Complementarity of electrophoretic, mass spectrometric, and gene sequencing techniques for the diagnosis and characterization of congenital disorders of glycosylation

Arnaud Bruneel, Sophie Cholet, Valérie Drouin-Garraud, Marie-line Jacquemont, Aline Cano, André Mégarbané, Coralie Ruel, David Cheillan, Thierry Dupré, Sandrine Vuillaumier-Barrot, et al.

Complementarity of electrophoretic, mass spectrometric and gene sequencing techniques for the screening and characterization of congenital disorders of glycosylation

Arnaud Bruneel[1,2], Sophie Cholet[3], Valérie Drouin-Garraud[4], Marie-Line Jacquemont[5], Aline Cano[6], André Mégarbané[7], Coralie Ruel[3,8], David Cheillan[9], Thierry Dupré[1], Sandrine Vuillaumier-Barrot[1], Nathalie Seta[1,10], François Fenaille[3].

[1]AP-HP, Biochimie Métabolique et Cellulaire, Hôpital Bichat-Claude Bernard, Paris, France

[2]INSERM UMR-1193 "Mécanismes cellulaires et moléculaires de l'adaptation au stress et cancérogenèse", Université Paris-Sud, Châtenay-Malabry, France

[3]Service de Pharmacologie et d'Immunoanalyse, Laboratoire d'Etude du Métabolisme des Médicaments, CEA, INRA, Université Paris Saclay, MetaboHUB, F-91191 Gif-sur-Yvette, France.

[4]Génétique Médicale, CHU Rouen, Rouen, France.

[5]Génétique Médicale, CHU - Groupe hospitalier Sud-Réunion, Saint-Pierre, La Réunion, France

[6]Centre de Référence des Maladies Héréditaires du Métabolisme, CHU la Timone-Marseille, Marseille, France.

[7] Institut Jérôme Lejeune, Paris, France.

[8]Proteins and Nanotechnology in Analytical Science (PNAS), CNRS, Université Paris-Sud, Châtenay-Malabry, France.

[9]Service de Biochimie et Biologie Moléculaire Grand Est, UM Pathologies Métaboliques, Erythrocytaires et Dépistage Périnatal; Centre de Biologie et de Pathologie Est, Groupement Hospitalier Est – Hospices Civils de Lyon, BRON, France.

[10]Paris Descartes University

Corresponding author:
Arnaud Bruneel
Arnaud.bruneel@aphp.fr
AP-HP, Biochimie Métabolique et Cellulaire
Hôpital Bichat,
46 rue Henri Huchard
75018, Paris - France.

Keywords: Congenital disorders of glycosylation (CDG); mass spectrometry; SLC35A2-CDG; SLC35A3-CDG; two-dimensional gel electrophoresis.

Abstract

Congenital disorders of glycosylation (CDG) are rare autosomal genetic diseases affecting the glycosylation of proteins and lipids. Since CDG-related clinical symptoms are classically extremely variable and non-specific, a combination of electrophoretic, mass spectrometric and gene sequencing techniques is often mandatory for obtaining a definitive CDG diagnosis, as well as identifying causative gene mutations and deciphering the underlying biochemical mechanisms. Here, we illustrate the potential of integrating data from capillary electrophoresis of transferrin, two-dimensional electrophoresis of N- and O-glycoproteins, mass spectrometry analyses of total serum N-linked glycans and mucin core1 O-glycosylated apolipoprotein C-III for the determination of various culprit CDG gene mutations. “Step-by-step” diagnosis pathways of four new CDG cases, including MGAT2-CDG, ATP6V0A2-CDG, SLC35A2-CDG and SLC35A3-CDG, are described as illustrative examples.

1. Introduction

Glycosylation is a complex post-translational modification of proteins and lipids taking place within the secretory pathway in endoplasmic reticulum (ER) and/or Golgi apparatus (GA). Protein glycosylation can be N- or O-linked depending on the attachment site to the protein backbone. N-linked glycans are attached to the amide group of an asparagine residue within an NX-S/T motif (X being any amino acid except proline), while O-linked glycans are commonly linked to the hydroxyl group of serine or threonine residues [1]. Congenital disorders of glycosylation (CDGs) are rare autosomal recessive inherited diseases sharing diverse and variable multi-organ clinical symptoms [2]. CDGs with abnormal protein N-glycosylation are classically sub-grouped as type I (CDG-I) or type II (CDG-II) according to the affected biosynthetic steps. In CDG-I, the defect alters the lipid-linked oligosaccharide synthesis or its transfer to protein backbone in the ER leading to under-occupancy of N-glycosylation sites (but with normally structured glycans). In CDG-II, the defect affects the maturation of protein-linked oligosaccharide in the GA leading to the inappropriate production of N-glycans intermediate motifs [3]. Since CDGs are accompanied by significant changes in glycoproteins electric charge and/or molecular weight (Mw), various electrophoresis methods can be used for the routine screening of these diseases. Isoelectric focusing of serum transferrin (Trf IEF) has been historically employed to detect CDGs [4]. Indeed, using either gel-based or capillary electrophoresis (CE) methods, Trf IEF assays are still considered as the methods of choice for CDG screening. By separating Trf glycoforms according to the number of negatively charged terminal sialic acid (SA) residues, it efficiently evidences and distinguishes tetra-, bi- and a-sialo Trf in CDG-I (two, one or zero bi-antennary mature N-glycan chains) and tetra-, tri-, bi-, mono- and a-sialo Trf in CDG-II (two bi-antennary chains lacking 0 to 2 terminal SA) [4, 5]. SDS-PAGE followed by Western-blot using specific antibodies targeting various serum N-glycoproteins can also be successfully

used to detect the loss of entire N-glycan chains linked to CDG-I but does not allow to detect the subtle Mw variations associated with the majority of CDG-II [6]. Besides these two techniques, we have recently reported that two-dimensional electrophoresis (2-DE), by coupling glycoform separations according to both their charge and Mw, can represent a particularly relevant tool for the detection of all types of CDGs [7]. Furthermore, 2-DE can be used to screen for O-glycosylation defects by monitoring the mucin core1 O-glycoprotein apolipoprotein C-III (apoC-III), thus enabling the detection and characterization of CDG-II related to protein defects impacting overall GA homeostasis [8].

Once evidenced using electrophoresis methods, CDG-I related enzymatic defects are now commonly determined using targeted genes or whole exome sequencing (WES) approaches [9]. In case of CDG-II-characteristic electrophoresis profile, since potentially involved molecular defects remain numerous and poorly understood, mass spectrometry (MS) techniques applied to enzymatically-removed N-glycans or to entire apoC-III constitute valuable tools for going deeper into glycan structural elucidation [8, 10]. The precise characterization of abnormal N- and O-glycan structures can provide precious clues for the delineation of related molecular deficiencies and further identification of 'culprit' gene mutations. Also, these glycomic tools can be very useful to corroborate suspected CDG-II causative mutations issued from targeted genes sequencing panels or WES data [9].

In this work, we illustrate the value and the necessity of combining electrophoresis, MS-based, and gene sequencing approaches for obtaining a definitive CDG diagnosis and getting deeper insight into the underlying biochemical mechanisms of type-II CDGs. This is exemplified in different manners using 4 distinct CDG-II patients intentionally chosen including MGAT2-CDG, ATP6V0A2-CDG, SLC35A2-CDG and SLC35A3-CDG.

2. Materials and Methods

2.1 Patient's samples

The 4 CDG patient serum samples analyzed in this work were sent to our laboratory for CDG screening (Pt.1 and Pt.2) or for targeted genes sequencing/WES data validation (Pt.3 and Pt.4). All results were compared to those from non-CDG patients ('controls').

2.2 CDG screening electrophoresis methods

SDS-PAGE Western-blot analyses of 3 serum glycoproteins i.e., transferrin (Trf), haptoglobin (Hpt) and acid α1-glycoprotein (AGP) were conducted as previously described [6]. Capillary electrophoresis based separation of serum transferrin glycoforms was carried out as previously described [11] using CZE method (Sebia Capillarys® CDT). Two-dimensional electrophoresis (2-DE) of serum N-glycoproteins i.e., α-antitrypsin (AAT), Trf and Hpt and of mucin core1 O-glycosylated apolipoprotein C-III were conducted as described in [7] and [8], respectively.

2.3 Mass spectrometry-based profiling of serum N-glycans

Sample processing for N-glycomic profiling of the serum samples was carried out essentially as described previously [12]. The samples (5µL) were diluted in 20 mM sodium phosphate buffer (pH 7.4) and 10 mM dithiothreitol solutions, and then heated at 95°C for 5 minutes. N-glycosidase F digestion was then performed overnight at 37°C (Roche Diagnostics, Meylan, France). After acidification, proteins were precipitated using ice-cold ethanol. Released N-glycans were purified using porous graphitic carbon solid phase extraction cartridges (Thermo Scientific, les Ulis, France). The native N-glycans were subsequently permethylated and purified on a C18 spin-column (Thermo Scientific) before analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The dried

permethylated samples were resuspended in 10 µL of a 50% methanol solution. 0.5 µL of the suspension was then spotted on the MALDI target and thoroughly mixed with 0.5 µL of 2,5-dihydroxybenzoic acid solution (10 mg/mL in 50% methanol containing 10 mM sodium acetate). Glycan analyses were performed on an UltrafleXtreme instrument (Bruker Daltonics, Bremen, Germany) operating in the reflectron positive ion mode. Manual assignment of glycan sequences was done from MS and MS/MS data on the basis of previously identified structures [12] and with the help of GlycoWorkBench software [13].

2.4. MALDI-TOF analysis of mucin core1 O-glycoforms of apoC-III

MALDI-TOF MS analysis of intact apoC-III was performed essentially as described before [8][14]. Briefly, 1 µL of serum was diluted in 15 µL of water/acetonitrile (95:5) containing 0.1% trifluoroacetic acid (TFA), allowed to stand 1h at room temperature and then purified/desalted by ZipTip C4 (Merck Millipore, Darmstadt, Germany). ApoC-III was eluted with 5 µL of 70% acetonitrile containing 0.1% TFA, 0.5 µL of this solution was then thoroughly mixed on-target with 0.5 µL of a saturated solution of sinapinic acid in water/acetonitrile (50:50) containing 0.1% TFA. MS analyses were performed on a Bruker Ultraflextreme MALDI-TOF/TOF instrument (Bruker Daltonics, Bremen, Germany) equipped with a smartbeam-II laser. Mass spectra of intact apoC-III were collected at 2 kHz laser repetition rate in the positive linear ion mode, using a 20 kV acceleration voltage and an extraction delay of 250 ns. All spectra were obtained by accumulating ~1000 laser shots over the 5000-20000 *m/z* range.

2.5 Molecular studies

Direct Sanger sequencing of MGAT2 and ATP6V0A2 genes was done respectively for Pt.1 and Pt.2 since electrophoretic and MS profiles, as well as clinical data (Pt.2), were pointing to

these genes. For Pt.3, targeted sequencing was performed on MiSeq sequencer (Illumina) with a custom kit of 17 CDG-II genes (Agilent). For Pt.4, WES was performed on NextSeq (Illumina) with a Med-Exome kit (Nimblegen) and analyzed with Polyweb Imagine (www.polyweb.fr). Filtering of WES data was done as follows: homozygote variants (consanguineous parents); <1% frequency; supposed recessive mode of inheritance; variant not known in dbSNP; deleterious as determined with SIFT and Polyphen. Sanger sequencing for all index cases and their parents confirmed all identified causative variants.

3. Results

3.1 Brief clinical description of CDG patients

The CDG patients were 3 girls (Pt.1 to Pt.3) and one boy (Pt.4). Only Pt.1 was issued from consanguineous parents. Ages at diagnosis ranged from 7 months to 14 years old. Since "all is clinically possible in CDG", the majority of reported clinical symptoms were rather unspecific, including severe to moderate mental retardation, dysmorphic features, failure to thrive (all patients), microcephaly (Pt.1, Pt.2 and Pt.4), cardiac malformations (Pt.2 and Pt.3), encephalopathy (Pt.2 and Pt.4), hyperlaxity (Pt.2 and Pt.3) and epilepsy (Pt.4). Some more specific additional clinical symptoms, such as large anterior fontanel and *Cutis laxa* (skin anomaly with redundant skinfolds and abnormal elasticity) were reported in Pt.2 (Supporting Information Fig. S1).

3.2 SDS-PAGE Western-blots of N-glycoproteins and transferrin CE patterns

While 1D SDS-PAGE Western-blot analysis could not distinguished Pt.2 to Pt.4 from the healthy subject (not shown), profiles of Trf, AAT and AGP from Pt.1 both harbored markedly lower bands with absence of expected ones (Fig.1A). Since the Mw differences could be easily detected by a rather poorly resolving method, these results strongly suggested an unusual complete deficiency in the early steps of the enzymatic maturation of N-glycans, after the transfer of the oligosaccharide chain. Also, a CDG-I was rapidly excluded since profiles did not evoke absence of entire N-glycan chains that would typically result in additional lower mass bands co-existing with the normal ones (Fig.1A).

When mainly separated according to electric charge using CE, Trf profiles of Pt.2 to Pt.4 both showed more or less elevated proportions of hyposialylated glycoforms (3-sialo to 0-sialo) highly suggestive of CDG-II (Fig.1B). Noticeably, CE Trf profile of Pt.3 showed rather normal proportions of the 3- and 2-sialo glycoforms but discreetly elevated levels of the 1-

sialo and 0-sialo- glycoforms, always absent in samples from control subjects. For Pt.4, 3- and 2-sialo glycoforms percentages were clearly elevated while 1-sialo and 0-sialo-glycoforms were undetectable. Due to a lack of serum sample, this analysis could unfortunately not be performed for Pt.1. In this particular case, analysis by 2-dimensional electrophoresis was privileged over CE analysis of Trf.

3.3 Two-dimensional electrophoresis (2-DE) profiles of N- and O-glycoproteins

When compared to control (Fig.2A), 2-DE profile of Pt.1 showed high levels of abnormal less acidic glycoforms of markedly lower masses for all the tested serum N-glycoproteins (Fig.2B), probably highlighting the presence of presumably incomplete and partially sialylated glycan structures. 2-DE patterns of Pt.2 to Pt.4 (Fig.2C-E) all exhibited less acidic (i.e., less sialylated) additional glycoforms ('cathodic shift' but without significant Mw differences) but still co-existing with important proportions of correctly glycosylated protein species. By contrast to Pt.2 and Pt.4 who showed relatively discrete abnormalities, Pt.3 profile was characterized by several additional abnormal spots (with up to 5 additional spots for Hpt) (Fig.2D). For Pt.4 (Fig.2E), Hpt spots were not detectable (as a probable consequence of intra-vascular hemolysis) depriving us from this very sensitive 2-DE biomarker of CDG [7].

Concerning apoC-III mucin core1 O-glycosylation, dedicated 2-DE patterns (Fig.3) showed abnormalities only for Pt.2 and Pt.3. The apoC-III pattern of Pt.1 (Fig.3B) could not be distinguished from the control (Fig.3A) indicating that O-glycosylation was not affected in this individual. In Pt.2 (Fig.3C), the level of the bi-sialylated glycoform (substituted with N-acetylgalactosamine-galactose carrying 2 SA) proved severely decreased. When detected along N-glycans defects, such marked abnormality tended to suggest a global disturbance in GA homeostasis. For Pt.3 (Fig.3D), an unusual asialylated apoC-III glycoform could be detected at a rather low level (this particular species is always absent in control individuals)

suggesting a partial defect in the linkage of the galactose (Gal) and/or the first sialic acid (SA) residue. ApoC-III pattern of Pt.4 (Fig.3E) appeared normal suggesting that only the GA-located N-glycans maturation pathway was affected in this individual. The same conclusions can be drawn from the MALDI-TOF analyses of apoC-III (Supporting Information Fig. S2), while no specific additional information was brought by this MS-based technique.

3.4 MS studies of serum N-glycans

MALDI-TOF analysis of N-linked glycans resulted in the detection of almost 50 distinct mass signals in the serum sample from the healthy subject (Fig.4A), corresponding to the generally recognized glycan structures [10, 15] with bi-antennary bi-sialylated structures (*m/z* 2792.4) as the most abundant species accounting for ~30% of the total N-glycan pool. Besides, bi-antennary mono-sialylated (*m/z* 2431.2), and fully sialylated tri-antennary and tetra-antennary N-glycans (*m/z* 3602.8 and *m/z* 4413.1, respectively) were also identified among the protruding structures (Fig.4A). MALDI-TOF profiles of permethylated total serum N-glycans were obtained for the 4 patients and compared with that of the healthy control. The spectrum of Pt.1 (Fig.4B) was highly characteristic with mono-antennary sialylated N-glycans as most abundant species (*m/z* 1981.9 and 2156). The presence of such highly dominant N-glycan structures lacking one antenna was highly evocable of deficiency in mannosyl glycoprotein N-acetylglucosaminyl transferase II (MGAT2). N-glycan profiles of Pt.2 and Pt.3 (Fig.4C-D) shared some similarities with accumulation of partially sialylated bi-antennary N-glycans (*m/z* 2431.2 and *m/z* 2605.3), of N-glycans lacking both terminal SA and galactose residues (*m/z* 1835.9, *m/z* 2040.0 and *m/z* 2227.1), and of oligomannose structures (*m/z* 1579.8 and *m/z* 1783.9). Pt.4 presented a rather atypical N-glycan pattern (Fig.4E). Although this sample predominantly contained bi-antennary mono- and bi-sialylated species as in samples from healthy subjects, it showed abundant mono-antennary sialylated structure at *m/z* 1981.9 as

well as oligomannose structures at *m/z* 1579.8 and *m/z* 1783.9. Furthermore, the total absence of polyantennary structures at *m/z* 3241.8 and *m/z* 3602.8, whose synthesis is initiated by linkages of N-acetylglucosamine (GlcNAc) residues, was also noticeable in this sample.

3.5 Molecular studies

Pt.1 carried a homozygote deletion of one base (c.693delA) on MGAT2 encoding gene in tight agreement with electrophoresis and MS studies of glycoproteins and N-glycans. Both consanguineous parents were heterozygote. Pt.2 carried a non-sense homozygote variant (c. p.Val44X) on ATP6V0A2 encoding gene. Both non-consanguineous parents were heterozygote. For Pt.3, a missense heterozygote variant (c.262G>C) was found on *SLC35A2* gene on one chromosome X, *de novo* inherited, as it was absent in parents. In this girl, normal X chromosome was determined to be totally inactivated. *SLC35A2* encodes for UDP-Gal transporter, which is required for in the entry of activated galactose into the GA lumen [16]. For Pt.4, WES was performed and mutations on the 3 following genes were filtered: *SLC35A3* encoding for the UDP-GlcNAc transporter, *PTPN23* (ciliogenesis control) and *TBX22* (transcription factor). Presence of a missense homozygote variant (c.797A>T p.Asp266Val) on *SLC35A3* was retained as causative. Both parents were heterozygote. The protein product of this gene is necessary for the entry of activated GlcNAc into the GA [16]. In summary, these genetic results led to the unambiguous diagnosis of MGAT2-CDG (Pt.1), ATP6V0A2-CDG (Pt.2), SLC35A2-CDG (Pt.3) and SLC35A3-CDG for Pt.4.

4. Discussion

Pt.1 showed non-specific but severe neurological and multi-organ clinical symptoms motivating CDG screening. SDS-PAGE Western-blot and 2-DE profiles of serum glycoproteins from this patient proved extremely altered with almost the complete loss of complex sialylated oligosaccharide structures leading to the appearance of unusual species with both markedly lower Mw and cathodic shift when compared to control individuals. Since analysis of apoC-III glycoform profile proved normal, we assumed that the deficiency did not alter mucin core1 O-glycosylation but was probably restricted to N-glycosylation. MS analysis of serum N-glycans confirmed our preliminary assumption and demonstrated a strong accumulation of a mono-antennary mono-sialylated glycan lacking one [GlcNAc-Gal-SA] trisaccharidic arm (+/- fucosylated; *m/z* 1981.9 and 2156.1) (Supporting Information Fig. S3). Such particular N-glycan accumulation can be caused by mutation(s) in the *MGAT2* gene encoding UDP-GlcNAc:α6-D-mannoside β1,2-N-acetylglucosaminyltransferase II. This enzyme localized in the Golgi membrane adds the second N-acetylglucosamine of bi-antennary complex-type chains, the most common N-glycans in serum glycoproteins. Molecular sequencing of the corresponding gene unambiguously corroborated the diagnosis of MGAT2-CDG. Noticeably, this particular type-II CDG was the first one described in 1990 [17]. Trace amounts of bi-antennary bi-sialylated N-glycans (*m/z* 2792.4) indicated the almost complete inactivation of this enzyme, which is in sharp contrast with the majority of other type-II CDGs where causative mutations typically preserved significant residual enzyme/protein activity [2].

Pt.2 presented with non-specific and moderate neurological and dysmorphic features associated to a peculiar skin disease i.e., '*Cutis laxa*' (Supporting Information Fig. S1). In this individual, CE profile of serum Trf demonstrated a characteristic CDG-II profile with high levels of hyposialylated glycoforms. 2-DE of serum N-glycoproteins and apoC-III both

showed abnormal hyposialylation, suggesting a global defect in N- and mucin core1 O-glycosylation pathways. MALDI-TOF MS analysis of N-linked glycans corroborated the abnormal accumulation of mono-sialylated species (*m/z* 2431.2 and 2605.3) but also revealed the presence of truncated structures lacking galactose residues on one or both arms (*m/z* 1835.9; 2040.0; 2227.1). Altogether, these abnormalities suggested a global defect of GA, and more precisely of terminal parts of this organelle where galactosylation and sialylation of N- and mucin core1 O-glycans occur [1]. Among CDG linked to GA homeostasis disturbances, emerging ones are notably COG-CDG (Conserved Oligomeric Golgi) [18] and ATP6V0A2-CDG [19]. Since ATPV0A2 gene encodes the α2 subunit of the vesicular ATPase, which plays a pivotal role in GA pH regulation through proton translocation [20], its association with N- and O-glycosylation defects was presumably linked to abnormal pH in medial and trans-Golgi, thus impacting the enzymatic activity of galactosyl- and sialyltransferases. Furthermore, to our knowledge, all ATP6V0A2-CDG cases described to date shared clinical *cutis laxa* [21]. For all these reasons, we sequenced the *ATP6V0A2* gene and formally diagnosed ATP6V0A2-CDG in this individual.

In Pt.3, CE-based Trf analysis showed an atypical CDG-II pattern with subtle but significant levels of asialylated and monosialylated glycoforms (absent in the serum of control individuals). 2-DE analysis of the N-glycoproteins corroborated the presence of hyposialylated glycoforms with numerous small additional cathodical spots. Although not evident on the corresponding MALDI-TOF mass spectrum, 2-DE analysis of apoC-III mucin core1 O-glycosylation revealed an abnormal tiny spot potentially corresponding to asialylated apoC-III harboring GalNAc-Gal disaccharide (hyposialylation) and/or only one GalNAc residue (hypogalactosylation). MS profile of N-glycans shared similarities with that of Pt.2 in terms of hyposialylation (*m/z* 2431.2) and hypogalactosylation (*m/z* 1661.8; 1835.9; 2040.0; 2227.1). Nevertheless, hyposialylation looked less marked in Pt.3, notably with the absence of

the oligosaccharidic structure at *m/z* 2605.3. Furthermore, the glycan moiety lacking the 2 sub-terminal galactoses (*m/z* 1661.8) was specifically retrieved in Pt.3 suggesting a higher level of hypogalactosylation of N-glycans. These results, when combined with non-specific clinical symptoms, did not evidently suggest a culprit CDG-related gene. COG subunits as well as ATP6V0A2 encoding genes were sequenced but gene sequencing did not evidence any mutation. Then, sequencing of a panel of targeted genes was performed and revealed one *de novo* mutation in SLC35A2 gene carried by one X chromosome of the related girl (the second one being completely inactivated). In light of this result, N-glycan data could be re-interpreted and, as described above, observed levels of hypogalactosylation, putatively associated to hypogalactosylation of apoC-III, finally appeared highly coherent with the diagnosis of SCL35A2-CDG. Indeed, SLC35A2 protein corresponds to the Golgi UDP-Gal transporter allowing the entry of 'activated' galactose into the GA lumen before its addition to N- and O-oligosaccharide elongating chains (Supporting Information Fig. S3) [16]. Since nucleotide sugar transporters have been suggested to cooperate with each other [22], it could be speculated that SLC35A2 mutation could also impact CMP-SA transporter (SLC35A1) functions in Pt.3 putatively explaining the observed associated hyposialylation.

Pt.4 showed atypical CDG-II Trf CE profile with high levels of 3- and 2-sialo glycoforms but without detectable 1- and 0-sialo glycoforms. 2-DE of N-glycoproteins further confirmed hyposialylation of Trf but proved poorly informative for AAT (one faint additional spot) while Hpt was undetectable. 2-DE of apoC-III did not evidence any specific abnormalities suggesting that observed Trf defects did not translate into mucin core1 O-glycosylation. By contrast with 2-DE, MS-based analysis of serum N-glycans proved highly informative revealing strong accumulation of mono-antennary mono-sialylated structures lacking one complete [GlcNAc-Gal-SA] arm (*m/z* 1981.9 and 2156.1). Strikingly, these MGAT2-CDG-typical structures (as found in Pt.1) were here associated to significant levels of normal

complex-type biantennary bi-sialylated N-glycan species at *m/z* 2792.4 and 2966.4 (at very low levels in Pt.1) (Supporting Information Fig. S3). In addition, atypical hybrid and oligomannose structures at *m/z* 2390.2, 1783.9 and 1579.8 were also detected at significant levels suggesting an associated deficiency in mannosidase II enzymatic activity (Supporting Information Fig. S3). Lastly, total absence of tri-antennary structures at *m/z* 3241.8 and 3602.8 was also evidenced in Pt.4. Considering these rather complex glycosylation profiles associated with non-specific clinical symptoms, no immediate conclusion regarding any particular causative gene mutations could be drawn. Sequencing of MGAT2 and mannosidase II encoding genes, as well as targeted genes panel sequencing, were performed but both proved unsuccessful. Then, a WES was performed and interestingly revealed relevant mutations in SLC35A3 gene encoding for the Golgi UDP-N-acetylglucosamine (UDP-GlcNAc) transporter. Since UDP-GlcNAc is the substrate of *MGAT2*-encoded enzyme and of other GlcNAc transferases involved in the synthesis of polyantennary N-glycans (Supporting Information Fig. S3) [16], observed accumulations of mono-antennary sialylated N-glycans at *m/z* 1981.9 and 2156.1 along with the absence of polyantennary structures, were first strong arguments for causative SLC35A3 mutations. Furthermore, the absence of any mucin core1 O-glycosylation defect on apoC-III (glycan motifs devoided of any GlcNAc residues), as well as the reported interaction of UDP-GlcNAc transporter with mannosidase II and various GlcNAc transferases [23, 24], were in strong agreement with our experimental results. Altogether these data confirmed the unambiguous diagnosis of SLC35A3-CDG in Pt.4. To our knowledge, this patient is only the second case (among 10 described patients) of SLC35A3-CDG presenting with clear serum glycans abnormalities [25, 26].

In this paper, we show through four specific examples that a panel of complementary tools is often mandatory to obtain a definitive CDG diagnosis, and for getting deeper insight into the

underlying biochemical mechanisms. Thus, glycomic tools (from electrophoresis to MS) and gene sequencing approaches represent essential analytical partners in the intricate diagnosis pathway of these inherited diseases. Capillary and/or gel electrophoresis techniques can be regarded as efficient first-line approaches for CDG screening. In case of CDG suspicion, MS-based analyses can be performed to confirm the preliminary diagnosis but also to highlight potential abnormal accumulation of particular glycan species. Targeted gene sequencing can be used for validating the potentially incriminated genes. If no obvious or consistent hypothesis can be drawn from the electrophoretic and mass spectrometric data, whole exome sequencing would be performed to get new perspectives.

Acknowledgements

This work was supported by grant ANR-15RAR3-0004-06 under the frame of E-RARE-3, the ERA-Net for Research on Rare Diseases. This work was also supported by the Commissariat à l'Energie Atomique et aux Energies Alternatives and the MetaboHUB infrastructure (ANR-11-INBS-0010 grant).

Figure legends:

Figure 1: Western-blot of N-glycoproteins and CE of serum transferrin, A) SDS-PAGE Western-blot analysis of serum transferrin (Trf), haptoglobin (Hpt) and acid α1-glycoprotein (AGP) from (left to right) 2 controls (T), Pt.1 and CDG-I. **B)** Capillary electrophoresis serum transferrin patterns of one control, Pt.2, Pt.3 and Pt.4.

Figure 2: 2-DE patterns of N-glycoproteins

2-DE patterns of serum α1-antitrypsin (AAT), haptoglobin (Hpt) and transferrin (Trf) from one control, Pt.1, Pt.2, Pt.3 and Pt.4. In all patterns, non-glycosylated albumin (Alb) can be detected in the 'background' constituting a useful landmark (vertical dotted line) for accurate visual interpretation. Plain arrows indicate additional abnormal spots corresponding to hyposialylated glycoforms. Vertical dotted arrow indicates the tetra-sialylated (4-sialo) transferrin glycoform.

Figure 3: 2-DE patterns of mucin core1 O-glycosylated apoC-III

2-DE patterns of serum apoC-III from one control, Pt.1, Pt.2, Pt.3 and Pt.4. In control pattern, apoC-III glycoforms structures corresponding to each spot are detailed: square, N-acetylgalactosamine; circle, galactose; diamond square, sialic acid. In pattern of Pt.3, an unusual spot is indicated (arrow), probably corresponding to one between two putative schematized glycan structures. Dotted line (in control and Pt.3) is a landmark for accurate interpretation of spots corresponding to asialylated structures. Vertical dotted arrow in profile of Pt.2 schematizes the decreased percentage of di-sialylated apoC-III.

Figure 4: MALDI-TOF mass spectra of permethylated N-glycans released from serum samples.

Measurements were performed in the positive-ion mode and all ions are present in sodiated form. Green circles, mannose; yellow circles, galactose; blue squares, N-acetyl glucosamine; red triangles, fucose; purple diamonds, sialic acid.

Legends of supplementary Figures

Supporting Information Figure S1. Picture of abdominal '*Cutis laxa*' as observed in patient 2.

Supporting Information Figure S2. MALDI-TOF mass spectra of apoC-III. Species were detected as [M+H]+ ions. For patients 2 and 3, protein species were detected under their oxidized form (ox). Yellow squares, GalNAc; yellow circles, Gal; purple diamonds, Neu5Ac.

Supporting Information Figure S3. Simplified view of GA-located N-glycan biosynthesis pathway (maturation of protein-linked N-glycans). Given *m/z* values are those of the corresponding permethylated N-glycans as measured by MALDI-TOF MS ([M + Na]+ species). Numbered stars are related to the four patients studied (plain stars: accumulation; channeled stars: absence).

6. References

1. Moremen, K.W., Tiemeyer, M., Nairn, A.V., Nat. Rev. Mol. Cell. Biol. 2012, 13(7), 448-62.

2. Peanne, R., de Lonlay, P., Foulquier, F., Kornak, U., Lefeber, D.J., Morava, E., Perez, B., Seta, N., Thiel, C., Van Schaftingen, E., Matthijs, G., Jaeken, J., Eur. J. Med. Genet. 2017.

3. Jaeken, J., Ann. NY. Acad. Sci. 2010, 1214, 190-8.

4. Jaeken, J., van Eijk, H.G., van der Heul, C., Corbeel, L., Eeckels, R., Eggermont, E., Clin. Chim. Acta 1984, 144(2-3), 245-7.

5. Lefeber, D.J., Morava, E., Jaeken, J., J. Inherit. Metab. Dis. 2011, 34(4), 849-52.

6. Seta, N., Barnier, A., Hochedez, F., Besnard, M.A., Durand, G.,. Clin. Chim. Acta 1996, 254(2), 131-40.

7. Bruneel, A., Habarou, F., Stojkovic, T., Plouviez, G., Bougas, L., Guillemet, F., Brient, N., Henry, D., Dupre, T., Vuillaumier-Barrot, S., Seta, N., Clin. Chim. Acta 2017, 470, 70-74.

8. Yen-Nicolay, S., Boursier, C., Rio, M., Lefeber, D.J., Pilon, A., Seta, N., Bruneel, A., Proteomics Clin. Appl. 2015, 9(7-8), 787-93.

9. Timal, S., Hoischen, A., Lehle, L., Adamowicz, M., Huijben, K., Sykut-Cegielska, J., Paprocka, J., Jamroz, E., van Spronsen, F.J., Korner, C., Gilissen, C., Rodenburg, R.J., Eidhof, I., Van den Heuvel, L., Thiel, C., Wevers, R.A., Morava, E., Veltman, J., Lefeber, D.J., Hum. Mol. Genet. 2012, 21(19), 4151-61.

10. Guillard, M., Morava, E., van Delft, F.L., Hague, R., Korner, C., Adamowicz, M., Wevers, R.A., Lefeber, D.J., Clin. Chem. 2011, 57(4), 593-602.

11. Parente, F., Ah Mew, N., Jaeken, J., Gilfix, BM., Clin. Chim. Acta 2010, 411(1-2), 64-6.

12. Goyallon, A., Cholet, S., Chapelle, M., Junot, C., Fenaille, F., Rapid Commun. Mass Spectrom. 2015, 29(6), 461-73.

13. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., Haslam, S.M., J. Proteome Res. 2008, **7**(4), 1650-9.

14. Ruel, C., Morani, M., Bruneel, A., Junot, C., Taverna, M., Fenaille, F., Tran, N.T., J. Chromatogr. A, 2018, 1532, 238-245.

15. Biskup, K., Braicu, E.I., Sehouli, J., Fotopoulou, C., Tauber, R., Berger, M., Blanchard, V., J. Proteome Res. 2013, 12(9), 4056-63.

16. Caffaro, C.E., Hirschberg, C.B., Acc. Chem. Res. 2006, 39(11), 805-12.

17. Stibler, H., Jaeken, J., Arch. Dis. Child 1990, 65(1), 107-11.

18. Zeevaert, R., Foulquier, F., Jaeken, J., Matthijs, G., Mol. Genet. Metab. 2008, 93(1), 15-21.

19. Kornak, U., Reynders, E., Dimopoulou, A., van Reeuwijk, J., Fischer, B., Rajab, A., Budde, B., Nurnberg, P., Foulquier, F., Lefeber, D., Urban, Z., Gruenewald, S., Annaert, W., Brunner, H.G., van Bokhoven, H., Wevers, R., Morava, E., Matthijs, G., Van Maldergem, L., Mundlos, S., Nat. Genet. 2008, 40(1), 32-4.

20. Guillard, M., Dimopoulou, A., Fischer, B., Morava, E., Lefeber, D.J., Kornak, U., Wevers, R.A., Biochim. Biophys. Acta 2009, 1792(9), 903-14.

21. Mohamed, M., Kouwenberg, D., Gardeitchik, T., Kornak, U., Wevers, R.A., Morava, E., J. Inherit. Metab. Dis. 2011, 34(4), 907-16.

22. Hadley, B., Maggioni, A., Ashikov, A., Day, C.J., Haselhorst, T., Tiralongo, J., Comput. Struct. Biotechnol. J. 2014, 10(16), 23-32.

23. Slusarewicz, P., Nilsson, T., Hui, N., Watson, R., Warren, G., J. Cell Biol. 1994 124(4), 405-13.

24. Maszczak-Seneczko, D., Sosicka, P., Kaczmarek, B., Majkowski, M., Luzarowski, M., Olczak, T., Olczak, M., J. Biol. Chem. 2015, 290(25), 15475-86.

25. Edvardson, S., Ashikov, A., Jalas, C., Sturiale, L., Shaag, A., Fedick, A., Treff, N.R., Garozzo, D., Gerardy-Schahn, R., Elpeleg, O., J. Med. Genet. 2013, 50(11), 733-9.

26. Edmondson, A.C., Bedoukian, E.C., Deardorff, M.A., McDonald-McGinn, D.M., Li, X., He, M., Zackai, E.H., Am. J. Med. Genet. A 2017, 173(10), 2758-2762.

Figure 1

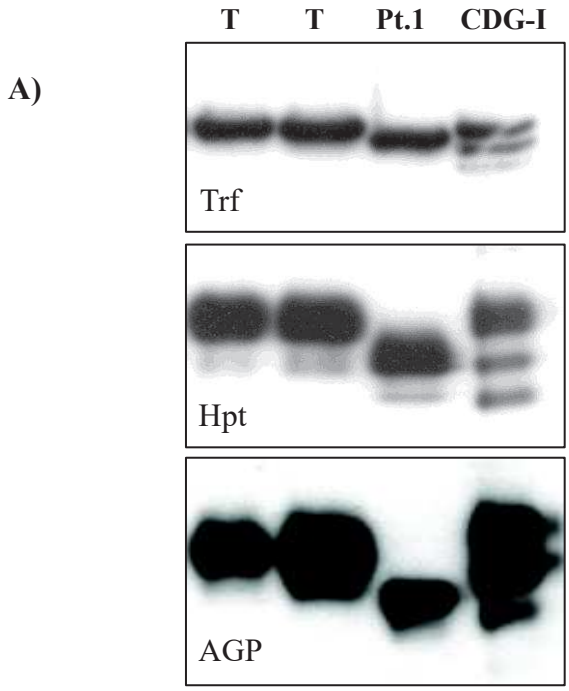

A)
T
T
Pt.1
CDG-I
Trf
Hpt
AGP
B)
Control
5-sialo
4-sialo
3-sialo
2-sialo
Pt.2
1-sialo
Pt.3
0-sialo
Pt.4

Figure 2

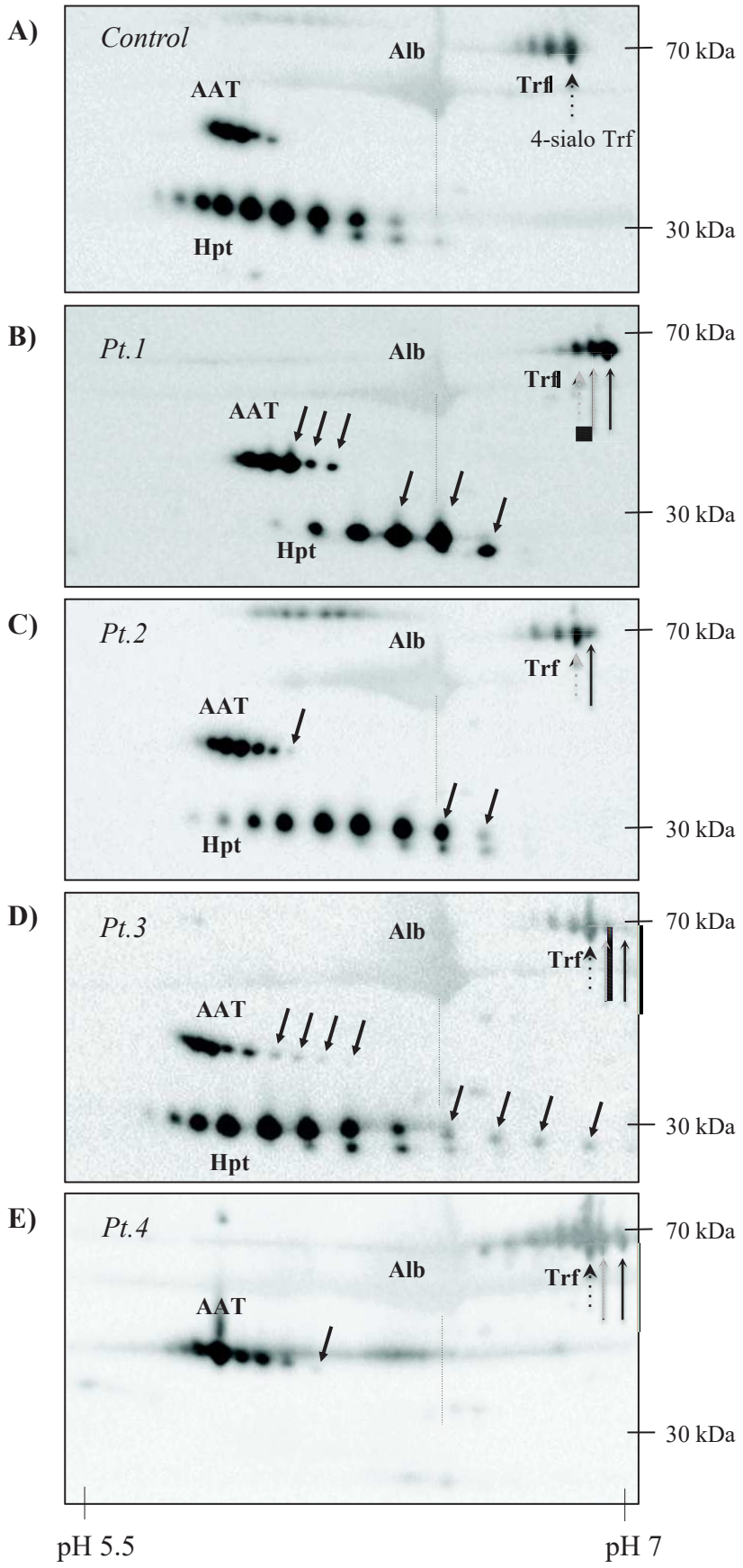

A)
Control
Alb
AAT
Trfl
4-sialo Trf
Hpt
70 kDa
30 kDa
B)
Pt.1
Alb
AAT
Trfl
Hpt
70 kDa
30 kDa
C)
Pt.2
Alb
AAT
Trf
Hpt
70 kDa
30 kDa
D)
Pt.3
Alb
AAT
Trf
Hpt
70 kDa
30 kDa
E)
Pt.4
Alb
AAT
Trf
70 kDa
30 kDa
pH 5.5
pH 7

Figure 3

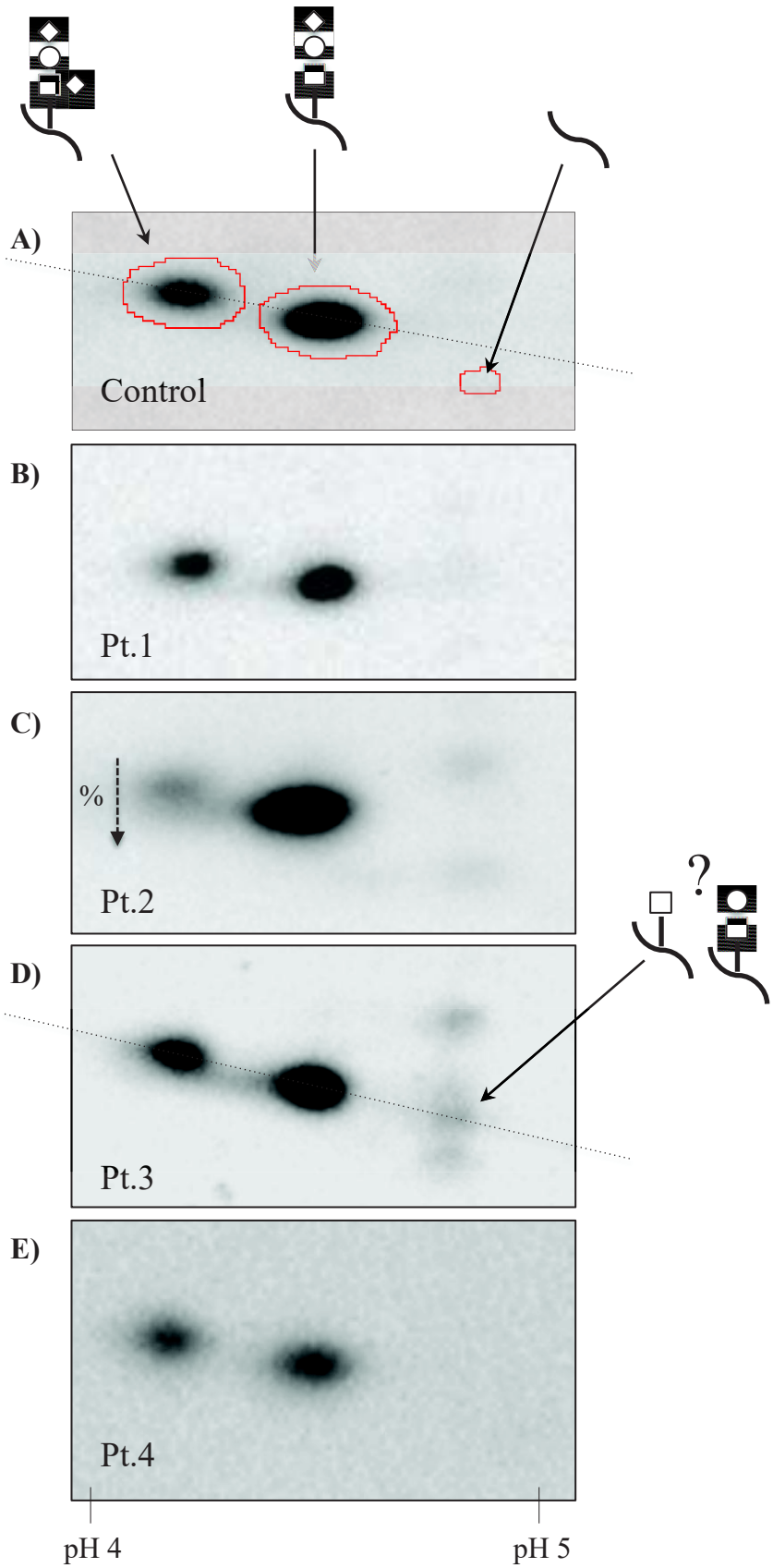
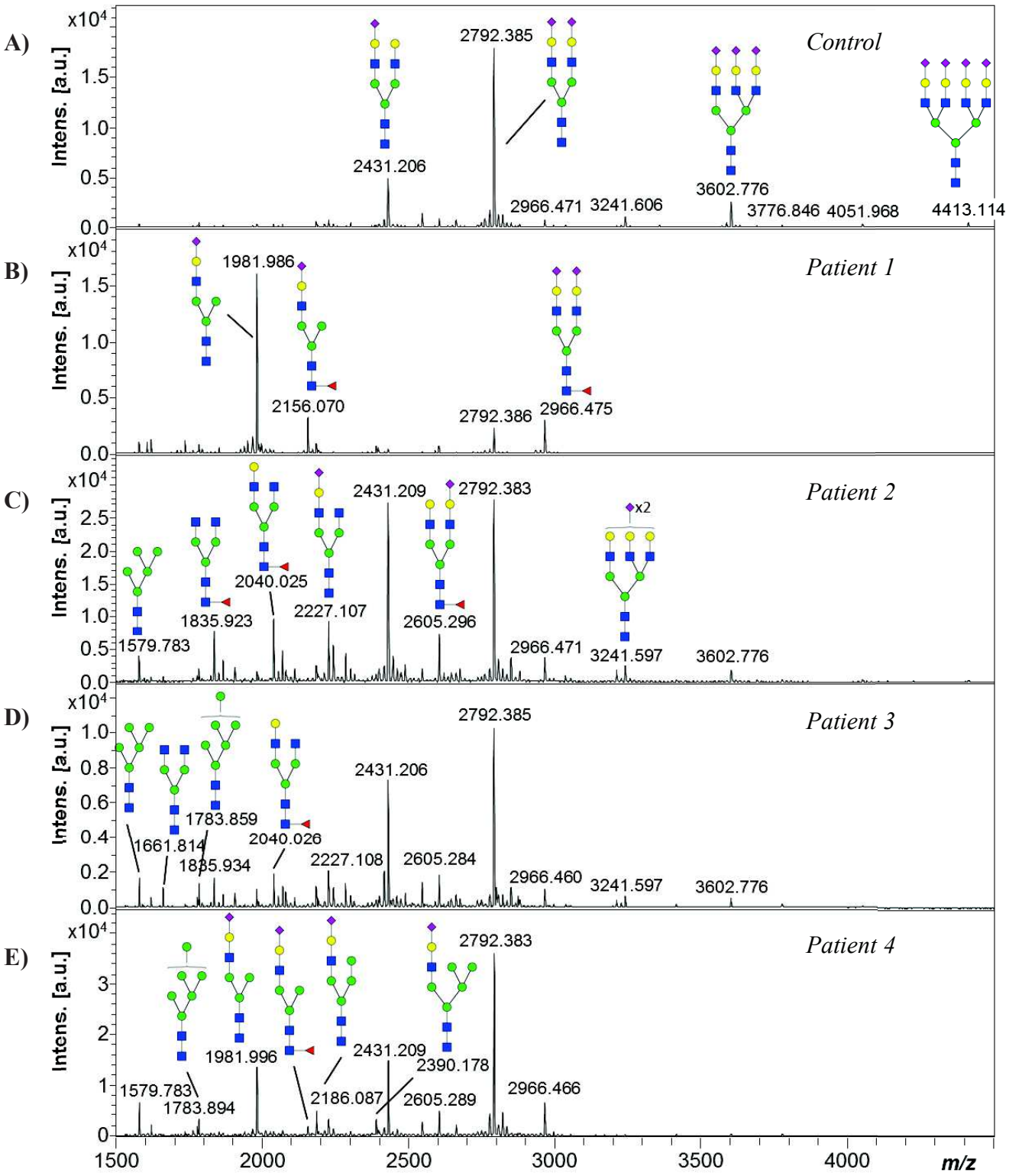


Figure 4


A)
Control
Intens. [a.u.]
x10^4
2792.385
2431.206
3602.776
2966.471
3241.606
3776.846
4051.968
4413.114
B)
Patient 1
1981.986
2156.070
2792.386
2966.475
C)
Patient 2
2431.209
2792.383
x2
2040.025
1835.923
2227.107
2605.296
1579.783
2966.471
3241.597
3602.776
D)
Patient 3
2792.385
2431.206
1783.859
1661.814
2040.026
1835.934
2227.108
2605.284
2966.460
3241.597
3602.776
E)
Patient 4
2792.383
1981.996
2431.209
2390.178
1579.783
1783.894
2186.087
2605.289
2966.466
1500
2000
2500
3000
3500
4000
m/z