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

Received January 15, 2018
Revised May 9, 2018
Accepted May 25, 2018

Research Article

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

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 nonspecific, 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 particular and new CDG cases, including MGAT2-CDG, ATP6V0A2-CDG, SLC35A2-CDG, and SLC35A3-CDG, are described as illustrative examples.

Keywords:
2DE / congenital disorders of glycosylation / MS / SLC35A2-CDG / SLC35A3-CDG
DOI 10.1002/elps.201800021

Additional supporting information may be found online in the Supporting Information section at the end of the article.

1 Introduction

Glycosylation is a complex posttranslational 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 multiorgan 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 ER or in

Correspondence: Dr. Arnaud Bruneel, AP-HP, Biochimie Métabolique et Cellulaire, Hôpital Bichat, 46 rue Henri Huchard, F-75018 Paris, France
Fax: +33-1-40-25-88-21
E-mail: arnaud.bruneel@aphp.fr

Abbreviations: AAT, α-antitrypsin; **AGP**, acid α1-glycoprotein; **apoC-III**, apolipoprotein C-III; **CDG**, congenital disorders of glycosylation; **ER**, endoplasmic reticulum; **GA**, golgi apparatus; **Hpt**, haptoglobin,; **MALDI-TOF MS**, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; **MGAT2**, mannosyl glycoprotein *N*-acetylglucosaminyl transferase II; **Mw**, molecular weight; **SA**, sialic acid; **TFA**, trifluoroacetic acid; **Trf**, transferrin; **Trf IEF**, isoelectric focusing of serum transferrin; **UDP-GlcNAc**, Uridine diphosphate *N*-acetylglucosamine ; **WES**, whole exome sequencing

*Additional corresponding author: Dr. François Fenaille
E-mail: francois.fenaille@cea.fr

Color Online: See the article online to view Figs. 3–5 in color.

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 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 zero to two 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 2DE, 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, 2DE 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 the case of CDG-II-characteristic electrophoresis profile, since potentially involved molecular defects remain numerous and poorly understood, 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].

As discussed above, CDG are often difficult to clinically and accurately diagnose because they can affect many organs and functions, and lack symptom uniformity. In this study, 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 four distinct but representative and new 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 four CDG patient serum samples analyzed in this work were sent to our laboratory for CDG screening (Patients 1 and 2) or for targeted genes sequencing/WES data validation (Patients 3 and 4). All results were compared to those from non-CDG patients ('controls').

2.2 CDG screening electrophoresis methods

SDS-PAGE Western-blot analyses of three 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). 2DE of serum *N*-glycoproteins, i.e., α-antitrypsin (AAT), Trf, and Hpt, and of mucin core1 *O*-glycosylated apoC-III, were conducted as described in [7] and [8], respectively.

2.3 MS-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 (5 μL) were diluted in 20 mM sodium phosphate buffer (pH 7.4) and 10 mM dithiothreitol solutions, and then heated at 95°C for 5 min. *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 MALDI-TOF–MS. The dried permethylated samples were resuspended in 10 μL of a 50% methanol solution. A total of 0.5 μL 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 1 h 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 Patient 1 and Patient 2 since electrophoretic and MS profiles, as well as clinical data (Patient 2), were pointing to these genes. For Patient 3, targeted sequencing was performed on MiSeq sequencer (Illumina) with a custom kit of 17 CDG-II genes (Agilent). For Patient 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; and 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 three girls (Patient 1–3) and one boy (Patient 4). Only Patient 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 (Patients 1, 2 and 4), cardiac malformations (Patients 2 and 3), encephalopathy (Patients 2 and 4), hyperlaxity (Patients 2 and 3), and epilepsy (Patient 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 Patient 2 (Supporting Information Fig. 1).

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

While 1D SDS-PAGE Western-blot analysis could not distinguish Patients 2–4 from the healthy subject (Supporting Information Fig. 2), profiles of Trf, Hpt, and AGP from Patient 1 harbored markedly lower bands in the 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 Patients 2–4 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 Patient 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 Patient 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 Patient 1. In this particular case, analysis by 2DE was privileged over CE analysis of Trf.

3.3 2DE profiles of *N*- and *O*-glycoproteins

When compared to control (Fig. 2A), 2DE profile of Patient 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. 2DE patterns of Patients 2–4 (Fig. 2C–E) all exhibited less acidic (i.e., less sialylated) additional glycoforms ('cathodic shift' but without significant Mw differences) but still coexist with important proportions of correctly glycosylated protein species. By contrast to Patient 2 and Patient 4, who showed relatively discrete abnormalities, Patient 3 profile was characterized by several additional abnormal spots (with up to five additional spots for Hpt; Fig. 2D). For Patient 4 (Fig. 2E), Hpt spots were not detectable (as a probable consequence of intravascular hemolysis) depriving us from this very sensitive 2DE biomarker of CDG [7].

Concerning apoC-III mucin core1 *O*-glycosylation, dedicated 2DE patterns (Fig. 3) showed abnormalities only for Patients 2 and 3. The apoC-III pattern of Patient 1 (Fig. 3B) could not be distinguished from the control (Fig. 3A) indicating that *O*-glycosylation was not affected in this individual. In Patient 2 (Fig. 3C), the level of the bi-sialylated glycoform (substituted with *N*-acetylgalactosamine-galactose

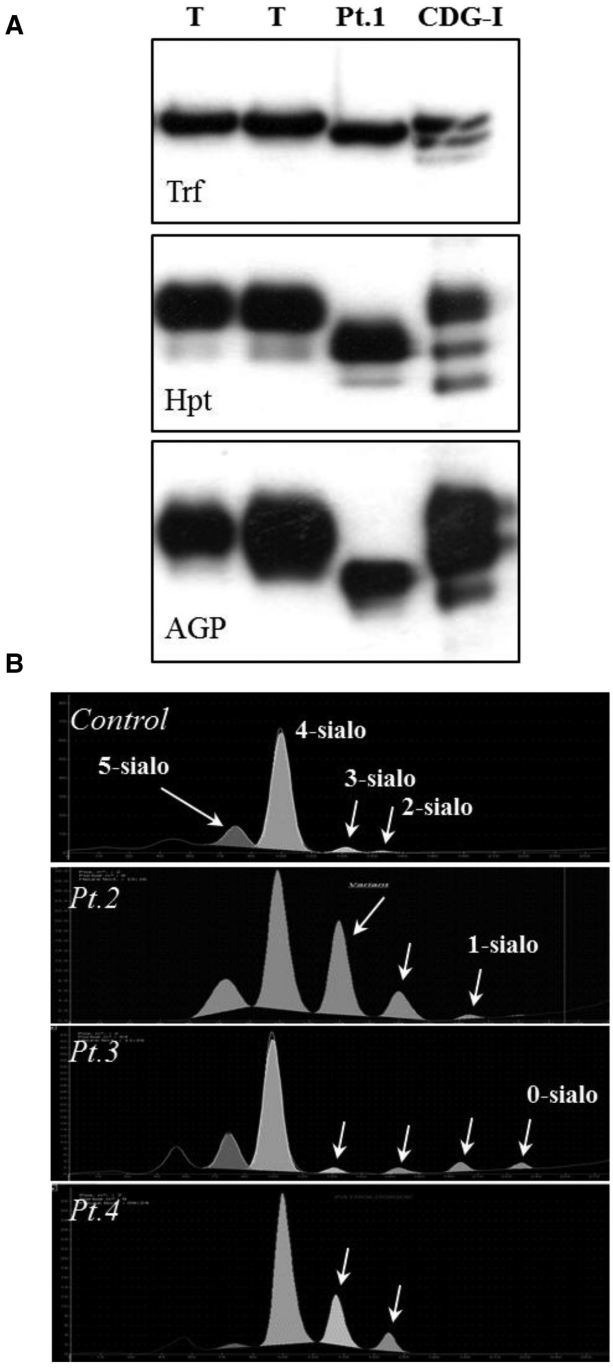

Figure 1. Western-blot of *N*-glycoproteins and CE of serum transferrin (A) SDS-PAGE Western-blot analysis of serum transferrin (Trf), haptoglobin (Hpt), and acid α1-glycoprotein (AGP) from (left to right) two controls (T), Patient 1 and CDG-I. (B) CE serum transferrin patterns of one control, and Patients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

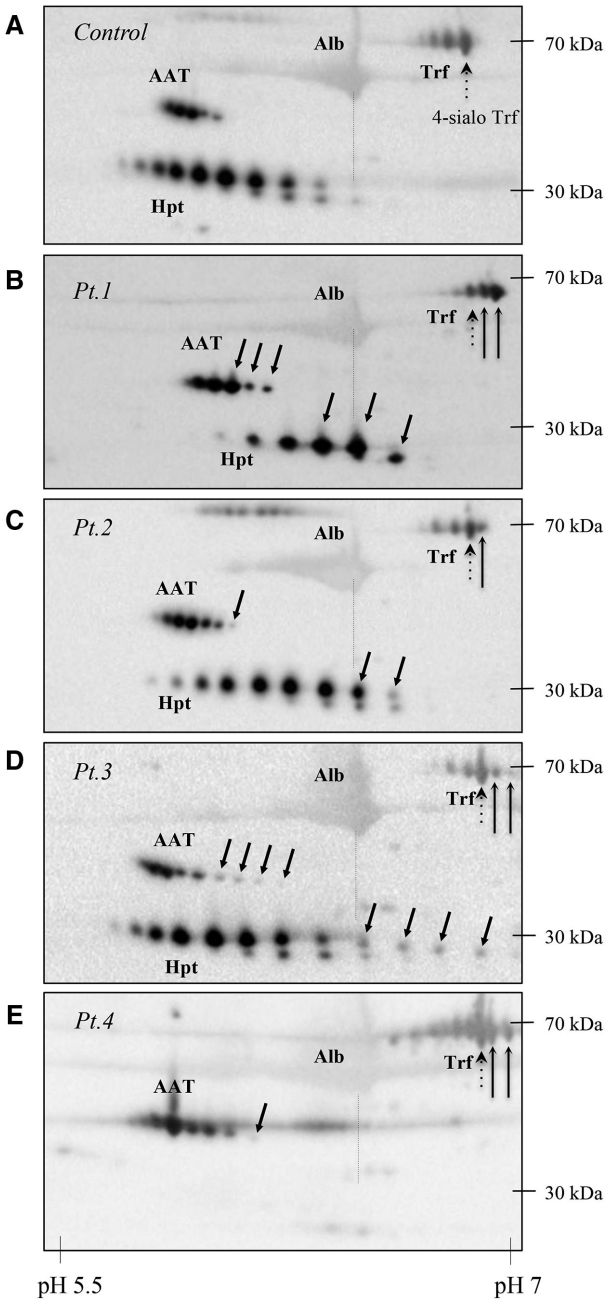

Figure 2. 2DE patterns of *N*-glycoproteins. 2DE patterns of serum α1-antitrypsin (AAT), haptoglobin (Hpt) and transferrin (Trf) from one control, and Patients 1–4. In all patterns, non-glycosylated albumin (Alb) can be detected in the 'background' constituting a useful landmark (vertical dotted line) 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.

carrying two SA) proved severely decreased. When detected along *N*-glycans defects, such marked abnormalities tend to suggest a global disturbance in GA homeostasis. For Patient 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 SA residue. ApoC-III pattern of Patient 4 (Fig. 3E) appeared normal suggesting that only the GA-located *N*-glycans

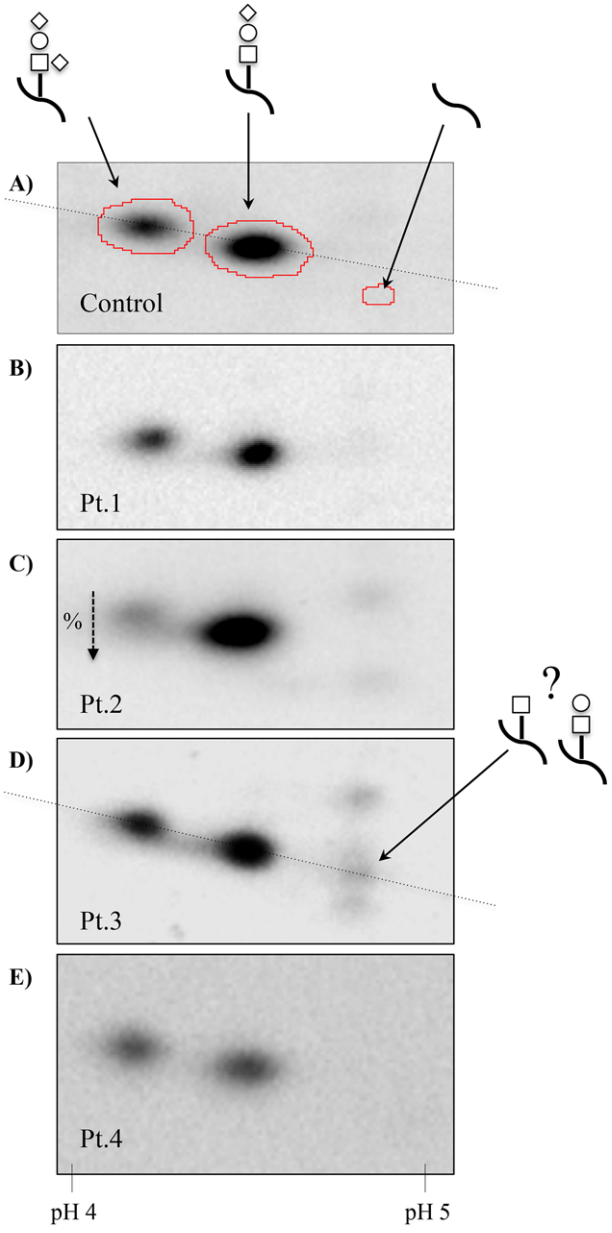

Figure 3. 2DE patterns of mucin core1 *O*-glycosylated apoC-III. 2DE patterns of serum apoC-III from one control, and Patients 1–4. In control pattern, apoC-III glycoforms structures corresponding to each spot are detailed: square, *N*-acetylgalactosamine; circle, galactose; diamond square, SA. In pattern of Patient 3, an unusual spot is indicated (arrow), probably corresponding to one between two putative schematized glycan structures. Dotted line (in control and Patient 3) is a landmark for accurate interpretation of spots corresponding to asialylated structures. Vertical dotted arrow in profile of Patient 2 schematizes the decreased percentage of di-sialylated apoC-III.

maturation pathway was affected in this individual. The same conclusions can be drawn from the MALDI-TOF analyses of apoC-III (Supporting Information Fig. 3), 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 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 four patients and compared with that of the healthy control. The spectrum of Patient 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 Patients 2 and 3 (Fig. 4C-D) shared some similarities with accumulation of partially sialylated bi-antennary *N*-glycans (*m/z* 2431.2 and 2605.3), of *N*-glycans lacking both terminal SA and galactose residues (*m/z* 1835.9, 2040.0, and 2227.1), and of oligomannose structures (*m/z* 1579.8 and 1783.9). Patient 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 1783.9. Furthermore, the total absence of polyantennary structures at *m/z* 3241.8 and 3602.8, whose synthesis is initiated by linkages of *N*-acetylglucosamine (GlcNAc) residues, was also noticeable in this sample.

3.5 Molecular studies

Patient 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. Patient 2 carried a non-sense homozygote variant (c. p.Val44X) on ATP6V0A2 encoding gene. Both nonconsanguineous parents were heterozygote. For Patient 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 the entry of activated galactose into the GA lumen [16]. For Patient 4, WES was performed and mutations on the three following genes were filtered: *SLC35A3* encoding for the golgi UDP-*N*-acetylglucosamine (UDP-GlcNAc) transporter, *PTPN23* (ciliogenesis control), and *TBX22* (transcription factor). Presence of a missense

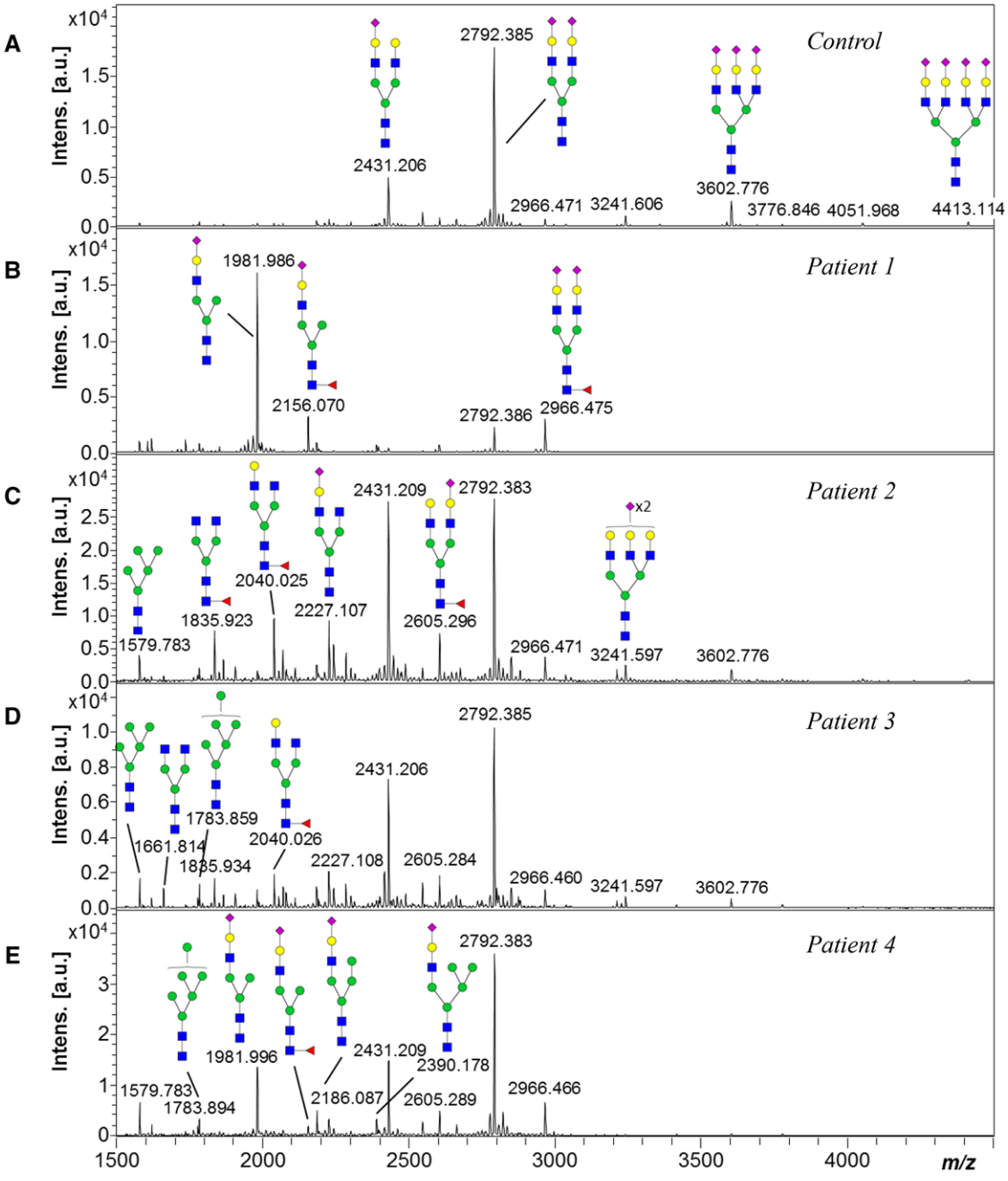

Figure 4. MALDI-TOF mass spectra of permethylated *N*-glycans released from serum samples. Measurements were performed in the positive-ion mode and all ions are present in sodiated form. Green circles, mannose; yellow circles, galactose; blue squares, *N*-acetyl glucosamine; red triangles, fucose; purple diamonds, SA.

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 (Patient 1), ATP6V0A2-CDG (Patient 2), SLC35A2-CDG (Patient 3), and SLC35A3-CDG (Patient 4).

4 Discussion

Patient 1 showed nonspecific but severe neurological and multiorgan clinical symptoms motivating CDG screening. SDS-PAGE Western blot and 2DE 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. 4). 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].

Patient 2 presented with nonspecific and moderate neurological and dysmorphic features associated to a peculiar skin disease, i.e., '*Cutis laxa*' (Supporting Information Fig. 1). In this individual, CE profile of serum Trf demonstrated a characteristic CDG-II profile with high levels of hyposialylated glycoforms. 2DE 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 Patient 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). 2DE 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, 2DE 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 Patient 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 Patient 3, notably with the absence of the oligosaccharidic structure at *m/z* 2605.3. Furthermore, the glycan moiety lacking the two subterminal galactoses (*m/z* 1661.8) was specifically retrieved in Patient 3 suggesting a higher level of hypogalactosylation of *N*-glycans. These results, when combined with nonspecific 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. 3) [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 Patient 3 putatively explaining the observed associated hyposialylation.

Patient 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. 2DE of *N*-glycoproteins further confirmed hyposialylation of Trf but proved poorly informative for AAT (one faint additional spot) while Hpt was undetectable. One might notice slight differences between the CE and 2DE Trf profiles for this patient. This is probably linked to the fact that the CE method employed to analyze Trf glycosylation efficiently quantifies Trf glycoforms containing up to five SA residues but does not sensitively detect hypersialylated glycoforms with less than six SA residues [11]. In the opposite, 2DE, especially when combined with an antibody-based detection, can reveal all the Trf sialylated glycoforms with equivalent sensitivity.

2DE 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 2DE, MS-based analysis of serum *N*-glycans proved highly informative revealing strong accumulation of mono-antennary monosialylated structures lacking one complete [GlcNAc-Gal-SA] arm (*m/z* 1981.9 and 2156.1). Strikingly, these MGAT2-CDG-typical structures (as found in Patient 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 Patient 1; Supporting Information Fig. 4). 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. 4). Lastly, total absence of tri-antennary structures at *m/z* 3241.8 and 3602.8 was also evidenced in Patient 4. Considering these rather complex glycosylation profiles associated with nonspecific 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, was performed but both proved unsuccessful. Then, a WES was performed and interestingly revealed relevant mutations in SLC35A3 gene encoding for the 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. 4) [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 Patient 4. To our knowledge, this patient is only the second case (among ten described patients) of SLC35A3-CDG presenting with clear serum glycans abnormalities [25, 26].

The usefulness and complementarity of the presented electrophoretic, mass spectrometric, and gene sequencing methods can be highlighted at different levels through the four detailed CDG cases. First, CE analysis of serum Trf allows fast (10 min per sample), simple and robust first-line CDG screening (as exemplified for Patients 2–4). Although we describe only CDG-II cases in this paper, this technique can also efficiently detect CDG-I [11]. The main limitation of this efficient CE approach is that it cannot be applied to EDTA plasma and dried blood spot samples. When serum or heparinized plasma is not available or in insufficient quantity (as for Patient 1; minimal volume = 150 μL), glycoproteins from EDTA plasma samples or dried blood spots can be conveniently analyzed by 2DE. Furthermore, this technique monitors simultaneously the potential occurrence of *N*-glycosylation defects on up to three distinct glycoproteins, thus strengthening the sensitivity and specificity of the CDG screening (Patients 1–4). Despite a rather poor ability to distinguish protein glycoforms, SDS-PAGE Western-blot analysis of *N*-glycoproteins can easily be applied in a straightforward routine manner to all sample types for evidencing CDG-I as well as peculiar CDG-II such as MGAT2-CDG (Patient 1).

In parallel, 2DE of apoC-III efficiently determines if mucin core1 *O*-glycosylation is impacted concomitantly to *N*-glycosylation, thus suggesting an overall defect in GA homeostasis (Patient 2) or a defect common to both *N*- and *O*-glycosylation pathways (Patient 3). In the case of Patients 1 and 4, apoC-III electrophoresis analysis did not show particular abnormality orientating toward specific *N*-glycosylation defects. Additionally, apoC-III patterns have been showed to distinguish between ATP6V0A2-CDG and COG-CDG [8]. Altogether, when considering the very low prevalence of CDG, we assume that such fast, simple, and inexpensive electrophoretic techniques are mandatory prior to MS-based glycan study.

By furnishing accurate information about accumulated or lacking *N*-glycan structures, MALDI-TOF analysis of enzymatically released total serum *N*-glycans can be, in some instances, directly indicative of the CDG-II causative genetic defect. For instance, the intense peak at *m/z* 1981.9 can be considered as 'diagnostic' of MGAT2-CDG (Patient 1). Also, global *N*-glycan profiling is essential to get insight into the electrophoretic patterns, by providing unsurpassed structural information regarding both negatively charged (sialylated) and neutral fucosylated or galactosylated *N*-glycan species (Patients 2 and 3). However, as illustrated for Patients 2–4, MS-based *N*-glycan profiles are often complex and potentially difficult to interpret (especially for nonexpert people), which might prevent its use as single screening method in routine conditions. In addition, this technique does not easily distinguish CDG-I and CDG-II [27], and might be potentially confounded by some external nongenetic causes of *N*-glycosylation modification, such as liver cirrhosis that could produce CDG-II-like *N*-glycan profiles [28, 29]. Table 1 summarizes some of the main features of these different electrophoretic approaches used to screen for *N*-glycosylation disorders.

Table 1. Analytical characteristics of the electrophoretic and MS-based approaches used in this study for the screening and diagnosis of *N*-glycosylation disorders

	Time-to-result	Ease of use	Diagnostic power	Information provided
Western blot SDS-PAGE	~48 h	+++	+	+
CE	~10 min	+++	+++	++
2DE	~48 h	++	++(+)	++
MALDI-TOF	~48 h–72 h[a)]	+	++(+)	+++

a) Under the conditions of the present study [12], but method throughput can be substantially improved [30, 31].

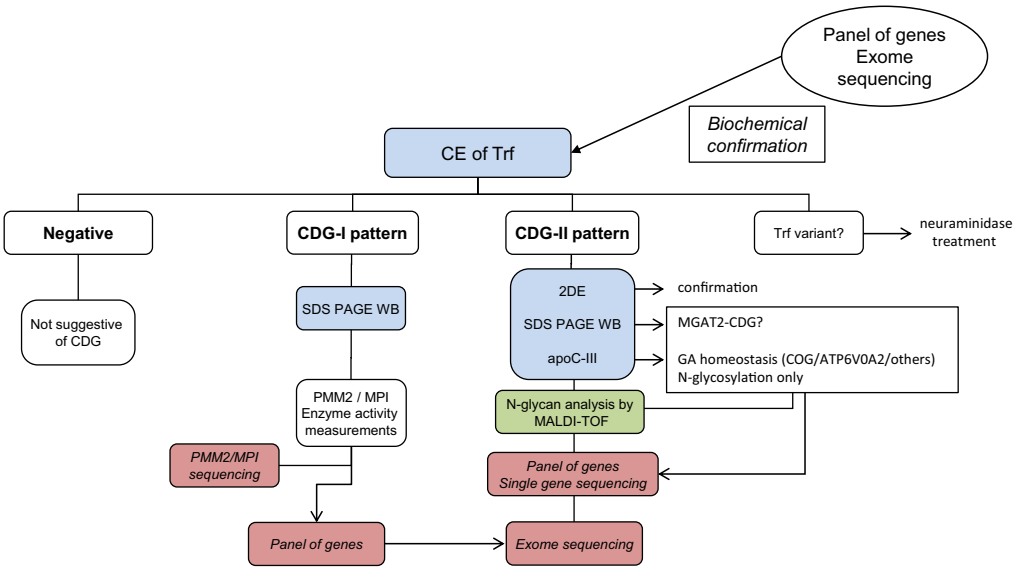

Figure 5. Proposed analytical flowchart for CDG diagnosis.

Concerning molecular studies, targeted sequencing of single or panel of genes often suffices to corroborate suspected CDG causative mutations (Patients 1–3). When unsuccessful or if preliminary information regarding potential culprit genes is lacking, WES can reveal (mutated) gene candidates whose causality can eventually be asserted by reinterpreting glycosylation-related data such as MS-based total *N*-glycan profiles (Patient 4). Finally, it should be noticed that specialized biochemical laboratories are now increasingly solicited for validating the causality of initially unsuspected CDG-related mutations issued from WES data.

To summarize, 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. Altogether, these data led us to propose the analytical flowchart from Fig. 5 for CDG diagnosis. 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, WES would be performed to get new perspectives.

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).

The authors have declared no conflict of interest.

5 References

[1] Moremen, K. W., Tiemeyer, M., Nairn, A. V., *Nat. Rev. Mol. Cell. Biol.* 2012, *13*, 448–462.

[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 https://doi.org/10.1016/j.ejmg.2017.10.012.

[3] Jaeken, J., *Ann. NY. Acad. Sci.* 2010, *1214*, 190–198.

[4] Jaeken, J., van Eijk, H. G., van der Heul, C., Corbeel, L., Eeckels, R., Eggermont, E., *Clin. Chim. Acta* 1984, *144*, 245–247.

[5] Lefeber, D. J., Morava, E., Jaeken, J., *J. Inherit. Metab. Dis.* 2011, *34*, 849–852.

[6] Seta, N., Barnier, A., Hochedez, F., Besnard, M. A., Durand, G., *Clin. Chim. Acta* 1996, *254*, 131–140.

[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*, 787–793.

[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*, 4151–4161.

[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*, 593–602.

[11] Parente, F., Ah Mew, N., Jaeken, J., Gilfix, B. M., *Clin. Chim. Acta* 2010, *411*, 64–66.

[12] Goyallon, A., Cholet, S., Chapelle, M., Junot, C., Fenaille, F., *Rapid Commun. Mass Spectrom.* 2015, *29*, 461–473.

[13] Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., Haslam, S. M., *J. Proteome Res.* 2008, *7*, 1650–1659.

[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*, 4056–4063.

[16] Caffaro, C. E., Hirschberg, C. B., *Acc. Chem. Res.* 2006, *39*, 805–812.

[17] Stibler, H., Jaeken, J., *Arch. Dis. Child* 1990, *65*, 107–111.

[18] Zeevaert, R., Foulquier, F., Jaeken, J., Matthijs, G., *Mol. Genet. Metab.* 2008, *93*, 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*, 32–34.

[20] Guillard, M., Dimopoulou, A., Fischer, B., Morava, E., Lefeber, D. J., Kornak, U., Wevers, R. A., *Biochim. Biophys. Acta* 2009, *1792*, 903–914.

[21] Mohamed, M., Kouwenberg, D., Gardeitchik, T., Kornak, U., Wevers, R. A., Morava, E., *J. Inherit. Metab. Dis.* 2011, *34*, 907–916.

[22] Hadley, B., Maggioni, A., Ashikov, A., Day, C. J., Haselhorst, T., Tiralongo, J., *Comput. Struct. Biotechnol. J.* 2014, *10*, 23–32.

[23] Slusarewicz, P., Nilsson, T., Hui, N., Watson, R., Warren, G., *J. Cell Biol.* 1994 *124*, 405–413.

[24] Maszczak-Seneczko, D., Sosicka, P., Kaczmarek, B., Majkowski, M., Luzarowski, M., Olczak, T., Olczak, M., *J. Biol. Chem.* 2015, *290*, 15475–15486.

[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*, 733–739.

[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*, 2758–2762.

[27] Zhang, W., James, P. M., Ng, B. G., Li, X., Xia, B., Rong, J., Asif, G., Raymond, K., Jones, M. A., Hegde, M., Ju, T., Cummings, R. D., Clarkson, K., Wood, T., Boerkoel, C. F., Freeze, H. H., He, M., *Clin. Chem.* 2016, *62*, 208–217.

[28] Morelle, W., Flahaut, C., Michalski, J. C., Louvet, A., Mathurin, P., Klein, A.. *Glycobiology* 2006, *16*, 281–293.

[29] Klein, A., Carre, Y., Louvet, A., Michalski, J. C., Morelle, W.. *Proteomics Clin. Appl.* 2010, *4*, 379–393.

[30] Kang, P., Mechref, Y., Novotny, M. V.. *Rapid Commun. Mass Spectrom.* 2008, *22*, 721–734.

[31] Shubhakar, A., Kozak, R. P., Reiding, K. R., Royle, L., Spencer, D. I., Fernandes, D. L., Wuhrer, M., *Anal. Chem.* 2016, *88*, 8562–8569.

6 Addendum

[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, Gif-sur-Yvette, France
[4]Génétique Médicale, CHU Rouen, Rouen, France
[5]Génétique Médicale, CHU-Groupe hospitalier Sud-Réunion, 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, Paris, France

