Mol. Life Sci.*, vol. 71, no. 13, pp. 2403–2427, 2014.

[71] E. R. Barton, “The ABCs of IGF-I isoforms: Impact on muscle hypertrophy and implications for repair,” *Appl. Physiol. Nutr. Metab.*, vol. 31, no. 6, pp. 791–797, 2006.

[72] Q. Gao, W. Sun, M. Ballegeer, C. Libert, and W. Chen, “ Predominant contribution of cis- regulatory divergence in the evolution of mouse alternative splicing ,” *Mol. Syst. Biol.*, vol. 11, no. 7, p. 816, 2015.

[73] G. G. Yang S, Alnaqeeb M, Simpson H, “Cloning and characterization of an IGF-1 isoform expressed in skeletal muscle subjected to stretch,” *J Muscle Res Cell Motil*, vol. 17, no. 4, pp. 487–95, 1996.

[74] D. L. Black, “Mechanisms of Alternative Pre-Messenger RNA Splicing,” *Annu. Rev. Biochem.*, vol. 72, pp. 291–336, 2003.

[75] B. S.Miller, M. J. Khosravi, M. C. Patterson, C A Conover, “IGF system in children with congenital disorders of glycosylation,” *Clin. Endocrinol. (Oxf).*, vol. 70, no. 6, pp. 892–897, 2009.

[76] P. F. Collett-Solberg and P. Cohen, “The role of the insulin-like growth factor

binding proteins and the IGFBP proteases in modulating IGF action," *Endocrinol. Metab. Clin. North Am.*, vol. 25, no. 3, pp. 591–614, 1996.

[77] S. M. Firth and R. C. Baxter, "Cellular actions of the insulin-like growth factor binding proteins," *Endocr. Rev.*, vol. 23, no. 6, pp. 824–854, 2002.

[78] J. B. M. Janosi, P. A. Ramsland, M. R. Mott, S. M. Firth, R. C. Baxter, and P. J. D. Delhanty, "The acid-labile subunit of the serum insulin-like growth factor-binding protein complexes. Structural determination by molecular modeling and electron microscopy," *J. Biol. Chem.*, vol. 274, no. 33, pp. 23328–23332, 1999.

[79] S. R.Holman, R. C. Baxter, "Insulin-like growth factor binding protein-3: factors affecting binary and ternary complex formation," *Growth Regul*, vol. 6, no. 1, pp. 42–7, 1996.

[80] Y. R. Boisclair, R. P. Rhoads, I. Ueki, J. Wang, and G. T. Ooi, "The acid-labile subunit (ALS) of the 150 kDa IGF-binding protein complex: An important but forgotten component of the circulating IGF system," *J. Endocrinol.*, vol. 170, no. 1, pp. 63–70, 2001.

AIMS OF THE THESIS

The primary aim of this Thesis was to investigate how N-glycosylation might influence the IGF-1 system.

The research objectives include:

a) The investigation of the cellular mechanisms controlling IGF-1 prohormones production and secretion.

b) The evaluation of N-glycosylation inhibition by tunicamycin (TUN) or by genetic knockdown of phosphomannomutase 2 (*PMM2*) on myoblast differentiation and IGF-1R signalling pathways activation.

c) The determination of the impact of N-glycosylation defects found in Congenital Disorders of Glycosylation (CDG) on the IGF-1R expression level and activation, and IGF-1Ea prohormone N-glycosylation and IGF-1 secretion.

CHAPTER 1

The intrinsically disordered E-domains regulate the IGF-1 prohormones stability, subcellular localisation and secretion

Giosuè Annibalini[1]*, Serena Contarelli[1], Mauro De Santi[1], Roberta Saltarelli[1], Laura Di Patria[1], Michele Guescini[1], Anna Villarini[2], Giorgio Brandi[1], Vilberto Stocchi[1] and Elena Barbieri[1,3]

[1]Department of Biomolecular Sciences, University of Urbino Carlo Bo, 61029 Urbino, Italy.
[2]Research Department, Fondazione IRCCS Istituto Nazionale dei Tumori, 20133 Milan, Italy.
[3]IIM, Interuniversity Institute of Myology, 61029 Urbino, Italy.

Published in Scientific Reports, 8(1):8919, 2018 (Article Open Access Published: 11 June 2018)

Abstract

Insulin-like growth factor-1 (IGF-1) is synthesised as a prohormone (proIGF-1) requiring enzymatic activity to yield the mature IGF-1. Three proIGF-1s are encoded by alternatively spliced IGF-1 mRNAs: proIGF-1Ea, proIGF-1Eb and proIGF-1Ec. These proIGF-1s have a common IGF-1 mature sequence but different E-domains. The structure of the E-domains has not been resolved, and their molecular functions are still unclear.

Here, we show that E-domains are Intrinsically Disordered Regions that have distinct regulatory functions on proIGF-1s production. In particular, we identified a highly conserved N-glycosylation site in the Ea-domain, which regulated intracellular proIGF-1Ea level preventing its proteasome-mediated degradation. The inhibition of N-glycosylation by tunicamycin or glucose starvation markedly reduced proIGF-1Ea

and mature IGF-1 production. Interestingly, 2-deoxyglucose, a glucose and mannose analogue, increased proIGF-1Ea and mature IGF-1 levels, probably leading to an accumulation of an under-glycosylated proIGF-1Ea that was still stable and efficiently secreted. The proIGF-1Eb and proIGF-1Ec were devoid of N-glycosylation sites, and hence their production was unaffected by N-glycosylation inhibitors. Moreover, we demonstrated that alternative Eb- and Ec-domains controlled the subcellular localisation of proIGF-1s, leading to the nuclear accumulation of both proIGF-1Eb and proIGF-1Ec.

Our results demonstrated that E-domains are regulatory elements that control IGF-1 production and secretion.

Introduction

Insulin-like growth factor-1 (IGF-1) is a growth factor with multiple roles in various aspects of normal and pathological growth and differentiation[1-2]. The translation of the IGF-1 gene gives rise to an immature IGF-1 peptide, which has a signal peptide at the 5' end of the gene, a core region and a C-terminal E-domain extension. The passage of the polypeptide into the endoplasmic reticulum (ER) removes the signal peptide, while the nascent IGF-1 prohormone (proIGF-1) is emerging, retaining the E-domain. Conversion of proIGF-1 to mature peptide requires the endoproteolytic cleavage of the E-domain by proprotein convertases, such as furin, which processes proproteins at highly conserved, unique pentabasic motif[3-4].

Due to alternative splicing of terminal exon 5 of the IGF-1 gene, three distinct proIGF-1s might exist: proIGF-1Ea, proIGF-1Eb and proIGF-1Ec[3-5]. These prohormones have the same IGF-1 mature sequence of 70 amino acids (aa) but different E-domains. In particular, the human Ea-domain is composed of 35 aa; the first 16 aa of Ea-domain are common in all E-domains, while 19 aa are unique to this isoform. The human Ea-domain contains a potential N-glycosylation site, N92, which follows the consensus sequence motif for N-glycosylation, NX(S/T) (where X can be any amino acid except proline). Accordingly, both unglycosylated proIGF-1Ea (11.7 kDa) and glycosylated proIGF-1Ea (~17-22 kDa) were found in normal and IGF-1-overexpressing cells[6-7]. The human Eb- and Ec-domains contain the 16 common aa and 61 and 24 additional isoform-specific aa respectively, with a predicted molecular weight of 16.5 kDa for proIGF-1Eb and 12.5 kDa for proIGF-1Ec. The human Eb- and Ec-domains lack potential N-linked glycosylation consensus sequences[3-4].

Previously termed "inactive precursors", proIGF-1s are currently recognised as stable intermediates of posttranslational processing. Accordingly, under physiological condition mammalian tissues mainly produced the glycosylated proIGF-1Ea[6]. Moreover, several studies demonstrated that proIGF-1s remained unprocessed in cultured cells, whether endogenously expressed IGF-1 isoforms (HepG2, K562 and HeLa cells[8-9]) or that were exogenously transfected with IGF-1 isoforms (HEK293 cells[7]). More interestingly, a differential expression pattern of the proIGF-1s was reported in normal *versus* cancer tissues[8,10-11].

Whether the alternative E-domains might regulate proIGF-1s is still an open question[4]. In mouse skeletal muscle, viral delivery of IGF-1Ea and IGF-1Ec, but not mature IGF-1, increases muscle mass. Hence, the E-domains are necessary to promote the local production of IGF-1 and functional hypertrophy[12].

We and others have recently shown that the distribution of intrinsic disorder propensity within the amino acid sequence of mature IGF-1 is markedly different compared to E-domains[5,13]. In particular, bioinformatic analysis of proIGF-1 structures showed that the E-domains were putative intrinsically disordered regions (IDRs). IDRs are regions within proteins that exhibit high flexibility and may lack a secondary or tertiary structure[14]. It is worth mentioning that also the other two members of the IGF family, proinsulin and proIGF-2, possess IDRs, although the degree of disorder across the IGF family varies significantly[13]. Despite the fact that amino acid sequences of E-domains are less conserved than those of mature IGF-1, we demonstrated that the disordered propensity of E-domains has been strongly conserved[5]. In fact, IDRs can tolerate a higher number of mutations without substantial loss of flexibility and function[15-16].

IDRs may facilitate the regulation of protein function through various mechanisms. For example, owing to their conformational flexibility, IDRs have a high propensity to undergo posttranslational modifications, such as acetylation, glycosylation, methylation, or phosphorylation[15,17]. IDRs might also control protein half-life by efficiently engaging proteins to the proteasome[18-19]. Moreover, studies have identified IDRs as enriched in the alternatively spliced protein segments, indicating that protein isoforms may display functional diversity due to the alteration of tissue-specific and species-specific modules within these regions[20].

In this study, we analysed the structural proprieties of proIGF-1s, using a combination of bioinformatics analyses and limited proteolysis. Site-direct mutagenesis and inhibition of N-glycosylation were used to evaluate the role of glycosylation on proIGF-1s regulation in terms of stability and secretion. Finally, we investigated the role of alternative E-domains on proIGF-1s subcellular localisation. Our results show that the alternative disordered E-domains affect distinct aspects of proIGF-1s regulation including protein stability, localisation and secretion. Thus, E-domains may represent novel targets to control proIGF-1s and, by extension, mature IGF-1 production.

Results

Disorder propensity of Human proIGF-1s

We used the D^2P^2 platform (http://d2p2.pro/) and limited proteolysis to predict intrinsically disordered regions of human proIGF-1Ea (ENSP00000416811), proIGF-1Eb (ENSP00000302665) and proIGF-1Ec (ENSP00000376638)[21-22] (Fig. 1). Figure 1A shows the plot generated by the D^2P^2 platform: this analysis showed that the mature IGF-1 is mostly ordered, while all E-domains were predicted to contain disordered residues. Moreover, the Eb-domain also contains two predicted molecular recognition features (MoRFs) and two phosphorylation sites.

Subsequently, we used limited proteolysis to identify the regions of the polypeptide chain mostly prone to proteolysis and thus the sites of high flexibility or local unfolding[22] (Fig. 1B). The supernatant of HEK293 cells enriched with glycosylated proIGF-1Ea, proIGF-1Eb, proIGF-1Ec and mature IGF-1 were digested with trypsin, loaded on SDS-PAGE gels and probed with the anti-mature IGF-1 antibody. As shown in Figure 1B, all proIGF-1s were sensitive to trypsin digestion while mature IGF-1 was significantly more resistant. Digestion of mature IGF-1 required long term incubation with trypsin (i.e. > 45 minutes) (Supplementary Fig. S1B). Similar results were obtained using proteinase K digestion of proIGF-1Ea (Supplementary Fig. S1C). These data show that proIGF-1s are composed of both protein structural domain, i.e. the mature IGF-1, and intrinsically disordered regions, i.e. the C-terminal E-domains.

Intracellular IGF-1 is mainly expressed as proIGF-1Ea, not mature IGF-1

Using RT-PCR analyses, we previously demonstrated that skeletal muscles, adipose tissues and liver of several mammalian species mainly expressed the IGF-1Ea isoform, which represents about 90% of IGF-1 transcripts[5]. The first goal of the present study was to examine the protein forms endogenously produced in these tissues. Immunoblotting of protein lysates using the anti-mature IGF-1 antibody showed a distinct ~17 kDa band, most likely representing glycosylated proIGF-1Ea, in all samples analysed (Fig. 2A). Notably, the band corresponding to mature IGF-1 (~7 kDa) was not found in naïve tissues, in agreement with Durzynska, J., et al.[6]

To further confirm the presence of proIGF-1s, we used specific antibody directed against the common E-domain region of proIGF-1s (RSVRAQRHTD). The antibody specificity towards E-domain region was checked using HEK293 cells overexpressing IGF-1 isoforms (Supplementary Fig. S2). As shown in Figure 2B, two bands of ~12 kDa and ~17 kDa were detected with the anti E-domain antibody, corresponding respectively to the molecular weight of unglycosylated and glycosylated proIGF-1Ea. As expected, no band corresponding to the molecular size of Ea-domain (~4 kDa) was detected in the lysate of HEK293 cells overexpressing IGF-1Ea or tissues confirming that the E-domain was not cleaved intracellularly from the IGF-1 mature protein. Subsequently, we moved to a cell-based system to improve IGF-1 detection and to control the IGF-1 isoforms produced. We recently demonstrated that in HEK293 cells over-expressing IGF-1 isoforms the proIGF-1s are the main forms produced intracellularly, and for proIGF-1Ea both unglycosylated (~12 kDa) and glycosylated (~17 kDa) forms were detected[7]. Hence the tissue expression pattern of IGF-1 was recapitulated in our HEK293 cell-based transient gene expression system.

We next examined the effects of E-domains on proIGF-1s stability and secretion, starting with the predominant isoform produced in normal tissues, i.e. the proIGF-1Ea.

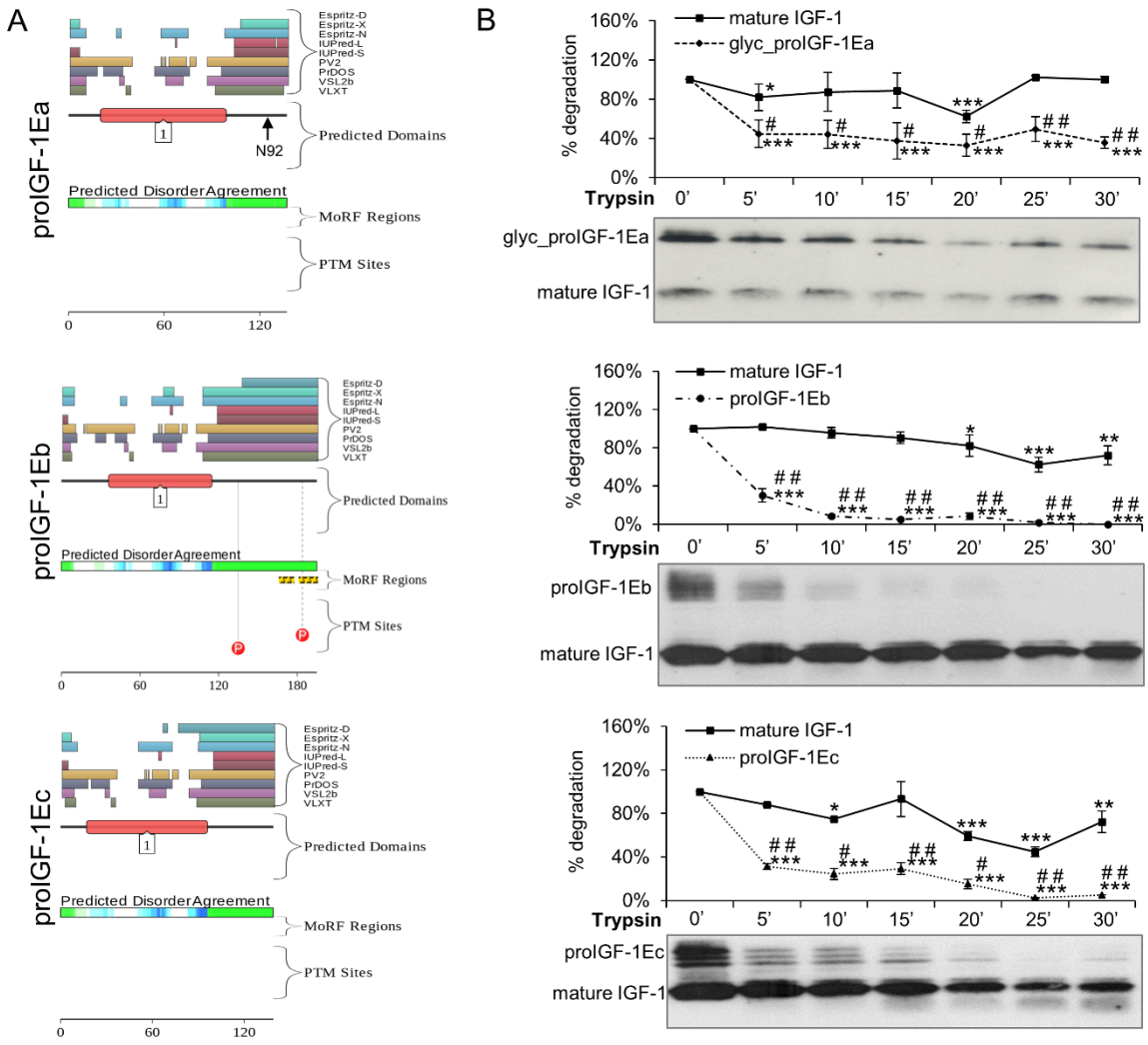


Figure 1. Evaluation of the intrinsic disorder propensity of human proIGF-1s predicted using the D^2P^2 platform (A) and limited proteolysis (B). (A) Analysis of the intrinsically disordered regions of human proIGF-1Ea (ENSP00000416811), proIGF-1Eb (ENSP00000302665) and proIGF-1Ec (ENSP00000376638) sequences by the D^2P^2 platform (http://d2p2.pro/). The red box corresponds to the insulin-like domain; the green-and-white bar in the middle of the plot shows the predicted disorder agreement among nine predictors, with the green parts corresponding to the portions of sequence where at least 75% of the predictors agreed. Yellow bars show the location of the predicted disorder-based binding sites (molecular recognition features, MoRFs), whereas red circles at the bottom of the plot show the location of putative phosphorylation sites. Position of the N-glycosylation site of Ea-domain (N92) is also indicated. (B) Limited proteolysis of proIGF-1s. Cell culture supernatants of IGF-1Ea-, IGF-1Eb- or IGF-1Ec-transfected HEK293 cells were concentrated using Amicon Ultra 3K centrifugal filters and incubated with 0.2 µM trypsin at 37°C for different times. Reactions were removed over a time-course and the digested products were loaded on 12% SDS-PAGE and analysed by western blotting with an anti-mature IGF-1 antibody. Results are means ± SEM (n = 3). Repeated measures ANOVA, # (p<0.01) and ## (p<0.0001) significantly different compared to mature IGF-1; *(p<0.05), **(p<0.001) and ***(p<0.0001) significantly different compared to the 0-minute time point. Cropped blots are shown. Uncropped blots are presented in Supplementary Fig. S1A.

Glycosylation is necessary to stabilise proIGF-1Ea and regulate mature IGF-1 secretion

Multiple sequence alignment of vertebrate Ea-domain showed that the N-glycosylation site of proIGF-1Ea has been conserved from teleosts to mammals (Fig. 3A). This strong evolutionary conservation led us to hypothesise that glycosylation could play an important role in the regulation of proIGF-1Ea.

Site-direct mutation of this glycosylation site (IGF-1Ea^{N92D}) resulted in a ~12 kDa band, corresponding to the size of unglycosylated proIGF-1Ea (Fig. 3B). Western blotting analysis showed that the intracellular level of unglycosylated proIGF-1Ea was significantly lower in IGF-1Ea^{N92D}-transfected HEK293 cells compared to wild-type IGF-1Ea (IGF-1Ea^{WT}) (Fig. 3B). Moreover, contrary to IGF-1Ea^{WT}, the IGF-1Ea^{N92D}-transfected HEK293 cells did not secrete IGF-1 (both glycosylated proIGF-1Ea and mature IGF-1) (Fig. 3C). Notably, similar GFP fluorescence intensity (Fig. 3D) and total IGF-1 mRNA (Fig. 3E; p= 0.314) were found in IGF-1Ea^{WT}- and IGF-1Ea^{N92D}- transfected HEK293 cells, ruling out the possibility that the marked reduction in the protein levels between the two constructs were due to different transfection efficiency. Thus, these data suggest that glycosylation of proIGF-1Ea is required for efficient IGF-1 production and secretion. Subsequently, we wondered whether direct inhibition of N-glycosylation by tunicamycin (Tun) might interfere with IGF-1 production. Notably, the band corresponding to glycosylated proIGF-1Ea disappeared after treatment with Tun (Fig. 4A). Moreover, the analysis of cell culture supernatants of the IGF-1Ea-transfected HEK293 cells showed that Tun treatment completely abrogated the glycosylated proIGF-1Ea secretion and markedly reduced the secretion of mature IGF-1 (Fig. 4B). The marked reduction of glycosylated proIGF-1Ea after Tun treatment was not due to general suppression of transcription, as shown by total IGF-1 mRNA quantification (Supplementary Fig. 4C), or general protein synthesis inhibition, as shown by co-transfection of GFP (Supplementary Fig. 4D). Notably, conditioned media from IGF-1Ea-transfected HEK293 cells treated with Tun was unable to activate the IGF-1 receptor (IGF-1R) and downstream phosphorylation of ERK1/2 and AKT of MCF-7 breast cancer cells (Fig. 4C).

Similar results were obtained by blocking N-glycosylation by glucose withdrawal although the effect was less specific since glucose starvation also slightly decreased cell number and GFP-dependent fluorescence (Supplementary Fig. 5).

Collectively, these results documented that the interference with Ea-domain glycosylation resulted in a dramatic decrease of intracellular proIGF-1Ea level and hence proIGF-1Ea and mature IGF-1 secretion.

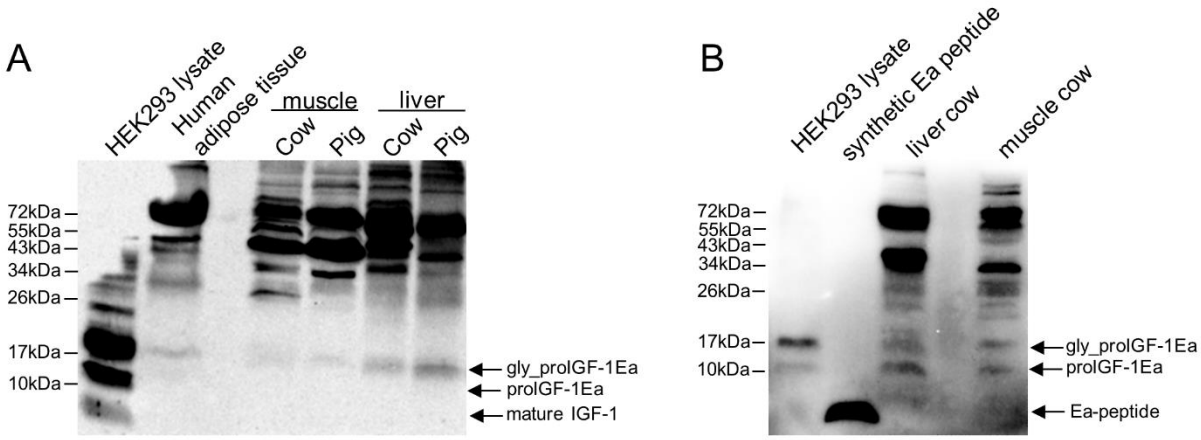


Figure 2. Immunoblotting of several mammalian tissues using an antibody directed against mature IGF-1 sequence (A) or common region of E-peptides (B). (A) Protein lysates (80 µg) were subjected to 12% SDS-PAGE and immunoblotted with anti-mature IGF-1 antibody. A band at a molecular weight around 17 kDa, most likely representing glycosylated proIGF-1Ea, was detected in all the tissue samples tested. A cell lysate of HEK293 overexpressing IGF-1Ea was used has a positive control. The band corresponding to mature IGF-1 (~7 kDa) was not found in tissues and was detectable in HEK293 overexpressing IGF-1Ea only after long exposure of the blots. (B) Immunoblotting of liver and muscle cow lysate using an antibody directed against E-domains. Two bands at a molecular weight around 12 kDa and 17 kDa were detected in all the tissue samples tested, most likely representing the unglycosylated and glycosylated proIGF-1Ea respectively. No band at a molecular weight of Ea peptide (~4 kDa) was detected. A cell lysate of HEK293 overexpressing IGF-1Ea and synthetic human Ea peptide were used as positive controls.

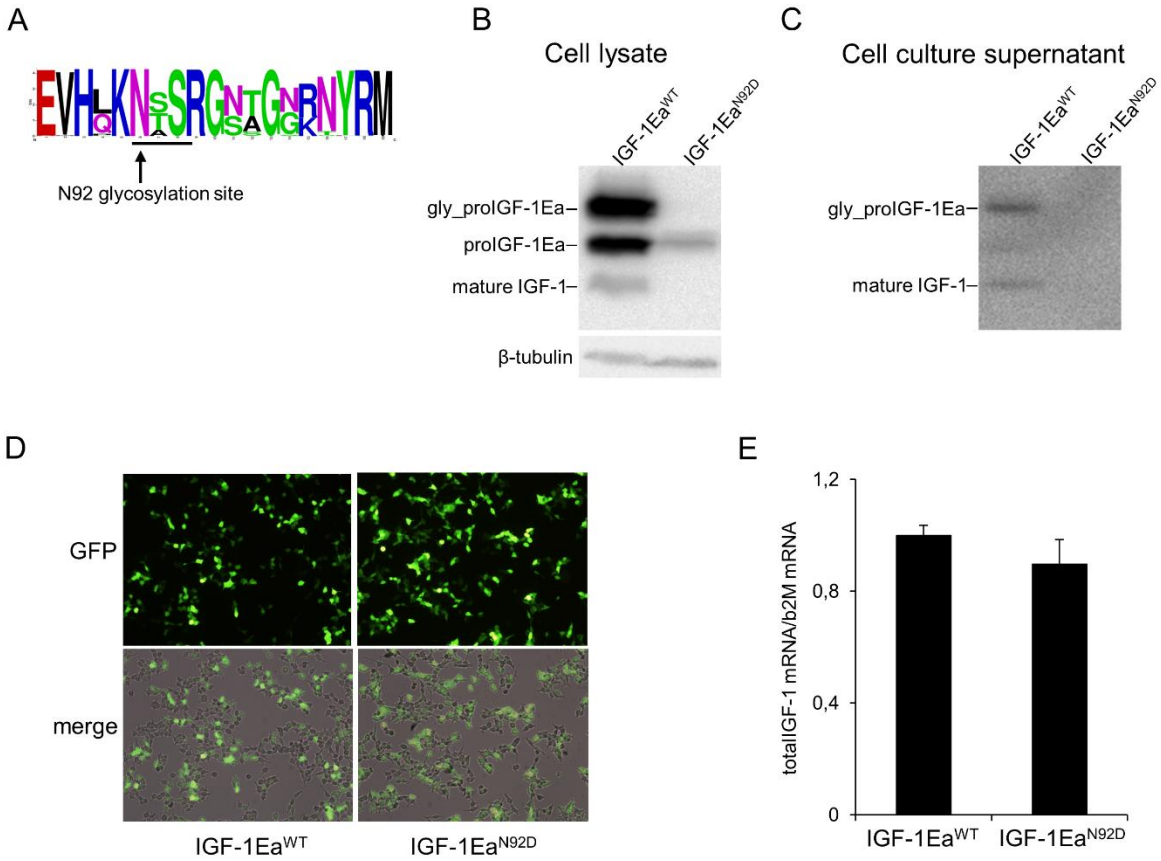


Figure 3. Conservation of the N-glycosylation site of Ea-domain (N92) (A) and effect of site-direct mutation of this N-glycosylation site (IGF-1Ea^{N92D} mutant) on intracellular (B) or extracellular (C) proIGF-1Ea production. Comparison of transfection efficiency between wild-type (IGF-1Ea^{WT}) and IGF-1Ea^{N92D} constructs (D and E). (A) WebLogo of Ea-domain sequences obtained from UniProt database. The relative frequency plots of amino acids of 250 E-domain sequences obtained from UniProt database is shown. The consensus sequence motif for N-glycosylation NX(S/T) (where X can be any amino acid except proline) is underline. The WebLogo was produced using the web server at http://weblogo.berkeley.edu/logo.cgi. (B and C) Transfection of HEK293 cells with IGF-1Ea^{WT} or IGF-1Ea^{N92D}. IGF-1Ea^{WT} and IGF-1Ea^{N92D} were transiently expressed in HEK293 cells. After 24 h the cell lysates (B) and cell culture supernatants (C) were analysed by western blot using an antibody directed against mature IGF-1 sequence. There was no significant difference in GFP fluorescence intensity (10x magnification) (D) and total IGF-1 mRNA quantity (p= 0.314) (E) between the IGF-1Ea^{WT} and IGF-1Ea^{N92D} constructs. Cropped blots are shown. Uncropped blots are presented in Supplementary Fig. S3.

The turnover of unglycosylated proIGF-1Ea is faster than glycosylated proIGF-1Ea and depends on proteasome activity

The inhibition of proIGF-1Ea glycosylation by mutation of the glycosylation site (proIGF-1Ea^{N92D}) (Fig. 3B), Tun treatment (Fig. 4A) or glucose starvation (Supplementary Fig. 5A), did not determine a concomitant increase of unglycosylated proIGF-1Ea. One possibility is that the unglycosylated proIGF-1Ea is rapidly degraded. Thus, we next sought to examine the role of glycosylation in the stability of proIGF-1Ea. In presence of cycloheximide (CHX), a protein synthesis inhibitor, the turnover rate of unglycosylated proIGF-1Ea was faster than glycosylated one (Fig. 5A). To verify the involvement of 26S proteasome machinery on unglycosylated proIGF-1Ea degradation, we subsequently treated IGF-1Ea-transfected HEK293 cells with proteasome inhibitor MG132. As shown in Figure 5B, we found an increase of unglycosylated proIGF-1Ea, while glycosylated proIGF-1Ea was only marginally affected by the proteasome inhibitor. These results demonstrated that unglycosylated proIGF-1Ea was unstable and degraded faster than glycosylated proIGF-1Ea.

Besides the increase of unglycosylated proIGF-1Ea, treatment with MG132 also promoted an accumulation of a band of ~23 kDa (indicated with an asterisk in Fig. 5B). This band was approximately twice the molecular weight of the unglycosylated proIGF-1Ea monomer (11.7 kDa) and was detected with both the anti-mature IGF-1 antibody (Fig. 5B) and the anti-E-domain antibody (Supplementary Fig. S6E). Thus, we hypothesised that inhibition of the proteasome leads to unglycosylated proIGF-1Ea accumulation and dimerisation. In support of this hypothesis, we found that proteasome inhibition by MG132 increased the cytoplasmic level of the ~23 kDa protein (Fig. 5C). Moreover, PNGase deglycosylation of the supernatant of HEK293 cells enriched with glycosylated proIGF-1Ea increased both the unglycosylated proIGF-1Ea and the putative ~23 kDa dimer in the cell culture supernatants (Fig. 5D). Taken together, these data showed that addition of N-glycan on N92 site of Ea-domain prevented the degradation of proIGF-1Ea by the proteasome, probably overcoming the folding limitation of Ea-domain and the tendency of unglycosylated proIGF-1Ea to self-aggregate.

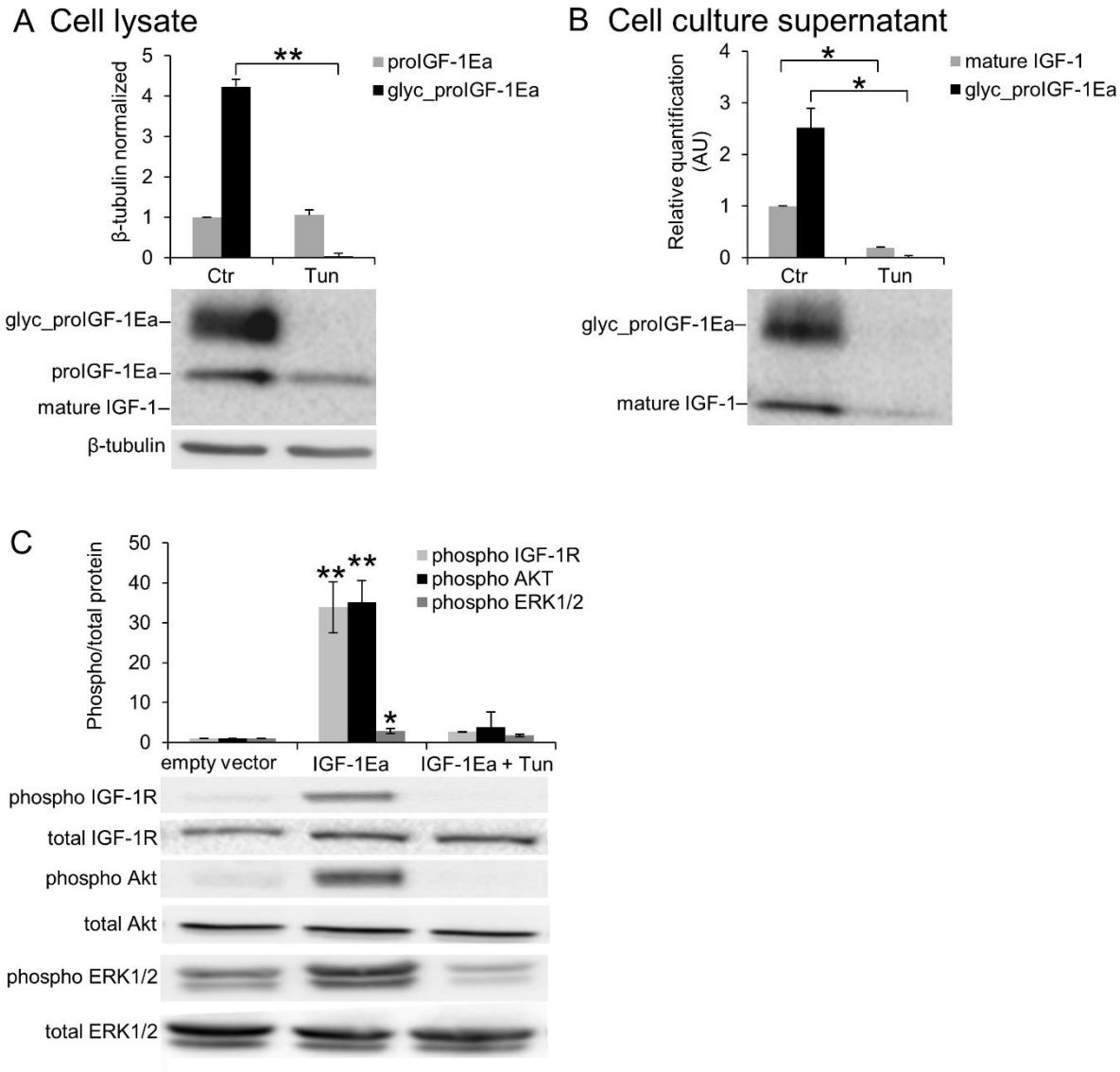


Figure 4. Effect of tunicamycin (Tun) treatment on proIGF-1Ea glycosylation. (A and B) IGF-1Ea was transiently expressed in HEK293 cells in the presence of 0.1 µg/ml of Tun. After 24 h the cell lysates (A) and cell culture supernatants (B) were analysed by western blot and relative expression level of glycosylatyed proIGF-1Ea, unglycosylated proIGF-1Ea and mature IGF-1 was calculated. The band at a molecular weight of ~17 kDa, corresponding to glycosylated proIGF-1Ea, disappeared in presence of Tun in cell lysates (A) and in cell culture supernatants (B). The band corresponding to mature IGF-1 (~7 kDa) was markedly reduced in cell culture supernatants after Tun treatment (B). (C) Phosphorylation of IGF-1R, AKT and ERK1/2 after treatment of MCF-7 cells with cell culture supernatants from IGF-1Ea-transfected HEK293 cells treated with Tun. The phosphorylation of the IGF-1R pathway was markedly reduced by Tun treatment. β-tubulin was used as a loading control for the cell lysates. Results are means ± SEM (n = 3); T-test or a one-way ANOVA was used to evaluate statistical significance (*p < 0.01, **p<0.0001). Cropped blots are shown. Uncropped blots are presented in Supplementary Fig. S4.

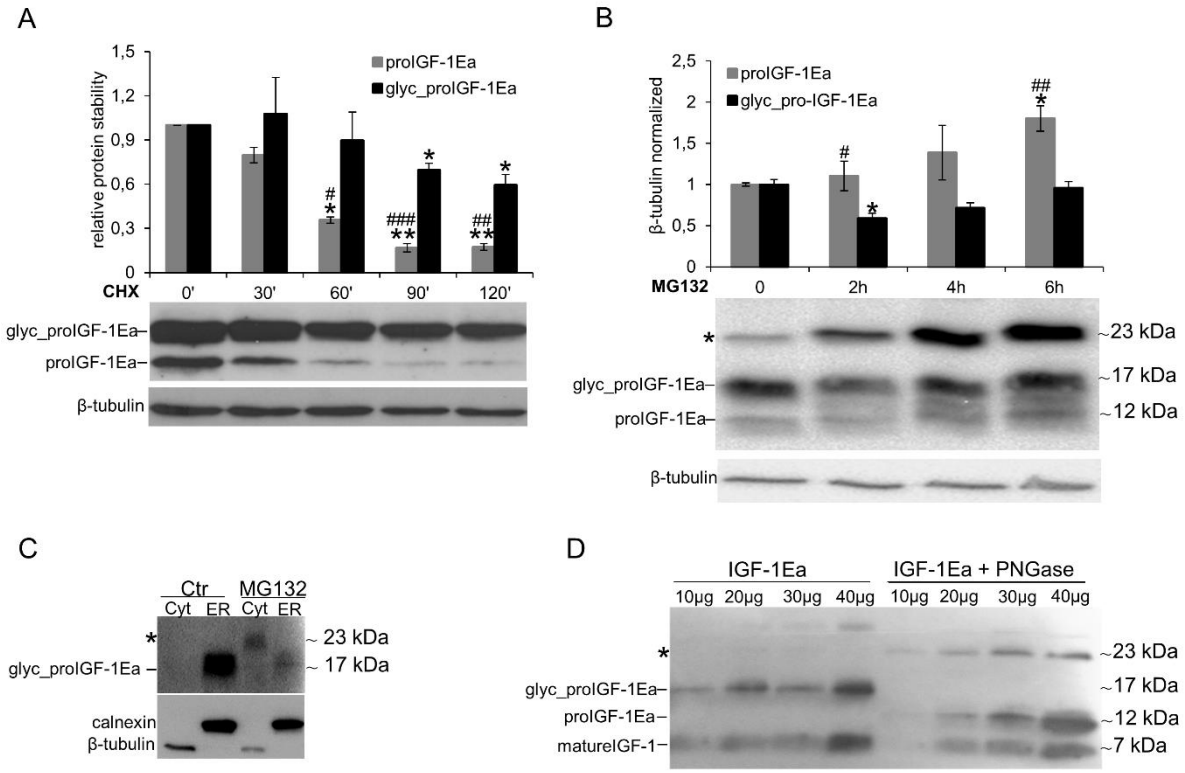


Figure 5. Analysis of unglycosylated and glycosylated proIGF-1Ea turnover. (A and B) IGF-1Ea was transiently expressed in HEK293 cells in the presence of 25µg/ml of protein synthesis inhibitor cycloheximide (CHX) (A) or 10µM of the proteasome inhibitor MG132 (B) in a time-course experiment. Cells were collected at different time points, and relative expression level of glycosylated and unglycosylated proIGF-1Ea was calculated. After MG132 treatment intracellular accumulation of a ~23 kDa band was found (indicated with an asterisk in Fig. 5B), probably representing unglycosylated proIGF-1Ea dimer. (C) Cytosol (Cyt) and endoplasmic reticulum (ER) isolations of IGF-1Ea-transfected HEK293 cells treated with MG132. The band at a molecular weight around 23 kDa was detected only after MG132 treatment and only in the cytosolic fraction. (D) Deglycosylation of proIGF-1Ea enriched media using the *N*-Glycosidase F (PNGase F). The PNGase treatment determined an accumulation of the band at a molecular weight around 23 kDa probably corresponding to unglycosylated pro-IGF-1Ea dimer. Results are means ± SEM (n = 3). Repeated measures ANOVA, $^{\#}(p<0.05)$, $^{\#\#}(p<0.01)$ and $^{\#\#\#}(p<0.0001)$ significantly different compared to glycosylated proIGF-1Ea; *(p<0.01) and **(p<0.0001) significantly different compared to the 0-minute time point. β-tubulin was used as a loading control for the cell lysates and the cytosol separation; calnexin was used as a control for the ER separation. Cropped blots are shown. Uncropped blots are presented in Supplementary Fig. S6.

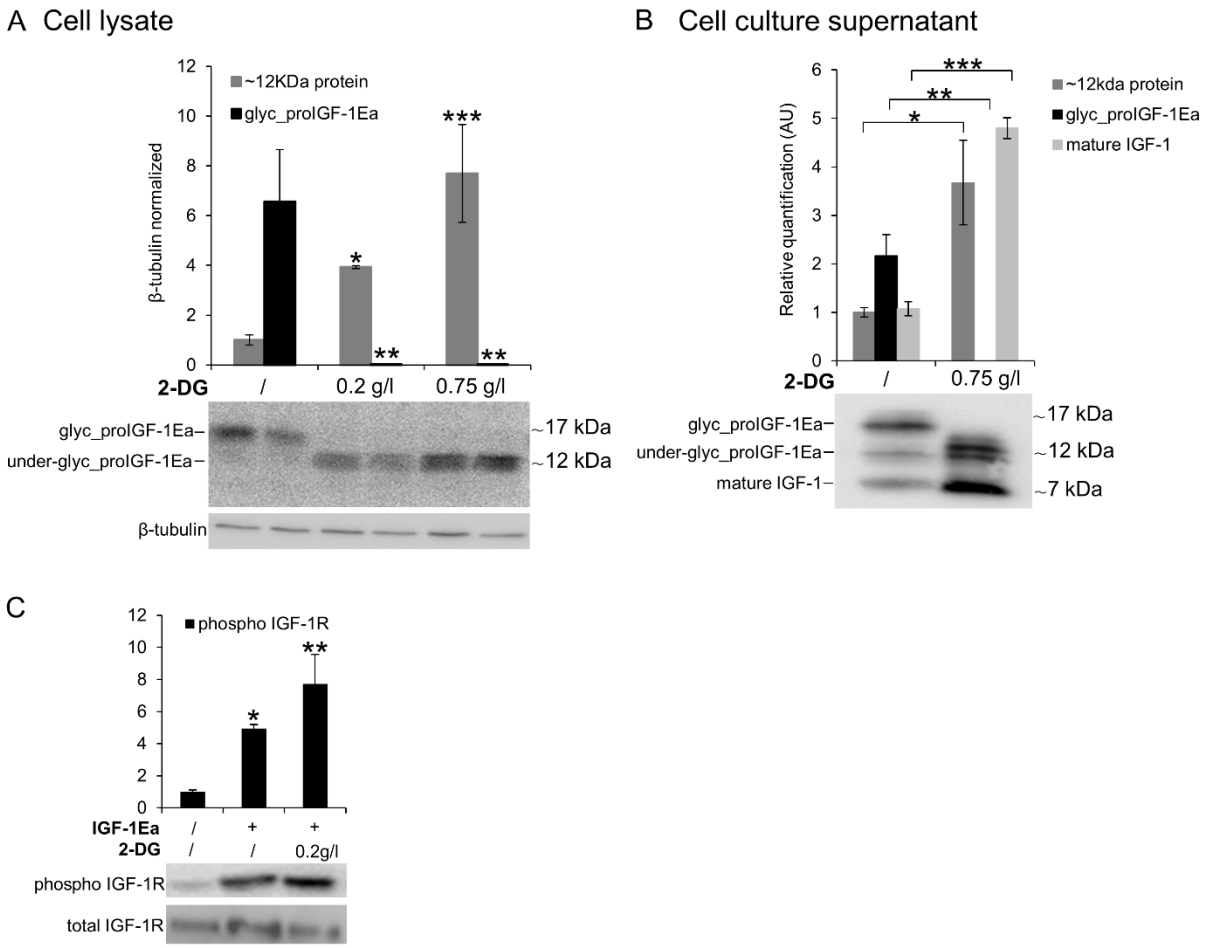


Figure 6. Effect of 2-Deoxyglucose (2-DG) on glycosylated proIGF-1Ea production. (A and B) IGF-1Ea was transiently expressed in HEK293 cells growth in low glucose medium (0.65 g/L) with or without 2-DG. After 24 h the cell lysates (A) and cell culture supernatants (B) were analysed by western blot and relative expression level of glycosylated proIGF-1Ea, unglycosylated proIGF-1Ea and mature IGF-1 was calculated. The band at a molecular weight around 17 kDa, corresponding to glycosylated proIGF-1Ea, disappeared in the presence of 2-DG in the cell lysates (A) and in the cell culture supernatants (B). 2-DG treatment determined also an intracellular (A) and supernatants (B) accumulation of a ~12 kDa band, probably representing an under-glycosylated proIGF-1Ea form. (C) Phosphorylation of IGF-1R after treatment of MCF-7 cells with conditioned media from IGF-1Ea-transfected HEK293 cells treated with 2-DG. Phosphorylation of IGF-1R increased both in 2-DG-treated and untreated cells. β-tubulin was used as a loading control for the cell lysates. Results are means ± SEM (n = 3); T-test or a one-way ANOVA was used to evaluate statistical significance (*p<0.05, **p< 0.01 and ***p<0.0001). Cropped blots are shown. Uncropped blots are presented in Supplementary Fig. S7.

The glucose analogue, 2-deoxyglucose, rescues the secretion defect of IGF-1 under low glucose conditions

Due to its inhibitory activity on glycolysis, 2-deoxyglucose (2-DG) has been shown to selectively kill cancer cells[23-24]. Because of its similarity to mannose, 2-DG is also known to interfere with N-linked glycosylation[24-26]. Notably, increasing evidence links 2-DG effects to the interference with glycosylation instead of glycolysis inhibition[27-29]. Prompted by this, we investigated whether 2-DG interfered with proIGF-1Ea glycosylation. IGF-1Ea-transfected HEK293 cells were grown in low glucose medium (0.65 g/L) in presence of increasing concentration of 2-DG (Fig. 6). As shown in Figure 6A, the band of glycosylated proIGF-1Ea (~17 kDa) disappeared after the addition of 0.2 g/L or 0.75 g/L of 2-DG. Notably, we also found a dose-dependent accumulation of a ~12 kDa band, both intracellularly (Fig. 6A) and extracellularly (Fig. 6B). Both the anti-mature IGF-1 antibody (Fig. 6A) and the anti-E-domain antibody (Supplementary Fig. 7D) recognised this ~12 kDa band. Thus, it is likely that the 12kD protein was not a non-specific product that appears after 2-DG treatment. Moreover, cell culture supernatant from IGF-1Ea-transfected HEK293 cells treated with 2-DG fully activated the IGF-1R of MCF-7 breast cancer cells (Fig. 6C). Thus, these data indicate that 2-DG interfered with the proIGF-1Ea glycosylation leading to an accumulation of a ~12 kDa protein, probably representing an aberrant under-glycosylated form of proIGF-1Ea, which was stable and efficiently secreted. These results are in line with a recent study showing that the effect of 2-DG on protein glycosylation was dose-dependent and, at doses similar to those used in the present study, 2-DG increased mannose incorporation into cellular glycoproteins instead of inhibits glycosylation[26]. Further studies are needed to identify the glycan structure of proIGF-1Ea that appears after 2-DG treatment and to understand how 2-DG rescues the secretion defect of proIGF-1Ea under low glucose conditions.

The presence of the alternative Eb- and Ec-domains hampers the proIGF-1Eb and proIGF-1Ec sensitivity to N-glycosylation inhibitors and determines the nuclear accumulation of prohormones

The proIGF-1Eb and proIGF-1Ec ran at the expected molecular weight of ~16.5 kDa and ~12.5 kDa in SDS-PAGE gels respectively, suggesting that these prohormones did not undergo posttranslational modification (Supplementary Fig. S2)[7]. Accordingly, unlikely proIGF-1Ea, proIGF-1Eb and proIGF-1Ec did not contain any potential glycosylation sites and migrated at the same molecular weight after treatment of HEK293 cells with Tun (Fig. 7A) or no glucose media (Supplementary Fig. S8C). Noteworthy, unlikely proIGF-1Ea, also the intracellular levels of proIGF-1Eb and proIGF-1Ec were unaffected by Tun treatment (Fig. 7A). Hence, the presence of the Eb- or Ec-domains, instead of Ea-domain, completely abrogated the response of proIGF-1s to glycosylation inhibitors.

Notably, previous studies demonstrated that both proIGF-1Eb and proIGF-1Ec might localise in the nucleus[30-32]. This led us to hypothesise that disordered Eb- and Ec-domains might control subcellular localisation of proIGF-1Eb and proIGF-1Ec. To test this hypothesis, we investigated the subcellular localisation of proIGF-1s by cytosol, ER and nucleus isolations (Fig. 7B) and immunofluorescence (Fig. 7C). In line with previously published studies[30-32], we found that proIGF-1Eb and proIGF-1Ec partially accumulated in the nucleus of HEK293-transfected cells, while proIGF-1Ea was mainly localised into the ER (Figs. 7B and 7C). Notably, proIGF-1Eb and proIGF-1Ec did not appear to be accumulated in the cytosol, and the molecular weight of proIGF-1Eb or proIGF-1Ec in the ER and nuclear fractions corresponded to that of proIGF-1s without signal peptide (5.3 kDa) (Fig. 7B). These data suggest that proIGF-1Eb and proIGF-1Ec translocate to the nucleus from the ER and not from the cytoplasm. Accordingly, treatment with the proteasome inhibitor MG132 increased the nuclear accumulation of proIGF-1Eb and proIGF-1Ec, while their cytoplasmic level remained undetectable (Fig. 7D). Several substrates of the nuclear ubiquitin-proteasome system have been identified to date, showing that the nuclear ubiquitin-proteasome system is a key quality-control mechanism that rapidly eliminates unfolded or damaged proteins[33]. In this regard, the nuclear degradation of proIGF-1Eb and proIGF-1Ec might represent a mechanism for controlling the

intracellular level of these highly disordered proteins; however, further studies are needed to clarify this point.

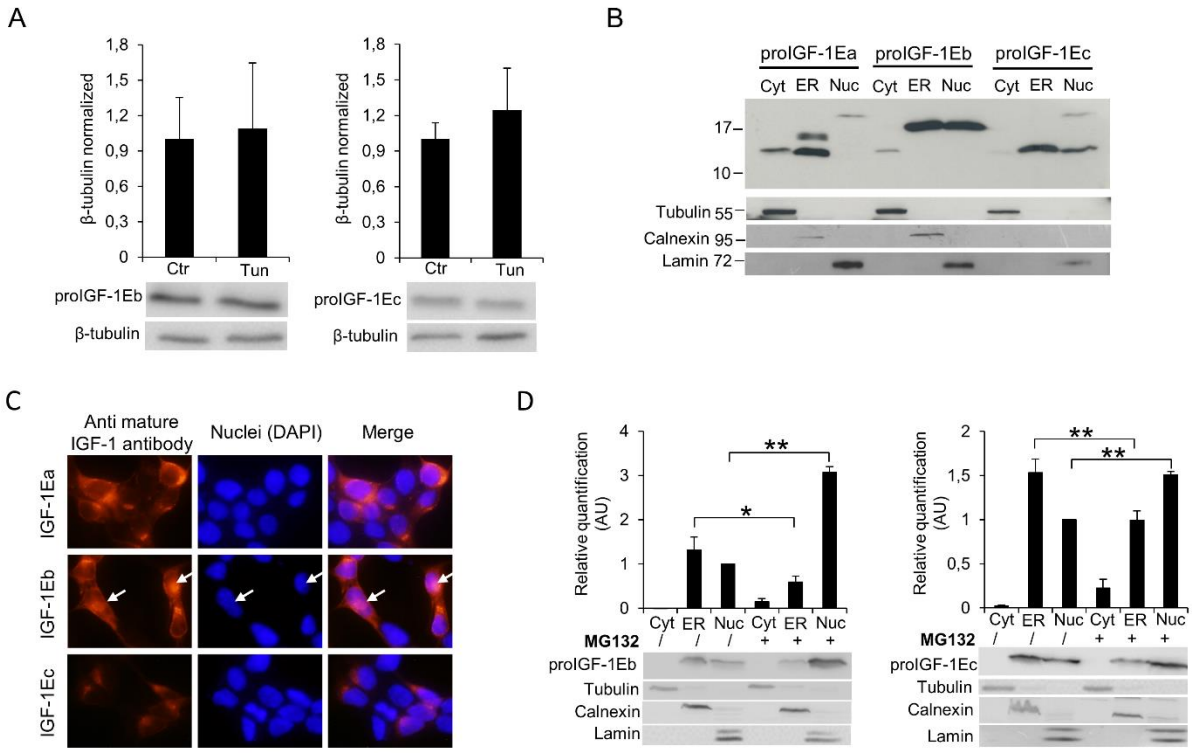


Figure 7. Effect of inhibition of N-glycosylation using Tun on IGF-1Eb (A, left panel) and IGF-1Ec (A, right panel) production and subcellular localisation of proIGF-1s in untreated (B and C) or MG132 treated HEK293 cells (D). (A) IGF-1Eb or IGF-1Ec were transiently expressed in HEK293 cells in the presence of 0.1 µg /ml of Tun. After 24 h the cell lysates were analysed by western blot and relative expression level of proIGF-1Eb (left panel) and proIGF-1Ec (right panel) were calculated. The molecular weight and the expression level of proIGF-1Eb (p=0.8) and proIGF-1Ec (p=0.6) were unaffected by Tun treatment. Results are means ± SEM (n = 3); T-test was used to evaluate statistical significance. (B) Subcellular localization of IGF-1 isoforms analysed by cytosol, ER and nucleus isolations or (C) immunofluorescence staining of IGF-1Ea-, IGF-1Eb- or IGF-1Ec-transfected HEK293 cells. ProIGF-1Ea was mainly localised in the cytosol fraction (unglycosylated proIGF-1Ea) and the ER fraction (both unglycosylated and glycosylated proIGF-1Ea) while proIGF-1Eb and proIGF-1Ec were mainly localised in the nuclear fraction and ER. (D) Subcellular localisation of proIGF-E1b (left panel) or proIGF-1Ec (right panel) after treatment with 10µM of the proteasome inhibitor MG132 for 6h. Results are means ± SEM (n = 3); a two-way ANOVA was used to evaluate statistical significance (*p<0.01 and **p< 0.001). β-tubulin was used as a loading control for the cell lysates and the cytosol separation; calnexin as a control for the ER separation and lamin as a control for the nucleus separation. DAPI nuclear stain (blue). Cropped blots are shown. Uncropped blots are presented in Supplementary Fig. S8.

Discussion

The mechanism by which IGF-1 production is regulated remains largely unknown[4]. In this study, we show the functional role of disordered E-peptides in the regulation of proIGF-1s production.

It is now well established that most peptide hormones and growth factors are initially synthesised as prohormones that are converted to active forms by endoproteolysis at specific sites[34]. Accordingly, the growth factor IGF-1 is produced as prohormone which contains a C-terminal domain, i.e. the E-domain, cleaved by furin convertase[3]. Evidence has been provided that only a small portion of proIGF-1 cleavage occurs intracellularly; hence most proIGF-1 might be converted to mature form at the cell surface membrane or extracellularly[6-7]. The results of our study confirm these findings showing that intracellular IGF-1 is mainly expressed as proIGF-1, not mature IGF-1, both *in vitro* and *in vivo* (Fig. 2 and Supplementary Fig. S2).

The amino acid composition of mature IGF-1 markedly differs compared to E-domains[5,35]. In particular, the E-domains were enriched in disorder-promoting amino acids, and by bioinformatic analyses, we demonstrated that the E-domains are IDRs, which were also confirmed by the limited proteolysis approach (Fig. 1 and Supplementary Fig. S1C). Therefore, the proIGF-1s consist of two different parts: the mature IGF-1, with a well-organised structure, and the flexible C-terminal E-domains. Notably, we verified that the disorder features of E-domains are conserved both across species and for all the three E-domains generated by alternative splicing[5].

Many studies show that the conformational plasticity associated with intrinsic disorder provides IDRs with a complementary functional repertoire of ordered domains[15,17], especially if these IDRs are at the protein termini[36]. Therefore, we next focused on the potential functional role of C-terminal Ea-, Eb- and Ec-domains.

The Ea-domain contains a highly conserved N-glycosylation site (N92) (Fig. 3A), which is heavily glycosylated with sugars comprising over 30% of the total mass of the proIGF-1Ea (Fig. 3B; see also[6-7]). Here, we demonstrated that the inhibition of N-glycosylation by site-directed mutagenesis (Fig. 3B), Tun treatment (Fig. 4A) or glucose withdrawal (Supplementary Fig. S5A), blocked the intracellular production of glycosylated proIGF-1Ea. The secretion of glycosylated proIGF-1Ea and mature IGF-1 was also substantially affected by the inhibition of N-glycosylation (Fig. 3C,

Fig. 4B and Supplementary Fig. S5B). Accordingly, the conditioned media from IGF-1Ea-transfected cells treated with Tun were unable to activate the IGF-1R pathway in human MCF-7 breast cancer cell line (Fig. 4C). It is important to highlight that the unglycosylated proIGF-1Ea was not accumulated after N-glycosylation inhibition. In this line, we demonstrated that unglycosylated proIGF-1Ea had a faster turnover rate compared to glycosylated proIGF-1Ea (Fig. 5A). The proteasome inhibitors MG132 partially rescued the accumulation defect of unglycosylated proIGF-1Ea, but also markedly increased the production of a ~23 kDa band, probably representing unglycosylated proIGF-1Ea dimer (Figs. 5B, 5C and 5D). N-glycosylation influences the folding and trafficking of many glycoproteins ensuring protein solubility and minimising aggregation[37-39]. Accordingly, our work indicates that addition of N-glycan on Ea-domain prevents the degradation of proIGF-1Ea by the proteasome, probably overcoming the folding limitation of unglycosylated proIGF-1Ea and its entry into the ER-associated degradation (ERAD) pathway[40-41]. This finding has interesting functional implications since the IGF-1Ea isoform is widely expressed in normal tissues as well as in various tumour cells[5,8,10,42]. Thus, interfering with proIGF-1Ea glycosylation provides a means to regulate IGF-1 production. In this regard, we also investigated the effect of 2-DG on proIGF-1Ea glycosylation. 2-DG, a non-metabolizable glucose analogue, is one of the most frequently used antiglycolytic agents and it has come under increasing scrutiny as a therapeutic agent, especially in cancer[24]. However, since 2-DG also mimics mannose, might also interfere with N-linked glycosylation[24,26]. In fact, recent studies showed that interference with N-linked glycosylation rather than glycolysis as the predominant mechanism by which 2-DG inhibits cancer cell growth under normoxia[27-29]. Here, we demonstrated that 2-DG interferes with proIGF-1Ea glycosylation preventing the formation of the normal, highly glycosylated proIGF-1Ea (~17 kDa) (Fig. 6A). However, 2-DG also determined a concomitant dose-dependent increase of a ~12 kDa protein that probably represented an aberrant under-glycosylated proIGF-1Ea form. Notably, this protein was stable and efficiently secreted (Fig. 6B). Accordingly, despite the loss of highly glycosylated proIGF-1Ea after the 2-DG treatment, the condition media of these cells fully activated the IGF-1R of human breast cancer cell line MCF-7 (Fig. 6C). Thus, these data suggested that 2-DG interfere with normal proIGF-1Ea glycosylation process leading to an accumulation of an under-glycosylated form which is still able to fold and secreted by the cells. These results

are in line with a recent study showing that the main effect of low, non-toxic and pharmacologically relevant concentration of 2-DG was to increase the incorporation of mannose on the glycan structure instead of inhibiting glycosylation[26]. Thus, 2-DG alters the glycosylation process favouring the synthesis of glycoproteins with incomplete or defective oligosaccharide chains. Taking advantage of the increased glucose uptake that occurs in most tumours, treatment with 2-DG might represent a new diagnostic and therapeutic tool to discriminate normal and cancer cells based on recognition of 2DG induced altered glycoproteins[26]. Further studies are needed to characterise the glycan structure of the 2-DG- induced under-glycosylated proIGF-1Ea and to test its potential role as a cancer biomarker.

We finally investigated if N-glycosylation has a functional role in proIGF-1Eb or proIGF-1Ec production. We should point out that in these prohormones two alternative E-domains, i.e. the Eb- and Ec-domains, replace the Ea-domain[3-5]. The IGF-1Eb and IGF-1Ec isoforms are not so widely expressed as IGF-1Ea, although their levels increased under specific conditions/stimuli (e.g. cancer[10,43] or exercise[44]) and are also expressed in a species-specific manner suggesting isoform-specific functions[5]. Accordingly, in the present study, we demonstrated that the behaviour of proIGF-1Eb and proIGF-1Ec entirely differed compared to proIGF-1Ea. In particular, since both Eb- and Ec-domains lacked N-glycosylation sites, the production of proIGF-1Eb and proIGF-1Ec was unaffected by inhibition of N-glycosylation (Fig. 7A and Supplementary Fig. S8C). Moreover, the Eb- and Ec-domains determined the partial nuclear localisation of proIGF-1Eb and proIGF-1Ec (Figs. 7B and 7C; see also[30-32]). The mechanisms by which proIGF-1Eb and proIGF-1Ec enter into the nucleus after trafficking through the ER is still unknown[31]. As described for Ea-domain, the presence of the disordered Eb- and Ec- tails probably represents an obstacle for their correct folding and traffic through the ER. Accordingly, we previously demonstrated that the secretion of IGF-1 (both proIGF-1s and mature IGF-1) was significantly lower in the cell overexpressing IGF-1Eb or IGF-1Ec isoforms compared to IGF-1Ea-transfected cells[7]. Furthermore, the inhibition of proteasome by MG132 determined nuclear accumulation of proIGF-1Eb and proIGF-1Ec (Fig. 7D) demonstrating that the production of these prohormones was controlled through their subcellular localization.

In conclusion, our data suggested that alternative E-domains act as flexible tails controlling proIGF-1s and mature IGF-1 production. In particular, we demonstrated

that N-linked glycosylation regulates the stability and secretion of proIGF-1Ea, probably ensuring proper prohormone folding and favouring its passage through the secretory pathway. Interference with proIGF-1Ea N-glycosylation (e.g. mutation of N-glycosylation site or Tun, glucose starvation and 2-DG treatment) directly affects protein IGF-1 level. The splice variants IGF-1Eb and IGF-1Ec encode the alternative proIGF-1Eb and proIGF-1Ec forms, which were insensitive to modulation of glycosylation. Notably, the Eb- and Ec- disordered tails promoted the nuclear accumulation of proIGF-1Eb and proIGF-1Ec, and thus directly affected the efficiency of proIGF-1s secretion.

Thus, disordered E-domains play a crucial role in the structure, regulation, localisation and functioning of IGF-1.

Methods

Tissue sampling and cell cultures

Human and animal tissues were previously obtained and analysed for mRNA extraction and IGF-1 isoforms quantification as described in Annibalini et al.[5]. From the same tissues stored at -80°C, about 30 mg were used for protein extraction and subsequent western blotting analysis. Freshly frozen normal human adipose (2 males and 2 females, mean age 52 +/− 12 years) samples were provided by the complex structure of biomarkers (DOSMM) of National Cancer Institute of Milan. All subjects provided written informed consent before archival tissue samples. Pig (Sus scrofa; 3 males) and cow (Bos Taurus; 3 females) tissues were collected at local slaughter during routine meat inspection. The age of the animals ranged from 2 to 5 years. The study was conducted and reported in accordance with standards for reporting of diagnostic accuracy (STARD) requirements. The experimental protocols were in accordance with the Guide for the Care and Use of Laboratory Animals by Ministero della Sanità D.L. 116 (1992) and approved by the University of Urbino “Carlo Bo” Committee. The HEK293 and MCF-7 cell lines were obtained from the American Type Culture Collection (ATCC, Rockville, MD, USA). The cells lines were cultured in DMEM media supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 1x MEM Non-essential Amino Acid Solution, 0.1 mg/ml streptomycin and 0.1 U/L penicillin. Cells were maintained in a humidified incubator (5% CO_2) at 37 °C. All cell culture materials were purchased from Sigma-Aldrich (St. Louis, MO, USA).

Protein extraction and western blotting analysis

Tissues and cells were processed for western blot analysis as previously reported[7]. Briefly, protein extracts were prepared by homogenizing with a Polytron homogenizer (KINEMATICA AG, Switzerland) in lysis buffer containing: 20 mM HEPES (pH 7.9), 25% v/v glycerol, 0.42 M NaCl, 0.2 mM EDTA, 1.5 mM MgCl2, 0.5% v/v Nonidet P-40, 1 mM DTT, 1 mM Naf, 1 mM Na_3VO_4, and 1× complete protease inhibitor cocktail (Roche Diagnostics Ltd, Mannheim, Germany). The lysates were frozen and thawed twice and clarified by centrifugation at 12000 rpm for 10 minutes at 4°C. Protein concentration in each sample was determined using the Bradford colourimetric assay. An equal amount of total proteins was fractionated

by SDS-PAGE on a 15% polyacrylamide gel and then transferred to PVDF or nitrocellulose membranes (Bio-Rad Laboratories Inc). The membranes were incubated overnight at 4°C with the primary antibodies directed towards: IGF-1 (1:2000; no. 500P11), purchased from Peprotech (Rocky Hill, New Jersey, USA), proIGF-1s (1:2000; no. PA5-19382) purchased from Invitrogen (Carlsbad, California, USA), β-tubulin (1:2000; no. 2146), lamin A/C (1:2000; no. 4777), calnexin (1:2000; no. 2679) phospho-IGF-1 Receptor β (1:2000; no. 3024), IGF-1 Receptor β (1:2000; no. 3027), phospho-p44/42 (ERK1/2) (1:2000; no. 9101), p44/42 (ERK1/2) (1:2000; no. 9102), phospho-Akt (Ser473) (1:2000; no. 9271) and Akt (1:2000; no. 9272) were purchased from Cell Signaling Technology (Beverly, MA, USA). Membranes were washed and incubated with appropriate secondary HRP-conjugated antibodies (Bio-Rad Laboratories Inc) at room temperature for 1 hour. After TBS-T washing, protein bands were visualised using Clarity Western ECL Substrate (Bio-Rad Laboratories Inc) and were quantified using Fluor-S MAX System (Bio-Rad Laboratories Inc) equipped with Quantity One software. β-tubulin was used for normalisation.

Cell culture transfection assays

Transient cell transfection was carried out with TransIT-X2® Transfection Reagent (Mirus Bio, Madison, WI, USA) with plasmid constructs containing sequences encoding proIGF-1s as previously described[7]. Each plasmid contained DNA encoding the class 1 IGF-1 48-amino acid signal peptide, the mature 70-amino acid IGF-1 peptide, the first 16 aa common in all E-domains, and C-terminal sequences encoding either the Ea (19 aa), the Eb (61 aa) or the Ec (24 aa) domain. Where indicated, to interfere with N-glycosylation, cells were grown in glucose-depleted medium or treated with 0.1 µg/ml of Tunicamycin (Tun) (Sigma-Aldrich, St. Louis, MO, USA) or 2-Deoxyglucose (2-DG) (Sigma-Aldrich, St. Louis, MO, USA) for 24h. The supernatants of HEK293-transfected cells treated with Tun or 2-DG were also used to treat the MCF-7 cells for 1h to evaluate their effects on IGF-1R, AKT and ERK1/2 phosphorylation. Afterwards, the MCF-7 cells were washed with PBS and lysed for western blotting analysis. Where indicated, the protein synthesis inhibitor cycloheximide (CHX) (Sigma-Aldrich) and the proteasome inhibitor MG132 (Sigma-Aldrich) were added to HEK293 cells cultured with a final concentration of 25µg/ml

and 10μM respectively in a time-course experiment. The cells were then collected at indicated time points, lysed and prepared to western blotting analysis.

The efficiency of transfection was estimated by GFP-dependent fluorescence and by real-time RT-PCR for total IGF-1 mRNA levels at 24 hours after transfection, as previously described[7].

Deglycosylation of proIGF-1Ea was performed by incubation of proIGF-1Ea enriched media with 2500 U of PNGase F (NewEngland Biolabs) for 3 hours at 37°C, according to manufacturer's recommendations. Aliquot of proIGF-1Ea supernatant incubated with equal volume of PNGase assay reaction buffer without the enzyme PNGase F was used as a control.

Limited proteolysis

Supernatants of IGF-1Ea-, IGF-1Eb- or IGF-1Ec-transfected HEK293 cells were concentrated using an Amicon Ultra 3K centrifugal filter unit (Merck Millipore, Billerica, MA, USA) and subjected to limited proteolysis. Limited proteolysis was performed by enzymatic digestion at 37°C adding bovine trypsin (Sigma-Aldrich, Italy) or proteinase K (Sigma-Aldrich, Italy) to protein extract at a ratio 1:100 enzyme/substrate (w/w). Reactions were removed over a time-course, and the digested products were quenched with SDS sample buffer before SDS-PAGE and analysed by western blotting. The quantification of bands intensity was normalised to the "no trypsin" samples of each set.

Conservation of Ea-domain N-glycosylation site

BLASTP was used to align protein sequence of human Ea-domain (EVHLKNASRGSAGNKNYRM) against the entire UniProt database using E-threshold of 0.01[45]. The obtained UniProt hits (250 sequences with a minimal sequence identity of ~60%) were aligned using MUSCLE[46], and WebLogo (http://weblogo.berkeley.edu/logo.cgi) was used to create relative frequency plots.

Site-direct mutagenesis

Site-direct mutagenesis was used to generate the proIGF-1Ea^{N92D} mutant lacking the glycosylation site. The composition of PCR mutagenesis reactions run on an Applied Biosystems SimpliAmp Thermal Cycler were as follows: 25 µl of 2× Platinum SuperFi PCR Master Mix (Thermo Fisher); 10 µM of each primer (N92D-F 5'-ATTTGAAGGACGCAAGTAGAGGGAG-3'; 5'- N92D-R TTGCGTCCTTCAAATGTACTTCCTTC-3') 1 ng plasmid contained DNA encoding proIGF-1Ea and H_2O to 50 µl. Following PCR, the reactions were incubated with 1 µl (20 U) of Anza^{TM} 10 DpnI (Invitrogen) for 1 h to selectively digest the methylated parent plasmids, and the resulting PCR products were purified with GenElute PCR Clean-Up Kit (Sigma-Aldrich). 2 µl of the purified PCR reactions were transformed into electrocompetent E. coli XL1-Blue cells with selection for resistance to ampicillin (100 µg/ml), and successful mutagenesis was confirmed by sequencing of plasmid DNA.

Subcellular localisation analysis

The cytosol, ER and nucleus isolations were performed as described in[47]. Briefly, cells were treated with permeabilisation buffer for 5 minutes and then were centrifuged at 3000 g for 5 min to collect the cytosol fraction in the soluble fraction. The pellet was then washed, subjected to lysis buffer for 5 minutes and centrifuged again at 3000 g for 5 min to collect the ER in the soluble fraction, whereas the pellet represented the nucleus fraction. Finally, all the samples were clarified at 7500 g for 10 minutes to remove cell debris and transferred to clean tubes for further analysis. The HEK293 cells subjected to immunofluorescence were seeded in 4-well chamber slide at a density of $5x10^4$ cells/well, incubated overnight and transfected for IGF-1 isoforms expression as previously described. After overnight incubation, cells were fixed with 4% paraformaldehyde for 15 minutes, permeabilised with 0.2% TritonX-100, blocked with 5% of goat serum and incubated overnight at 4°C with the anti-mature IGF-1 antibody (Prepotech). Next, cells were incubated 1 hour with an anti-rabbit-PE conjugated antibody, stained with DAPI, mounted with Fluoreshield (Sigma) and photographed with a fluorescence microscope (ZEISS AxioVert A.1).

Statistical analysis

Data are represented as mean ± SEM of at least three independent experiments. Statistical analyses were performed using repeated measures ANOVA or one-way ANOVA as appropriate, followed by Bonferroni's multiple comparison post hoc tests. A *p*-value <0.05 was considered statistically significant.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its Supplementary Information files or upon reasonable request.

References

1 Dyer, A. H., Vahdatpour, C., Sanfeliu, A. & Tropea, D. The role of Insulin-Like Growth Factor 1 (IGF-1) in brain development, maturation and neuroplasticity. *Neuroscience* **325**, 89-99, doi:10.1016/j.neuroscience.2016.03.056 (2016).

2 Christopoulos, P. F., Msaouel, P. & Koutsilieris, M. The role of the insulin-like growth factor-1 system in breast cancer. *Mol Cancer* **14**, 43, doi:10.1186/s12943-015-0291-7 (2015).

3 Barton, E. R. The ABCs of IGF-I isoforms: impact on muscle hypertrophy and implications for repair. *Appl Physiol Nutr Metab* **31**, 791-797, doi:10.1139/h06-054 (2006).

4 Philippou, A., Maridaki, M., Pneumaticos, S. & Koutsilieris, M. The complexity of the IGF1 gene splicing, posttranslational modification and bioactivity. *Mol Med* **20**, 202-214, doi:10.2119/molmed.2014.00011 (2014).

5 Annibalini, G. *et al.* MIR retroposon exonization promotes evolutionary variability and generates species-specific expression of IGF-1 splice variants. *Biochim Biophys Acta* **1859**, 757-768, doi:10.1016/j.bbagrm.2016.03.014 (2016).

6 Durzynska, J., Philippou, A., Brisson, B. K., Nguyen-McCarty, M. & Barton, E. R. The pro-forms of insulin-like growth factor I (IGF-I) are predominant in skeletal muscle and alter IGF-I receptor activation. *Endocrinology* **154**, 1215-1224, doi:10.1210/en.2012-1992 (2013).

7 De Santi, M. *et al.* Human IGF1 pro-forms induce breast cancer cell proliferation via the IGF1 receptor. *Cell Oncol (Dordr)* **39**, 149-159, doi:10.1007/s13402-015-0263-3 (2016).

8 Durzynska, J. & Barton, E. IGF expression in HPV-related and HPV-unrelated human cancer cells. *Oncol Rep* **32**, 893-900, doi:10.3892/or.2014.3329 (2014).

9 Pastural, E. *et al.* RIZ1 repression is associated with insulin-like growth factor-1 signaling activation in chronic myeloid leukemia cell lines. *Oncogene* **26**, 1586-1594, doi:10.1038/sj.onc.1209959 (2007).

10 Christopoulos, P. F., Philippou, A. & Koutsilieris, M. Pattern of IGF-1 variants' expression in human cancer cell lines using a novel q-RT-PCR approach. *Anticancer Res* **35**, 107-115, doi:35/1/107 (2015).

11 Alexandraki, K. I. *et al.* IGF-IEc expression is increased in secondary compared to primary foci in neuroendocrine neoplasms. *Oncotarget* **8**, 79003-79011, doi:10.18632/oncotarget.20743 (2017).

12 Barton, E. R., DeMeo, J. & Lei, H. The insulin-like growth factor (IGF)-I E-peptides are required for isoform-specific gene expression and muscle hypertrophy after local IGF-I production. *J Appl Physiol (1985)* **108**, 1069-1076, doi:10.1152/japplphysiol.01308.2009 (2010).

13 Redwan, E. M., Linjawi, M. H. & Uversky, V. N. Looking at the carcinogenicity of human insulin analogues via the intrinsic disorder prism. *Sci Rep* **6**, 23320, doi:10.1038/srep23320 (2016).

14 Lieutaud, P. *et al.* How disordered is my protein and what is its disorder for? A guide through the "dark side" of the protein universe. *Intrinsically Disord Proteins* **4**, e1259708, doi:10.1080/21690707.2016.1259708 (2016).

15 Babu, M. M. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem Soc Trans* **44**, 1185-1200, doi:10.1042/BST20160172 (2016).

16 Brown, C. J., Johnson, A. K., Dunker, A. K. & Daughdrill, G. W. Evolution and disorder. *Curr Opin Struct Biol* **21**, 441-446, doi:10.1016/j.sbi.2011.02.005 (2011).

17 Dunker, A. K., Bondos, S. E., Huang, F. & Oldfield, C. J. Intrinsically disordered proteins and multicellular organisms. *Semin Cell Dev Biol* **37**, 44-55, doi:10.1016/j.semcdb.2014.09.025 (2015).

18 Ciechanover, A. Intracellular protein degradation: from a vague idea through the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Neurodegener Dis* **10**, 7-22, doi:10.1159/000334283 (2012).

19 Pauwels, K., Lebrun, P. & Tompa, P. To be disordered or not to be disordered: is that still a question for proteins in the cell? *Cell Mol Life Sci* **74**, 3185-3204, doi:10.1007/s00018-017-2561-6 (2017).

20 Buljan, M. *et al.* Tissue-specific splicing of disordered segments that embed binding motifs rewires protein interaction networks. *Mol Cell* **46**, 871-883, doi:10.1016/j.molcel.2012.05.039 (2012).

21 Oates, M. E. *et al.* D(2)P(2): database of disordered protein predictions. *Nucleic Acids Res* **41**, D508-516, doi:10.1093/nar/gks1226 (2013).

22 Fontana, A., de Laureto, P. P., Spolaore, B. & Frare, E. Identifying disordered regions in proteins by limited proteolysis. *Methods Mol Biol* **896**, 297-318, doi:10.1007/978-1-4614-3704-8_20 (2012).

23 Boutrid, H. *et al.* Targeting hypoxia, a novel treatment for advanced retinoblastoma. *Invest Ophthalmol Vis Sci* **49**, 2799-2805, doi:10.1167/iovs.08-1751 (2008).

24 Xi, H., Kurtoglu, M. & Lampidis, T. J. The wonders of 2-deoxy-D-glucose. *IUBMB Life* **66**, 110-121, doi:10.1002/iub.1251 (2014).

25 Parodi, A. J. Protein glucosylation and its role in protein folding. *Annu Rev Biochem* **69**, 69-93, doi:10.1146/annurev.biochem.69.1.69 (2000).

26 Ahadova, A., Gebert, J., von Knebel Doeberitz, M., Kopitz, J. & Kloor, M. Dose-dependent effect of 2-deoxy-D-glucose on glycoprotein mannosylation in cancer cells. *IUBMB Life* **67**, 218-226, doi:10.1002/iub.1364 (2015).

27 Kurtoglu, M. *et al.* Under normoxia, 2-deoxy-D-glucose elicits cell death in select tumor types not by inhibition of glycolysis but by interfering with N-linked glycosylation. *Mol Cancer Ther* **6**, 3049-3058, doi:10.1158/1535-7163.MCT-07-0310 (2007).

28 Liu, H. *et al.* Conversion of 2-deoxyglucose-induced growth inhibition to cell death in normoxic tumor cells. *Cancer Chemother Pharmacol* **72**, 251-262, doi:10.1007/s00280-013-2193-y (2013).

29 Gu, L. *et al.* Low dose of 2-deoxy-D-glucose kills acute lymphoblastic leukemia cells and reverses glucocorticoid resistance via N-linked glycosylation inhibition under normoxia. *Oncotarget* **8**, 30978-30991, doi:10.18632/oncotarget.16046 (2017).

30 Peng, Q. *et al.* The nuclear localization of MGF receptor in osteoblasts under mechanical stimulation. *Mol Cell Biochem* **369**, 147-156, doi:10.1007/s11010-012-1377-9 (2012).

31 Tan, D. S., Cook, A. & Chew, S. L. Nucleolar localization of an isoform of the IGF-I precursor. *BMC Cell Biol* **3**, 17 (2002).

32 Durzynska, J., Wardzinski, A., Koczorowska, M., Gozdzicka-Jozefiak, A. & Barton, E. R. Human Eb peptide: not just a by-product of pre-pro-IGF1b processing? *Horm Metab Res* **45**, 415-422, doi:10.1055/s-0032-1331699 (2013).

33 von Mikecz, A. The nuclear ubiquitin-proteasome system. *J Cell Sci* **119**, 1977-1984, doi:10.1242/jcs.03008 (2006).

34 Chretien, M. How the prohormone theory solved two important controversies in hormonal and neural Peptide biosynthesis. *Front Endocrinol (Lausanne)* **4**, 148, doi:10.3389/fendo.2013.00148 (2013).

35 Hede, M. S. *et al.* E-peptides control bioavailability of IGF-1. *PLoS One* **7**, e51152 (2012).

36 Uversky, V. N. The most important thing is the tail: multitudinous functionalities of intrinsically disordered protein termini. *FEBS Lett* **587**, 1891-1901, doi:10.1016/j.febslet.2013.04.042 (2013).

37 Helenius, A. & Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* **73**, 1019-1049, doi:10.1146/annurev.biochem.73.011303.073752 (2004).

38 Skropeta, D. The effect of individual N-glycans on enzyme activity. *Bioorg Med Chem* **17**, 2645-2653, doi:10.1016/j.bmc.2009.02.037 (2009).

39 Mitra, N., Sinha, S., Ramya, T. N. & Surolia, A. N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem Sci* **31**, 156-163, doi:10.1016/j.tibs.2006.01.003 (2006).

40 Smith, M. H., Ploegh, H. L. & Weissman, J. S. Road to ruin: targeting proteins for degradation in the endoplasmic reticulum. *Science* **334**, 1086-1090, doi:10.1126/science.1209235 (2011).

41 Tsai, B., Ye, Y. & Rapoport, T. A. Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nat Rev Mol Cell Biol* **3**, 246-255, doi:10.1038/nrm780 (2002).

42 Pelosi, M., Alfo, M., Martella, F., Pappalardo, E. & Musaro, A. Finite mixture clustering of human tissues with different levels of IGF-1 splice variants mRNA transcripts. *BMC Bioinformatics* **16**, 289, doi:10.1186/s12859-015-0689-7 (2015).

43 Pickard, A., Durzynska, J., McCance, D. J. & Barton, E. R. The IGF axis in HPV associated cancers. *Mutat Res Rev Mutat Res* **772**, 67-77, doi:10.1016/j.mrrev.2017.01.002 (2017).

44 Vassilakos, G., Philippou, A. & Koutsilieris, M. Identification of the IGF-1 processing product human Ec/rodent Eb peptide in various tissues: Evidence for its differential regulation after exercise-induced muscle damage in humans. *Growth Horm IGF Res* **32**, 22-28, doi:10.1016/j.ghir.2016.11.001 (2017).

45 Magrane, M. UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)* **2011**, bar009, doi:10.1093/database/bar009 (2011).

46 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

47 Jagannathan, S., Nwosu, C. & Nicchitta, C. V. Analyzing mRNA localization to the endoplasmic reticulum via cell fractionation. *Methods Mol Biol* **714**, 301-321, doi:10.1007/978-1-61779-005-8_19 (2011).

Acknowledgements

This research was supported by the Italian Ministry of Health, “Ricerca Finalizzata 2011” (grant number: GR-2011-02348423) and “Progetti di ricerca DISB 2017”, granted by the Department of Biomolecular Sciences, University of Urbino Carlo Bo, Italy.

We thank Dr. F. Berrino for his careful and critical reading of our discussion paper.

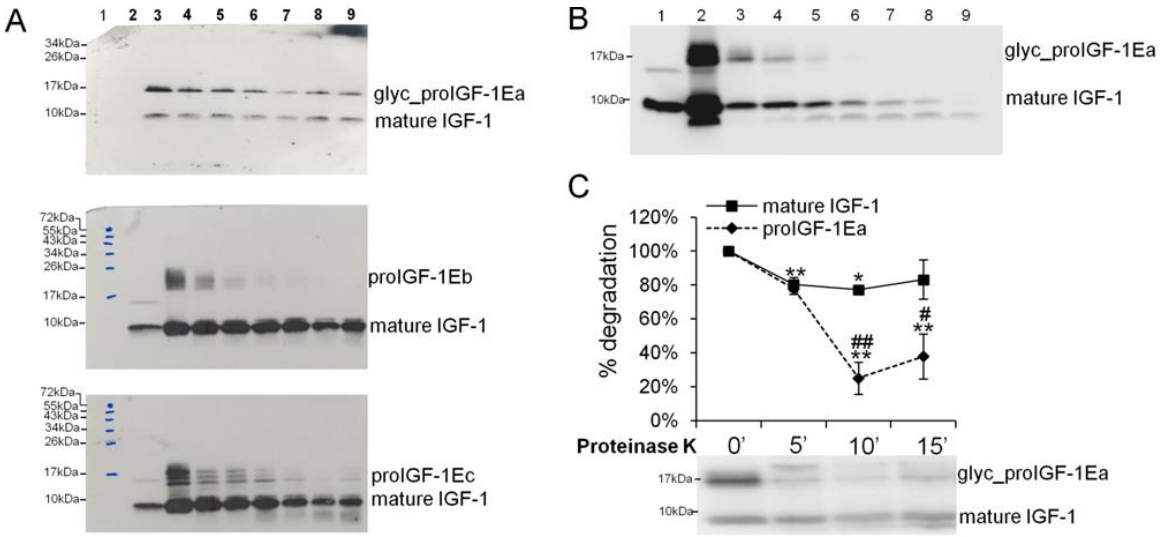
Author Contributions

G.A., S.C. and M.D.S. conceived and designed the experiments. G.A., S.C., M.D.S., R.S. and L.D.P. performed the experiments. G.A., M.G. and E.B. analysed the data and performed the statistical analysis. G.A. wrote the paper. A.V., G.B., V.S. and E.B. provided overall direction to the project and revised the manuscript. All authors reviewed, discussed and approved the manuscript.

Competing Interests

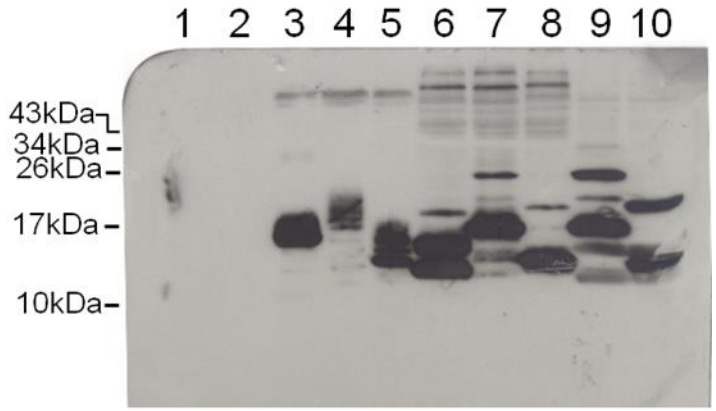
The authors declare that there is no conflict of interest regarding the publication of this article.

Supplementary Information



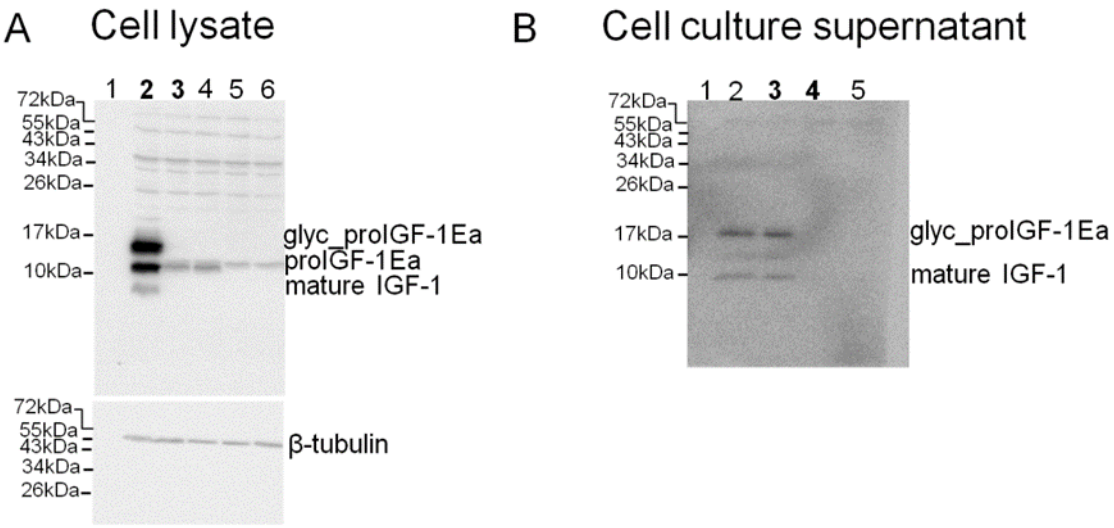
Supplementary Figure S1. Limited proteolysis of proIGF-1s with trypsin, original uncropped blots (A). Limited proteolysis of mature and glycosylated proIGF-1Ea with long-term trypsin incubation (B), or proteinase K digestion (C).

Cell culture supernatants of IGF-1Ea-, IGF-1Eb- or IGF-1Ec-transfected HEK293 cells were concentrated using an Amicon Ultra 3K centrifugal filter and incubated with trypsin (A and B) or proteinase K (C) at 37°C. Reactions were removed over a time-course, and the digested products were loaded on 12% SDS-PAGE and analysed by western blotting with an anti-mature IGF-1 antibody. Sample names are as follows: Figure S1A: 1: PageRuler Prestained Protein Ladder; 2: 25 ng of recombinant mature IGF-1 (Sigma-Aldrich I3769): 3-9: supernatants of IGF-1Ea-, IGF-1Eb- or IGF-1Ec-transfected HEK293 cells incubated with trypsin at 37°C for 0, 5, 10, 15, 20, 25 and 30 minutes. Figure S1B: 1: 25 ng of recombinant mature IGF-1 (Sigma-Aldrich I3769): 2: supernatants of IGF-1Ea transfected HEK293 cells (no trypsin); 3-9 supernatants of IGF-1Ea transfected HEK293 cells incubated with trypsin at 37°C for 0, 15, 30, 45, 60, 90 and 120 minutes. Figure S1C: Results are means ± SEM (n = 3). Repeated measures ANOVA, # ($p<0.01$) and ## ($p<0.0001$) significantly different compared to mature IGF-1; *($p<0.05$) and **($p<0.001$) significantly different compared to 0-minute time point. Samples that were included in cropped blots are indicated with bold numbers.



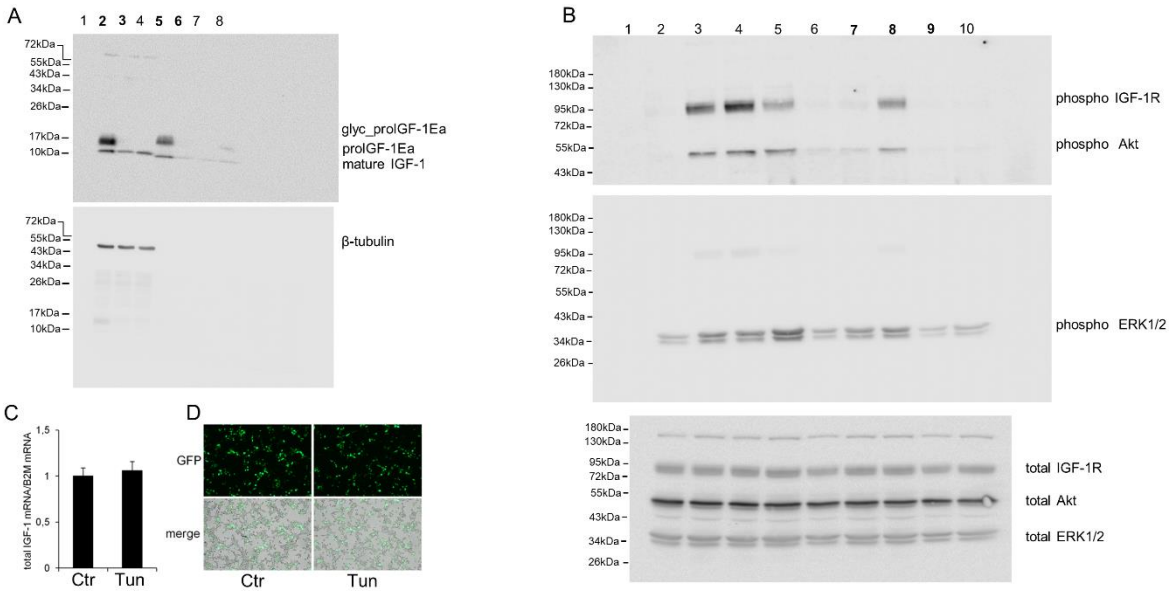
Supplementary Figure S2: Immunoblotting using an antibody directed against the common E-domain region of proIGF-1s (RSVRAQRHTD) (Invitrogen no. PA5-19382).

HEK293 cells were transfected with specific constructs and supernatants (lanes 3-5), and cell lysates (lanes 6-10) were analysed after 24 hours by western blot. Sample names are as follows: 1: PageRuler Prestained Protein Ladder; 2: 25 ng of recombinant mature IGF-1 (Sigma-Aldrich I3769); 3: cell culture supernatant from IGF-1Ea-transfected HEK293 cells; 4: cell culture supernatant from IGF-1Eb-transfected HEK293 cells; 5: cell culture supernatant from IGF-1Ec-transfected HEK293; 6: cell lysate from IGF-1Ea-transfected HEK293 cells; 7 and 9: cell lysate from IGF-1Eb-transfected HEK293 cells (technical replicates); 8 and 10: cell lysate from IGF-1Ec-transfected HEK293 cells (technical replicates).



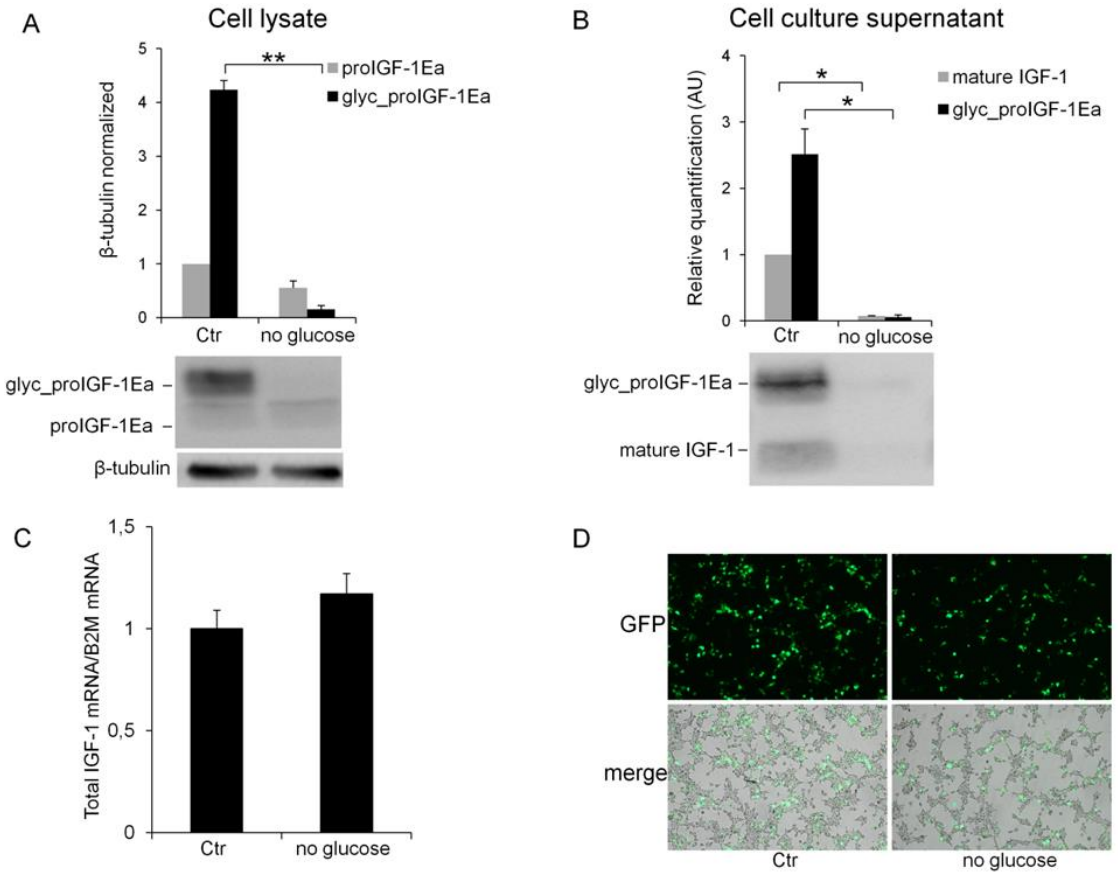
Supplementary Figure S3. Effect of Ea-domain N-glycosylation site mutation (IGF-1Ea^{N92D} mutant) on intracellular (A) or extracellular (B) proIGF-1Ea production. Original uncropped blots.

IGF-1Ea^{WT} and IGF-1Ea^{N92D} were transiently expressed in HEK293 cells. After 24 h the cell lysates (A) and cell culture supernatants (B) were analysed by western blot using an antibody directed against mature IGF-1 sequence. Sample names are as follows: Figure S3A; 1: PageRuler Prestained Protein Ladder; 2: cell lysate from IGF-1Ea^{WT}-transfected HEK293 cells; 3-6: cell lysate from IGF-1Ea^{N92D}-transfected HEK293 (biological replicates). Figure S3B; 1:PageRuler Prestained Protein Ladder; 2-3: cell culture supernatant from IGF-1Ea^{WT}-transfected HEK293 cells (biological replicates); 4-5: cell culture supernatant from IGF-1Ea^{N92D}-transfected HEK293 (biological replicates). Samples that were included in cropped blots are indicated with bold numbers.



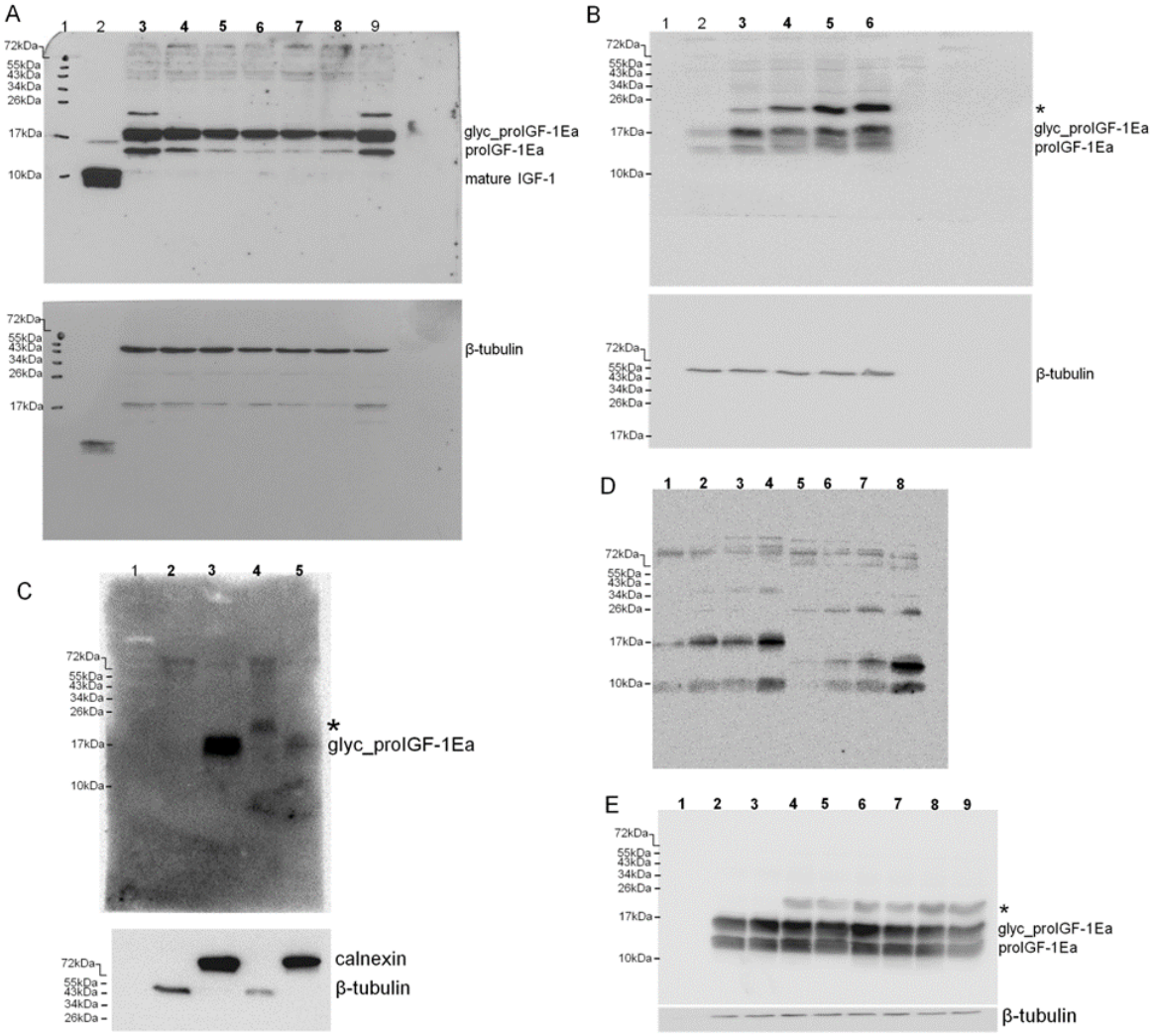
Supplementary Figure S4. Effects of tunicamycin (Tun) on proIGF-1Ea glycosylation, original uncropped blots (A and B). Total IGF-1 mRNA quantification (C) and GFP fluorescence (D) in IGF-1Ea-transfected HEK293 cells treated with Tun.

(A) IGF-1Ea was transiently expressed in HEK293 cells in the presence of 0.1 µg/ml of Tun. After 24 h the cell lysates (lanes 2-4) and cell culture supernatants (lanes 5-8) were analysed by western blot using an antibody directed against mature IGF-1 sequence. Sample names of Figure S4A are as follows: 1: PageRuler Prestained Protein Ladder; 2: cell lysate from untreated IGF-1Ea-transfected HEK293 cells; 3-4: cell lysate from IGF-1Ea-transfected HEK293 treated with 0.1 µg/ml of Tun (biological replicates); 5: cell culture supernatant from untreated IGF-1Ea-transfected HEK293 cells; 6-8: cell culture supernatant from IGF-1Ea-transfected HEK293 treated with 0.1 µg/ml of Tun (biological replicates). (B) Phosphorylation of IGF-1R, AKT and ERK1/2 after treatment of MCF-7 cells with cell culture supernatants from IGF-1Ea-transfected HEK293 cells treated with Tun. Sample names of Figure S4B are as follows: 1: PageRuler Prestained Protein Ladder; 2 and 7: treatment of MCF-7 cells with cell culture supernatants from empty vector-transfected HEK293 cells (biological replicates); 3-4: treatment of MCF7 cells with 25 ng of recombinant mature IGF-1 (biological replicates); 5 and 8: treatment of MCF-7 cells with cell culture supernatants from IGF-1Ea-transfected HEK293 cells (biological replicates); 6, 9 and 10: treatment of MCF-7 cells with cell culture supernatants from IGF-1Ea-transfected HEK293 treated with 0.1 µg/ml of Tun (biological replicates). (C and D) Comparison of transfection efficiency between IGF-1Ea-transfected HEK293 cells treated and untreated with Tun. No significant difference in total IGF-1 mRNA quantity (p=0.634) (C) or GFP fluorescence intensity (10x magnification) (D) was found. Samples that were included in cropped blots are indicated with bold numbers.



Supplementary Figure S5. Effects of glucose withdrawal on proIGF-1Ea glycosylation.

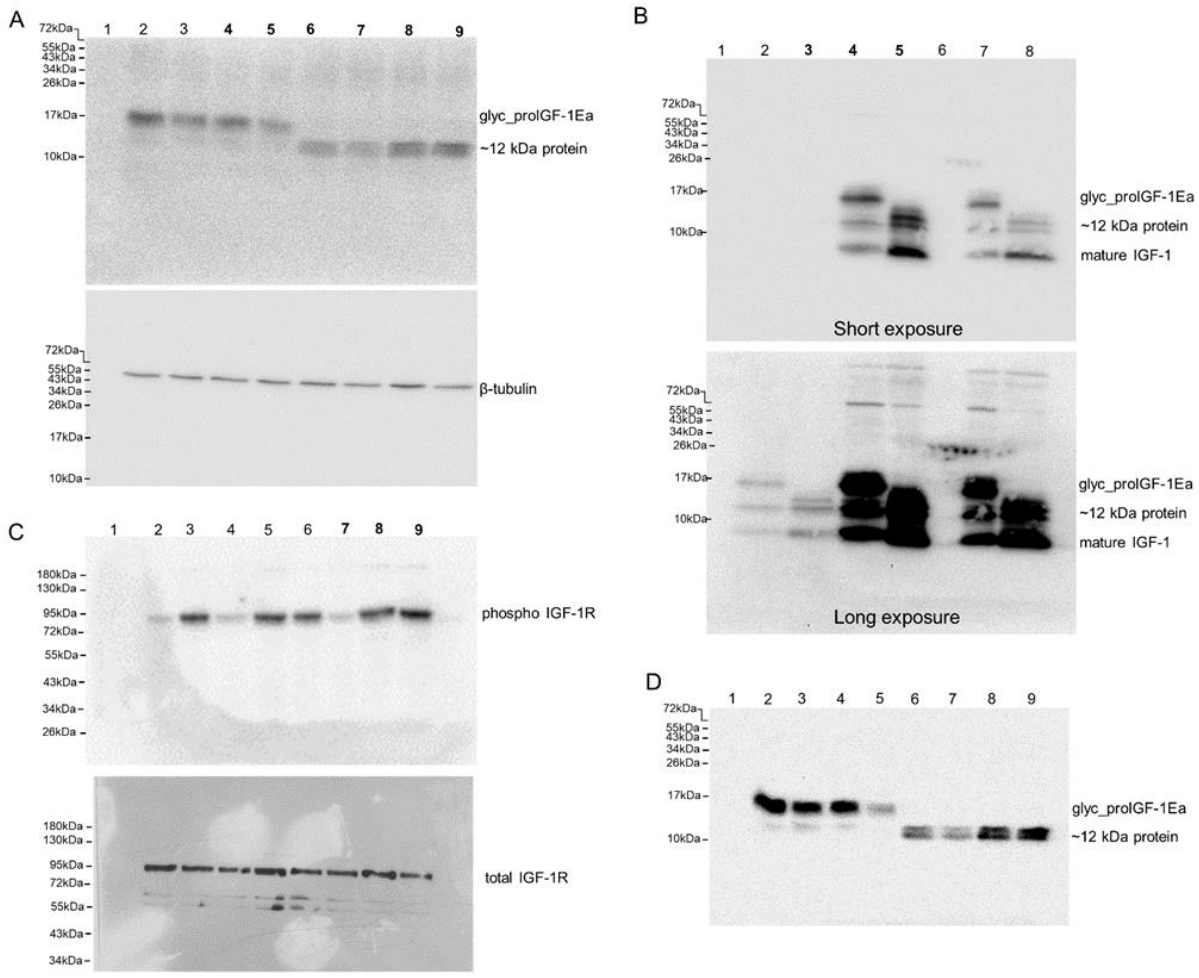
IGF-1Ea was transiently expressed in HEK293 cells in glucose-depleted medium (no glucose). After 24 h the cell lysates (A) and cell culture supernatants (B) were analysed by western blot and relative expression level of glycosylated proIGF-1Ea, unglycosylated proIGF-1Ea and mature IGF-1 was calculated. The band at a molecular weight around 17 kDa, corresponding to glycosylated proIGF-1A, disappeared in the absence of glucose in cell lysates (A) and the culture supernatants (B). The band corresponding to mature IGF-1 (~7kDa) was markedly reduced in the culture supernatants after glucose withdrawal (B). Results are means ± SEM (n= 3); T-test was used to evaluate statistical significance (*$p < 0.01$, **$p<0.0001$). β-tubulin was used as a loading control for the cell lysates. (C) Total IGF-1 mRNA quantification and (D) GFP fluorescence (10x magnification) in IGF-1Ea-transfected HEK293 cells grown in glucose-depleted medium. The mRNA expression of total IGF-1 (C) was unaffected by glucose withdrawal while glucose-depleted medium caused a slight decrease in cell number and GFP fluorescence (D).



Supplementary Figure S6. Analysis of unglycosylated and glycosylated proIGF-1Ea turnover, original uncropped blots (A, B, C and D). Immunoblotting of IGF-1Ea-transfected HEK293 treated with MG132 using an antibody directed against the common E-domain region of proIGF-1s (E).

(A and B) IGF-1Ea was transiently expressed in HEK293 cells in the presence of 25µg/ml of protein synthesis inhibitor cycloheximide (CHX) (A) or 10µM of the proteasome inhibitor MG132 (B) in a time-course experiment. Cytosol (Cyt) and endoplasmic reticulum (ER) isolations of IGF-1Ea-transfected HEK293 cells treated with MG132 were shown in figure S6C. Deglycosylation of proIGF-1Ea enriched media using the *N*-Glycosidase F (PNGase F) was displayed in figure S6D. Immunoblotting of IGF-1Ea-transfected HEK293 treated with 10µM MG132 using an antibody directed against common region of E-peptides (E).Sample names of Figure S6A are as follows: 1: PageRuler Prestained Protein Ladder; 2: 25 ng of recombinant mature IGF-1 (Sigma-Aldrich I3769); 3-7: cell lysate from IGF-1Ea-transfected HEK293 treated with 25µg/ml of CHX for 0', 30', 60', 90' and 120' respectively; 8: cell lysate from IGF-1Ea-transfected HEK293 co-treated with CHX and MG132 (10µM) for four hours; 9: technical replicate of sample 3. Sample names of Figure S6B are as follows: 1: PageRuler Prestained Protein Ladder; 2: cell lysate from untreated IGF-1Ea-transfected HEK293 cells; 3-6: cell lysate from IGF-1Ea-transfected HEK293 treated with 10µM of MG132 for 0, 2, 4 and 6 hours. After MG132 treatment intracellular accumulation of a ~23kDa band was found (indicated with an asterisk in figure S6B), probably representing unglycosylated proIGF-1Ea dimer. Sample names of Figure S6C are as follows: 1: PageRuler Prestained Protein Ladder; 2: Cyt fraction of IGF-1Ea-transfected HEK293 cells; 3: ER fraction of IGF-1Ea-transfected HEK293 cells; 4: Cyt fraction of IGF-1Ea-transfected HEK293 cells treated with 10µM of MG132 for 6h; 5: ER of IGF-1Ea-

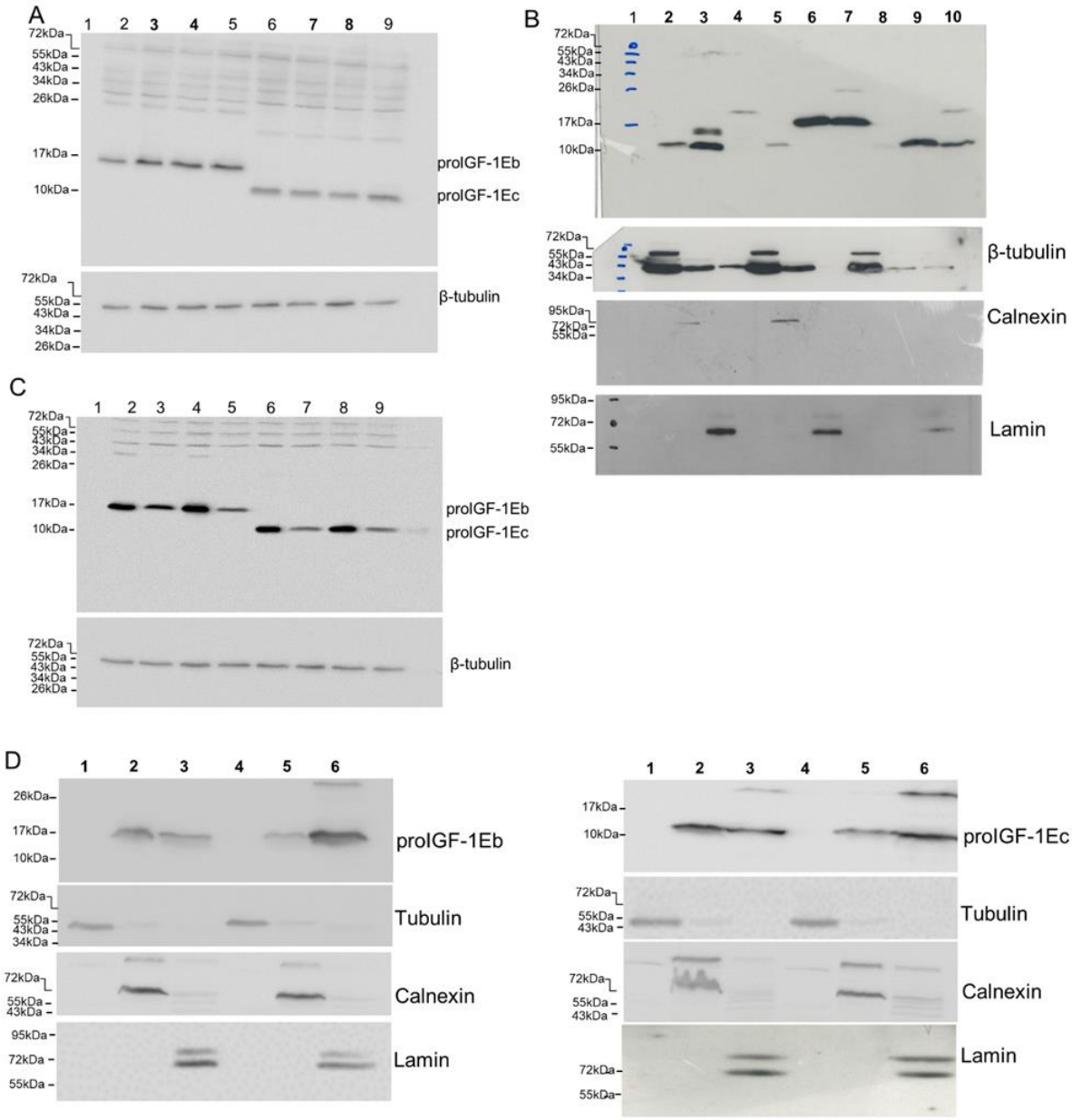
transfected HEK293 cells treated with 10µM of MG132 for 6h. Sample names of Figure S6D are as follows: 1-4: 10, 20, 30 and 40 µg of cell culture supernatants of IGF-1Ea-transfected HEK293; 5-8: 10, 20, 30 and 40 µg of cell culture supernatants of IGF-1Ea-transfected HEK293 after PNGase deglycosylation. Sample names of Figure S6E are as follows: 1: PageRuler Prestained Protein Ladder; 2-3: cell lysate from untreated IGF-1Ea-transfected HEK293 cells (technical replicates); 4-5: cell lysate from IGF-1Ea-transfected HEK293 treated with 10µM of MG132 for 2 hours (technical replicates); 6-7: cell lysate from IGF-1Ea-transfected HEK293 treated with 10µM of MG132 for 4 hours (technical replicates); 8-9: cell lysate from IGF-1Ea-transfected HEK293 treated with 10µM of MG132 for 6 hours (technical replicates). Samples that were included in cropped blots are indicated with bold numbers.



Supplementary Figure S7. Effect of 2-Deoxyglucose (2-DG) on glycosylated proIGF-1Ea production, original uncropped blots (A, B and C). Immunoblotting of IGF-1Ea-transfected HEK293 treated with 2-DG using an antibody directed against the common E-domain region of proIGF-1s (D).

IGF-1Ea was transiently expressed in HEK293 cells grown in low glucose medium (0.65 g/L) with or without 2-DG. After 24 h the cell lysates (A) and cell culture supernatants (B) were analysed by western blot using an antibody directed against mature IGF-1 sequence. Phosphorylation of IGF-1R after treatment of MCF-7 cells with cell culture supernatants from IGF-1Ea-transfected HEK293 cells treated or not with 0.2 g/L of 2-DG (C). Immunoblotting of IGF-1Ea-transfected HEK293 treated with 2-DG using an antibody directed against common region of E-peptides (D). Sample names of Figure S7A and S7D are as follows: 1: PageRuler Prestained Protein Ladder; 2-5: cell lysate from IGF-1Ea-transfected HEK293 (biological replicates); 6-7: cell lysate from IGF-1Ea-transfected HEK293 treated with 0.2 g/L of 2-DG (biological replicates); 8-9: cell lysate from IGF-1Ea-transfected HEK293 treated

with 0.75 g/L of 2-DG (biological replicates). Sample names of Figure S7B are as follows: 1: PageRuler Prestained Protein Ladder; 2, 4 and 7: cell culture supernatants from IGF-1Ea-transfected HEK293 (biological replicates); 3, 5 and 8: cell culture supernatants from IGF-1Ea-transfected HEK293 treated with 0.2 g/L of 2-DG (biological replicates). Sample names of Figure S7C are as follows: 1: PageRuler Prestained Protein Ladder; 2, 4 and 7: treatment of MCF-7 cells with cell culture supernatants from empty vector-transfected HEK293 cells (biological replicates); 3: treatment of MCF-7 cells with 25 ng of recombinant mature IGF-1; 5 and 8: treatment of MCF-7 cells with cell culture supernatants from IGF-1Ea-transfected HEK293 cells (biological replicates); 6 and 9 treatment of MCF-7 cells with cell culture supernatants from IGF-1Ea-transfected HEK293 cells treated with 0.2 g/L of 2-DG (biological replicates). Samples that were included in cropped blots are indicated with bold numbers.



Supplementary Figure S8. Effect of N-glycosylation inhibitor tunicamycin (Tun) on IGF-1Eb and IGF-1Ec production (A) and subcellular localisation of proIGF-1s in untreated (B) or MG132 treated HEK293 cells (D), original uncropped blots. Effects of glucose withdrawal on proIGF-1Eb and proIGF-1Ec production (C).

(A) IGF-1Eb (lanes 2-5) or IGF-1Ec (lanes 6-9) was transiently expressed in HEK293 cells in the presence of 0.1 µg /ml of tunicamycin (Tun). After 24 h the cell lysates were analysed by western blot using an antibody directed against mature IGF-1 sequence. Sample names of Figure S8A are as follows: 1: PageRuler Prestained Protein Ladder; 2-3: cell lysate from IGF-1Eb-transfected HEK293 (biological replicates); 4-5: cell lysate from IGF-1Eb-transfected HEK293 treated with 0.1 µg/ml Tun (biological replicates); 6-7: cell lysate from IGF-1Ec-transfected HEK293 (biological replicates); 8-9: cell lysate from IGF-1Ec-transfected HEK293 treated with 0.1 µg/ml Tun (biological replicates). β-tubulin was used as a loading control for the cell lysates. (B) Subcellular localisation of IGF-1 isoforms analysed by cytosol (Cyt), endoplasmic reticulum (ER) and nucleus isolations. Sample names of Figure S8B are as follows: 1: PageRuler Prestained Protein Ladder; 2, 5 and 8: Cyt fraction of HEK293 cells transfected with IGF-1Ea, IGF-1Eb and IGF-1Ec isoforms respectively; 3, 6 and 9: ER fraction of HEK293 cells transfected with IGF-1Ea, IGF-1Eb and IGF-1Ec isoforms

respectively; 4, 7 and 10: nuclear fraction of HEK293 cells transfected with IGF-1Ea, IGF-1Eb and IGF-1Ec isoforms respectively. β-tubulin was used as a loading control for the Cyt separation; calnexin as a control for the ER separation and lamin as a control for the nucleus separation. Samples that were included in cropped blots are indicated with bold numbers. (C) IGF-1Eb (lanes 2-5) or IGF-1Ec (lanes 6-9) was transiently expressed in HEK293 cells grown in normal or glucose-depleted medium (no glucose). After 24 h the cell lysates were analysed by western blot using an antibody directed against mature IGF-1 sequence. β-tubulin was used as a loading control for the cell lysates. Sample names of Figure S8C are as follows: 1: PageRuler Prestained Protein Ladder; 2 and 4: cell lysate from IGF-1Eb-transfected HEK293 cells (biological replicates); 3-5: cell lysate from IGF-1Eb-transfected HEK293 cells grown in glucose-depleted medium (biological replicates); 6-8: cell lysate from IGF-1Ec-transfected HEK293 cells (biological replicates); 7-9: cell lysate from IGF-1Ec-transfected HEK293 cells grown in glucose-depleted medium (biological replicates). (D) Subcellular localisation of proIGF-1Eb (left panel) or proIGF-1Ec (right panel) after treatment with 10μM of the proteasome inhibitor MG132 for 6h. Sample names of Figure S8D are as follows: 1, 2 and 3: Cyt, ER and nuclear fraction of HEK293 cells transfected with IGF-1Eb (left panel) or IGF-1Ec (right panel); 4, 5 and 6 Cyt, ER and nuclear fraction of HEK293 cells transfected with IGF-1Eb (left panel) or IGF-1Ec (right panel) and treated with the proteasome inhibitor MG132. β-tubulin was used as a loading control for the Cyt separation; calnexin as a control for the ER separation and lamin as a control for the nucleus separation.

CHAPTER 2

Inhibition of N-glycosylation impaired myoblast differentiation and IGF-1 receptor signalling pathways activation

Abstract

The number of skeletal muscle diseases shown to have a direct connection to defects in N-glycosylation has increased dramatically over the past several years. Patients suffering from some forms of congenital disorders of glycosylation (CDG), rare genetic diseases caused by abnormalities in glycan biosynthesis, present progressive muscle pathologies characterized by hypotonia, muscle weakness and increased muscle damage. The most common CDG, PMM2-CDG, also called CDG-Ia, results from mutations in phosphomannomutase 2 (PMM2), an indispensable enzyme that catalyses an early step of the N-glycosylation pathway. However, PMM2-CDG patients seldom undergo muscle biopsies, thus our knowledge of muscle glycosylation in these patients is very limited. In the present study, we used the C2C12 cell culture model to investigate the effects on myoblast differentiation of N-glycosylation inhibition by tunicamycin (TUN) or by genetic knockdown of PMM2. Treatment of C2C12 cells with non-cytotoxic doses of TUN (0.01µg/ml) inhibited myoblast differentiation and disrupted the coordinated temporal expression of myogenic regulator genes *Ccnd1*, *MyoD*, *Myogenin* and *Mrf4*. Accordingly, the cell proliferation marker PCNA increased in TUN-treated myoblasts and was accompanied by a reduction in myotube formation marker MF20. Interestingly, both transient (siRNA) and stable (CRISPR/Cas9) knockdown of the *PMM2* gene markedly decreased the differentiation of C2C12 cells, determining effects on myoblast differentiation markers resembling those of TUN administration. Control (CTR) myotubes also showed an increase in high mannose (ConA), complex mannose (PHA-L) and fucose (AAL) reactivities compared to myoblasts, while lectin binding decreased both in TUN-treated and CRISPR-PMM2 knockdown myotubes indicating a N-glycosylation deficiency. Finally, N-glycosylation inhibition by TUN or PMM2 knockdown decreased insulin-like growth factor 1 receptor (IGF-1R) and

downstream activation of AKT and dysregulated IGF-1 production compared to CTR cells.

Taken together, these results show that N-glycosylation plays a key role in myoblast differentiation and suggest that defects in skeletal muscle N-glycosylation could be a contributing factor to muscle-related symptoms commonly found in CDG patients.

Introduction

Myogenesis is a complex and highly regulated process that requires myoblast proliferation, alignment of cells, and subsequent fusion into multinucleated myotubes [1].

Many steps of myogenesis can be replicated *in vitro* using primary myogenic cells or C2C12 myoblasts, which in the absence of mitogenic stimuli form multinucleated myotubes [1].

The myogenic program is tightly coupled to a complex network of signal transduction pathways regulating the repression or expression of myogenic regulatory factors (MRFs), a group of basic helix-loop-helix transcription factors that include: *MyoD, Myf5, Myogenin, and Mrf4* [1,2]. In addition to myogenic factors, myogenesis involves other molecular actors such as insulin-like growth factor-1 (IGF-1) [3]. The primary effects of IGF-1 are mediated by binding to the IGF-1 receptor (IGF-1R), which is a ligand-activated receptor tyrosine kinase.

IGF-1R signals through the phosphatidylinositol 3 kinase (PI3K)/AKT pathway which promotes muscle protein synthesis and inhibits muscle protein degradation [4]. Recent studies have shown that *IGF-1R* is translated into a proreceptor form and its maturation requires N-linked glycosylation for the correct processing and transport to the cell surface [5-7]. Interestingly, we have recently demonstrated that IGF-1 is also synthesised as an N-glycosylated pro-hormone and N-glycosylation ensures proper pro-hormone folding and favors the passage of IGF-1 through the secretory pathway [8]. Thus, the IGF-1 signalling pathway seems to be particularly sensitive to the occurrence of N-glycosylation.

Growing evidence points to the role of protein glycosylation in myoblast fusion and differentiation [9-12]. For example, R.L. Gundry et al. identified 128 glycosylated proteins on the C2C12 myoblast cell surface, including the insulin receptor and IGF-1R [13]. Interestingly, several of these glycoproteins were modulated after the differentiation of myoblasts into myotubes, suggesting that temporal- and site-specific glycosylation is important for skeletal muscle cells development, differentiation and function. Accordingly, following myoblast differentiation, increased binding of lectins with different carbohydrate specificities was shown in both human [14] and mouse [15] cells.

However, unlike the extensively studied O-mannose glycans on α-Dystroglycan, the role of muscle protein N-glycosylation remains poorly characterized. Early evidence suggested that pharmacological inhibition of glycosylation by Tunicamycin (TUN), an antibiotic that interferes with the transfer of N-acetylglucosamine-1-phosphate from uridine 5-diphosphate-N-acetylglucosamine to dolichol monophosphate in the first step of glycoprotein synthesis, blocked the fusion of embryonic quail muscle cells [16]. Moreover, hypotonia, proximal muscle weakness and poor growth are present in some types of congenital disorder of glycosylation (CDG), rare genetic diseases in which glycosylation pathways are defective [17]. The most common CDG is PMM2-CDG, also called CDG-Ia, resulting from mutations in phosphomannomutase 2 enzyme (*PMM2*) that converts mannose-6-phosphate to mannose-1-phosphate, the obligatory precursor for GDP-mannose production and N-linked glycosylation [18]. Defects in the N-glycosylation pathway of PMM2-CDG patients cause hypoglycosylation of different glycoproteins, leading to a variety of symptoms affecting multiple systems. In PMM2-CDG subjects, myopathy has not been reported explicitly, although almost all CDGs are characterized by the presence of neurological involvement with muscular hypotonia, failure to thrive and severe psychomotor retardation [19].

Considering that the role of N-glycosylation in muscle differentiation process remains poorly characterized to date, the aim of this study was to analyse the effects of N-glycosylation inhibition by TUN treatment or PMM2-downregulation on C2C12 myoblasts and myotubes. Furthermore, based on the hypothesis that N-glycosylation inhibition may impair the functioning of key myogenic factors, lectin binding and IGF-1R pathway activation was also analysed.

Materials and Methods

Cell cultures

Wild type and CRISPR-PMM2 C2C12 mouse adherent myoblasts were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS), 2 mM glutamine, penicillin (100 U/mL) and streptomycin (100 µg/mL), and maintained in a 5% CO_2 atmosphere at 37 °C. To induce C2C12 myogenic differentiation, the C2C12 at 80–90% confluence were transferred to a differentiation medium containing 2% horse serum, as previously described [20]. The differentiation medium was maintained for 5-6 days. For TUN (cat. no. T7765; Sigma-Aldrich) treatment, C2C12 myoblasts were incubated with 0.01 µg/ml of TUN for 24 h and then in differentiation medium for the 5-6 days. For IGF-1R activation, after 48 h of serum-free culture in twelve-well plates, cells were stimulated with [100 ng/ml] IGF-1 (cat. no. 50437-MNAY, SinoBiological) for 1 h before lysis. We found this stimulation condition to be the most accurate to assess both AKT and ERK1/2 activation. C2C12 concentration and cell viability were determined by the LUNA-II™ Automated Cell Counter (Logos Biosystems, Twin Helix) with trypan blue staining. Cytotoxicity was calculated as follows: Cytotoxicity (%) = [(dead cell number/total cells) x 100].

Fusion index scoring

C2C12 cell cultures were washed with phosphate buffered saline (PBS), fixed with 100% methanol for 10 min at room temperature, and then washed again with PBS. Cells were then incubated with 1 mL (in a twelve well plate) of Giemsa's Azur-Eosin-Methylene Blue solution (PanReac AppliChem) diluted 1/15 in PBS for 10 min at room temperature as previously described [21]. Finally, cells were washed four times with distilled water before collecting six random images per well (20X objective) using a digital camera (ColorView, Soft Imaging Systems) adapted to an inverted microscope (Olympus CKX41). Myotubes can be identified by a dark purple color, while nuclei can be distinguished by a light blue stain. Fusion index (FI) was determined by counting the number of nuclei in differentiated myotubes using ImageJ software and expressed as a percentage of the total number of nuclei.

Cellular senescence quantification

Cellular senescence was evaluated via detection of the senescence-associated beta-galactosidase (SA-βgal) activity in wild type and CRISPR-PMM2 cells. Wild type cells were also treated with 600 µM of hydrogen peroxide (H_2O_2) for 2 h and served as positive control [22]. Myoblasts at 80–90% confluence were washed twice with PBS, fixed for 5 min at room temperature with 4% paraformaldehyde and washed again three times in PBS. Cells were then incubated overnight at 37°C without CO_2 with freshly prepared staining solution at pH 6 according to Chen et al. [22]. After the incubation, six random images per well (10X objective) were collected using a digital camera (ColorView, Soft Imaging Systems) adapted to an inverted microscope (Olympus CKX41). The senescent cells responding to the SA-βgal activity (identified by a blue ring usually around the nucleus) were manually counted with the ImageJ software and the ratio of senescent cells was calculated.

RNA extraction, cDNA synthesis, and qRT-PCR

Total RNA was extracted and purified using the Omega Bio-Tek E.Z.N.A.™ Total RNA kit (VWR International) according to the manufacturer's instructions. The amount of RNA was assessed with SpectraMax® QuickDrop™ Micro-Volume Spectrophotometer (Molecular Devices, CaRli biotec) and the complementary DNA was synthesized from 500 ng of total RNA using Takara PrimeScript™RT Master Mix (Takara Bio Inc., Diatech Lab Line Srl). Subsequently, quantitative real-time PCR was performed with 2 µl of cDNA and 300 nM of each primer in the Applied Biosystems StepOnePlus™ Real-Time PCR System using SYBR Select Master Mix (Applied Biosystems). The real-time PCR conditions were as follows: 50°C for 2 min, 95°C for 2 min followed by 40 cycles of three-steps at 95°C for 15 s, 60°C for 15 s and 72°C for 30 s. The relative mRNA expression of target genes was normalized to *GAPDH* internal control. The genes of interest and the sequence of the specific primer used in real-time RT-PCR quantification are listed in the supplementary data Table S1.

Cell ELISA assay

TUN-treated C2C12 cells or CRISPR-PMM2 were seeded in a 96 well plate at a density of 1.0 x 10^4 cells/well. After removing the culture medium, the cells were washed with PBS and then 200 µl of 4% paraformaldehyde was added to each well and incubated for 15 min at room temperature. After permeabilization (PBS plus 0.1% Triton X-100), the wells were blocked at room temperature for 2h by adding 4% BSA solution. Then, 200 µl of primary antibody IGF-1R solution (1:2000; n. 3027 Cell Signaling Technology) was added to each well and incubated overnight at 4 °C with low speed shaking. After three times of washing, 100 µl of secondary antibody solution was added to the wells and incubated for 2h a room temperature. The procedure was continued with 4 times of washing. The color development was initiated with 100 µl of TMB solution for 20 min. Ultimately, the reaction was stopped by adding 50 µl of sulphuric acid (0.15 M) to each well.

Immunoblotting and Lectin Blotting

Cells were lysed adding 40-60 µl of lysis buffer containing: 20 mM HEPES (pH 7.9), 25% v/v glycerol, 0.42 M NaCl, 0.2 mM EDTA, 1.5 mM $MgCl_2$, 0.5% v/v Nonidet P-40, 1 mM DTT, 1 mM Naf, 1 mM Na_3VO_4, and 1X complete protease inhibitor cocktail (Roche Diagnostics). The lysates were frozen and thawed twice and clarified by centrifugation at 12000 rpm for 10 min at 4 °C. Protein concentration in each sample was determined using the Bradford colorimetric assay (Bio-Rad Laboratories) and the DU-640 UV Spectrophotometer (Beckman Coulter). The protein samples (20µg total proteins) were electrophoresed through 10% SDS-PAGE, and then transferred to nitrocellulose or PVDF membranes (Bio-Rad Laboratories) for immunoblotting. Primary antibodies against IGF-1 Receptor β (1:2000; n. 3027 Cell Signaling Technology), phospho-AKT (Ser473) (1:2000; n. 9271 Cell Signaling Technology), AKT (1:2000; n. 9272 Cell Signaling Technology), phospho-p44/42 (ERK1/2) (1:2000; n. 9101 Cell Signaling Technology), p44/42 (ERK1/2) (1:2000; n. 9102 Cell Signaling Technology), PCNA (1:5000 Millipore MAB 424R) and MF20 (1:500, DSHB) were incubated overnight at 4 °C. For lectin blotting, membranes were probed with biotinylated Concanavalin A (ConA, 1:1000), *Phaseolus vulgaris* leucoagglutinin (PHA-L, 1:200) and *Aleuria aurantia* (AAL, 1:400) lectins (Vector laboratories, D.B.A. Italia) at room temperature while shaking

for 1 h. After washes, the membranes were incubated with the appropriate horseradish peroxidase (HRP)-conjugated secondary antibody (Bio-Rad Laboratories) at room temperature for 1 h and were then washed three times. Blots were developed using Clarity Western ECL Substrate (Bio-Rad Laboratories) and quantified using the ChemiDoc MP (Bio-Rad Laboratories) equipped with Image Lab software.

siRNA and CRISPR

Transient cell transfection with 25 nM of Silencer Select® PMM2 siRNA (4390771, assay ID s79427; Thermo scientific) or nonsense scrambled siRNA (4390843; Thermo scientific) was carried out with the TransIT-X2® Transfection Reagent (Mirus Bio, TEMA ricerca) according to the manufacturer's instructions. Transfection and gene knockdown efficiency were assessed by qRT-PCR. For CRISPR/Cas9 silencing, 5.0×10^4 C2C12 cells/well were seeded in twelve well plates and transfected with 1 µg of the PMM2 Double Nickase plasmid (cat. no. sc-424790-NIC; Santa Cruz Biotechnology, D.B.A. Italia) or the corresponding Double Nickase Control plasmid (cat. no. sc-437281; Santa Cruz Biotechnology, D.B.A. Italia) using the TransIT-X2® Transfection Reagent (Mirus Bio, TEMA ricerca) for 48 h. Following the manufacturer's protocol, selection was performed with puromycin (cat. no. P9620; Sigma-Aldrich), and clones were selected and analyzed using qRT-PCR and PMM2 enzymatic activity assay.

The PMM2 enzymatic activity assay

Wild type C2C12 cells and CRISPR-PMM2 cells were washed twice in PBS, scraped, pelleted by centrifugation and resuspended in lysis buffer (20 mM Hepes, 25 mM KCl, 1 mM dithiothreitol and 1x protease inhibitor mixture) and then briefly sonicated and incubated at 4°C for 15 min. The cell lysates were then centrifuged at 15000 rpm for 10 min at 4°C to remove cellular debris. The protein concentration of each freshly prepared cell lysate was determined with the Bradford Reagent (Sigma-Aldrich). PMM2 enzymatic activity was assayed as reported by Van Schaftingen (1995) with slight modifications. Briefly, PMM2 activity was measured at 32°C in Hepes 20 mM pH 7.5 containing $MgCl_2$ 5 mM and $NADP^+$ 0.25 mM, with 0.3 mM Mannose-1-P as the substrate, 1µM Glc-1,6-P_2 as the activator, 2.8 U/ml

glucose 6-phosphate dehydrogenase, 3.7 U/ml phosphoglucose isomerase and 3.9 µg/ml phosphomannose isomerase. The reaction was initiated by adding 50 µg/mL of cell lysates. The activity was monitored spectrophotometrically for 60 min at 340 nm, recording the reduction of $NADP^+$ to NADPH. One unit is the amount of enzyme that catalyzes the conversion of 1 µmol of substrate per min under this condition.

Statistical analysis

Data are represented as mean ± SEM of at least three independent experiments. Statistical analyses were performed using the *t*-test or two-way ANOVA as appropriate, followed by Bonferroni's multiple comparison post hoc tests. A p-value < 0.05 was considered statistically significant.

Results

TUN treatment and C2C12 myoblast differentiation

C2C12 myoblasts remained viable for up to 24-48 h when incubated in the lowest concentration (0.01 µg/ml) of TUN. The cells appeared healthy by phase-contrast microscopy and retained the capacity to exclude trypan blue (cytotoxicity CTR= 1.7% ± 1.8%; TUN 0.01 µg/ml= 1.8% ± 0.5; p=0.92). At the highest concentration (0.05 µg/ml), distinct cytotoxic effects were apparent, and the cells became highly vacuolated and started to detach from the dish surfaces (cytotoxicity CTR= 1.7% ± 1.8%; TUN 0.05 µg/ml= 50.5% ± 3.3%; $p < 0.0001$). A TUN concentration of 0.01 µg/ml was therefore used for the subsequent experiments. To evaluate the effect of TUN on myotubes formation, TUN was added to proliferating C2C12 myoblasts 24 h before and during 5-6 days of differentiation. Representative photomicrographs from 5-6 days of differentiation of CTR C2C12 myoblasts into myotubes showed long multinucleated cells with a fusion index of 53 ± 8 % (Fig. 1A) and a mean value of 6.3 ± 1.1 nuclei per cell. The TUN-treated myoblasts continued to proliferate until the cells reached confluent monolayer; however, myoblast fusion was markedly inhibited (Fig. 1A). We then examined the mRNA levels of the myogenic markers *Ccnd1*, *MyoD*, *Myogenin* and *Mrf4* during differentiation (Fig. 1B). As expected, *Ccnd1* was expressed in CTR myoblasts and downregulated in myotubes, while *MyoD* and *Myogenin* showed the opposite trend. TUN-treated myoblasts failed to downregulate *Ccnd1* mRNA after induction of differentiation and its expression remained higher in TUN-treated myotubes than it did in control myotubes. Conversely, TUN-treated myoblasts expressed lower levels of *MyoD*, *Myogenin* and *Mrf4* mRNAs than CTR and the level of *MyoD* and *Myogenin* also remained lower also in TUN-treated myotubes than in CTR. Finally, at the protein level, we found that the TUN treatment slightly increased PCNA protein expression in myoblasts while the protein expression level of myosin heavy chain (MF20) in myotubes markedly decreased (Fig. 1C).

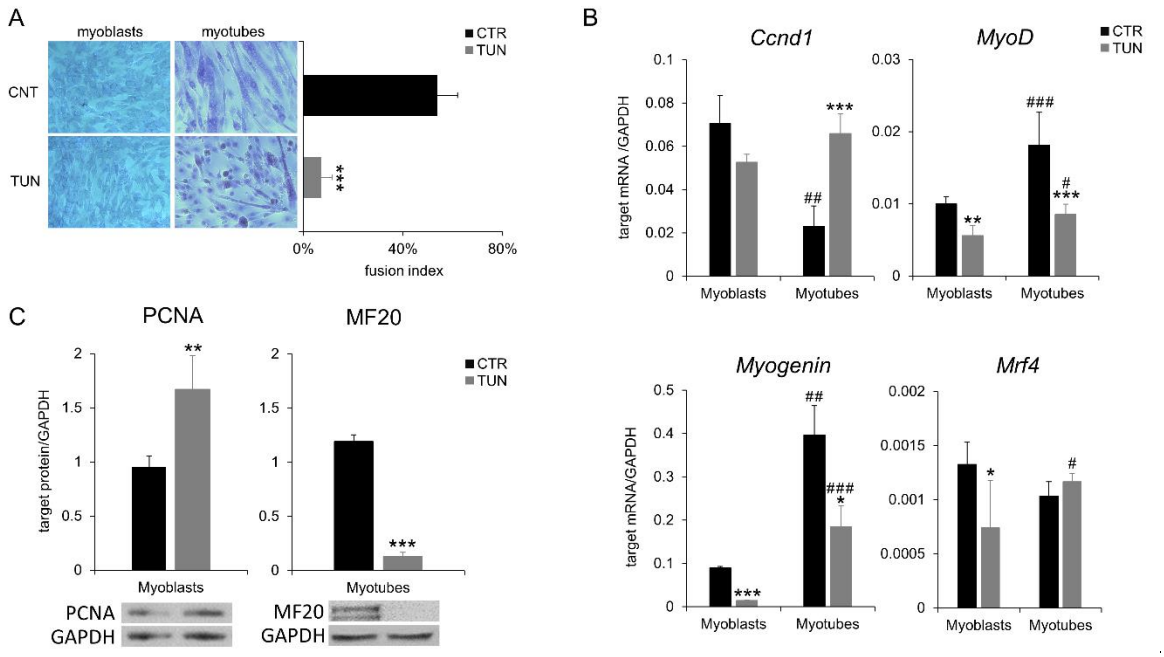


Figure 1. TUN treatment and C2C12 differentiation. Effects of TUN treatment (0.01 µg/ml) on myotubes formation (A); *Ccnd1*, *MyoD*, *Myogenin* and *Mrf4* mRNA expression (B); and PCNA and MF20 protein expression (C). * Significantly different compared to CTR; #Significantly different compared to myoblasts; * and # p<0.05; ## and ** p<0.001; *** and ### p<0.0001.

PMM2 siRNA or PMM2 CRISPR downregulation and C2C12 myoblast differentiation

We subsequently analyzed the effect of PMM2 downregulation on C2C12 differentiation. C2C12 myoblasts were transiently transfected with *PMM2* siRNA for 24 h, and for the following 5-6 days in the differentiation medium. Expression analysis of *PMM2*-siRNA-transfected cells showed a *PMM2* mRNA reduction of 4.5 ± 2.1 in myoblasts and 5.7 ± 1.7 in myotubes obtained 2-3 days after the induction of differentiation compared to CTR ($p<0.0001$). On the contrary, the *PMM2* mRNA returned to the baseline level in siRNA-treated myotubes after 5-6 days of differentiation ($p = 0.43$), probably due to a low transfection efficiency of siRNA in fully formed myotubes and/or reduced viability of the *PMM2*-siRNA-transfected cells. Thus, the effects of *PMM2* downregulation on C2C12 differentiation were examined in myoblasts and myotubes harvested after 2-3 days of differentiation. Cell fusion and the formation of small myotubes were evident 2-3 days after induction (Fig. 2A) and, as expected, the myogenic index was lower than that which was found in myotubes differentiated for 5-6 days (fusion index of 36 ± 1 %) (compare Fig. 1A and 2A). *PMM2*-siRNA-transfected myotubes showed a myogenic

index of 17 ± 2 %, and hence a slight reduction compared to CTR-siRNA-transfected myotubes. As we had observed in TUN-treated C2C12 cells, the mRNA knockdown of *PMM2* increased *Ccnd1* expression in myoblasts and reduced *MyoD*, *Myogenin*, and *Mrf4* mRNA levels in myotubes (Fig. 2B). Accordingly, *PMM2* gene downregulation also decreased MF20 protein expression in myotubes more than it did in mock cells (Fig. 2C).

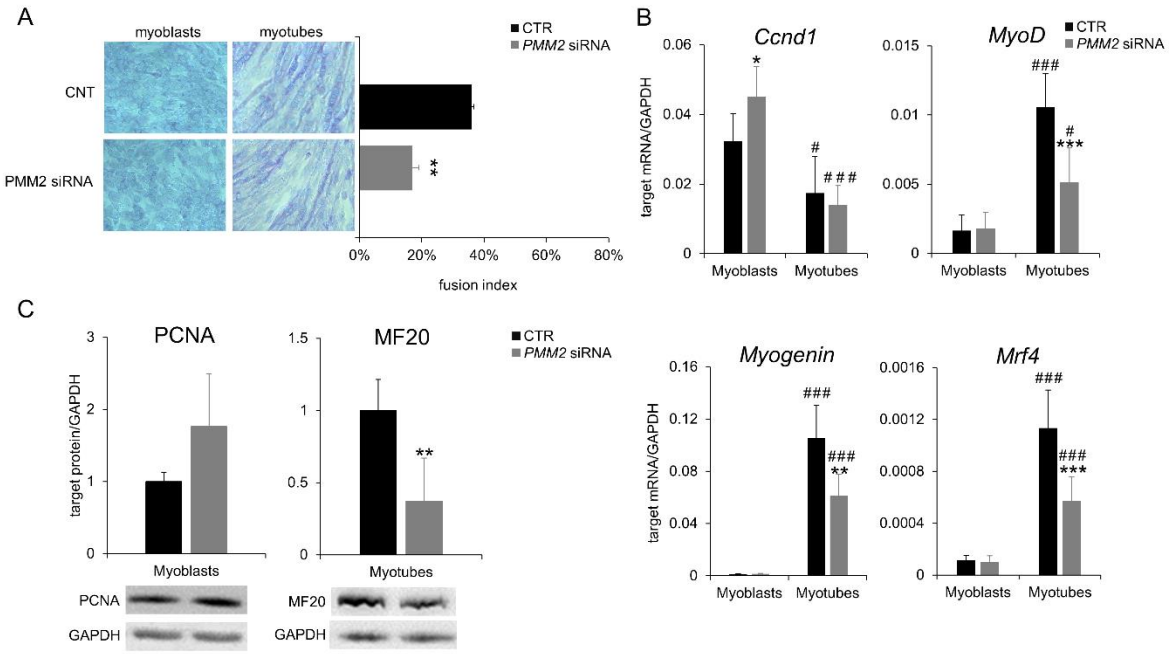


Figure 2. PMM2 siRNA transfection and C2C12 differentiation. Effects of siRNA-induced PMM2 downregulation on myotubes formation (A); *Ccnd1*, *MyoD*, *Myogenin* and *Mrf4* mRNA expression (B); and PCNA and MF20 protein expression (C). Myotubes were harvested after 2-3 days of differentiation in order to maintain a sufficiently PMM2-downregulation. *Significantly different compared to CTR; #Significantly different compared to myoblasts; * and # p<0.05; ## and ** p<0.001; *** and ### p<0.0001.

In order to confirm that downregulation of *PMM2* was responsible for the inhibition of C2C12 differentiation, we also tried to knock out *PMM2* in C2C12 by using a commercially available CRISPR/Cas9 Double-Nickase plasmid against *PMM2*, but we were unable to obtain complete PMM2 knockouts. However, we did obtain clones showing a marked reduction in *PMM2* mRNA compared to the corresponding CTR CRISPR/Cas9 clones (10.7± 2.8 fold reduction; $p < 0.0001$). Accordingly, CRISPR-PMM2 cells showed an almost complete inhibition of PMM2 enzyme activity (CTR = 1.64 ± 0.4 mU/mg of proteins; CRISPR-PMM2 = 0.34±0.25 mU/mg of proteins; $p < 0.01$). Moreover, unlike *PMM2*-siRNA downregulation, CRISPR-PMM2 cells showed a long-term reduction in *PMM2* mRNA, including for cells

harvested 5-6 days after induction of differentiation (16.9 ± 4.0 fold reduction; $p < 0.0001$). Thus, all the experiments with CRISPR-PMM2 cells took place after 5-6 days of differentiation when control cells showed fully formed myotubes (Fig 3A). As had occurred after TUN treatment and siRNA-*PMM2* downregulation, CRISPR-PMM2 cells failed to form myotubes, showing a marked reduction in the myogenic index (Fig. 3A). Accordingly, we found an increased *Ccnd1* mRNA level in CRISPR-PMM2 myoblasts and a marked reduction in *MyoD*, *Myogenin*, and *Mrf4* mRNA (Fig. 3B) and MF20 protein expression (Fig. 3C) in myotubes. The analysis of SA-βgal activity showed no difference between CRISPR-PMM2 and CTR cells (% senescence CTT= 2.5 ± 2.5; CRISPR-PMM2= 2.2 ± 2.1; p= 0.85), suggesting that the impaired differentiation is not attributable to a senescent condition due to CRISPR-PMM2 clone selections.

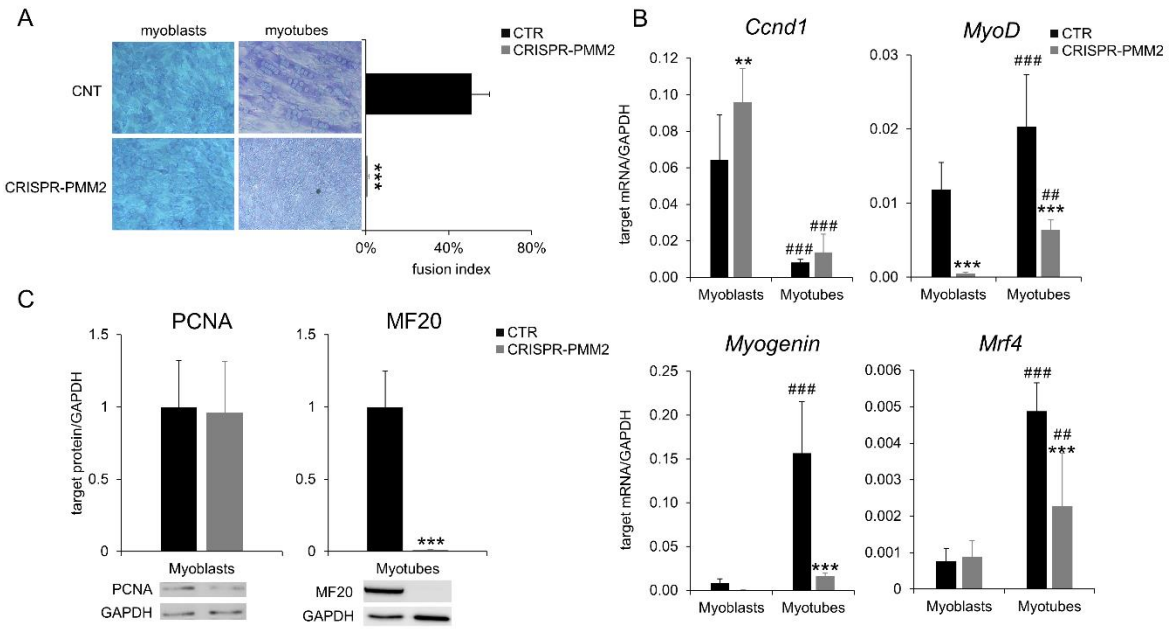


Figure 3. CRISPR/Cas9 PMM2 downregulation and C2C12 differentiation. Effects of PMM2 knockdown on myotubes formation (A); *Ccnd1*, *MyoD*, *Myogenin* and *Mrf4* mRNA expression (B); and PCNA and MF20 protein expression (C). *Significantly different compared to CTR; #Significantly different compared to myoblasts; ** and ## p<0.001; *** and ### p<0.0001.

Effects of TUN treatment and CRISPR/Cas9 PMM2 downregulation on lectin binding

Subsequently, we used Concanavalin A (ConA), *Phaseolus vulgaris* leucoagglutinin (PHA-L) and *Aleuria aurantia* (AAL) lectins to analyze glycosylation of C2C12 myoblasts and differentiated myotubes (Fig. 4). Con A and PHA-L recognize high mannose and complex type N-glycans, respectively, while AAL binds to fucose linked (α -1,6) to N-acetylglucosamine or to fucose linked (α -1,3) to N-acetyllactosamine. As shown in Figure 4, we found a significant increase in Con A and PHA-L reactivity in myotubes compared to myoblasts, showing a rise in high mannose and complex N-glycans following differentiation. Moreover, we observed increased levels of fucosylated proteins (i.e. increased AAL reactivity) in myotubes compared to myoblasts. Importantly, the lectin binding profile of TUN-treated and CRISPR-PMM2 myotubes markedly differed compared to CTR myotubes, showing a general reduction of N-glycosylation.

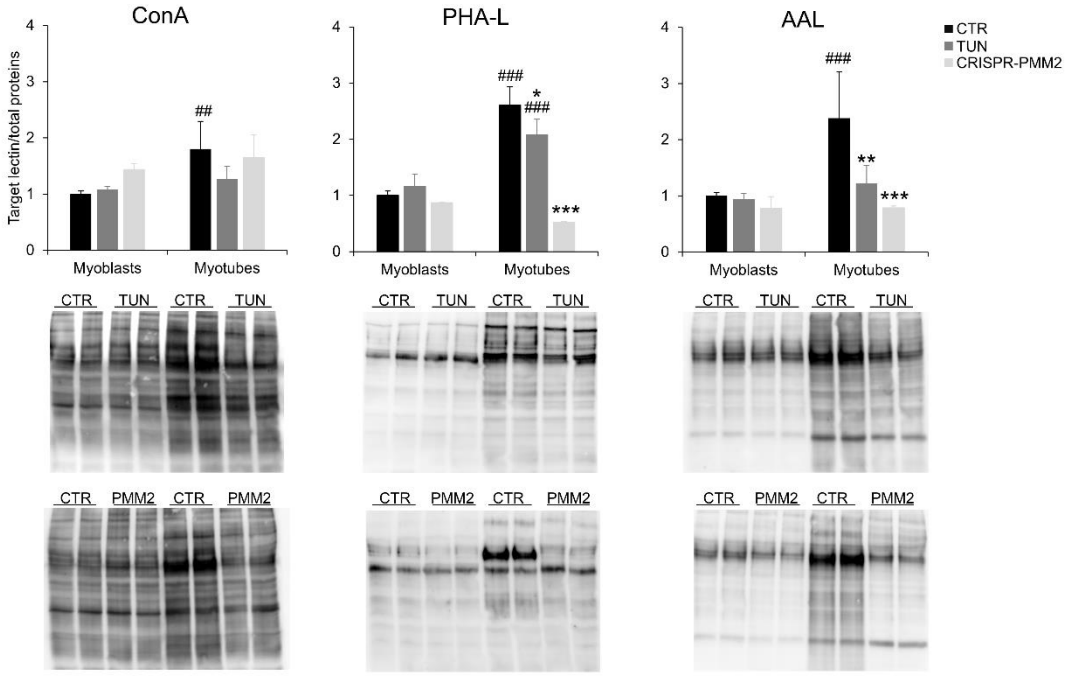


Figure 4. Lectin binding profile of TUN-treated and CRISPR-PMM2 cells. Analysis of lectin binding in C2C12 treated with TUN or in CRISPR-PMM2 myoblasts and myotubes. ConA, Concanavalin A; PHA-L, *Phaseolus vulgaris* leucoagglutinin; AAL, *Aleuria aurantia* Lectin. Con A and PHA-L recognize high mannose and complex type N-glycans, respectively, while AAL binds to fucose linked (α -1,6) to N-acetylglucosamine or to fucose linked (α -1,3) to N-acetyllactosamine. *Significantly different compared to CTR; #Significantly different compared to myoblasts; * p<0.05; ** and ## p<0.01, *** and ### p<0.0001.

Effects of glycosylation inhibition on C2C12 IGF-1R expression and IGF-1 pathway activation

As shown in Figure 5, TUN treatment (Fig. 5A) or *PMM2* gene downregulation by siRNA (Fig. 5B) or CRISPR/Cas9 (Fig. 5C) decreased the IGF-1R protein level in myoblasts and myotubes. Accordingly, using IGF-1R cell-ELISA quantification we found that both TUN-treated and CRISPR-PMM2 cells had reduced total IGF-1R level (Fig. 5D). Moreover, the IGF-1-induced activation of AKT was reduced in TUN-treated myoblasts and slightly, but not significantly, also in CRISPR-PMM2 cells (Fig. 5E). Finally, we analysed the *IGF-1* mRNA level and protein secretion during C2C12 differentiation and after N-glycosylation inhibition by TUN (Fig. 6A and 6B) or CRISPR-PMM2 (Fig. 6C and 6D). As expected, the *IGF-1* mRNA level and protein secretion were higher in myotubes than in myoblasts. Unexpectedly, we found a higher level of *IGF-1* mRNA expression in TUN-treated myotubes (Fig. 6A) and CRISPR-PMM2 myoblasts and myotubes than in CTR (Fig. 6C). However, the increased *IGF-1* mRNA found in TUN-treated myotubes was not associated with higher IGF-1 protein secretion, which, on the contrary, decreased slightly (Fig. 6B). Conversely, both *IGF-1* mRNA and protein level increased in CRISPR-PMM2 cells (Fig. 6D).

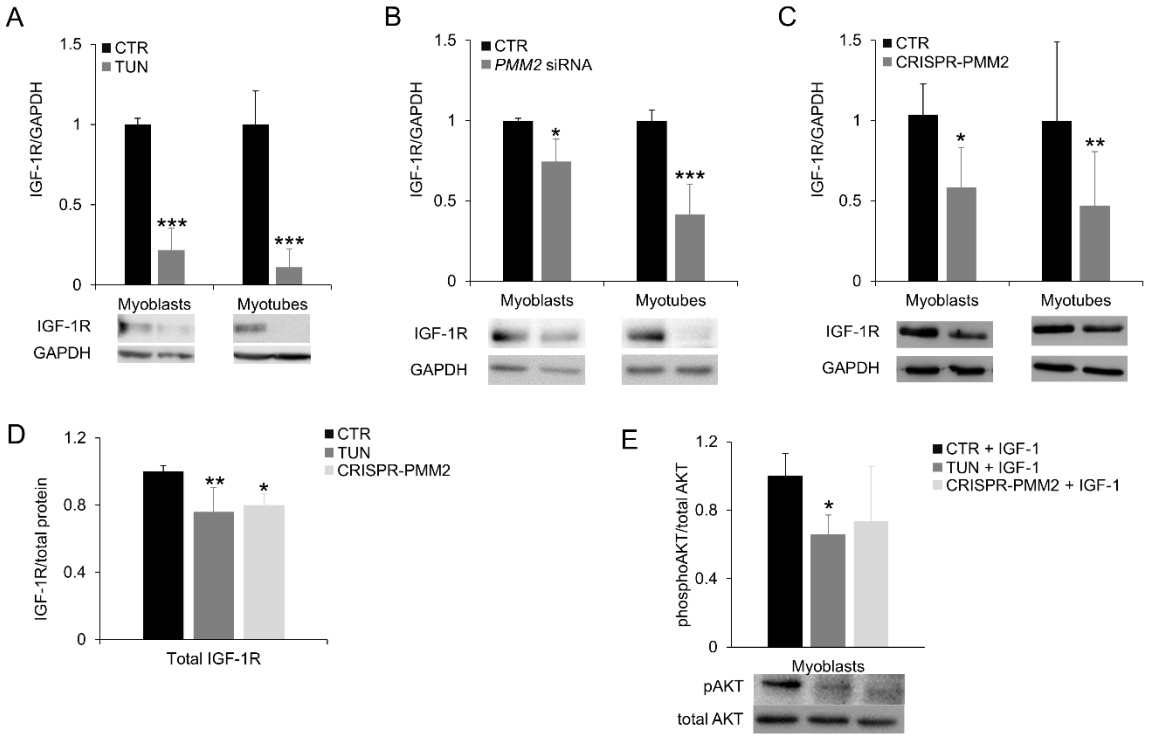


Figure 5. IGF-1R expression and IGF-1-induced activation in TUN-treated C2C12 cells and after *PMM2* gene downregulation by siRNA or CRISPR/Cas9. Effect of TUN treatment (A), siRNA- (B) or CRISPR/Cas9-induced (C) PMM2 downregulation on IGF-1R protein expression measured by western blotting or cell ELISA (D). IGF-1-induced AKT phosphorylation in TUN-treated myoblasts or CRISPR-PMM2 cells (E). *Significantly different compared to CTR; #Significantly different compared to myoblasts; * $p<0.05$, ** $p<0.01$, *** $p<0.0001$.

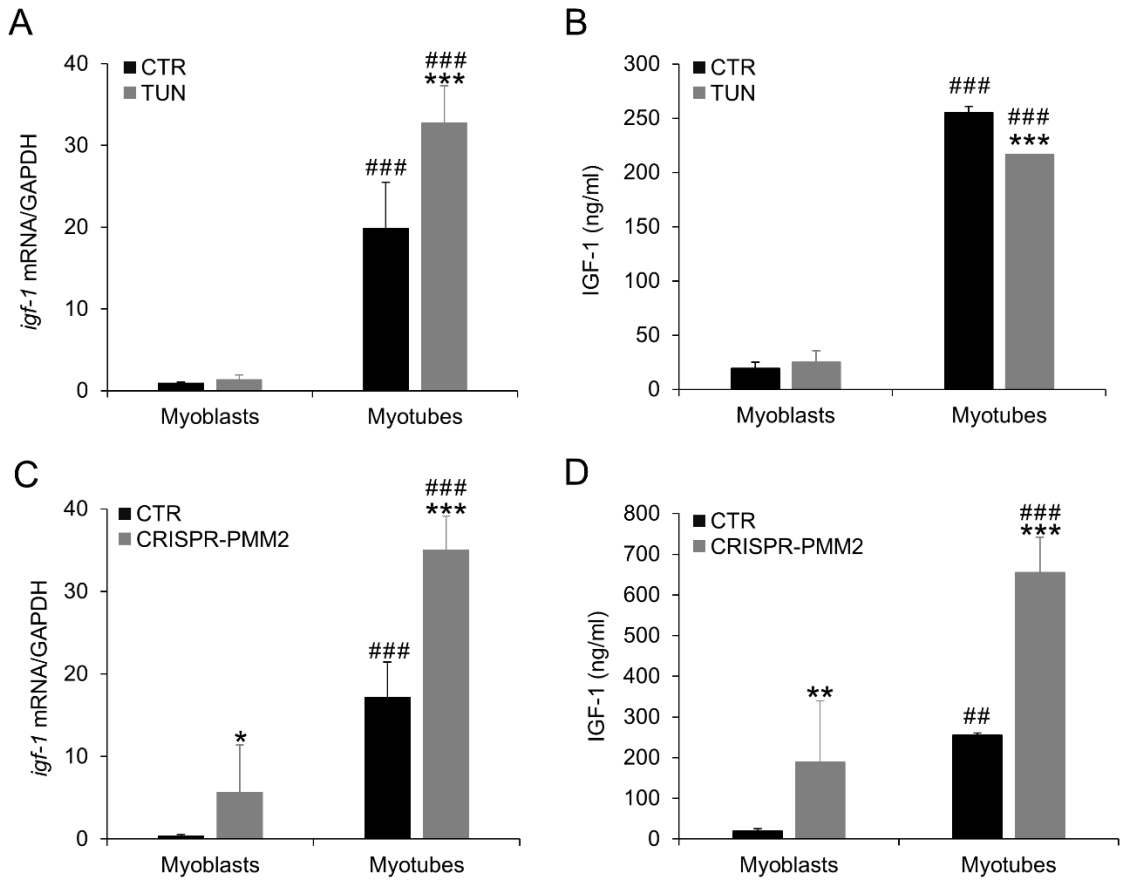


Figure 6. IGF-1 expression in TUN-treated and CRISPR-PMM2 cells. Effect of TUN treatment on *IGF-1* mRNA expression level (A) and protein secretion (B). Effect of CRISPR/Cas9-induced PMM2 downregulation on *IGF-1* mRNA expression level (C) and protein secretion (D). *Significantly different compared to CTR; #Significantly different compared to myoblasts; * $p<0.05$, ** and ## $p<0.001$, *** and ### $p<0.0001$.

Discussion

In the present study, we used two different approaches based on TUN-treatment and PMM2 downregulation to investigate the effects of N-glycosylation inhibition on C2C12 myoblast differentiation. TUN, as an inhibitor of N-acetylglucosamine transferases, blocks the N-glycosylation pathway at the first step, resulting in the synthesis of glycoproteins deficient in N-linked glycans [23]. Here, we found that chronic treatment with TUN (concentration 0.01 μg/ml) markedly reduced C2C12 differentiation. In particular, the differentiation index was markedly reduced in TUN-treated C2C12 cells and the coordinated temporal expression of myogenic regulator genes *Ccnd1*, *MyoD*, *Myogenin* and *Mrf4*, was also disrupted. Accordingly, the levels of the proliferation marker PCNA increased in TUN-treated myoblasts and those of the myogenic marker MF20 decreased in TUN-treated myotubes compared to CTR. To confirm the efficacy of TUN-treatment on N-glycosylation inhibition, we determined the binding of ConA, PHA-L and ALL lectins. In untreated cells, we observed increased binding of ConA, PHA-L and ALL lectins in myotubes compared to myoblasts. This finding is in agreement with other studies, suggesting that specific changes in muscle cells glycosylation occur following differentiation [14,15]. Given the specificities of these lectins, glycoproteins on the myotube surface bear higher (ConA) and complex (PHA-L) mannose N-glycans and fucosylated (ALL) residues compared to undifferentiated myoblasts. Interestingly, TUN-treated myotubes showed a marked reduction in ConA, PHA-L and ALL binding compared to CTR cells, with a pattern of lectin-binding more similar to that of myoblasts. It has previously been shown that doses from 1.0 to 10 μg/ml of TUN are toxic in C2C12 cells and induce ER stress pathways activation, inhibit mitosis and increase apoptosis in C2C12 [24, 25]. However, the effects of high-dose of TUN on myoblasts differentiation are still unclear [16, 25]. Here, we found that low doses of TUN (i.e. concentrations ≤ 0.01 μg/ml) did not have toxic effects on C2C12 myoblasts but rather increased the myoblast proliferation marker PCNA. Our findings suggest that the inhibition of C2C12 differentiation induced by non-toxic doses of TUN was mainly due to reduced protein N-glycosylation instead of secondary effects connected with TUN-treatment (i.e. increased ER stress and apoptosis), although further studies are required to verify this point.

We subsequently analysed the effect of the downregulation of PMM2, an enzyme involved in the initial steps of N-glycosylation, on C2C12 differentiation. PMM2 is the most frequently mutated enzyme among the CDG subtypes and CDG patients have hypomorphic alleles because complete loss of PMM2 function is lethal [26]. As occurred with the TUN treatment, it was found that PMM2 downregulation by siRNA or CRISPR/Cas9 system markedly decreased C2C12 differentiation markers. Moreover, we also found a reduction in ConA, PHA-L and ALL lectin binding in CRISPR-PMM2 myotubes compared to CTR cells. PMM2-CDG is associated with a broad range of clinical symptoms although muscular hypotonia is a common feature in these patients. PMM2-CDG patients seldom have muscle biopsies performed and few case reports show slight myopathic alterations with fiber size variations and myofibrillar disarrays [27] and mildly elevated creatine kinase (CK) levels [28]. Interestingly, a recent study showed that muscle-specific PMM2 knockdown caused qualitative movement defects in a *Drosophila* PMM2-CDG model [29]. Thus, although it is likely that several different mechanisms may be involved in the hypotonia found in most PMM2-CDG patients, defects in skeletal muscle could be a contributing factor.

We then investigated whether the decreased C2C12 differentiation found after N-glycosylation inhibition corresponded to a decrease in IGF-1R signaling pathway activation. IGF-1R plays an important role in normal muscle cells growth and differentiation and disrupted IGF-1R signaling may lead to impaired skeletal muscle development [30]. Recently, it has been shown that both IGF-1 and IGF-1R pro-hormones are glycosylated proteins [5-8]. Notably, posttranslational N-glycosylation modification plays an essential function in protein folding, stability, and intracellular trafficking of both IGF-1 and IGF-1R prohormones [5,8]. Here we found that both TUN treatment and *PMM2* gene downregulation by siRNA or CRISPR decreased IGF-1R level in C2C12 cells. We also observed a marked reduction in AKT activation in TUN-treated cells and a slight, although not significant, decrease in CRISPR-PMM2. Subsequently, we investigated the regulation of the main activator of IGF-1R, namely IGF-1. *IGF-1* mRNA and protein secretion increased with myoblast differentiation. Surprisingly, both TUN-treated myotubes and CRISPR-PMM2 myoblasts and myotubes showed an increase in *IGF-1* mRNA levels compared to CTR cells. We hypothesize that the increase in *IGF-1* mRNA levels might be a compensatory consequence of IGF-1R suppression due to N-glycosylation

inhibition [31]. Another possibility is that N-glycosylation disrupts the tightly regulated expression pattern of IGF-1 thus favoring the proliferative rather than the differentiative actions of this growth factor [32,33]. It is important to note that the effect of TUN and CRISPR-PMM2 knockdown on IGF-1 protein secretion differed markedly. Indeed, despite the increased *IGF-1* mRNA level in TUN-treated myotubes, the protein secretion decreased compared to CTR cells. On the other hand, both *IGF-1* mRNA and protein secretion increased in CRISPR-PMM2 cells. The results obtained in TUN-treated cells are consistent with our previous findings, which showed that IGF-1 secretion was extremely sensitive to N-glycosylation inhibition [8]. The discrepancies between the effects obtained with TUN-treated and CRISPR-PMM2 cells on IGF-1 protein secretion might be accounted for by the presence of a heterogeneous cell population in our CRISPR-PMM2 downregulation model. We have attempted to knock out PMM2 in C2C12 cells using the CRISPR-Cas9 Double-Nickase-mediated system but were not able to obtain complete PMM2 knockout clones. Nevertheless, in a heterogeneous population of stable knockdown PMM2 cells, we ensured that *PMM2* transcripts and enzyme activity were sufficiently suppressed. Further studies conducted with PMM2 knockout cells and/or primary myoblasts isolated from PMM2-patients are needed to confirm the findings of our study. The presence of a heterogeneous population of stable knockdown PMM2 cells and the increase in IGF-1 secretion found in CRISPR-PMM2 cells also likely account for the smaller decrease in AKT activation found in these cells compared to those treated with TUN.

In conclusion, these results highlight the important role played by N-glycosylation in the regulation of myoblast differentiation. Our findings showed that both TUN-treatment and PMM2 downregulation inhibited myoblast fusion and interfered with the early stage of the myogenic program. N-glycosylation inhibition also affected the IGF-1R pathway reducing the IGF-1R expression, inhibiting the AKT activation and deregulating *IGF-1* mRNA expression and protein secretion. Overall, these results offer new insights into the impairment of myogenic differentiation capacity in the context of pathological N-glycosylation deficiencies.

Supplementary Information

Table S1. Primers used in real-time RT-PCR quantification.

Gene	Primer forward (5'-3')	Primer reverse (5'-3')
Ccnd1	CTTCCTCTCCAAAATGCCAG	TGGAGGGTGGGTTGGAAAT
MyoD	TTCTTCACCACTCCTCTGAC	GCCGTGAGAGTCGTCTTAT
Myogenin	CAACCCAGGAGATCATTG	CATATCCTCCACCGTGATGC
MRF4	GTGGCCAAGTGTTTCGGAT	AAAGGCGCTGAAGACTGC
PMM2	TTGCCTGAGACACCTGGAAC	AGATCCTGCGTGTGTCTTCG
IGF-1	GCTAAGGCTCCAGCATTCG	TCCGGAAGCAACACTCATCC
GAPDH	GGCAAATTCAACGGCACAGT	CGCTCCTGGAAGATGGTGAT

Note: *Ccnd1*, cyclin D1; *MyoD*, myogenic differentiation 1; *MRF4*, myogenic factor 6; *Myf5*, myogenic factor 5; *PMM2*, phosphomannomutase 2; *IGF-1*, insulin-like growth factor-1; *GAPDH*, gliceraldeyde-3-phosphate dehydrogenase.

References

[1] C. F. Bentzinger, Y. X. Wang, and M. A. Rudnicki, “Building Muscle: Molecular Regulation of Myogenesis,” Cold Spring Harb Perspect Biol., vol. 4, no. 2, p. a008342, 2012.

[2] T. Kislinger, A. O. Gramolini, Y. Pan, K. Rahman, D. H. MacLennan, and A. Emili, “Proteome dynamics during C2C12 myoblast differentiation,” Mol. Cell. Proteomics, vol. 4, no. 7, pp. 887–901, 2005.

[3] T. Yoshida and P. Delafontaine, “Mechanisms of IGF-1-Mediated Regulation of Skeletal Muscle Hypertrophy and Atrophy,” Cells, vol. 9, no. 9, pp. 1–25, 2020.

[4] D. J. Glass, “PI3 kinase regulation of skeletal muscle hypertrophy and atrophy.,” Curr. Top. Microbiol. Immunol., no. 346, pp. 267–278, 2010.

[5] E. Klaver et al., “Selective inhibition of N-linked glycosylation impairs receptor tyrosine kinase processing,” DMM Dis. Model. Mech., vol. 12, no. 6, 2019.

[6] K. W. Siddals, E. Marshman, M. Westwood, and J. M. Gibson, “Abrogation of insulin-like growth factor-I (IGF-I) and insulin action by mevalonic acid depletion: Synergy between protein prenylation and receptor glycosylation pathways,” J. Biol. Chem., vol. 279, no. 37, pp. 38353–38359, 2004.

[7] H. M. Itkonen and I. G. Mills, “N-Linked Glycosylation Supports Cross-Talk between Receptor Tyrosine Kinases and Androgen Receptor,” PLoS One, vol. 8, no. 5, pp. 1–10, 2013.

[8] G. Annibalini, S. Contarelli, M. De Santi, R. Saltarelli, L. Di Patria, M. Guescini, A. Villarini, G. Brandi, V. Stocchi & E. Barbieri, “The intrinsically disordered E-domains regulate the IGF-1 prohormones stability, subcellular localisation and secretion,” Sci. Rep., vol. 8, no. 1, 2018.

[9] C. Vergé, A. Bouchatal, F. Chirat, Y. Guérardel, A. Maftah, and J. M. Petit, “Involvement of ST6Gal I-mediated α2,6 sialylation in myoblast proliferation and differentiation,” FEBS Open Bio, vol. 10, no. 1, pp. 56–69, 2020.

[10] M. Janot, A. Audfray, C. Loriol, A. Germot, A. Maftah, and F. Dupuy, “Glycogenome expression dynamics during mouse C2C12 myoblast differentiation suggests a sequential reorganization of membrane glycoconjugates,” BMC Genomics, vol. 10, p. 483, 2009.

[11] V. Grassot, A. Da Silva, J. Saliba, A. Maftah, F. Dupuy, and J. M. Petit, “Highlights of glycosylation and adhesion related genes involved in myogenesis,” BMC Genomics, vol. 15, no. 1, pp. 1–17, 2014.

[12] P. T. Martin and H. H. Freeze, “Glycobiology of neuromuscular disorders,” Glycobiology, vol. 13, no. 8, pp. 67–75, 2003.

[13] R. L. Gundry et al., “The mouse C2C12 myoblast cell surface N-linked glycoproteome: Identification, glycosite occupancy, and membrane orientation,” Mol. Cell. Proteomics, vol. 8, no. 11, pp. 2555–2569, 2009.

[14] B. J. McMorran, M. C. Miceli, and L. G. Baum, “Lectin-binding characterizes the healthy human skeletal muscle glycophenotype and identifies disease-specific changes in dystrophic muscle,” Glycobiology, vol. 27, no. 12, pp. 1134–1143, 2017.

[15] B. J. McMorran et al., “Differentiation-related glycan epitopes identify discrete domains of the muscle glycocalyx,” Glycobiology, vol. 26, no. 10, pp. 1120–1132, 2016.

[16] J. B. Parent and K. Olden, “Inhibition of fusion of embryonic muscle cells in culture by tunicamycin is prevented by leupeptin,” J. Supramol. Cell. Biochem., vol. 15, no. Suppl.5, pp. 199–204, 1981.

[17] S. R. Harris, “Congenital hypotonia: Clinical and developmental assessment,” Dev. Med. Child Neurol., vol. 50, no. 12, pp. 889–892, 2008.

[18] R. Péanne et al., “Congenital disorders of glycosylation (CDG): Quo vadis?,” Eur. J. Med. Genet., vol. 61, no. 11, pp. 643–663, 2018.

[19] P. V. Cabrera et al., “High throughput screening for compounds that alter muscle cell glycosylation identifies new role for N-glycans in regulating sarcolemmal protein abundance and laminin binding,” J. Biol. Chem., vol. 287, no. 27, pp. 22759–22770, 2012.

[20] P. Sestili, E. Barbieri, C. Martinelli, M. Battistelli, M. Guescini, L. Vallorani, L. Casadei, A. D’Emilio, E. Falcieri, G. Piccoli, D. Agostini, G. Annibalini, M. Paolillo, A.M. Gioacchini, and V. Stocchi, “Creatine supplementation prevents the inhibition of myogenic differentiation in oxidatively injured C2C12 murine myoblasts.,” Mol. Nutr. Food Res., vol. 53, no. 9, pp. 1187–1204, 2009.

[21] E. Archer-Lahlou, C. Lan, and R. T. Jagoe, “Physiological culture conditions alter myotube morphology and responses to atrophy treatments: implications for in vitro research on muscle wasting,” Physiol. Rep., vol. 6, no. 12, pp. 1–14, 2018.

[22] J.-H. Chen, S. E. Ozanne, C. N. Hales, “Methods of Cellular Senescence Induction Using Oxidative Stress,” Biol. Aging Springer Link, no. 371, pp. 179–189, 2007.

[23] L. Lehle and W. Tanner, “The specific site of tunicamycin inhibition in the formation of dolichol-bound N-acetylglucosamine derivatives,” FEBS Lett., vol. 71, no. 1, pp. 167–170, 1976.

[24] A. P. Lightfoot, R. S. Morgan, J. E. Parkes, A. Thoma, and A. Leslie, “17-(Allylamino)-17demethoxygeldanamycin reduces Endoplasmic Reticulum (ER) stress-induced mitochondrial dysfunction in C2C12 myotubes”, 2018.

[25] K. Nakanishi, N. Dohmae, and N. Morishima, “Endoplasmic reticulum stress increases myofiber formation in vitro,” FASEB J., vol. 21, no. 11, pp. 2994–3003, 2007.

[26] C. Thiel, T. Lübke, G. Matthijs, K. von Figura, and C. Körner, “Targeted Disruption of the Mouse Phosphomannomutase 2 Gene Causes Early Embryonic Lethality,” Mol. Cell. Biol., vol. 26, no. 15, pp. 5615–5620, 2006.

[27] E. Aronica et al., “Congenital disorder of glycosylation type Ia: A clinicopathological report of a newborn infant with cerebellar pathology,” Acta Neuropathol., vol. 109, no. 4, pp. 433–442, 2005.

[28] R. H. Wu et al., “Atrial septal defect in a patient with congenital disorder of glycosylation type 1a: A case report,” J. Med. Case Rep., vol. 12, no. 1, pp. 1–7, 2018.

[29] W. M. Parkinson, B. M. Dookwah, M. L. Dear, C. L. Gatto, K. Aoki, M. Tiemeyer, K. Broadie, “Synaptic roles for phosphomannomutase type 2 in a new Drosophila congenital disorder of glycosylation disease model,” Dis. Model. Mech, vol. 9(5):513-27, 2016.

[30] S. S. Ahmad, K. Ahmad, E. J. Lee, Y.-H. Lee and I. Choi, “Implications of Insulin-Like Growth Factor-1 in Skeletal Muscle and Various Diseases,” Cells, vol. 9, no. 1773, 2020.

[31] H. Zhang, A. M. Pelzer, D. T. Kiang, and D. Yee, “Down-regulation of type I insulin-like growth factor receptor increases sensitivity of breast cancer cells to insulin,” Cancer Res., vol. 67, no. 1, pp. 391–397, 2007.

[32] S. A. Coolican, D. S. Samuel, D. Z. Ewton, F. J. McWade, and J. R. Florini, “The mitogenic and myogenic actions of insulin-like growth factors utilize distinct signaling pathways,” J. Biol. Chem., vol. 272, no. 10, pp. 6653–6662, 1997.

[33] F. Hakuno et al., “Constitutive expression of insulin receptor substrate (irs)-1 inhibits myogenic differentiation through nuclear exclusion of foxo1 in l6 myoblasts,” PLoS One, vol. 6, no. 10, 2011.

CHAPTER 3

Defective insulin-like growth factor 1 (IGF-1) pro-hormone N-glycosylation and reduced IGF-1 receptor signalling activation in congenital disorders of glycosylation

Abstract

Insulin-like growth factor-1 (IGF-1) signalling pathway plays important roles in growth, development, and metabolism. The IGF-1 receptor (IGF-1R) mediates the effects of IGF-1 and low IGF-1 production or IGF-1R defects lead to a large heterogeneous group of conditions associated with growth retardation. Both IGF-1 and IGF-1R are synthesized as peptide precursors containing highly conserved N-glycosylation sites. Recent evidence highlights the pivotal role of N-glycosylation in the control of IGF-1 and IGF-1R stability and correct precursor processing.

The aim of this study was to analyze if N-glycosylation defects found in Congenital Disorders of Glycosylation (CDG) impact the proper IGF-1 and IGF-1R propeptide processing and therefore cause molecular defects in IGF-1 signal transduction pathways.

We found a reduced IGF1-R expression level in ALG3-CDG, ALG8-CDG and in some PMM2-CDG fibroblasts obtained from different patients (n=7) compared to control (CTR; n=5) cells. In addition, IGF-1-induced IGF-1R activation was lower in most PMM2-CDG fibroblasts and was associated with decreased ERK1/2 phosphorylation compared to CTR, while the Akt activation did not differ significantly. Lectin-binding analyses showed increased fucosylation in CDG fibroblasts. Furthermore, CDG fibroblasts had increased mRNA expression of Endoplasmic Reticulum (ER) stress-related genes CHOP, uXBP1 and ATF4. Finally, we demonstrated that ALG3-CDG, ALG8-CDG, GMPPB-CDG and some PMM2-CDG fibroblasts presented reduced N-glycosylation of IGF-1Ea prohormone which is associated with lower IGF-1 secretion in the cell culture media. Thus, this result provides new evidence of a direct link between N-glycosylation defects and reduced circulating IGF-1 level found in CDG patients. Moreover, this data suggests that not only systemic but also local IGF-1 production might be impaired in CDG patients.

In conclusion, the results of the present study provide new evidence of a direct link between N-glycosylation defects found in CDG and the IGF-1 system. Further studies are warranted to investigate whether the IGF-1/IGF-1R signaling deficiency found in CDG might be related to growth-related problems associated with CDG.

Introduction

Insulin-like growth factor-1 (IGF-1) system consists of ligand (IGF-1), receptor (IGF-1R) and a family of binding proteins (IGFBPs) [1]. Binding of IGF-1 to IGF-1R leads to autophosphorylation of intracellular tyrosine residues, which in turn actives the phosphatidyl-inositide 3-kinase (PI3K)/AKT and mitogen-activated protein kinase (MAPK)/ERK signaling pathways, determining growth-promoting effects [2]. Thus, defects in either IGF-1 or its receptor can result in poor pre- and post-natal growth [2, 3].

Besides growth phenotype characterized by low birth weight, failure to thrive and microcephaly, other features have been frequently reported in patients with IGF-1R deletions, including skeletal malformation, intellectual impairment, cardiac defects and facial dysmorphism [4-6]. The IGF-1R is synthesized as a single polypeptide precursor (proreceptor), which undergoes proteolytic cleavage into α (130 KDa) and β (97 kDa) chains and forms a tetramer (α-β-β-α), with the extracellular α-subunits involved in ligand binding and the intracellular β-subunits carrying the intrinsic tyrosine kinase function necessary for signal transduction [7]. IGF-1R α and β subunits contain 11 and 5 N-glycosylation sites, respectively [8]. Thus, the α-β-β-α tetramer structure may enclose 32 glycosylation positions [9]. Several studies demonstrated that proper N-glycosylation of IGF-1R proreceptor is required for correct IGF-1R maturation and transport to the cell surface [9 -11].

On the other hand, IGF-1, in its mature form, is a non-glycosylated polypeptide of 7.6 kDa. However, as many peptide hormones, IGF-1 is first synthesised as a part of a larger prohormone. This prohormone contains the mature IGF-1 sequence, which is responsible for binding and activation of the IGF-1R, and a C-terminal intrinsically disordered domain named E-peptide [12]. Although different IGF-1 isoforms might be produced by alternative splicing, we recently demonstrated that the IGF-1Ea pro-hormone, which contains a 35 amino acids long Ea-peptide, is the most ubiquitously expressed isoform [12,13]. More interestingly, intracellular IGF-1 is mainly expressed as proIGF-1Ea of ~17-22 kDa, not mature IGF-1 [14]. The human Ea-peptide contains a highly conserved N-glycosylation site (N92) which is heavily glycosylated with sugars comprising over 30% of the total proIGF-1Ea mass. Notably, the proIGF-1Ea glycosylation ensures proper prohormone folding and favours its passage through the secretory pathway, accordingly the interference with

the Ea-peptide glycosylation causes a rapid degradation of the unglycosylated IGF-1Ea prohormone by the proteasome and blunts mature IGF-1 secretion [13]. Once in the bloodstream, IGF-1 is carried primarily as part of a ternary complex with the IGF binding protein type 3 (IGFBP-3) and the acid-labile subunit (ALS). This complex prolongs circulating IGF-1 half-life from 10–12 minutes to about 15 hours, thus regulating serum levels of IGF-1 [15]. Both IGFBP-3 and ALS are glycosylated proteins. IGFBP-3 contains three N-glycosylation sites [1,16] and it is found in the bloodstream in the glycosylated state with a molecular weight between 40 and 44 kDa. ALS is a soluble glycoprotein of 85 kDa with approximately 20 kDa made up of N-linked oligosaccharides [17]. Although the role of IGFBP-3 and ALS N-glycosylation is still unclear, it seems to contribute to the correct formation of the 150 kDa ternary complex, essential for transport and stabilization of IGF-1 [18,19]. Enzymes and proteins involved in the proteolytic maturation of the IGF-1proreceptor and IGF-1 prohormone, such as furin convertase, represent another important indirect association between the IGF-1 system and N-glycosylation. Indeed, blockade of N-glycosylation reduces the convertase activity and the ability to process the IGF-1 proreceptor [20].

All together, these studies provide clear evidence for a relationship between N-glycosylation and IGF-1 system. Not surprisingly, inhibition of the N-linked glycosylation synthetic pathway has gained increased research interest in an effort to disrupt IGF-1R signaling in the context of cancer [21].

In contrast, much less is known about the impact of N-glycosylation genetic defects, such as those found in Congenital Disorders of Glycosylation (CDG), on the IGF-1 system. CDG are a family of rare metabolic and inherited multisystem disorders characterized by impaired glycosylation in the formation of proteins and lipids. Glycosylation defects affect all organs and accordingly, CDG show a wide spectrum of manifestations and severity including developmental delay, failure to thrive, hypotonia, neurological anomalies and cognitive delay [22]. CDG are classified as type I or II on the basis of the position of the defect in the glycosylation pathway. CDG type-I results from defects in the assembly of the sugar side chain in the cytoplasm and endoplasmic reticulum (ER) [23] and is hallmarked by the absence of one or more N-glycans on glycoproteins. Altered glycosylation in CDG leads to protein misfolding and mild increase of ER-stress markers in patient fibroblasts [24].

Growing evidence suggests that insufficient level and/or activity of growth factors might contribute to some of clinical characteristics of CDG patients [25]. For example, G.S. Dhaunsi found that CDG type-I leukocytes had decreased IGF-1R level compared to control. Moreover, children with PMM2-CDG, the most common form of CDG with more than 900 cases identified worldwide [3,22], have decreased levels of ALS, IGFBP-3, IGF-1 and IGF-2, and ternary complex formation compared with age-matched controls [1]. A finding also confirmed in the recently developed mouse model of PMM2-CDG [26].

The molecular mechanisms related to the IGF-1 system deficiency found in PMM2-CDG patients are currently unknown. Therefore, in the present study we tested if the defective N-glycosylation in different CDG types might have a direct effect in the impairment of IGF-1R signaling activation and proper IGF-1Ea prohormone N-glycosylation.

Methods

CDG-fibroblasts

Primary fibroblasts derived from patients affected by ALG3-CDG, PGAP2-CDG, ALG8-CDG, GMPPB-CDG and control (CTR) fibroblasts from age-matched volunteers were kindly provided by the Meyer Children's Hospital (Florence, Italy). PMM2-CDG fibroblasts were obtained from the Giannina Gaslini Institute-Telethon Network of Genetic Biobanks (Genoa, Italy) [27]. Clinical features of these patients have been partially previously published [28], all subjects experience typical signs and symptoms of CDG including failure to growth and neurologic abnormalities (Table 1). Gene variants and growth hormone (GH), IGF-1 and IGFBP-3 serum levels were also reported where available. Fibroblasts were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS), 2 mM glutamine, penicillin (100 U/mL) and streptomycin (100 µg/mL), and maintained in a 5% CO_2 atmosphere at 37 °C. For IGF-1R activation, after 24 h of serum-free culture in twelve well plates, cells were stimulated with IGF-1 (100 ng/ml) (cat. no. I3769-50UG, Sigma-Aldrich) for 30 or 60 min before lysis.

For IGF-1 gene overexpression, fibroblasts were transfected using the Human Dermal Fibroblast Nucleofector™ Kit (cat no. VPD1001; Lonza) using the Nucleofector™ Device (Lonza) following the instruction manual. Briefly, 1 x 10^6 fibroblasts were resuspended in the nucleofector solution, combined with 2.5 µg of plasmid construct contains the class 1 IGF-1Ea isoforms [29], transferred to the cuvette supplied and finally transfected with the U-23 program of the Nucleofector™ Device. After electroporation, cells were mixed with 500 µl of DMEM and immediately plated into a six-well plate. The efficiency of transfection was estimated by GFP-dependent fluorescence and by real-time RT-PCR for total IGF-1 mRNA levels at 48 h after transfection, as previously described [29]. The Fibroblasts concentration and cell viability were determined by the LUNA-II™ Automated Cell Counter (Logos Biosystems, Twin Helix) with trypan blue staining.

RNA extraction, cDNA synthesis and quantitative Real-Time polymerase chain reaction

Total RNA was extracted and genes of interest analysed by real-time PCR as previously described [30]. Briefly, the expression of ER stress related genes was monitored by qPCR using TB Green Premix Ex TaqII Mastermix (Takara Bio Europe, France), in a RotorGene 6000 instrument (Corbett life science, Sydney, Australia). As template, 2 ng of total RNA used for cDNA synthesis was used in each PCR tube, and a non-template control was included for each primer pair reaction as negative control. All amplification reactions were performed in duplicate. The amplification conditions were: 95°C for 10 min, 95°C for 10 s and 60°C for 50 s for 40 cycles. At the end of each run, a melting curve analysis from 65°C to 95°C was performed to exclude the presence of non-specific products or primer dimers. The data were normalized against the reference gene GAPDH (glyceraldehyde-3-phosphate dehydrogenase). The relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method [31].

Western blot and lectin blot analyses

Fibroblasts were processed for western blot analysis as previously reported [29]. The protein samples (20-40 µg total proteins) were electrophoresed with 10% SDS-PAGE, and then transferred to nitrocellulose or PVDF membranes (Bio-Rad Laboratories) for immunoblotting. Primary antibodies against phospho IGF-1 Receptor β (1:200; n. 3024 Cell Signaling Technology); IGF-1 Receptor β (1:2000; n. 3027 Cell Signaling Technology), phospho-Akt (Ser473) (1:2000; n. 9271 Cell Signaling Technology), Akt (1:2000; n. 9272 Cell Signaling Technology), phospho-p44/42 (ERK1/2) (1:2000; n. 9101 Cell Signaling Technology), p44/42 (ERK1/2) (1:2000; n. 9102 Cell Signaling Technology) and IGF-1 (1:2000; no. 500P11 Peprotech) were incubated overnight at 4 °C. For lectin blotting, membranes were probed with biotinylated Concanavalin A (ConA, 1:1000), *Phaseolus vulgaris* leucoagglutinin (PHA-L, 1:200) and *Aleuria aurantia* (AAL, 1:400) lectins (Vector laboratories, D.B.A. Italia) at room temperature while shaking for 1 h. After washes, the membranes were incubated with the appropriate horseradish peroxidase (HRP)-conjugated secondary antibody (Bio-Rad Laboratories) at room temperature for 1 h and were then washed three times. Blots were developed using Clarity Western ECL

Substrate (Bio-Rad Laboratories) and were quantified using the ChemiDoc MP (Bio-Rad Laboratories) equipped with Image Lab software.

ELISA Assay

For the quantitative determination of IGF1 concentrations in transfected fibroblasts supernatants, a commercially available ELISA kit was used according to the manufacturer's instructions (Human IGF-I/IGF-1 DuoSet ELISA cat no. DY291-05; R&D Systems). Data were acquired at a wavelength of 450 nm using a microplate reader (Model 680 microplate reader, Biorad).

Statistical analysis

Data are represented as mean ± SEM of at least three independent experiments. For western blot, lectin blot and ELISA assay, statistical analyses were performed using the t-test or one-way ANOVA as appropriate, followed by Bonferroni's multiple comparison post hoc tests. Unpaired t-test with Welch correction was used to compare ER-stress related genes fold change values between control and CDG fibroblasts. The statistical tests were performed using SPSS (IBM SPSS Statistics for Windows, Version 20.0, IBM Corp.) and GraphPad Prism version 5 (GraphPad Software, Inc., La Jolla, CA, USA). A p value ≤ 0.05 was considered statistically significant.

Patient code/sex	Age at last follow-up	Clinical findings	Growth failure	GH (NV 0.09 – 1.95 ng/ml)	IGF-1 (NV 95-312 ng/ml)	IGFBP-3 (NV 1.89 – 7.33 µg/ml)	Gene variants	Diagnosis
1/M	5 years	Spastic quadriparesis, severe developmental delay, acquired microcephaly, (<4SD)	yes	8.61 ↑↑↑	52 ↓↓	2.09	Homozygous ALG3 c. 1A>G p.(Met1?)	ALG3-CDG
2/F	Deceased at 6 years for respiratory failure during pneumonia	Spastic quadriparesis, severe cognitive impairment, acquired microcephaly, (<4SD), facial dysmorphism, severe scoliosis, distal arthrogryposis, hepatomegaly	yes	NA	NA	NA	Compound heterozygous ALG3 c. 101 G>A p. (Arg35His) / c. 165 C>T p. (Gly55Gly)	ALG3-CDG
3/M	10 years	Mild hypotonia, facial dysmorphism, psychomotor delay, autism	no	0.07 ↓	80 ↓	4.45	Compound heterozygous ALG8 c. 122 G>A p. (Arg41Gln) / [c. 446 T>G p. (Leu149Arg) + c. 980 C>G p. (Thr327Arg)]	ALG8-CDG
4/M	5 years	Severe hypotonia, epileptic encephalopathy, psychomotor delay, hyperphosphatasia	no	1.41	116	3.16	Compound heterozygous PGAP2 c.549dupA p. (Gln184Thrfs*26) / c. 686 C>T p. (Ala229Val)	PGAP2-CDG
5/M	13 months	Severe hypotonia, congenital muscular dystrophy, facial dysmorphism, epileptic encephalopathy, distal arthrogryposis	-	planned	planned	planned	Compound heterozygous GMPPB c. 95 C>T p. (Pro32Leu)/ c. 931 C>T p. (Arg311Cys)	GMPPB-CDG
20706/M	36 years	Dysmorphic features, microcephaly, cerebellar atrophy, peripheral neuropathy, severe intellectual disability	+	np	np	np	V129M/R141H	PMM2 _p1-CDG
20707/M	29 years		+	np	np	np	V129M/R141H	PMM2_p2-CDG
21197/M								PMM2_p3-CDG
21213/M	29 years	Dysmorphic features, microcephaly, cerebellar atrophy, peripheral neuropathy, severe intellectual disability	+	planned	planned	planned	V231M/R141H	PMM2 _p4-CDG
22433/M								PMM2_p5-CDG
22540/F								PMM2_p6-CDG
23167/M	15 years	Dysmorphic features, microcephaly, cerebellar atrophy, peripheral neuropathy, moderate intellectual disability	+	planned	141.4	planned	R141/N216I	PMM2_p7-CDG

Table 1: Clinical features of CDG patients analysed in the present study [28].

Results

Lectin binding analysis and ER-stress related markers in CDG fibroblasts

We used Concanavalin A (ConA), *Phaseolus vulgaris* leucoagglutinin (PHA-L) and *Aleuria aurantia* lectins (AAL) to analyze N-glycosylation of CDG-fibroblasts (Fig. 1). ConA and PHA-L recognize high mannose and complex type N-glycans, respectively, while AAL binds to fucose linked (α -1,6) to N-acetylglucosamine or to fucose linked (α -1,3) to N-acetyllactosamine. As shown in Fig. 1, AAL reactivity was increased in CDG fibroblasts compared to CTR. Howeverver, subsequent post hoc analyses did not reveal any significant difference between CDG-subtypes and CTR cells. PHA-L binding showed a nonsignificant trend to decrease in CDG fibroblasts compared to CTR (p=0.38) while ConA reactivity did not change (p=0.92). Subsequently, quantitative RT-PCR was carried out to quantify the expression level of selected ER-stress related genes: *CHOP*, *sXBP1*, *uXBP1*, *MAP1LC3B*, *HSPA5*, *CEBPB*, *CHAC1* and *ATF4*. As shown in Fig. 2, *CHOP*, *uXBP1* and *ATF4* mRNA levels increased significantly in CDG fibroblasts compared to CTR.

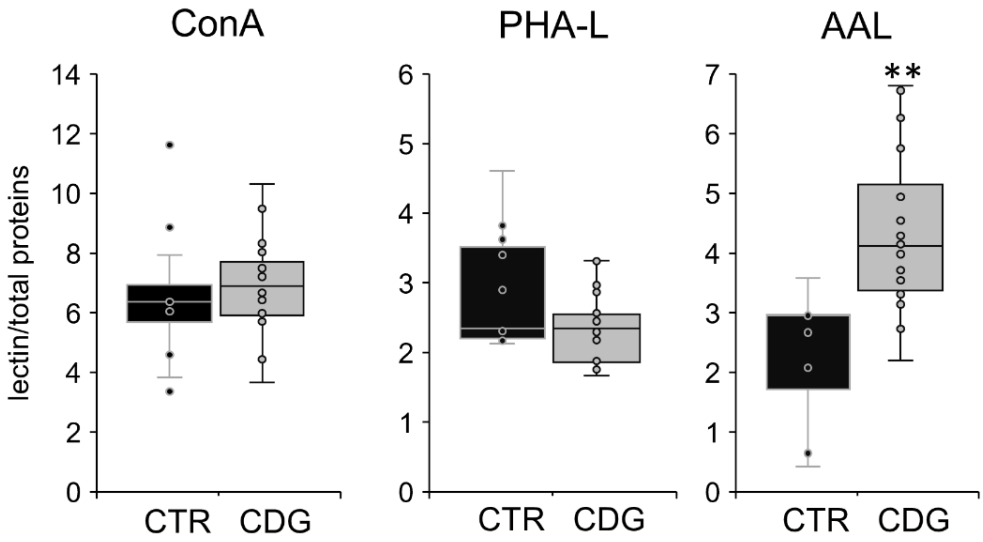


Figure 1. Lectin-binding analysis. Concanavalin A (ConA), Phaseolus vulgaris leucoagglutinin (PHA-L) and Aleuria aurantia (AAL) binding to CTR and CDG fibroblasts. Con A and PHA-L recognize high mannose and complex type N-glycans, respectively, while AAL binds to fucose linked (α -1,6) to N-acetylglucosamine or to fucose linked (α -1,3) to N-acetyllactosamine. *significantly differed compared to CTR, **p< 0.01.

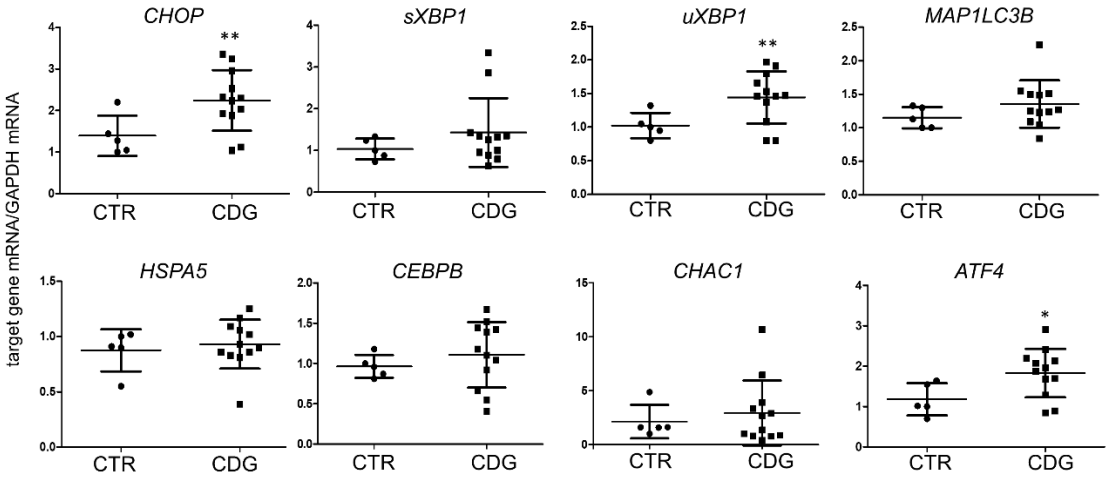


Figure 2. ER-stress markers mRNA expression in CDG fibroblasts. Quantification of the expression level of selective ER-stress related genes: *CHOP*, *sXBP1*, *uXBP1*, *MAP1LC3B*, *HSPA5*, *CEBPB*, *CHAC1* and *ATF4*. *CHOP*, *uXBP1* and *ATF4* mRNA levels were significantly upregulated in CDG fibroblasts compared to CTR (Unpaired T-test with Welch correction *p< 0.05, **p< 0.01).

IGF-1R expression and IGF1R pathway activation in CDG fibroblasts

As shown in Fig. 3A, both ALG3-CDG fibroblasts which derived from two different patients and ALG8-CDG cells revealed reduced IGF-1R expression levels compared to CTR fibroblasts. PMM2-CDG fibroblasts presented significant variability in IGF-1R expression level, which was reduced in about 50% of PMM2-CDG cells. The reduction of IGF-1R found in CDG fibroblasts was not followed by an IGF-1R proreceptor accumulation (Fig. 3B and 3C). Altogether, fibroblasts from both ALG-CDG (n=5) and PMM2-CDG (n=7) subtypes showed a reduction of IGF-1R level compared to CTR (n=5) (Fig. 3D). To evaluate if IGF-1R deficit observed in some PMM2-CDG was associated with a reduction of IGF-1R pathway activation, fibroblasts were treated with recombinant IGF-1 (100 ng/ml) for 30 and 60 min and the level of IGF-1R (Fig. 4A), Akt (Fig. 4B) and ERK1/2 (Fig. 4C) phosphorylation was quantified by western blotting. The IGF-1-induced activation of IGF-1R and ERK1/2 were reduced in most PMM2-CDG fibroblasts, while Akt phosphorylation did not change. Notably, in some PMM2-CDG the decreased response to IGF-1 treatment was mainly due to higher basal IGF-1R and ERK1/2 activation (e.g. PMM2_p2; PMM2_p3; PMM2_p5; PMM2_p6), instead of general inhibition of IGF-

1R and ERK1/2 activity (Fig. 4D). When PMM2-CDG fibroblasts derived from different patients were pooled together (n=7), both IGF-1R and ERK1/2 IGF-1-induced activation were reduced compared to CTR (Fig. 4E). Preliminary results obtained in some ALG-CDG fibroblasts also showed a downregulation of the IGF-1 pathway (not shown).

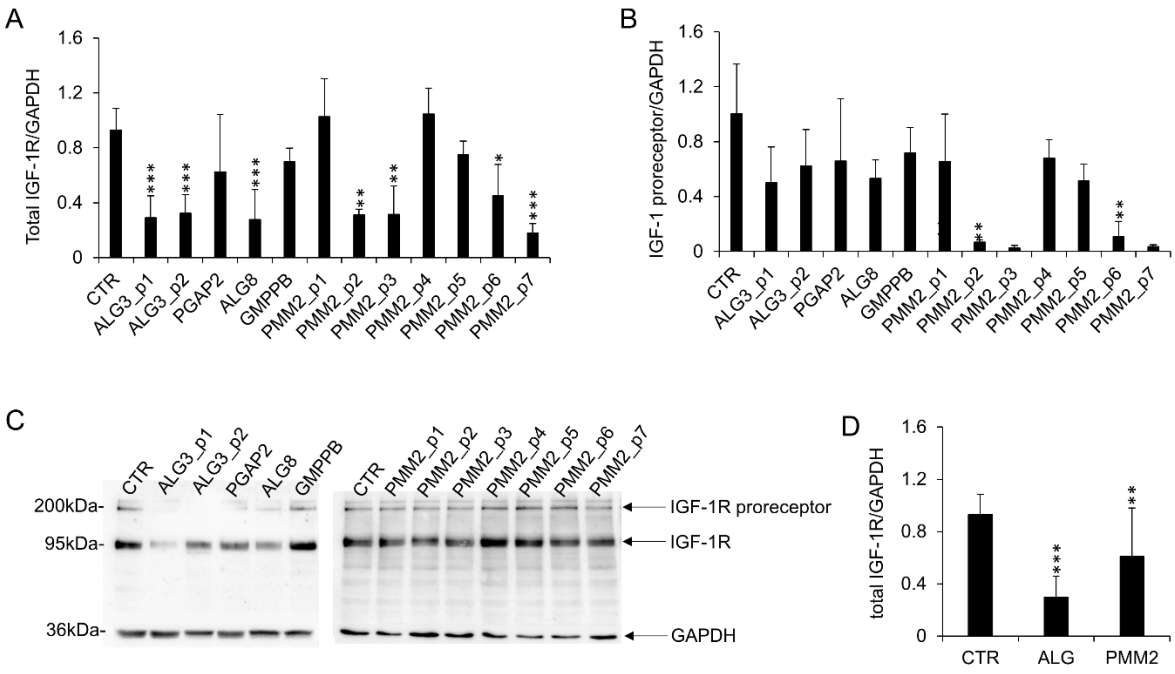


Figure 3. IGF-1R protein level in CDG fibroblasts. Relative expression level of IGF-1R (A) and IGF-1R prohormone (B) in different CDG subtypes quantified by western blot. Representative western blot showing expression level of IGF-1R (~97kDa), IGF-1R proreceptor (~200kDa) and GAPDH (~36kDa) (C). Relative expression level of IGF-1R in CTR (n=5), ALG-CDG (n=5) and PMM2-CDG (n=7) fibroblasts quantified by western blot (D). *significantly differed compared to CTR; ***p< 0.0001, **p< 0.001, *p< 0.05.

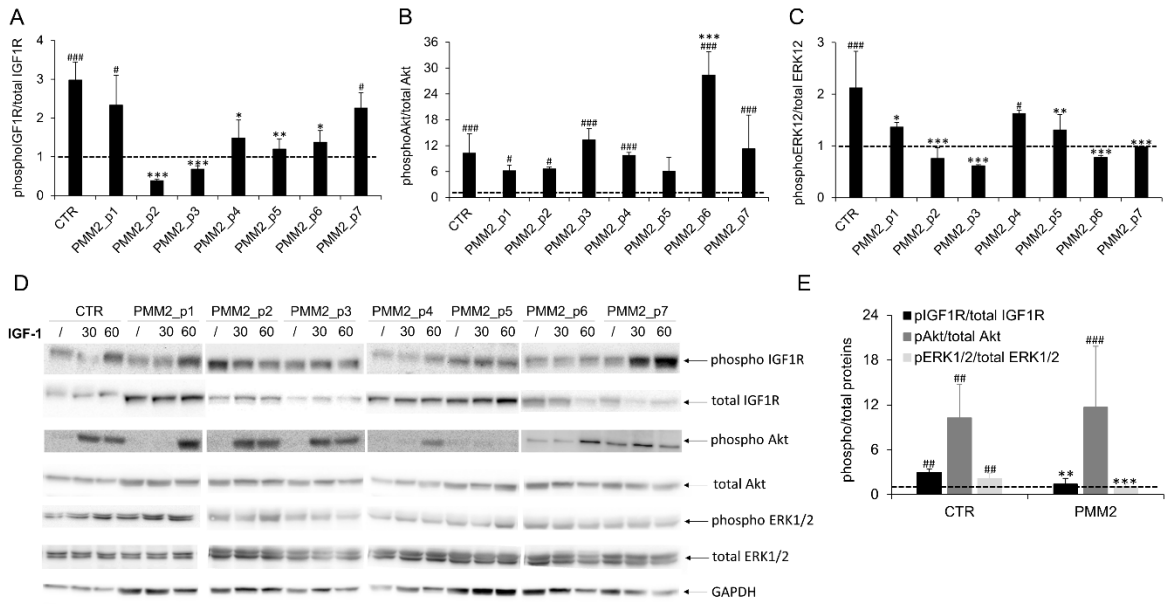


Figure 4. IGF-1R signalling pathway activation. Relative expression level of phosphorylated IGF-1R (A), Akt (B) and ERK1/2 (C) in PMM2-CDG fibroblasts obtained from different patients. Representative western blot showing expression level of phosphorylated and total IGF-1R, Akt, ERK1/2 and GAPDH (D). IGF-1R, Akt and ERK1/2 phosphorylation level in CTR (n=5) and PMM2-CDG (n=7) fibroblasts (E). *significantly differed compared to CTR; #significantly differed compared to untreated cells *** and ###p< 0.0001, **p< 0.001, *p< 0.01; #p< 0.05.

proIGF-1Ea N-glycosylation pattern and mature IGF-1 secretion in CDG fibroblasts

In order to evaluate the expression level of IGF-1 in CTR and CDG fibroblasts we conducted preliminary analyses on the IGF-1 mRNA and protein quantity. IGF-1 mRNA quantification by real-time RT-PCR showed that all fibroblasts had detectable levels of IGF-1 mRNA (mean CT value 30.5), however the IGF-1 protein level was too low to be detected by western blot or ELISA. We also tried to quantify the IGF-1 secretion in the fibroblast supernatants after concentration using 3kDa Spin Columns (by about 20-fold; Amicon® Ultra Merck Millipore), but the IGF-1 concentration still remained below the ELISA sensitivity (93.8 pg/ml). Subsequently, to assess the expression pattern and N-glycosylation status of IGF-1 protein, we transiently transfected fibroblasts with plasmid encoding the IGF-1Ea isoform or an empty vector using electroporation [29].

As shown in Figure 5A, two distinct bands, likely representing glycosylated (~17kDa) and unglycosylated (~12kDa) proIGF-1Ea, were detected in CTR and PGAP2 fibroblasts. On the other hand, the glycosylation pattern of IGF-1Ea prohormone

differed in CDG fibroblasts. In particular, the two fibroblasts from ALG3-CDG presented a distinct band of about 14 kDa, while ALG8 and GMPPB revealed a marked accumulation of 12 kDa proteins. All these bands likely represent underglycosylated forms of IGF-1Ea pro-hormone. Among PMM2-CDG fibroblasts analysed, there was a general preservation of IGF-1Ea N-glycosylation pattern, however the IGF-1 bands intensity in PMM2_p1, PMM2_p6 and PMM2_p7 CDG fibroblasts were lower compared to CTR. Subsequently, we quantified the IGF-1 protein level in the cell culture media of IGF-1Ea-transfected fibroblasts (Fig. 5B). We observed that both ALG3-CDG fibroblasts derived from two different patients, GMPPB-CDG, PMM2_p1-CDG and PMM2_p6-CDG showed a decreased IGF-1 secretion compared to CTR. IGF-1 plasma level was also quantified from some available patients (Table 1). Low plasma IGF-1 was found in ALG-3-CDG and ALG-8 patients (52 ng/ml and 80 ng/ml, respectively; n.v. 95-312 ng/ml), while normal level was observed in PGAP2 (116 ng/ml). Among PMM2 patients, the only available IGF-1 serum level was that of patient 7 (141.4 ng/ml) which fell into normal level.

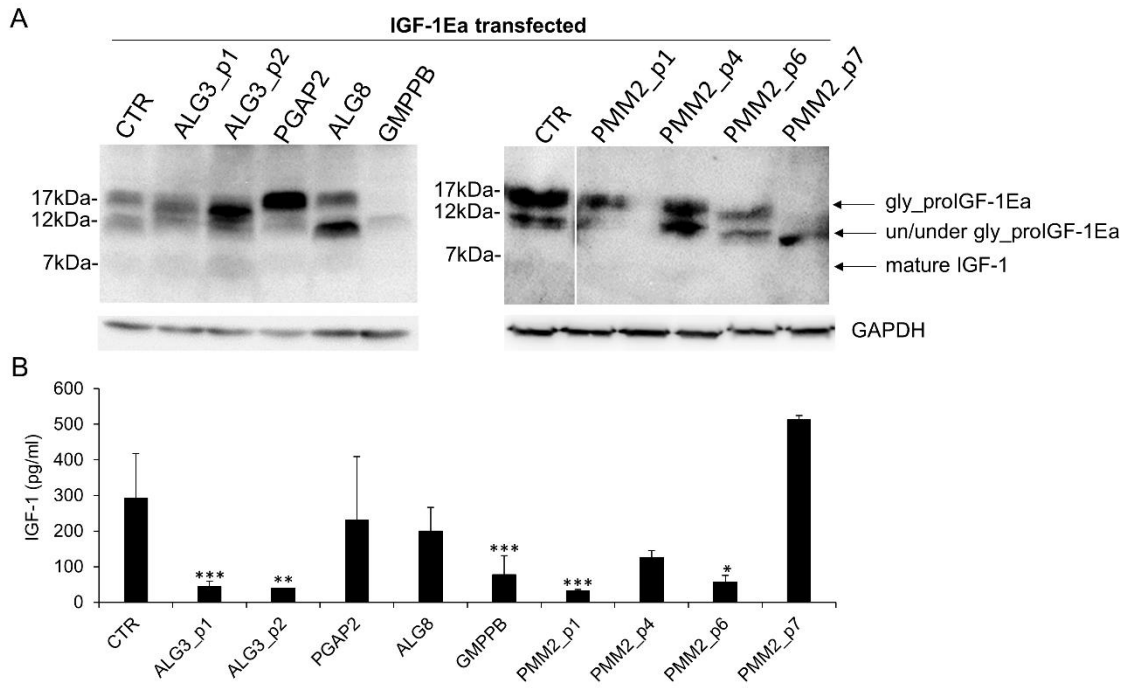


Figure 5. IGF-1Ea prohormone expression pattern and IGF-1 secretion. Immunoblotting of protein lysates from IGF-1Ea transfected fibroblasts using an antibody directed against the mature sequence of IGF-1 (A). Quantification of IGF-1 level in fibroblast supernatants measured by ELISA (B). gly_proIGF-1Ea: glycosylated IGF-1Ea prohormone; un/under gly_proIGF-1Ea: un/underglycosylated IGF-1Ea prohormone. *significantly differed compared to CTR; *** p< 0.0001, **p< 0.001, *p< 0.01.

Discussion

During human development, growth factors such as growth hormone (GH) and IGF-1 are known to play an important role to determine the growth and function of various organs and tissues [32]. Postnatal growth failure is common in children with CDG despite normal or elevated levels of GH [1]. Although some of the effects of GH are direct actions, most effects are mediated through the peptide hormone IGF-1 [33], [34]. An imbalance of the IGF system has been reported in patients and animal models of CDG [1,26]. In the serum, >90% of IGF-1 is found in a 150-kDa ternary complex formed by IGF-1, ALS and IGFBP-3, which is necessary to stabilize IGF-1 [1]. Inadequate glycosylation of ALS and IGFBP-3 has been described in PMM2-CDG patients which probably leads to a reduction in the formation of the ternary complex and at least partially explains the GH-independent reduction of circulating IGF-1 found in CDG.

The results of the present study provide new evidence of a direct link between N-glycosylation defects found in CDG and the IGF-1 system. Previous studies showed that N-glycosylation contributes to ligand binding, kinase activity, and turnover of IGF-1R and other receptor tyrosine kinases [35]. Accordingly, here we demonstrated that both IGF-1R level and its activity were reduced in primary fibroblasts from different CDG subtypes. In particular, western blot analysis revealed reduced levels of IGF-1R in ALG3-CDG, ALG-8-CDG and in some PMM2-CDG fibroblasts. We did not find an accumulation of IGF-1 proreceptor in these cells, indicating that defective N-glycosylation mainly affected IGF-1R proreceptor stability [25]. In PMM2-CDG we also analysed the IGF-1-induced IGF-1R pathway activation. Most PMM2-CDG fibroblasts showed reduced IGF-1R and ERK1/2 phosphorylation compared to CTR, while the Akt response did not differ significantly. These results are in agreement with those obtained by G.S. Dhaunsi who showed IGF-1-R deglycosylation in CDG type-I lymphocytes [25].

Considering that IGF-1R is the sole mediator of IGF-1 actions [36], the reduced level and impaired functioning of IGF-1R signaling might directly impact IGF-1-induced growth and differentiation responses. Intriguingly, symptoms of patients with heterozygous defects in IGF-1R, which include pre and postnatal growth retardation, microcephaly, cardiac defects and dysmorphic features show a significant overlap with those found in CDG. The fact that PMM2-CDG fibroblasts

from distinct patients showed different levels of IGF-1R protein reduction and IGF-1-induced activation adds another layer of complexity. In this regard, we also performed a lectin-binding analysis to profile CDG protein N-glycosylation. Binding of AAL lectin was increased in CDG fibroblasts compared to CTR, suggesting an increase in fucosylation in CDG cells, in agreement with [37]. However, we were unable to find significant differences between different CDG-subtypes. Moreover, we did not find a significantly different ConA- and PHA-L-reactivity between CDG and CTR fibroblasts. We also analysed the expression of genes related to ER-stress since previous studies showed chronic ER stress and unfolded protein response (UPR) in type I CDG [38]. Accordingly, we found a significant increase of several ER-stress related genes in CDG fibroblasts compared to CTR. Further studies are ongoing to find if IGF-1R level expression and its activation are correlated with specific gene variants or PMM2 enzyme residual activity.

Finally, we investigated the IGF-1Ea prohormone N-glycosylation pattern in CDG fibroblasts. In our previous work, we demonstrated that, under physiological conditions, intracellular IGF-1 is mainly expressed as a heavily N-glycosylated prohormone of 17-22 kDa. We were unfortunately not able to monitor the glycosylation status of endogenous IGF-1Ea in CTR and CDG fibroblasts due to very low protein expression. Thus, we transiently overexpressed the IGF-1Ea isoform in CDG fibroblasts and monitored N-glycosylation pattern by western blotting and IGF-1 secretion by ELISA. Two distinct bands, likely representing glycosylated and unglycosylated proIGF-1Ea, were detected in CTR, PGAP2 and most PMM2-CDG fibroblasts. On the other hand, the glycosylation pattern of IGF-1Ea prohormone differed in other CDG fibroblasts. In particular, the two fibroblasts from ALG3-CDG showed a distinct band of about 14 kDa while ALG8 and GMPPB showed a marked accumulation of 12 KDa proteins, all these bands likely represent underglycosylated forms of IGF-1Ea prohormone. Subsequently, we quantified the IGF-1 protein level in the cell culture supernatants, and we found that most CDG fibroblasts presented reduced accumulation of IGF-1 in the culture media compared to CTR. Thus, these data show that proper glycosylation of IGF-1Ea prohormone is impaired in CDG fibroblasts and is associated with decreased IGF-1 secretion. These data suggest that not only systemic but also local IGF-1 production might be compromised in CDG patients. This aspect might be relevant in the context of CDG disorders since both liver (endocrine) and locally produced IGF-1 are important for

normal growth [39]. Since, IGF-1 is not measured frequently in clinical practice the IGF-1 serum level was available only for four patients (ALG3-CDG, ALG8-CDG, PGAP2-CDG and PMM2_p7-CDG). Both ALG3-CDG and ALG8-CDG patients presented a marked reduction of IGF-1 plasma level however, the low patient numbers, limited the statistical analysis that could be performed.

In conclusion, we showed that primary fibroblasts from CDG patients had reduced levels of IGF-1R and IGF-1. These findings corroborate other studies that implicate N-glycosylation in the regulation of IGF-1 system. We also reported that distinct CDG subtypes have a different impact on the IGF-1 system suggesting that also the degree of IGF-1R and IGF-1 deficiency may differ in these patients. However, since the IGF-1 level is not quantified in routine clinical practice, at the moment, we are unable to support this hypothesis. Additional *in vivo* evidence, for example using the recently developed mouse model of PMM2, might help to clarify the impact of IGF-1 signalling pathways deficiency in CDG and to evaluate if potential therapeutics, such as recombinant IGF-1, may help to overcome at least some of the growth-related problems associated with CDG.

“Acknowledgements”: “The Cell Line and DNA Biobank from Patients Affected by Genetic Diseases”, member of the Telethon Network of Genetic Biobanks (project no. GTB18001), funded by Telethon Italy, provided us with specimens”. Thanks to the Professor Timothy Bloom for a critical reading of this work.

References

[1] S. Miller, M. J. Khosravi, M. C. Patterson, C. A. Conover, “IGF system in children with congenital disorders of glycosylation.,” *Clin. Endocrinol. (Oxf).*, vol. 70, no. 6, pp. 892–897, 2009.

[2] W. K. J. Klammt, R. Pfäffle, H. Werner, W. Kiess, “IGF signaling defects as causes of growth failure and 383 IUGR.,” *Trends Endocrinol Metab TEM*, vol. 19, no. 197–205., 2008.

[3] A. Varki et al., “Essentials of Glycobiology,” *3rd Ed. Cold Spring Harb. Cold Press, 2015-2017.*

[4] S. Jain, D. W. Golde, R. Bailey, M. E. Geffner, “Insulin-like growth factor-I resistance,” *Endocr Rev.*, vol. 19, no. 5, pp. 625–46, 1998.

[5] A. W. Ester. et al., “Two short children born small for gestational age with insulin-like growth factor 1 receptor haploinsufficiency illustrate the heterogeneity of its phenotype,” *J Clin Endocrinol Metab.*, vol. 94, no. 12, pp. 4717–27, 2009.

[6] D. C. M. Veenma et al., “Phenotype-genotype correlation in a familial IGF1R microdeletion case,” *J Med Genet.*, vol. 47, no. 7, pp. 492–8, 2010.

[7] T. E. Adams, V. C. Epa, T. P. Garrett, C. W. Ward, “Structure and function of the type 1 insulin-like growth factor receptor,” *Cell Mol Life Sci.*, vol. 57, no. 7, pp. 1050–93, 2000.

[8] A. Dricu, M. Carlberg, M. Wang, and O. Larsson, “Inhibition of N-linked glycosylation using tunicamycin causes cell death in malignant cells: Role of down-regulation of the insulin-like growth factor 1 receptor in induction of apoptosis,” *Cancer Res.*, vol. 57, no. 3, pp. 543–548, 1997.

[9] M. Carlberg *et al.*, “Mevalonic acid is limiting for N-linked glycosylation and translocation of the insulin-like growth factor-1 receptor to the cell surface. Evidence for a new link between 3-hydroxy-3-methylglutaryl-coenzyme A reductase and cell growth,” *J. Biol. Chem.*, vol. 271, no. 29, pp. 17453–17462, 1996.

[10] K. W. Siddals, E. Marshman, M. Westwood, and J. M. Gibson, “Abrogation of insulin-like growth factor-I (IGF-I) and insulin action by mevalonic acid depletion: Synergy between protein prenylation and receptor glycosylation pathways,” *J. Biol. Chem.*, vol. 279, no. 37, pp. 38353–38359, 2004.

[11] A. Dricu et al., “Expression of the insulin-like growth factor 1 receptor (IGF-1R) in breast cancer cells: evidence for a regulatory role of dolichyl phosphate in the transition from an intracellular to an extracellular IGF-1 pathway,” *Glycobiology*, vol. 9, pp. 571–579, 1999.

[12] G. Annibalini et al., “MIR retroposon exonization promotes evolutionary variability and generates species-specific expression of IGF-1 splice variants, ”*Biochim. Biophys. Acta - Gene Regul. Mech.*, vol. 1859, no. 5, pp. 757–768, 2016.

[13] J. Durzyńska et al., “The pro-forms of insulin-like growth factor I (IGF-I) are predominant in skeletal muscle and alter IGF-I receptor activation,” *Endocrinology*, vol. 154, no. 3, pp. 1215–24, 2013.

[14] G. Annibalini *et al.*, “The intrinsically disordered E-domains regulate the IGF-1 prohormones stability, subcellular localisation and secretion,” *Sci. Rep.*, vol. 8, no. 1, 2018.

[15] R. C. Bunn, J. L. Fowlkes, “Insulin-like growth factor binding protein proteolysis,” *Trends Endocrinol Metab.*, vol. 14, no. 4, pp. 176–81, 2003.

[16] P. F. Collett-Solberg and P. Cohen, “The role of the insulin-like growth factor binding proteins and the IGFBP proteases in modulating IGF action,” *Endocrinol. Metab. Clin. North Am.*, vol. 25, no. 3, pp. 591–614, 1996.

[17] J. B. M. Janosi, P. A. Ramsland, M. R. Mott, S. M. Firth, R. C. Baxter, and P. J. D. Delhanty, “The acid-labile subunit of the serum insulin-like growth factor-binding protein complexes. Structural determination by molecular modeling and electron microscopy,” *J. Biol. Chem.*, vol. 274, no. 33, pp. 23328–23332, 1999.

[18] S. M. Firth and R. C. Baxter, “Cellular actions of the insulin-like growth factor binding proteins,” *Endocr. Rev.*, vol. 23, no. 6, pp. 824–854, 2002.

[19] Y. R. Boisclair, R. P. Rhoads, I. Ueki, J. Wang, and G. T. Ooi, “The acid-labile subunit (ALS) of the 150 kDa IGF-binding protein complex: An important but forgotten component of the circulating IGF system,” *J. Endocrinol.*, vol. 170, no. 1, pp. 63–70, 2001.

[20] E. Klaver et al., “Selective inhibition of N-linked glycosylation impairs receptor tyrosine kinase processing,” *Dis Model Mech*, vol. 12, no. 6, p. dmm039602, 2019.

[21] J. N. Contessa et al., “Inhibition of N-linked glycosylation disrupts receptor tyrosine kinase signaling in tumor cells,” *Cancer Res.*, vol. 68, no. 10, pp. 3803–9, 2008.

[22] I. J. Chang, M. He, C. T. Lam, “Congenital disorders of glycosylation,” *Rev. Artic. Inborn Errors Metab. Ann. Transl. Med.*, vol. 6, no. 24, p. 477, 2018.

[23] J. Jaeken, G. Matthijs, “Congenital disorders of glycosylation: a rapidly expanding disease family,” *Annu Rev Genomics Hum Genet.*, vol. 8, pp. 261–78, 2007.

[24] M. R. Lecca et al., “Genome-wide analysis of the unfolded protein response in fibroblasts from congenital disorders of glycosylation type-I patients,” *FASEB J.*, vol. 19, no. 2, pp. 240–2, 2005.

[25] G. S. Dhaunsi, “Receptor-mediated selective impairment of insulin-like growth factor-1 activity in congenital disorders of glycosylation patients,” *Pediatr. Res.*, vol. 81, no. 3, pp. 526–530, 2017.

[26] B. Chan *et al.*, “A mouse model of a human congenital disorder of glycosylation caused by loss of PMM2,” *Hum. Mol. Genet.*, vol. 25, no. 11, pp. 2182–2193, 2016.

[27] M. Filocamo *et al.*, “Cell Line and DNA Biobank From Patients Affected by Genetic Diseases,” *Open J. Bioresour.*, vol. 1, no. January 1976, p. e2, 2014.

[28] R. Barone *et al.*, “A nationwide survey of PMM2-CDG in Italy: High frequency of a mild neurological variant associated with the L32R mutation,” *J. Neurol.*, vol. 262, no. 1, pp. 154–164, 2015.

[29] M. De Santi *et al.*, “Human IGF1 pro-forms induce breast cancer cell proliferation via the IGF1 receptor,” *Cell. Oncol.*, vol. 39, no. 2, pp. 149–159, 2016.

[30] L. Galluzzi et al., “Leishmania infantum induces mild unfolded protein response in infected macrophages,” PLoS One, vol. 11, no. 12, pp. 1–19, 2016.

[31] M. W. Pfaffl, “A new mathematical model for relative quantification in real-time RT-PCR.,” Acids Res., vol. 29, no. 9, pp. 2002-2007, 2001.

[32] L. Laviola, A. Natalicchio, and F. Giorgino, “The IGF-I Signaling Pathway,” *Curr. Pharm. Des.*, vol. 13, no. 7, pp. 663–669, 2007.

[33] D. Le Roith, C. Bondy, S. Yakar, J. L. Liu, and A. Butler, “The somatomedin hypothesis: 2001,” *Endocr. Rev.*, vol. 22, no. 1, pp. 53–74, 2001.

[34] D. Le Roith, L. Scavo, A. Butler, “What is the role of circulating IGF-I?,” *Trends Endocrinol Metab.*, vol. 12, no. 2, pp. 48–52, 2001.

[35] I. G. Ferreira, M. Pucci, G. Venturi, N. Malagolini, M. Chiricolo, and F. Dall’Olio, “Glycosylation as a main regulator of growth and death factor receptors signaling,” *Int. J. Mol. Sci.*, vol. 19, no. 2, 2018.

[36] S. Yakar, H. Werner, C. J. Rosen, “40 YEARS OF IGF1: Insulin-like growth factors: actions on the skeleton,” *J. Mol. Endocrinol.*, vol. 61, no. 1, pp. T115-T137, 2018.

[37] W. Van Dijk, C. Koeleman, B. Van het Hof, D. Poland, C. Jakobs, and J. Jaeken, “Increased α3-fucosylation of α1-acid glycoprotein in patients with

congenital disorder of glycosylation type IA (CDG-Ia),” *FEBS Lett.*, vol. 494, no. 3, pp. 232–235, 2001.

[38] J. Shang, C. Körner, H. Freeze H, M.A. Lehrman, “Extension of lipid-linked oligosaccharides is a high-priority aspect of the unfolded protein response: endoplasmic reticulum stress in Type I congenital disorder of glycosylation fibroblasts.,” *Glycobiology*, vol. 12., no. 5, pp. 307–17, 2002.

[39] J. E. Puche & A. Castilla-Cortázar, “Human conditions of insulin-like growth factor-I (IGF-I) deficiency,” *J. Transl. Med.*, vol. 10, no. 224, 2012.

CONCLUSIONS

During human development, growth factors such as GH and IGF-1 are known to play an important role to determine the growth and function of various organs and tissues. In keeping with its important physiological role, IGF-1 activity is controlled by temporal and spatial regulation of IGF-1, IGFBPs and IGF-1R productions. In this Thesis we provide new evidence regarding the role of N-glycosylation in the modulation of IGF-1 system components. Firstly, we demonstrated that intracellular IGF-1 is mainly expressed as a pro-hormone (proIGF-1Ea) and not mature IGF-1. The proIGF-1Ea is composed of protein structural domain, the mature IGF-1, and intrinsically disordered regions, the C-terminal Ea-domain which contains a highly conserved N-glycosylation site (N92). We demonstrate that N-glycosylation of Ea-domain ensures proIGF-1Ea folding and favouring its passage through the secretory pathway. On the other hand, the interference with proIGF-1Ea N-glycosylation causes a rapid degradation of the unglycosylated proIGF-1Ea by the proteasome and hence dramatically reduces the mature IGF-1 secretion.

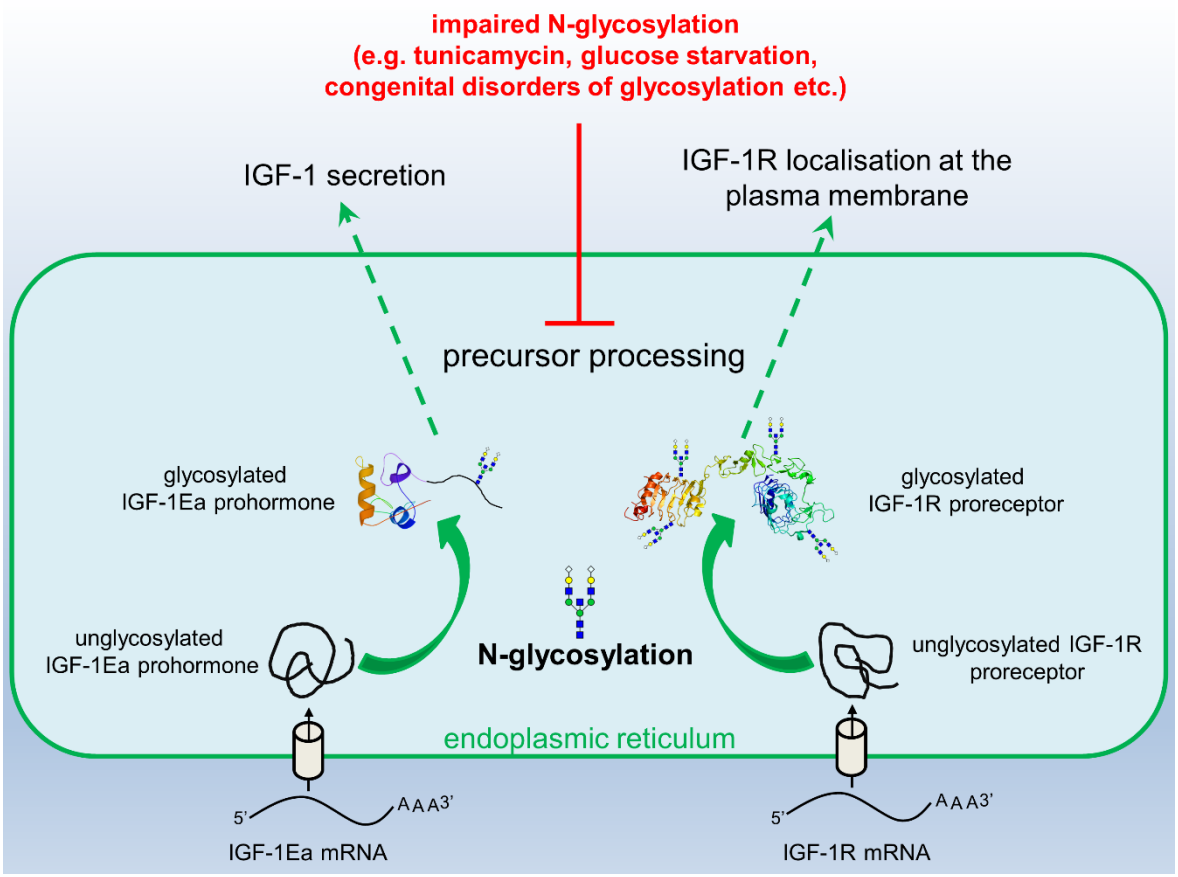
Using the C2C12 cell line, we demonstrate that the N-glycosylation inhibitor tunicamycin or the PMM2 knockdown, an indispensable enzyme that catalyses an early step of the N-glycosylation pathway, markedly reduces myoblast differentiation and disrupted the coordinated temporal expression of myogenic regulator genes. Interestingly, N-glycosylation inhibition decreases IGF-1R expression, the downstream signalling cascades and dysregulates IGF-1 production compared to CTR cells. These results offer new insights into the impairment of myogenic differentiation capacity in the context of pathological N-glycosylation deficiencies such as congenital disorders of glycosylation (CDG).

Finally, using primary fibroblasts from different CDG subtypes, we demonstrate that proper proIGF-1Ea N-glycosylation is impaired in some CDG fibroblasts and is associated with decreased IGF-1 secretion. Moreover, we find a reduced IGF1-R expression level, IGF-1-induced IGF-1R activation and lower ERK1/2 phosphorylation in CDG fibroblasts compared to CTR. These data suggest that both systemic IGF-1 production and local IGF-1 signalling might be impaired in CDG patients.

In conclusion, findings from this Thesis may represent a starting point to clarify the impact of N-glycosylation dysregulation on IGF-1 system and to test potential

therapeutics, such as recombinant IGF-1 and glycosylation promoter in the context of N-glycosylation deficiency.

Graphical Abstract



Proposed mechanism of the IGF-1Ea prohormone and the IGF-1R proreceptor regulation by N-glycosylation. The IGF-1Ea prohormone and IGF-1R proreceptor are synthesized and released into the lumen of the endoplasmic reticulum. Here, the newly created polypeptide is subjected to N-glycosylation which controls proper precursor processing and hence IGF-1 secretion and IGF-1R plasma membrane localisation. IGF-1Ea prohormone and IGF-1R proreceptor processing is impaired by the N-glycosylation inhibitor tunicamycin, glucose starvation or in pathological N-glycosylation deficiencies such as congenital disorders of glycosylation (CDG).