

Quantitative Assessment of Sialo-Glycoproteins and N-Glycans during Cardiomyogenic Differentiation of Human Induced Pluripotent Stem Cells

Sarah A. Konze,^[a, b] Samanta Cajic,^[c] Astrid Oberbeck,^[a, b] René Hennig,^[c, d] Andreas Pich,^[e]
Erdmann Rapp,^[c, d] and Falk F. R. Buettner*^[a, b]

[a] *Dr. S. A. Konze, A. Oberbeck, Prof. F. F. R. Buettner*
Institute of Clinical Biochemistry, Hannover Medical School, Carl-Neuberg-Straße 1, 30625
Hannover (Germany)
E-mail: buettner.falk@mh-hannover.de

[b] *Dr. S. A. Konze, A. Oberbeck, Prof. F. F. R. Buettner*
REBIRTH Cluster of Excellence, Hannover Medical School, 30625 Hannover (Germany)

[c] *S. Cajic, R. Hennig, Dr. E. Rapp*
Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106
Magdeburg (Germany)

[d] *R. Hennig, Dr. E. Rapp*
glyXera GmbH, Leipziger Straße 44, 39120 Magdeburg (Germany)

[e] *Prof. A. Pich*
Institute of Toxicology, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover
(Germany)

Supporting Information for this article can be found under <http://>

Abstract

Human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) may be used in regenerative medicine for individualized tissue transplants in future. For application in patients, the generated CMs have to be highly pure and well characterized. To overcome the prevalent scarcity of CM-specific markers, we quantitatively assessed cell surface exposed sialoglycoproteins and N-glycans of hiPSCs, CM progenitors and CMs derived thereof. Applying a combination of metabolic labeling and specific sialoglycoprotein capture, we could highly enrich and quantify membrane proteins during cardiomyogenic differentiation. Among them we identified a number of novel, putative biomarkers for hiPSC-CMs. Analysis of the N-glycome by capillary gel electrophoresis revealed three novel structures comprising β 1,3-linked galactose, α 2,6-linked sialic acid and complex fucosylation that were highly specific for hiPSCs. Bisecting GlcNAc structures strongly increased during differentiation and we propose that they are a characteristic feature of early, immature CMs.

Introduction

Coronary heart disease leading to heart failure is the most common cause of death in western industrialized countries.^[1] The human heart has only very limited regeneration capacity^[2;3] and transplants are scarce. Thus, one important aim in stem cell research is the replacement of damaged heart muscle with cardiomyocytes (CMs),^[4] which can be efficiently differentiated from human embryonic stem cells (hESCs).^[5] With the invention of the reprogramming technology,^[6;7] even cell replacement therapies by autologous CMs obtained from patient-derived induced pluripotent stem cells (iPSCs) became feasible.^[8] For the application of human pluripotent stem cell-derived cardiomyocytes (hPSC-CMs) in man, it must be assured that generated tissue grafts only contain pure CMs. Remaining stem cells could lead to teratoma formation after transplantation^[9;10] and contaminating cells types could affect the functionality of the produced graft.^[11] Pure populations of CMs will be also essential for drug discovery and drug safety testing.^[12] Cell surface-exposed biomolecules including proteins and glycans are ideally suited to be applied as markers for the characterization and purification of CMs. To date, only a few proteins have been described that are suitable for enrichment of hPSC-CMs. These include the activated leukocyte cell adhesion molecule (ALCAM),^[13] the elastin microfibril interface-located protein 2 (EMILIN2),^[14] the vascular cell adhesion molecule (VCAM1),^[15] and the signal regulatory protein (SIRPA).^[16] Of these markers, EMILIN2 was identified in a study by van Hoof *et al.* comparing plasma membrane proteins of hESCs with hESC-derived CMs and adult human cardiac tissue using stable isotope labeling of amino acids in cell culture (SILAC)-based quantitative proteomics after preparation of membrane fractions by ultracentrifugation.^[14] According to their uniprot (www.uniprot.org) entries, all of the above mentioned CM markers are glycosylated cell surface proteins, highlighting the potential of cell surface glycotopes for use as differentiation markers. The vast majority of cell surface proteins are N-glycosylated^[17] and often the outermost ends of these glycan chains contain N-acetylneuraminic acids (in the following referred to as sialic acids), thus making sialic acids ideal targets for capturing cell surface-

proteins.^[18] Various publications describe the selective capture and enrichment of surface exposed sialo-glycoproteins based on periodate oxidation of the glycan into an aldehyde, which can subsequently react with hydrazides or aminoxy compounds carrying affinity tags.^[19-27] By using that approach in combination with label-free quantification, Boheler *et al.* reported a comprehensive inventory of about 500 cell surface glycoproteins expressed by hPSCs with >100 surface proteins being not previously described on the surface of hPSCs.^[25] A similar approach called PAL (for periodate oxidation and aniline-catalyzed oxime ligation) even enables specific and efficient labeling of sialic acid containing glyco-conjugates under physiological conditions on living cells.^[22]

Besides glycoproteins, N-glycans are attractive targets to be used as markers for hPSC-CMs. The N-glycome of hPSCs has been the subject of a number of studies over the past years. Most prominent features of hPSC glycosylation are the abundance of high-mannose type^[28-31] and of α 1,2-fucosylated N-glycans^[30;32-35] as reviewed in Furukawa *et al.*^[36] and Berger *et al.*^[37] Sialylated glycans with complex fucosylation (one fucose at the chitobiose core and at least one antennae fucose) were increased in hPSCs compared to differentiated cell types.^[28] Similarly, Fujitani *et al.*^[31] showed enrichment of multiply fucosylated type, neutral triantennary type as well as bisecting and/or LacdiNAc type glycans in hPSCs. Our own group recently analyzed the N-glycome of hPSCs by quantitative capillary gel electrophoresis coupled with laser-induced fluorescence detection (xCGE-LIF) and we could confirm that high-mannose type N-glycans represent the major fraction of N-glycans on hPSCs.^[38] Also, as shown before,^[30;34] we exclusively identified glycans with α 2,6-linked sialic acid but none with α 2,3-linked sialic acid on hPSCs.^[38] In agreement to glycomic analyses of hPSCs, a recent glycomic comparison of hiPSCs, hiPSC-CMs and human cardiomyocytes (hCMCs) by Kawamura *et al.* revealed reduction of antennae fucosylation and increased expression of α 2,3-sialylation in hiPSC-CMs and hCMCs compared to hiPSCs.^[39] Furthermore, exposed N-acetylglucosamine (GlcNAc) was lowered in CMs whereas exposed galactose was increased. Interestingly, the level of high-

mannose type N-glycans was not altered in CMs.^[39] The above mentioned studies mainly applied MALDI-TOF^[28;30;31;39] or lectin array technologies^[34] for identification of glycan structures, whereas our analytical approach for glycan determination in the present study depends on xCGE-LIF.^[38] This technology has been widely applied by us and others for N-glycan analysis of various biological materials^[40-43] and was shown to be competitive to MS-based methods for glycan identification with the advantage to be performable at high throughput levels.^[44]

Membrane proteins are involved in the communication of a cell with its surrounding and their function may be modulated by glycosylation,^[17] which was e.g. impressively shown for the Notch receptor.^[45;46] Likewise, the presence or absence of terminal sialic acid have decisive effects on maintenance of pluripotency or differentiation.^[47] Moreover, enzymatic removal of terminal sialic acid and subsequent exposure of underlying β -galactopyranoside residues was shown to induce differentiation of hiPSCs towards the ectodermal lineage.^[48] Thus, besides marker discovery, unraveling the membrane glycoproteome and glycome of differentiating hPSCs is expected to reveal novel insight on signaling pathways affecting cardiomyogenesis.

A variety of protocols for differentiation of hPSCs into cardiomyocytes has been published in the past leading to increasing efficiencies of cardiomyogenesis nowadays.^[5] However, existing proteomic or glycomic studies of cardiomyogenic differentiation often rely on different approaches of cardiomyogenesis and it has to be considered that not only the purity of cardiomyocytes but also the protocol applied for differentiation might affect proteomic and glycomic signatures of these cells. For instance, the work presented by Kawamura *et al.*^[39] of the hiPSC glycome and a previous study from our group describing the whole cell proteome of hESCs and hiPSCs^[49] rely on an embryoid body (EB)-based differentiation approach. Historically, differentiation from EBs was the first method to derive cardiomyocytes from pluripotent stem cells and to date, a variety of different EB-based differentiation protocols exist as recently reviewed by Talkhabi *et al.*^[5] However, the efficiency of the EB-based differentiation approach is limited, although, for instance, manual selection of beating areas can increase the

purity of cardiomyocytes as performed by Kawamura *et al.* Moreover, due to the spatial organization of an EB,^[50] the cells display a relatively large heterogeneity, thus impairing their applicability in omics screens. In contrast, the present study applies a more recent monolayer-based differentiation procedure with temporal modulation of the Wnt signaling pathway as reported by Lian *et al.*^[51;52] This approach, also called GiWi (for GSK3 β inhibition and Wnt inhibition) method, can lead to up to 98% pure cardiomyocytes (depending on the applied hPSC line) and the cells are very homogenous due to the seeding in a monolayer.^[51;52] Since the latter approach is still relatively new, to the best of our knowledge, the proteome and glycome of cardiomyocytes produced by the GiWi method has not been studied so far.

In the present study we performed quantitative proteomics by SILAC upon PAL-based capture of sialylated glycoproteins to assess glycoproteomic alterations during cardiomyogenic differentiation of hiPSCs. Alterations of glycoprotein levels upon enrichment by PAL were moreover compared to non-enriched samples aiming at the identification of changes in protein sialylation. The glycoproteomic approach approach was supplemented by a global analysis of the N-glycome of the same samples by xCGE-LIF. Thereby, we present a global proteomic, sialo-glycoproteomic and glycomic characterization of hiPSCs in the undifferentiated state and during cardiomyogenic differentiation.

Results and Discussion

SILAC labeling is compatible with pluripotency and cardiomyogenic differentiation

For quantitative proteomic comparison between undifferentiated hiPSCs, hiPSC-derived cardiac progenitors and cardiomyocytes we decided to apply metabolic labeling by SILAC. Thus, hiPSCs were maintained and differentiated towards CMs under SILAC conditions (Figure 1). SILAC-labeled hiPSCs stained positive for the two pluripotency markers OCT-3/4 and SSEA-4 (Figure 2A) and qPCR analysis further revealed expression levels of pluripotency markers *OCT-3/4*, *SOX2*, *LIN28* and *NANOG* (Figure 2B) comparable to standard culture conditions. Moreover,

flow cytometry analysis of the cell surface marker SSEA-4 revealed 94.0%, 94.5% and 95.0% positive cells for the standard (light) and for medium and heavy SILAC-labeled hiPSCs, respectively (Figure 2C). Taken together, hiPSCs cultivated under SILAC-labeling conditions remained pluripotent and showed a similar expression profile of pluripotency markers as under standard culture conditions.

Upon cardiomyogenic differentiation of hiPSC, all three conditions (light, medium and heavy) yielded extended areas of spontaneously contracting cells (Additional Material Video 1 to 3). Immunofluorescence staining of these cells revealed comparable expression of the CM marker protein α -actinin (Figure 2D) and 50.5%, 48.1% and 44.2% of these cells, from light, medium and heavy conditions, respectively, were positive for cardiac troponin T (cTnT), as deduced from flow cytometry analysis using cTnT as a CM-specific marker (Figure 2E). qPCR analysis of the cardiac progenitor marker *NKX2.5* and of the CM-specific marker myosin heavy chain 6 (*MYH6*) revealed strong up-regulation of these factors upon 7 days of differentiation (d7), when no spontaneous contractions were visible yet, and their expression level remained equally high or even increased on d15 of differentiation (Figure 2F). This was observed for all three labeling conditions, indicating successful differentiation of hiPSCs towards CMs. In summary, we concluded that SILAC labeling did not affect pluripotency or cardiomyogenic differentiation of hiPSCs.

Determination of SILAC incorporation efficiencies

In order to scrutinize whether the cells were efficiently labeled with the different SILAC amino acids, we analyzed the samples separately (i.e. not as SILAC pools) and determined the overall incorporation efficiency (Supporting Information Figure S1 and S2). The median incorporation ratios of the different samples of d0, d7 and d15 for medium and heavy labeled amino acids in peptides were calculated as 90.0% \pm 1.5% and 91.0% \pm 0.6%, respectively. This analysis was performed with the 50% most intensive peptides of each sample in order to minimize the effect

of false positive signal detection (i.e. noise). Taken together the incorporation rates are sufficient for a proper quantitative proteomics analysis.

Quantitative proteomic and glycoproteomic analysis during cardiomyogenesis of hiPSCs

For quantitative proteomics during cardiomyogenesis of hiPSCs, three different stages during differentiation, d0 (hiPSCs), d7 (cardiac progenitors), and d15 (CMs) were compared by SILAC-based mass spectrometry (Figure 1). The main focus of our study was the analysis of alterations in the cell surface glycoproteome during differentiation and we decided to apply PAL^[22] for selective capture of sialo-glycoproteins. Additionally, we assessed in an independent approach the whole cell proteome of the above mentioned cellular stages (d0, d7 and d15) allowing to correlate changes of the sialo-glycoproteome to the whole cell proteome. Furthermore, to our knowledge, no global proteomic study showing quantitative alterations during hPSC cardiomyogenesis has been presented so far that uses the highly efficient cardiomyogenic differentiation protocol by Lian *et al.*^[51;52]

The whole cell proteomic and the sialo-glycoproteomic screen led to the identification of 2970 different proteins in total, provided that they were found in at least two independent sample pool replicates. Of these, 2091 were only found in the whole cell proteome analysis and 88 were exclusively identified in the sialo-glycoproteome, thus resulting in an overlap of 791 proteins (Figure 3A and Additional Material Table S1).

In order to confirm that the PAL-based purification procedure led to an increase of cell surface and extracellular proteins, we analyzed the gene ontology (GO) annotation (implemented in the Perseus software) of proteins exclusively found in the whole cell proteome or the sialo-glycoproteome or found in both (Figure 3B). The majority, i.e. 76.1%, of proteins exclusively found in the PAL analysis was annotated as “plasma membrane / cell surface” and further 11.4% were predicted to be located “extracellular”, whereas only 9.1% were predicted to be “cytosol / cytoplasm / cytoskeleton” (Figure 3B, left panel). In contrast, in the protein fraction that was

exclusively found in the analysis of the whole cell lysates, only 14.3% and 2.3% were annotated as “plasma membrane / cell surface” or “extracellular”, respectively, but a much larger proportion, i.e. 46.5% was annotated as “cytosol / cytoplasm / cytoskeleton” (Figure 3B middle panel). The protein fraction that was identified in the analysis of the whole cell lysates as well as in the PAL samples contained 35.0% proteins with “plasma membrane / cell surface” annotations, 7.5% were predicted to be located “extracellular” and 28.0% were annotated as “cytosol / cytoplasm / cytoskeleton” (Figure 3B, right panel). In summary, GO analysis for cellular compartments showed that sialo-glycoprotein purification by PAL led to a strong enrichment of membrane proteins or extracellular proteins. This strong enrichment of surface exposed proteins by PAL can explain the identification of numerous potentially low abundant proteins that were not identified in the whole cell proteome analysis, although the latter approach led to the identification of much more proteins in total.

We performed SILAC-based quantification and the numbers of proteins that were significantly regulated in the whole cell proteome or the sialo-glycoproteome, comparing the different time points (d0, d7, and d15) are presented in Tables 1 and 2. As expected, CM-related KEGG pathways including “cardiac muscle contraction” “arrhythmogenic right ventricular cardiomyopathy”, “hypertrophic cardiomyopathy” and “dilated cardiomyopathy” were enriched in the fraction of PAL-captured proteins being up-regulated >1.5-fold. Interestingly, proteins belonging to “focal adhesion” were enriched in the fraction being up-regulated >1.5-fold, whereas proteins belonging to the categories “cell adhesion molecules” and “adherens junctions” were enriched in the fraction down-regulated >1.5-fold (Figure 3C).

Whole cell proteomics reflects the cardiomyogenic differentiation process of iPSCs

Plotting the regulation ratios of whole cell lysate proteins for the three time point comparisons d7/d0, d15/d0 and d15/d7 versus their respective p-value (Volcano plots, Supporting Information Figure S3A-C) highlighted proteins that were substantially and significantly regulated during the

differentiation process. The d7/d0 comparison (Supporting Information Figure S3A) showed that known stem cell proteins^[53] such as developmental pluripotency-associated protein 4 (DPPA4), podocalyxin-like protein (PODXL), POU domain, class 5, transcription factor 1 (POU5F1, more commonly known as OCT-3/4), DNA (cytosine-5)-methyltransferase 3B (DNMT3B) and Sal-like protein 2 (SALL2) are among the significantly down-regulated proteins. Additionally, proteins that were according to gene ontology analysis (www.string-db.org, GO analysis biological function) annotated as involved in embryonic development and morphogenesis, such as DnaJ homolog subfamily B member 6 (DNAJB6), the RNA binding protein 19 (RBM19), telomere-associated protein RIF1, CAD protein, epidermal growth factor receptor (EGFR), ephrin type A receptor (EPHA1), DNA topoisomerase 2 alpha (TOP2A) were also significantly down-regulated on d7 compared to d0. In contrast, proteins involved in mesodermal commitment, namely mesoderm-specific transcript homolog protein (MEST)^[54] and neural cell adhesion molecule 1 (NCAM1),^[55] were up-regulated on d7 compared to d0. Moreover, according to gene ontology analysis using STRING a number of proteins important for cardiomyogenesis, i.e. desmoplakin (DSP), laminin subunit alpha (LAMA1), junction plakoglobin (JUP), the collagens COL1A1, COL5A1 and COL6A1, integrin alpha 5 (ITGA5), aminopeptidase N (ANPEP), ephrin-type B receptor 2 (EPHB2), jagged-1 (JAG1), N-cadherin (CDH2) and plakophilin-2 (PKP2) were up-regulated on d7 compared to d0, thus reflecting the early steps of cardiomyogenic differentiation. Notably, although on d7 we already detected strong up-regulation of the cardiac myosin transcript *MYH6* by qPCR (Figure 2F), no CM structure proteins such as myosins or troponins were among the significantly up-regulated proteins on d7. This might indicate that the emergence of CMs is already starting but is not reflected in the proteomic pattern. In contrast to this, in the d15/d0 comparison (Supporting Information Figure S3B), many of the strongly and significantly up-regulated proteins on d15 are structural components of the cardiac muscle,^[56] for instance, myosin light chain 3 (MYL3) and 4 (MYL4), myosin heavy chain 4 (MYH4), 6 (MYH6) and 7 (MYH7), the myosin binding protein MYBPC3 and the cardiac troponin I (TNNI) indicating

presence of fully functional CMs. Moreover, in this comparison again a number of proteins annotated as associated with cardiac development (according to GO analysis using STRING), namely the already mentioned NCAM1, LAMA5, PKP2, COL5A1, CDH2, JUP, ITGA5 and DSP as well as integrin beta 1 (ITGB1), calreticulin (CALR), prolow-density lipoprotein receptor-related protein 1 (LRP1/CD91), prelamin-A/C (LMNA), cAMP-dependent protein kinase catalytic subunit alpha (PRKACA) and the basement membrane-specific heparan sulfate proteoglycan core protein (HSPG2) were among the substantially up-regulated proteins. Concerning the down-regulated proteins in the comparison d15/d0, several proteins characteristic for pluripotent stem cells, such as the above mentioned RIF1, DNAJB6, CAD, OCT-3/4 and SALL2 as well as LIN28A and FKBP4,^[53] were among them. Finally, the d15/d7 comparison (Supporting Information Figure S3C) revealed that many proteins annotated by STRING as involved in “chromatin remodeling and replication” (e.g. DNA replication licensing factor (MCM3), homeobox protein MEIS1, DNA replication complex GINS protein PSF3 (GINS3), proliferating cell nuclear antigen (PCNA) and DNA polymerase delta catalytic subunit (POLD1)) were down-regulated on d15 compared to d7, but no stem cell markers were found in this fraction of proteins. Among the proteins that were up-regulated between d7 and d15, we again identified a large proportion of proteins belonging to cardiac development such as sarcoplasmic reticulum histidine-rich calcium-binding protein (HRC), LMNA, MYBPC3, creatine kinase M-type (CKM), tropomodulin-1 (TMOD1) and myosin MYL3. We furthermore screened our whole cell proteomic data set for the identification of known CM surface markers. Only in the d15/d7 comparison we found ALCAM,^[13] and EMILIN2,^[14] but the p-value for EMILIN2 regulation was below the significance threshold of 0.05. In summary, the proteomic screen clearly displays the process of cardiomyogenic differentiation and agrees with our previous work in which we applied a different protocol for cardiomyogenesis. However, whole cell proteomics revealed only few known surface markers and in the large data set it is laborious to screen for novel candidates.

PAL of iPSCs at different stages during cardiomyogenic differentiation confirmed known stem cell and CM cell surface markers

As shown before for whole cell lysates, the regulation ratios d7/d0, d15/d0 or d15/d7 of PAL-enriched proteins were plotted against the significance level (Volcano plots, Figure 4A-C) and as expected, cardiomyogenic differentiation of hiPSCs resulted in down-regulation of pluripotency-related proteins. Thus, we observed that in the d7/d0 as well as in the d15/d0 comparison, podocalyxin (PODXL), E-cadherin (CDH1), alkaline phosphatase (ALPL)^[57] and the more recently identified PSC-marker claudin-6 (CLDN6)^[58] were among the most strongly down-regulated proteins (Figure 4A and B). Furthermore, the aforementioned pluripotency-associated proteins EPHA1 and EGFR were significantly down-regulated in the d7/d0 comparison and, in case of EPHA1, in the d15/d0 comparison as well. We matched our findings to a recently published comprehensive repository of stem cell surface markers.^[25] Boheler *et al.* presented a list of 122 glycosylated proteins restricted to pluripotent cells (positive markers), of which we identified 27 proteins by our PAL approach (Table 3). Notably, for most of these proteins we also observed down-regulation upon differentiation, but a few candidates showed up-regulation >1.5-fold (EPHA7, HEPH [both significant] and CCKBR, CNTFR, LAMA1, LGR4, LRP4 and PTPRD [all non-significant]) in either the d7/d0, the d15/d0 or in both time point comparisons (highlighted in red in Table 3). Moreover, all of the six putative pluripotency markers reported by Prokhorova *et al.*,^[59] i.e. prominin 1 (PROM1), glypican-4 (GPC4), receptor-type tyrosine-protein phosphatase zeta (PTPRZ1), glycoprotein M6B (GPM6B), neuroligin-4 (NLG4) and the receptor tyrosine-protein kinase ErbB-2 were also identified in our study as down-regulated upon differentiation.

Additionally, using our approach, we could validate the up-regulation of previously described CM-specific cell surface proteins upon cardiomyogenic differentiation. These included SIRPA^[16] and Cadherin-2 (CDH2),^[56] which were up-regulated at d7 and d15 (latter not significantly) in comparison to d0 (Figure 4A,B), as well as VCAM1^[15] and EMILIN2^[14] which showed significant

and strong up-regulation in the d15/d7 comparison (Figure 4C). Furthermore the neural cell adhesion molecule NCAM1, which has been reported to define a population of multipotent mesoderm-committed cells^[55] is among the most significantly and strongly up-regulated proteins in the d7/0 and the d15/d0 comparison (Figure 4A and B). Additionally, the PDGF receptor alpha (PDGFRA), which was reported to be - in combination with FLK1 and CXCR4 - a marker of cardiovascular progenitor cells,^[60] was strongly and significantly up-regulated on d7 (Figure 4A). Appropriately, comparing d15/d0, PDGFRA was not regulated and the 15/d7 comparison even revealed down-regulation (although the significance threshold was slightly missed) (Figure 4C), indicating an only transient up-regulation during the differentiation of hiPSCs into CMs.

In summary, our results are in good concordance with previous studies on cell surface glycoproteins of hPSCs and CMs. Notably, this single analytical approach could confirm all relevant CM marker proteins (SIRPA, VCAM1, EMILIN2 and ALCAM) thereby showing the validity of the approach and its suitability to also identify novel marker proteins.

Identification of novel proteins being specific for hiPSCs, cardiac progenitors and CMs

As to the best of our knowledge, a PAL-based analysis of sialo-glycoproteins from CMs has not been performed before. Therefore we hypothesized that among the list of significantly up-regulated proteins there will be so far unknown CM-specific factors. Thus, we screened our dataset of substantially and significantly regulated proteins (Additional Material Table S2) in order to identify novel candidates being applied as markers for cardiac progenitors or CMs.

We observed that teneurin-4 (TENM4), the receptor-type tyrosine-protein phosphatase alpha (PTPRA) and protogenin (PRTG) are transiently up-regulated on d7 of cardiomyogenic differentiation and expression decreases again on d15. Tenm4 has been described as required for mesoderm induction in mESCs,^[61] and protogenin has been found in early mesodermal cells, the neuroepithelial cell layer of the brain and the trunk neural tube of developing chick embryos.^[62] The emergence of these proteins during early cardiomyogenic differentiation could

hint at a role in cardiac development and thus could make them novel cell surface markers for cardiovascular progenitors.

On d15 the glycoproteins agrin (AGRN), the choline transporter-like protein 2 (SLC44A2), EGF-like repeat and discoidin I-like domain-containing protein 3 (EDIL3), versican (VCAN), cadherin-13 (CDH13) and potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 4 (HCN4) as well as the basement membrane heparan sulfate proteoglycan core protein HSPG2 are strongly and significantly up-regulated compared to d0 and/or d7. Thus, these proteins could resemble novel markers for functional hPSC-CMs. While HCN4 is widely known as a marker for nodal/pacemaker type CMs^[63] its general expression on early embryonic CMs has, to the best of our knowledge, not been studied so far. SLC44A2 is known to be involved in autoimmune diseases causing hearing loss,^[64] but has not been described in the cardiovascular context before. Cadherin-13, in turn, was shown to be expressed in the heart and to be involved in adiponectin-mediated cardioprotection.^[65]

Taken together, our screen presents a comprehensive dataset with a number of promising candidates (as described above) for stage-specific markers but further studies are required to confirm or disprove specificity of these candidate proteins.

Comparison of differentially regulated proteins in the whole cell proteome and the sialoglycoproteome

We plotted the regulation ratios of the 791 proteins identified in both analytical approaches (PAL and whole cell proteome) against each other (Supporting Information Figure S4). These plots show that the majority of proteins are similarly regulated, for instance the pluripotency markers PODXL, ALPL and CLDN6 decrease in both, the PAL and whole cell proteome analysis when comparing differentiated cells to undifferentiated hiPSCs (Supporting Information Figure S4A,B). Similarly, the known CM cell surface proteins EMILIN2 and CDH2 as well as the mesoderm marker NCAM1 show an increase in differentiated cells compared to undifferentiated cells in

both the PAL and the whole cell proteome dataset (Supporting Information Figure S4A,B). However, some proteins show a differential regulation comparing the sialo-glycoproteomic analysis and the whole cell proteomic screen. For example, the recently published CM cell surface marker SIRPA^[16] as well as the cell surface proteoglycan syndecan-2 (SDC2) showed a substantial up-regulation in the d7/d0 PAL dataset, whereas their expression levels in the whole cell proteome analysis of d7/d0 slightly decreased (Supporting Information Figure S4A). A similar phenomenon was observed for the Neumann-Pick antigen C1 (NPC1) in the d15/d0 comparison (Supporting Information Figure S4B). This up-regulation during cardiomyogenesis, which was exclusively observed upon capture of sialo-glycoproteins, suggested an increase of sialylation of these proteins during cardiomyogenesis. It could further indicate accelerated protein membrane transport, which might also be caused by increased protein glycosylation.

N-glycosylation pattern considerably shifts during cardiomyogenesis

In order to gain a deeper understanding on alterations in glycosylation during cardiomyogenesis, we analyzed N-glycosylation of hiPSCs (d0), cardiovascular progenitors (d7) and CMs (d15) by xCGE-LIF. The three biological replicates for each time point were highly congruent (Supporting Information Figure S5) underlining the suitability of xCGE-LIF to dissect even minor differences in N-glycosylation during cardiomyogenic differentiation. Comparing the different time points (d0, d7, and d15), the samples revealed substantial alterations in the N-glycosylation pattern (Figure 5). We calculated the individual intensities of N-glycan peaks relative to the sum of all N-glycan peaks that could be quantified (all peaks with a signal-to-noise ratio above 9 in glyXtool™ software) by xCGE-LIF (relative quantification, Figure 6 and Supplemental Information Table S3). This presentation revealed a vast rearrangement of the cellular N-glycans [numbered [1] to [20] in Figure 6) during differentiation of hiPSCs into CMs. In agreement with previous studies, we observed that the most abundant glycan structures on hiPSCs are high-mannose type N-glycans (Figure 6A: [1], [2], and [3]; Supporting Information Table S3)^[29;30;38] and we observed

downregulation of Man7 and Man8 in hiPSC-CMs compared to hiPSCs. Interestingly, summing up all high-mannose peaks (Man5+Man6+Man7+Man8+Man9) ended up in 53%, 46% and 50% for d0, d7 and d15, respectively, thus revealed no pronounced differences between hiPSCs and hiPSC-CMs, which is in accordance to previous findings.^[39] Moreover, we showed that sialylation on hiPSCs exclusively occurred in α 2,6-linkage (Figure 6A: [4], [6] – [9]), a phenomenon that has been also shown before by us and others^[34;38;39;47] and we reconfirmed previous findings^[30;39] that N-glycans with α 2,3-sialylation emerged in differentiated hPSC (Figure 6B: [14] – [18]). In our experimental setup, differentiation resulted in decreasing levels of certain biantennary N-glycans with α 2,6-linked sialic acids (Figure 6A: [4], [6] – [9]). Sialylated structures in general comprised 19%, 24% and 15% on d0, d7 and d15, respectively. Upon removal of α 2,3-linked sialic acids with a specific sialidase, these values changed to 19%, 19% and 9% (Supporting Information Table S3). However, it has to be considered that sialylation might be influenced by changes in the underlying glycans. In agreement with previous findings,^[28;30;31;39] mono-sialylated and complex fucosylated (core-fucose plus at least one additional α 1,2- or α 1,3-linked antennae fucose) glycans (Figure 6A: [6] to [8]) were considerably down-regulated upon differentiation. Accordingly, antennae fucosylation in α 1,2-linkage has been reported by Tateno *et al.*^[34] in the context of glycan reversion upon reprogramming of somatic cells to pluripotent stem cells and was described as specific epitope for hiPSCs.^[30] Interestingly, biantennary N-glycans with β 1,3-linked terminal galactose (Figure 6A: [7] – [9]) were exclusively detectable in d0 hiPSCs and not on d7 or d15 of differentiation. In contrast, N-glycans containing β 1,4-linked galactose were detected at all three time points analyzed. The presence of a β 1,3 linked terminal galactose (type I LacNAc) on a neutral triantennary N-glycan was described as characteristic for hiPSCs.^[30] Furthermore, Hasehira *et al.* described the emergence of a type I LacNAc structure on an abundant O-glycan in hiPSCs.^[30] Natunen *et al.* reported that type I LacNAc structures on O-glycans are the epitopes recognized by the pluripotency marker antibodies Tra-1-60 and Tra-

1-81^[66] and Tateno *et al.* also observed increased binding of lectins recognizing β 1,3-linked galactose in hiPSCs compared to somatic cells, without further distinguishing between O- and N-glycans.^[34] However, we here describe three (Figure 6A: [7]-[9]) so far unrecognized, rather abundant biantennary N-glycans with β 1,3-linked terminal galactose, which were mono-sialylated in α 2,6-linkage and two of them (Figure 6A: [7], [8]) were even complex-fucosylated. We propose that these three features in combination, sialylation in α 2,6-linkage together with complex fucosylation and β 1,3-linked galactose (represented in Figure 7) might make these glycans rather exclusive to hPSCs and thereby attractive targets as hPSC-specific marker structures. Increased expression of the major beta 3 galactosyltransferase B3GALT5, which is involved in the synthesis of the glycosphingolipid-based pluripotency markers SSEA-3 and SSEA-4 has been observed in hPSCs^[34] and is likely to be responsible for the synthesis of the here mentioned glycans containing β 1,3-linked galactose.

Regarding glycan peaks that increased from d0 to d15, we observed an increase of α 2,3-sialylation (Figure 6B , [14]-[18]), which is not present in hiPSCs as mentioned above, as well as an increase of structures with antennae branching (Figure 6B, [17]-[20]). Bisecting GlcNAc structures (Figure 6B, [11]-[13]) represent a highly abundant glycan type in the d7 cardiac progenitors and d15 CMs, and were also reported in stem cell-derived CMs in a comprehensive study comprising a global analysis of the N-glycome of hiPSCs, hiPSC-CMs and commercially available human adult ventricular CMs.^[39] Interestingly, although α 2,6-sialylation appears to be characteristic for undifferentiated cells, one particular bisecting glycan carrying one α 2,6-residue was strongly up-regulated during differentiation (Figure 6B, [13]). This structure has also been mentioned before to be more abundant in hiPSC-CMs than in hiPSCs.^[39] Interestingly Kawamura *et al.* reported that bisecting glycans have not been identified on adult heart CMs and they hypothesized that the occurrence of bisecting glycans in hiPSC-CMs was due to contaminating cells from other lineages. However, in our study, the bisecting GlcNAc structures

were highly abundant in stem cell-derived cardiac progenitors and CMs. Different to the study by Kawamura *et al.*,^[39] who applied an EB-based differentiation protocol with cytokine-mediated modulation of the Wnt signaling pathway, we used a more recent and efficient small molecule-based monolayer approach.^[51;52] Thus, we assume that bisecting GlcNAc structures do not originate from contaminating cell types, but from early, immature CMs. Considering the previous data by Kawamura *et al.*,^[39] down-regulation of the respective enzymes (MGAT3, ST6GAL1) is expected to cause reduction of bisecting glycans during maturation of CMs. Additionally, bisecting GlcNAc structures, e.g. on E-cadherin, have been reported to inhibit metastasis formation by suppressing cell spreading, migration and proliferation.^[67] This also fits well with our observation that bisecting GlcNAc structures are increased in less proliferative differentiated cells.

Bisecting GlcNAc residues cannot be processed by the enzyme responsible for tri- and tetraantennary branching and conversely, tri- and tetraantennary structures are no substrates for MGAT3 (also referred to as GNT-III), which is responsible for the addition of the bisecting GlcNAc,^[68] as reviewed in Zhao *et al.*^[67] However, we observed that non-sialylated tri- and tetraantennary N-glycans with exposed galactose residues were only detectable in differentiating cells on d7 and d15, but not in undifferentiated cells [Figure 6B, [19], [20]], which is also in concordance with literature.^[39;48] However, the increase of both, bisecting and tri-/tetraantennary structures with exposed galactose during differentiation is not in discrepancy, because the two glycan types do not necessarily originate from the same cells in the differentiated cell population. Moreover it has been reported that the exposure of galactose, or rather the total absence of sialic acid, as hypothesized by Alisson-Silva *et al.*^[48] induces spontaneous differentiation of hiPSCs, fitting to our observation that these structures were only identified on differentiated cells. Taken together, our xCGE-LIF-based glycomic analysis reproduced most findings of previous studies that applied mass spectrometry or lectin-based technologies for glycan determination on hPSCs^[28;30;31;34] and CMs derived thereof, demonstrating the validity of our approach.^[39] Notably,

we additionally identified three novel glycan structures that combined the features, β 1,3-linked galactose, α 2,6-linked sialic acid and complex fucosylation on the very same glycan molecule and which were highly specific for hiPSCs. Moreover, we propose that bisecting GlcNAc structures emerge on immature hiPSC-derived cardiomyocytes

Conclusion

In summary, we here present a global analysis of the proteome, sialo-glycoproteome and N-glycome of human iPSCs, cardiovascular progenitors and early CMs derived thereof. Using *state-of-the-art* technologies we were able to reproduce earlier findings regarding cell surface markers of hiPSCs and hiPSC-CMs with relative ease and identified a set of novel surface molecules, both on the sialo-glycoproteomic as well as on the glycomic level. These are candidates to be further validated for their potential as stem cell or cardiomyocyte biomarkers enabling their envisioned application as targets for immunophenotyping or enrichment of desired cells types.

Experimental Section

Cell culture and SILAC labeling of hiPSCs: Cell culture reagents were purchased from Life Technologies (Carlsbad, CA, USA) unless stated otherwise. All cells were maintained at 37°C, 5% CO₂ and 85% relative humidity. Culture of hiPSCs was conducted under feeder-free conditions in mTeSRTM1 medium (STEMCELL Technologies, Grenoble, France) in T25 cell culture flasks (Greiner Bio-One) coated for at least 1h at 37°C with 2 mL of a 1:60 dilution of MatrigelTM matrix (BD Biosciences, Bedford, MA, USA) in DMEM. For routine passaging, cells were incubated for 7 min at 37°C with 1 mg/mL [w/v] Dispase (STEMCELL Technologies) and subsequently transferred in new MatrigelTM coated flasks. For SILAC labeling, hiPSCs were cultivated at least three passages in SILAC mTeSRTM1 (STEMCELL Technologies) on MatrigelTM. SILAC mTeSRTM1 was supplemented with 7×10^{-4} mol/l of the respective arginine (Arg) isotopologue (Arg-6 = L-[¹³C₆] Arg-HCl, referred to as “medium” Arg; Arg-10 = L-[¹³C₆¹⁵N₄] Arg-HCl, referred to as “heavy” Arg) and 5×10^{-4} mol/l of the respective lysine (Lys) isotopologue (Lys-4 = L-[²H₄] Lys-HCl, referred to as “medium” Lys; Lys-8 = L-[¹³C₆¹⁵N₂] Lys-HCl, referred to as “heavy” Lys) (Silantes, Munich, Germany). Standard mTeSRTM1 was used as “light” (control) condition.

Cardiomyogenic differentiation of hiPSCs: Cardiomyogenic differentiation from monolayers of hiPSCs was performed essentially as described by Lian *et al.*^[51;52] with the exception that for differentiation under SILAC conditions, the RPMI component was changed throughout the differentiation process to RPMI 1640 for SILAC supplemented with 1% [v/v] GlutamaxTM, 1.15 mM Arg-6 (“medium” condition) or Arg-10 (“heavy” condition) and 0.22 mM Lys-4 (“medium” condition) or Lys-8 (“heavy” condition) in the following referred to as SILAC RPMI. Briefly, SILAC-labeled or control hiPSCs were singularized by incubation with 0.5 mM EDTA/PBS for 10 min, then pelleted by centrifugation at 300xg for 5 min and seeded in a cell density of 9×10^5 cells/cm² on a 12-well-plate (Greiner Bio-One) in SILAC RPMI (medium, heavy) or standard

RPMI (light) supplemented with 5 μ M RI (day -3). On day -2 and day-1, the respective medium without RI was added. On day 0, when cells were fully confluent, the medium was replaced by (SILAC) RPMI / B27 medium minus insulin consisting of (SILAC) RPMI 1640 and 2% [v/v] B-27[®] Supplement minus insulin supplemented with the small molecule CHIR99021 (Selleckchem, Houston, TX, USA) in a final concentration of 8 μ M. The cells were cultivated for 24 h in the presence of CHIR99021. Then the medium was changed to (SILAC) RPMI / B27 minus insulin (day 1). On day 3, 5 μ M of the small molecule IWP-4 (Stemgent[®], Cambridge, MA, USA) in (SILAC) RPMI / B27 minus insulin were added and removed again on day 5. On day 7, the medium was replaced by (SILAC) RPMI 1640 with 2% [v/v] B-27[®] Supplement with insulin. Spontaneously contracting areas could be observed from day 10 onwards.

Experimental design: As depicted in Figure 1, three different time points of differentiation, i.e. d0 for the undifferentiated cells, d7 for cardiac progenitors and d15 for differentiated, spontaneously contracting CMs were studied. For each time point, three independent replicates were generated by cultivation under light, medium or heavy SILAC conditions, respectively. Cell lysates corresponding to equal amounts of protein from three different time points and with three different SILAC labels, e.g. d0 heavy, d7 medium and d15 light, were pooled (Supporting Information Table S4) and for analysis of the whole cell proteome directly precipitated or, for sialo-glycoprotein purification, processed according to the PAL method^[22] (Supporting Information Figure S6) and subsequently precipitated.

Cell harvest and lysis for whole cell proteomics and glycomics: Cells were harvested by release with 0.5 mM EDTA after washing twice with PBS and pelleted by centrifugation at 500xg for 5 min. The cell pellet was washed with PBS and the cell pellet was frozen at -80°C until further use. For cell lysis, 1 mL of RIPA buffer was used to lyse 5×10^6 of hiPSCs or differentiated cells. RIPA buffer comprised 50 mM Tris-HCl, pH=8 with 150 mM NaCl, 1% [v/v] NP-40 (Roche),

0.5% [w/v] sodium deoxycholate (Sigma Aldrich), and 1% [w/v] sodium dodecyl sulfate (SDS) (Serva, Heidelberg, Germany) supplemented with 1% HALT protease inhibitor (Thermo Fisher Scientific). Cells were sonicated two times using a Branson Sonifier 450 (settings: duty cycle 50%, output control 5). Cell debris were removed by centrifugation at 13000xg for 2 minutes and protein concentration of cell lysates was determined at 660 nm using the Pierce[®] 660nm protein assay reagent (Thermo Fisher Scientific) in a 96-well plate reader.

Periodate oxidation and aniline-catalyzed oxime ligation (PAL): Sialo-glycoproteins of (SILAC-labeled) hiPSCs, cardiac progenitors and CMs were assessed by PAL (periodate oxidation and aniline-catalyzed oxime ligation) essentially as described by Zeng *et al.*^[22] Briefly, for periodate oxidation, living cells on 12-well plates were incubated with 1 mM NaIO₄ and excess of periodate was neutralized by incubation with 1 mM glycerol. For oxime ligation, cells were incubated with 5% [v/v] FCS / 100 μM aminooxybiotin (Biotium, Hayward, CA, USA) / 10 mM aniline at pH = 6.7. After biotinylation, cells were pelleted by centrifugation and lysed in lysis buffer (150 mM NaCl, 20 mM Tris pH=7.4, 1% [v/v] Nonidet P40 (Roche) with 1% [v/v] HALT protease inhibitor (Thermo Fisher Scientific)) by sonication as described above. Clarified cell lysates of one sample of d0, d7 and d15, respectively, of different SILAC labels were pooled at equal amounts (100 μg protein, each) and all subsequent purification steps were carried out with the pooled lysates. Biotinylated sialo-glycoproteins were purified with 200 μL High Capacity Neutraavidin[®] Agarose Resin (Thermo Fisher Scientific) according to the manufacturer's instructions. As described by Tagwerker *et al.*,^[69] we applied harsher washing conditions for Neutraavidin[®]-bound proteins to efficiently remove unbound protein. The washing steps were as follows: three times for 1 min with lysis buffer incl. protease inhibitor, then 2x 10 min and 1x 1 min with 8M urea, 200 mM NaCl, 2% [w/v] SDS (Serva), 100 mM Tris, pH = 8 incl. 1% [v/v] HALT protease inhibitor, 2x 10 min and 1x 1 min with 8 M urea, 1,2 M NaCl, 0.2% [w/v] SDS, 100 mM Tris, 10% [v/v] 2-propanol, 10% [v/v] ethanol, pH = 8, 2x 10 min with 30 mM Tris, 2 M

thiourea, 7 M urea, 4% [w/v] CHAPS, pH = 8.5 incl. HALT protease inhibitor, 2x 10 min with lysis buffer incl. 50 mM dithiothreitol and 3x 1 min with 50 mM Tris pH = 6.8. To remove washing buffers between the washing steps, samples were centrifuged at 1000xg for 1 min. To elute biotinylated sialo-glycoproteins, NeutrAvidin beads were incubated with 2% [w/v] SDS, 3 mM biotin, in 100 mM Tris pH = 8.5 for 15 min at room temperature and for 10 min at 99°C and 850 rpm in a thermomixer and proteins were eluted by centrifugation.

SDS-PAGE, in gel digestion, LC-MS/MS and MS data analysis: Proteins were precipitated by adding the four-fold volume of acetone o/n at -20°C and subsequently pelleted by centrifugation at 13,000xg for 15 min at 4°C. The precipitated proteins were dissolved in Laemmli buffer (35 mM Tris-HCl pH = 6.8, 2.8% [w/v] SDS, 7% [v/v] glycerol, and 0.005% [w/v] bromophenol blue), incubated for 5 min at 95°C and separated by SDS-PAGE using 10% polyacrylamide gels with 5% stacking gels and subsequently stained for 16-18 h with Roti[®]-Blue Coomassie dye (Roth, Karlsruhe, Germany). In-gel digestion, LC-MS/MS and MS data analysis was performed essentially as described recently.^[49] We applied the MaxQuant proteomics software suite (v1.4.1.2)^[70] for identification and quantification of proteins and peak lists were searched against the uniprot homo sapiens database (uniprot-homo+sapiens, downloaded on 09/07/15 at www.uniprot.org).

Determination of SILAC incorporation efficiencies: The proportion of medium and heavy amino acid incorporation was assessed from individual hiPSC samples that were either medium or heavy labeled. Sample preparation, MS analysis and MS data analysis was carried out as described above. We used 50% of all peptides that were identified with the highest intensity in the respective samples. The incorporation efficiency was calculated as [%] = 100* (1-(1/(ratio +1))) for medium/light (M/L) or heavy/light (H/L) ratios of these peptides of the different samples.

Biostatistical analysis of MS data: Statistical significance testing of proteomics and glycoproteomics data was performed using student's t-test and a p-value of <0.05 was applied as threshold for significance. Venn diagram was constructed using the online tool BioVenn (<http://www.cmbi.ru.nl/cdd/biovenn/>).^[71] Gene ontology analysis for cellular compartments was performed using the implemented annotations in Perseus v.1.5.3.2. Gene ontology analysis for KEGG pathways was conducted with the online tool STRING (Search Tool for the Retrieval of Interacting Genes/Proteins; www.string-db.org)^[72] searching for *homo sapiens* proteins.

Multiplexed Capillary Gel Electrophoresis with Laser Induced Fluorescence Detection

(xCGE-LIF): Cell lysates were prepared as described above. Proteins were precipitated with a four-fold excess volume of ice cold acetone at -20 °C overnight. After centrifugation, protein pellets were dissolved in 2% [w/v] SDS in 1x PBS and further sample preparation for xCGE-LIF-based glycoanalysis was performed as described previously.^[38;73] Briefly, remaining SDS was neutralized with 8% [v/v] IGEPAL and proteins were digested in solution with peptide-N-glycosidase F (PNGase F from *Elizabethkingia meningoseptica*, BioReagent, Sigma Aldrich) to release the attached N-glycans. Released N-glycans were labeled with 8-aminopyrene-1,3,6-trisulfonic acid (APTS, Sigma Aldrich), separated by multiplexed capillary gel electrophoresis (xCGE) and monitored via laser induced fluorescence (LIF) detection. glyXtool™ software (glyXera, Magdeburg, Germany) was applied for xCGE-LIF data processing and normalization of migration times to an internal standard. N-glycan “fingerprints” (normalized electropherograms) were used for annotation of N-glycan peaks via migration time matching with the in-house N-glycan database. N-glycan structures were confirmed by extensive exoglycosidase digests and subsequent repeated analysis by xCGE-LIF. Following specific enzymes were used: $\alpha(2,3)$ sialidase (Sialidase S, recombinant from *Streptococcus pneumoniae*, expressed in *Escherichia coli*; Prozyme, Hayward, CA, U.S.A.), $\alpha(2,3,6,8)$ sialidase (Sialidase A, recombinant from *Arthrobacter ureafaciens*, expressed in *Escherichia coli*; Prozyme), $\alpha(1,3,4)$ fucosidase

(recombinant from *Xanthomonas*; QABio), $\alpha(1,2)$ fucosidase (recombinant from *Xanthomonas manihotis*, expressed in *Escherichia coli*; New England Biolabs), $\alpha(1,2,3,4,6)$ fucosidase (from bovine kidney; Prozyme), $\beta(1,3)$ galactosidase (recombinant from *Xanthomonas manihotis*, expressed in *Escherichia coli*; New England Biolabs), $\beta(1,4)$ galactosidase (recombinant from *Bacteroides fragilis*, expressed in *Escherichia coli*; New England Biolabs), $\beta(1,4,6)$ galactosidase (from Jack bean; Prozyme), $\beta(1,2,3,4,6)$ -N-acetylglucosaminidase (recombinant from *Streptococcus pneumoniae*, expressed in *Escherichia coli*; New England Biolabs), $\alpha(1,2,3,6)$ mannosidase (from Jack bean; Prozyme).

Flow cytometry: For the analysis of SSEA-4 or cardiac troponin T expression by flow cytometry, cells were washed twice with PBS and dissociated into single cells with 0.5 mM EDTA/PBS by incubation at 37°C for 10 min. For troponin T staining only, cells were fixed for 15 min at 4°C with 4% [w/v] paraformaldehyde (PFA) and permeabilized with 0.1% [v/v] TWEEN-20 in 2% [w/v] bovine serum albumin (BSA)/PBS. Cells were pelleted and resuspended in FACS buffer comprising 0.5% [w/v] BSA in PBS and 3×10^5 cells each were incubated with anti-SSEA-4 or anti-troponin T or the respective isotype controls (mouse IgG3 κ or IgG $_1$) and anti-mouse IgG (H+L)-Alexa Fluor[®]488 as secondary antibody in 0.5% [w/v] BSA/PBS (antibodies and dilutions are listed in Supporting Information Table S5). Antibody incubation was performed by shaking on a thermomixer for 45 min at 700 rpm and room temperature. After each staining step, cells were washed twice with 400 μ L 0.5% [w/v] BSA/PBS. Cells were resuspended in 0.5% [w/v] BSA/PBS and analyzed using a CyFlow ML flow cytometer (Partec GmbH, Münster, Germany) and Flowing Software 2 (Perttu Terho, Turku Centre for Biotechnology, Turku, Finland).

Immunofluorescence microscopy: For immunofluorescence microscopy (IF), cells were seeded onto Matrigel-coated 3.5 cm plastic dishes. hiPSCs were passaged as described above and grown for 2-4 days until they reached the desired density or colony size. hiPSCs or

differentiated cells from monolayer cultures were singularized with TrypLE Select (37°C, 5 min) after washing twice with PBS and further processed as described above.

For fixation, cells were washed with PBS and fixed with 4% [w/v] PFA in PBS for 30 min and blocked with 2% [w/v] BSA/PBS for 20 minutes. Incubation with antibodies was performed for 1 h at room temperature for each primary or secondary antibody. Between the different staining steps the cells were washed for 15 min with PBS. For staining of intracellular epitopes, cells were permeabilized in 0.1% [v/v] Triton X-100 for 15 min prior to blocking and incubation with the respective antibody. Cells were blocked in 2% [w/v] BSA/PBS with or without 0.1% [v/v] Triton X-100 for 20 min. Antibodies and dilutions used for IF are given in Supporting Information Table S5. All antibodies were diluted in 2% [w/v] BSA/PBS with or without 0.1% [v/v] Triton X-100. A freshly prepared 1:4000 dilution of Hoechst 32258 (AppliChem) with PBS was used to counterstain cell nuclei. Staining of the cytoskeleton was achieved by incubation for 20 min with a 1:80 dilution of Alexa Fluor® 546 phalloidin (Life Technologies) with PBS. Coverslips were mounted onto the plastic dishes after staining using one droplet of Dako fluorescence mounting medium (Agilent Technologies, Böblingen, Germany) and analyzed using a Zeiss Axiovert 200M microscope equipped with AxioCam MRm digital camera and Axio Vision software v4.7 (Zeiss, Oberkochen, Germany).

Isolation of RNA: For each sample, cells from one well of a 12-well plate were collected as described above and stored at -80°C until RNA preparation. Then, cell pellets were resuspended in 0.5 mL TRIzol® (Life Technologies). Extraction of RNA was achieved by adding 0.1 mL chloroform and centrifugation at 12,000xg. Supernatants were mixed with equal volumes of ice-cold 70% [v/v] ethanol and applied onto NucleoSpin® RNA II isolation columns (Macherey & Nagel, Düren, Germany) and subsequent steps were performed according to the manufacturer's instructions. RNA was eluted in sterile ddH₂O and stored at -80°C. The concentration of total RNA was determined at 260 nm using an UV spectrophotometer.

cDNA synthesis: cDNA synthesis was performed on 2.5 µg total RNA per sample in a volume of 50 µL per tube. Residual genomic DNA was removed by digestion with 2.5 U RQ1 RNase-free DNase M610A in 1x RQ1 DNase Reaction Buffer M198A (both from Promega) for 30 min at 37°C, then the reaction was stopped by adding RQ1 DNase stop solution (Promega) and incubating for 10 minutes at 70°C. RNA was reverse transcribed into cDNA using random hexamer primers (Life technologies) and 200 U RevertAid™ Premium Reverse Transcriptase in 1x RT Buffer (both Thermo Fisher Scientific) by incubating the samples for 10 min at 25°C, 50 min at 42°C and 10 min at 70°C. All reaction steps were performed in a Piko PCR cycler (Thermo Fisher Scientific). The synthesized cDNA was stored at -20°C and used in a 1:10 dilution for qPCR.

Quantitative real-time PCR (qPCR): Quantitative real-time PCR (qPCR) was performed on a 7500 Fast Real-time PCR System (Applied Biosystems, Foster City, CA, USA) in sealed 96-well optical reaction plates (Applied Biosystems) using 20 ng cDNA and 0.475 pmol of each primer (forward / reverse) in a mastermix with 4 nmol of each dNTP (dATP, dCTP, dTTP, dGTP), 10% [v/v] Maxima Hot Start Taq PCR buffer, 1.875 mM MgCl₂, 0.5 U Maxima Hot Start Taq DNA Polymerase (all from Thermo Scientific), 1% [v/v] ROX reference dye and 8% [v/v] of a 1:30000 dilution of SYBR Green Nucleic Acid Stain (both from Life Technologies) in a total volume of 20 µL. The PCR was run for 40 two-step cycles (60 s at 60°C, 15 s at 90 °C) after an initial denaturation step for 10 min at 95°C. Determination of product purity was performed by melt-curve analysis in an additional cycle (95°C, 15 s, 60°C 60 s, 95°C, 30 s, 60°C 15 s). Relative expression of target genes was determined by normalization to ACTB. CT values were determined automatically by the 7500 Software v2.0.5. Relative expression was calculated using the $\Delta\Delta CT$ values, with the relative expression being = $2^{-\Delta\Delta CT}$ depending on the evaluation to be

made. Sequences of primers used in qPCR analyses are listed in Supporting Information Table S6.

Acknowledgements

This work is supported by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for the Cluster of Excellence REBIRTH (From Regenerative Biology to Reconstructive Therapy, EXC 62/2). For ER and RH this work was supported by the European Union under the project “HighGlycan” (European Commission (EC) grant no. 278535). ER acknowledges support by the European Union under the project “HTP-GlycoMet” (European Commission (EC) grant no. 324400). For SC and ER this work was supported by the German Federal Ministry of Education and Research (BMBF) under the project “Die Golgi Glykan Fabrik 2.0” (grant identifier 031A557). The authors would like to thank Prof. Dr. Gerardy-Schahn, head of the Institute of Clinical Biology, Hannover Medical School (MHH) for providing general laboratory equipment and for carefully reading this manuscript. We also thank Prof. Dr. Scheper (Institute of Technical Chemistry, Leibniz University of Hannover) for providing the ROCK inhibitor Y27632, Prof. Dr. Martin and Dr. Haase (LEBAO, MHH) for providing hiPSCs and we would like to thank Karsten Heidrich for practical assistance.

Abbreviations

ACN, acetonitrile; AmBic, ammonium bicarbonate buffer; d, day; DTT, dithiothreitol; h, hour; hCBiPSC2, human cord blood derived induced pluripotent stem cell clone 2; hESCs, human embryonic stem cells; hiPSCs, human induced pluripotent stem cells; hPSCs, human pluripotent stem cells; hPSC-CMs, human pluripotent stem cell-derived cardiomyocytes; IF, immunofluorescence; min, minute; MYH6, myosin heavy chain 6 (α -MHC); PAL: periodate oxidation and aniline-catalyzed oxime ligation, qPCR, quantitative real-time polymerase chain

reaction; RI, Rho-associated coiled-coil kinase inhibitor Y27632; SILAC, stable isotope labeling by amino acids in cell culture, xCGE-LIF: multiplexed capillary gel electrophoresis and laser-induced fluorescence detection

Keywords

Proteomics; sialic acids; glycomics; human pluripotent stem cells; cardiomyocytes

Tables

Table 1. Numbers of significantly regulated proteins in the whole cell proteome analysis (2882 proteins in total).

timepoint comparison	significantly (p<0.05) up-regulated >1.5x	significantly (p<0.05) down-regulated >1.5x
d7/d0	235	213
d15/d0	299	198
d15/d7	172	76

Table 2. Numbers of significantly regulated proteins in the sialo-glycoproteome analysis (879 proteins in total).

timepoint comparison	significantly (p<0.05) up-regulated >1.5x	significantly (p<0.05) down-regulated >1.5x
d7/d0	55	56
d15/d0	42	51
d15/d7	30	33

Table 3. List of proteins previously proposed by Boheler *et al.* [25] as positive markers for PSCs and also identified by sialo-glycoproteomics in the present study.

Protein full name	Gene name	Regulation ratio d7/d0	Regulation ratio d15/d0
Gastrin/cholecystokinin type B receptor	CCKBR	2.74±2.25	0.44±0.37
Cadherin-3	CDH3	0.40±0.21	0.13±0.17
Cadherin EGF LAG seven-pass G-type receptor 2	CELSR2	0.32±0.28	0.26±0.43
Ciliary neurotrophic factor receptor subunit alpha	CNTFR	12.14±6.34	4.06±1.77
Ephrin type-A receptor 7	EPHA7	2.95*±0.19	1.09±0.64
Glycerophosphoinositol inositolphosphodiesterase	GDPD2	0.08±0.52	0.09*±0.53
Hephaestin	HEPH	1.85*±0.07	2.05±0.59
Immunoglobulin superfamily member 1	IGSF1	0.18*±0.06	0.48±0.49
Vascular endothelial growth factor receptor 2	KDR (CD309)	0.35±0.37	0.18±0.47
Laminin subunit alpha-1	LAMA1	1.61±0.56	2.59±4.14
Laminin subunit beta-2	LAMB2	1.05±0.33	1.41±0.42
Leucine-rich repeat-containing G-protein coupled receptor 4	LGR4	3.96±3.42	0.78±0.64
Leucine-rich repeats and immunoglobulin-like domains protein 1	LRIG1	0.21±0.46	0.13±0.50
Low-density lipoprotein receptor-related protein 4	LRP4	1.00±0.43	2.58±1.29
Leucine-rich repeat neuronal protein 1	LRRN1	0.19*±0.02	0.07*±0.06
Matrix metalloproteinase-15	MMP15	1.09±0.12	0.50±0.36
Nicalin	NCLN	0.64*±0.05	0.47±0.16
Neuroigin-4	NLGN4X	0.19*±0.05	0.15*±0.06
Atrial natriuretic peptide receptor 1;Guanylate cyclase	NPR1	0.70±0.17	0.89±0.06
Plexin domain-containing protein 2	PLXDC2	0.43*±0.06	0.68±1.42
Receptor-type tyrosine-protein phosphatase delta	PTPRD	1.87±2.16	1.96±2.54
Receptor-type tyrosine-protein phosphatase zeta	PTPRZ1	0.22*±0.08	0.09±0.18
Sodium-coupled neutral amino acid transporter 2	SLC38A2	0.33±0.42	0.25±0.44
Sodium-coupled neutral amino acid transporter 5	SLC38A5	0.08±0.53	0.14*±±0.49
Cationic amino acid transporter 3	SLC7A3	0.15*±0.06	0.11*±0.09
Kunitz-type protease inhibitor 1	SPINT1	0.30*±0.14	0.16*±0.14
Tumor necrosis factor receptor superfamily member 8	TNFRSF8	0.12*±0.04	0.07*±0.03

* indicates p<0.05 (after student's t-test)

Regulation ratios depicted in green: down-regulation (>1.5-fold) on d7 or d15 compared to d0.

Regulation ratios depicted in red: up-regulation (>1.5-fold) on d7 or d15 compared to d0.

Figure Legends

Figure 1. Experimental setup. (A) Experimental workflow for analysis of whole cellular proteins (a), sialo-glycoproteins (b) and N-glycans (c) of hiPSCs, cardiac progenitors and CMs. For whole cell proteomics (a) and sialo-glycoproteomics (b), cells were cultivated under three different SILAC conditions (light, medium, heavy) and harvested at three time points (d0, d7, d15) as indicated, resulting in a total of nine samples. For each biological replicate (n=3), three samples comprising time point d0, d7 and d15 as well as SILAC condition light, medium and heavy were pooled and analyzed by mass spectrometry (MS). For analysis of N-glycans (c), three biological replicates of individual samples of d0, d7 and d15, each, were analyzed by xCGE-LIF.

Figure 2. Quality control of hiPSCs and hiPSC-CMs. (A) Immunofluorescence microscopy of hiPSCs cultured under light, medium and heavy conditions detecting the pluripotency markers OCT-3/4 (green) and SSEA-4 (red). Nuclei were counterstained with Hoechst 33258. Scale bars represent 100 μ m. (B) qPCR assessing expression levels of the four pluripotency factors *OCT-3/4*, *SOX2*, *LIN28* and *NANOG* in exactly those three d0 replicates (light, medium and heavy) that were used for proteomic, sialo-glycoproteomic and glycomic analysis. (C) Flow cytometry analysis of the same three d0 replicates (light, medium and heavy) detecting the cell surface pluripotency marker SSEA-4. IgG3 κ was used as isotype control. Cells were stated positive for SSEA-4 if the fluorescence signal was higher than of 99% of the isotype control cells. (D) Immunofluorescence microscopy of d15 CMs cultured under three different SILAC conditions (light, medium and heavy) detecting the intracellular CM marker α -actinin. Nuclei were counterstained with Hoechst 33258, the cytoskeleton was visualized using phalloidin-Alexa488. Scale bars represent 100 μ m. (E) qPCR of the CM-specific gene products *MYH6* and *NKX2.5* of exactly the d0, d7 and d15 replicates used for the proteomic, sialo-glycoproteomic and glycomic

analysis. (F) Flow cytometry analysis of the three d15 replicates (L: light; M: medium; H: heavy) used for the proteomic, sialo-glycoproteomic and glycomic analysis showing expression of the CM marker cardiac troponin T (cTnT). IgG1 was used as isotype control. Cells were stated positive for cTnT if the fluorescence signal was higher than of 99% of the isotype control cells.

Figure 3. Comparative analysis of proteins identified by whole cell proteomics or sialo-glycoproteomics. (A) Numeric Venn diagram showing the overlap between proteins identified in the whole cell proteomic and the sialo-glycoproteomic screen. (B) Gene ontology analysis of cellular components for proteins identified by sialo-glycoproteomics only (left panel), by whole cell proteomics only (middle panel) or by both (right panel). (C) KEGG pathway enrichment analysis of proteins significantly up-regulated (left panel) or down-regulated (right panel) in at least one of the three time point comparisons d7/d0, d15/d0 or d15/d7. Blue bars indicate the numbers of proteins found for each KEGG category; purple dotted line represents the p-value for the enrichment. The Bonferroni model was applied for multiple testing correction of the p-value. Only pathways with a p-value <0.05 are presented.

Figure 4. Regulation of proteins during cardiomyogenic differentiation. (A) Volcano plot of all proteins of the d7/d0 comparison of sialo-glycoproteins purified by PAL. The log₂ mean regulation ratio is plotted against the negative decadic logarithm of the p-value obtained from Student's t-test. (B, C) Similar presentation as in (A), but for proteins of the d15/d0 (B) and d15/d7 (C) comparison of PAL-purified sialo-glycoproteins. Proteins names mentioned in the manuscript text are indicated in black (significantly regulated) or grey (non-significantly regulated). Putative markers for CMs or cardiac progenitors further referred to in the manuscript text are highlighted in red. All proteins significantly and substantially up- or down-regulated are listed in Additional Material Table S2.

Figure 5. d0, d7 and d15 specific N-glycosylation fingerprints. N-glycans were annotated via migration time matching with an in-house N-glycan database and confirmed by various exoglycosidase digests (see Experimental Section). Some prominent N-glycan structures that are differently distributed at d0, d7 and d15 are depicted; black boxed are N-glycan structures decreasing, and red boxed are the ones increasing with differentiation. Symbolic representation of N-glycan structures were drawn with GlycoWorkbench Version 1.1, following the guideline of the Consortium for Functional Glycomics.^[74] Linkage positions of sialic acid residues are indicated by differing angles.

Figure 6. Relative quantification of prominent N-glycan peaks from xCGE-LIF analyses.

Presented are % signal intensities of the 10 most prominent N-glycans that show a decrease (A) or increase (B) over the course of the differentiation, i.e. from d0 to d7, from d7 to d15 and/or from d0 to d15. Significance testing was performed using one-way ANOVA with Newman-Keuls post-test (as implemented in the software GraphPad Prism version 4.03). Non-significant changes (with a p-value >0.05) are indicated as n.s, all other changes are significant with p<0.05. Glycans are numbered in squared brackets to be easily referred to in the manuscript text and in Supporting Information Table S3. Symbolic representation of N-glycan structures are as described in the legend of Figure 5. Linkages of sialic acids are indicated as $\alpha 6$ for $\alpha 2,6$ -linkage and $\alpha 3$ for $\alpha 2,3$ -linkage. Galactose linkage is only highlighted in case of β -1,3-linkage as $\beta 3$. In all other cases, galactose linkage is β -1,4.

Figure 7. Novel hPSC N-glycans. Presentation of three N-glycans that are potentially characteristic and unique for hPSCs. These N-glycans are all biantennary with sialylation in $\alpha 2,6$ -linkage and $\beta 1,3$ -linked galactose and can further contain antennae fucosylation.

TOC (table of contents) graphic. Human induced pluripotent stem cells, cardiac progenitors and stem cell-derived cardiomyocytes were analyzed and compared at a sialo-glycoproteomic, proteomic and N-glycomic level. This approach led to the identification of novel potential biomarkers for these cell types.

Reference List

- [1.] N. Townsend, M. Nichols, P. Scarborough, M. Rayner, *Eur.Heart J.* **2015**, *36* 2696-2705.
- [2.] O. Bergmann, R. D. Bhardwaj, S. Bernard, S. Zdunek, F. Barnabe-Heider, S. Walsh, J. Zupicich, K. Alkass, B. A. Buchholz, H. Druid, S. Jovinge, J. Frisen, *Science* **2009**, *324* 98-102.
- [3.] J. Kajstura, M. Rota, D. Cappelletta, B. Ogorek, C. Arranto, Y. Bai, J. Ferreira-Martins, S. Signore, F. Sanada, A. Matsuda, J. Kostyla, M. V. Caballero, C. Fiorini, D. A. D'Alessandro, R. E. Michler, F. del Monte, T. Hosoda, M. A. Perrella, A. Leri, B. A. Buchholz, J. Loscalzo, P. Anversa, *Circulation* **2012**, *126* 1869-1881.
- [4.] M. A. Laflamme, C. E. Murry, *Nature* **2011**, *473* 326-335.
- [5.] M. Talkhabi, N. Aghdami, H. Baharvand, *Life Sci.* **2016**, *145* 98-113.
- [6.] K. Takahashi, S. Yamanaka, *Cell* **2006**, *126* 663-676.
- [7.] J. Yu, M. A. Vodyanik, K. Smuga-Otto, J. Antosiewicz-Bourget, J. L. Frane, S. Tian, J. Nie, G. A. Jonsdottir, V. Ruotti, R. Stewart, I. I. Slukvin, J. A. Thomson, *Science* **2007**, *318* 1917-1920.
- [8.] I. Y. Chen, E. Matsu, J. C. Wu, *Nat.Rev.Cardiol.* **2016**, *13* 333-349.
- [9.] T. J. Nelson, A. Martinez-Fernandez, S. Yamada, C. Perez-Terzic, Y. Ikeda, A. Terzic, *Circulation* **2009**, *120* 408-416.
- [10.] D. Wang, Y. Jin, C. Ding, F. Zhang, M. Chen, B. Yang, Q. Shan, J. Zou, K. Cao, *Ir.J.Med.Sci.* **2011**, *180* 379-385.

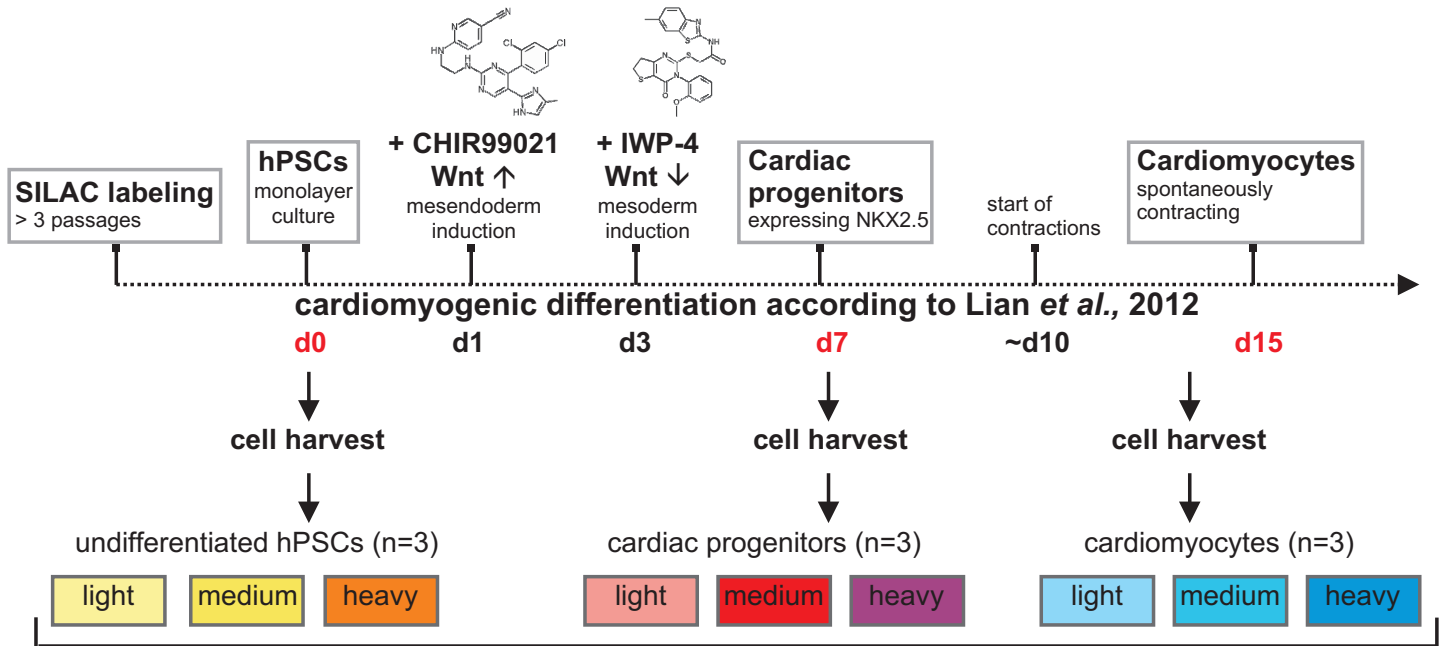
- [11.] M. Fuerstenau-Sharp, M. E. Zimmermann, K. Stark, N. Jentsch, M. Klingenstein, M. Drzymalski, S. Wagner, L. S. Maier, U. Hehr, A. Baessler, M. Fischer, C. Hengstenberg, *PLoS.One.* **2015**, *10* e0126596.
- [12.] S. R. Braam, R. Passier, C. L. Mummery, *Trends Pharmacol.Sci.* **2009**, *30* 536-545.
- [13.] W. Rust, T. Balakrishnan, R. Zweigerdt, *Regen.Med.* **2009**, *4* 225-237.
- [14.] D. Van Hoof, W. Dormeyer, S. R. Braam, R. Passier, J. Monshouwer-Kloots, D. Ward-van Oostwaard, A. J. Heck, J. Krijgsveld, C. L. Mummery, *J.Proteome.Res.* **2010**, *9* 1610-1618.
- [15.] H. Uosaki, H. Fukushima, A. Takeuchi, S. Matsuoka, N. Nakatsuji, S. Yamanaka, J. K. Yamashita, *PLoS.One.* **2011**, *6* e23657.
- [16.] N. C. Dubois, A. M. Craft, P. Sharma, D. A. Elliott, E. G. Stanley, A. G. Elefanty, A. Gramolini, G. Keller, *Nat.Biotechnol.* **2011**, *29* 1011-1018.
- [17.] K. W. Moremen, M. Tiemeyer, A. V. Nairn, *Nat.Rev.Mol.Cell Biol.* **2012**, *13* 448-462.
- [18.] M. Cohen, A. Varki, *OMICS.* **2010**, *14* 455-464.
- [19.] H. Zhang, X. J. Li, D. B. Martin, R. Aebersold, *Nat.Biotechnol.* **2003**, *21* 660-666.
- [20.] R. L. Gundry, M. Y. White, C. I. Murray, L. A. Kane, Q. Fu, B. A. Stanley, J. E. Van Eyk, *Curr.Protoc.Mol.Biol.* **2009**, *Chapter 10* Unit10.
- [21.] B. Wollscheid, D. Bausch-Fluck, C. Henderson, R. O'Brien, M. Bibel, R. Schiess, R. Aebersold, J. D. Watts, *Nat.Biotechnol.* **2009**, *27* 378-386.
- [22.] Y. Zeng, T. N. Ramya, A. Dirksen, P. E. Dawson, J. C. Paulson, *Nat.Methods* **2009**, *6* 207-209.
- [23.] G. Palmisano, S. E. Lendal, K. Engholm-Keller, R. Leth-Larsen, B. L. Parker, M. R. Larsen, *Nat.Protoc.* **2010**, *5* 1974-1982.
- [24.] A. Hofmann, B. Gerrits, A. Schmidt, T. Bock, D. Bausch-Fluck, R. Aebersold, B. Wollscheid, *Blood* **2010**, *116* e26-e34.
- [25.] K. R. Boheler, S. Bhattacharya, E. M. Kropp, S. Chuppa, D. R. Riordon, D. Bausch-Fluck, P. W. Burridge, J. C. Wu, R. P. Wersto, G. C. Chan, S. Rao, B. Wollscheid, R. L. Gundry, *Stem Cell Reports.* **2014**, *3* 185-203.
- [26.] E. M. Kropp, S. Bhattacharya, M. Waas, S. L. Chuppa, A. K. Hadjantonakis, K. R. Boheler, R. L. Gundry, *Proteomics Clin.Appl.* **2014**, *8* 603-609.
- [27.] D. Bausch-Fluck, A. Hofmann, T. Bock, A. P. Frei, F. Cerciello, A. Jacobs, H. Moest, U. Omasits, R. L. Gundry, C. Yoon, R. Schiess, A. Schmidt, P. Mirkowska, A. Hartlova, J. E. Van Eyk, J. P. Bourquin, R. Aebersold, K. R. Boheler, P. Zandstra, B. Wollscheid, *PLoS.One.* **2015**, *10* e0121314.

- [28.] T. Satomaa, A. Heiskanen, M. Mikkola, C. Olsson, M. Blomqvist, M. Tiittanen, T. Jaatinen, O. Aitio, A. Olonen, J. Helin, J. Hiltunen, J. Natunen, T. Tuuri, T. Otonkoski, J. Saarinen, J. Laine, *BMC.Cell Biol.* **2009**, *10* 42.
- [29.] H. J. An, P. Gip, J. Kim, S. Wu, K. W. Park, C. T. McVaugh, D. V. Schaffer, C. R. Bertozzi, C. B. Lebrilla, *Mol.Cell Proteomics* **2012**, *11* M111.
- [30.] K. Hasehira, H. Tateno, Y. Onuma, Y. Ito, M. Asashima, J. Hirabayashi, *Mol.Cell Proteomics* **2012**, *11* 1913-1923.
- [31.] N. Fujitani, J. Furukawa, K. Araki, T. Fujioka, Y. Takegawa, J. Piao, T. Nishioka, T. Tamura, T. Nikaido, M. Ito, Y. Nakamura, Y. Shinohara, *Proc.Natl.Acad.Sci.U.S.A* **2013**, *110* 2105-2110.
- [32.] Y. J. Liang, H. H. Kuo, C. H. Lin, Y. Y. Chen, B. C. Yang, Y. Y. Cheng, A. L. Yu, K. H. Khoo, J. Yu, *Proc.Natl.Acad.Sci.U.S.A* **2010**, *107* 22564-22569.
- [33.] Y. C. Wang, M. Nakagawa, I. Garitaonandia, I. Slavin, G. Altun, R. M. Lacharite, K. L. Nazor, H. T. Tran, C. L. Lynch, T. R. Leonardo, Y. Liu, S. E. Peterson, L. C. Laurent, S. Yamanaka, J. F. Loring, *Cell Res.* **2011**, *21* 1551-1563.
- [34.] H. Tateno, M. Toyota, S. Saito, Y. Onuma, Y. Ito, K. Hiemori, M. Fukumura, A. Matsushima, M. Nakanishi, K. Ohnuma, H. Akutsu, A. Umezawa, K. Horimoto, J. Hirabayashi, M. Asashima, *J.Biol.Chem.* **2011**, *286* 20345-20353.
- [35.] T. Ojima, E. Shibata, S. Saito, M. Toyoda, H. Nakajima, M. Yamazaki-Inoue, Y. Miyagawa, N. Kiyokawa, J. Fujimoto, T. Sato, A. Umezawa, *Sci.Rep.* **2015**, *5* 14988.
- [36.] J. Furukawa, K. Okada, Y. Shinohara, *Glycoconj.J.* **2016**, *33* 707-715.
- [37.] R. P. Berger, M. Dookwah, R. Steet, S. Dalton, *Bioessays* **2016**, *38* 1255-1265.
- [38.] C. T. Thiesler, S. Cajic, D. Hoffmann, C. Thiel, L. van Diepen, R. Hennig, M. Sgodda, R. Weissmann, U. Reichl, D. Steinemann, U. Diekmann, N. M. Huber, A. Oberbeck, T. Cantz, A. W. Kuss, C. Korner, A. Schambach, E. Rapp, F. F. Buettner, *Mol.Cell Proteomics.* **2016**.
- [39.] T. Kawamura, S. Miyagawa, S. Fukushima, N. Kashiya, A. Kawamura, E. Ito, A. Saito, A. Maeda, H. Eguchi, K. Toda, S. Miyagawa, H. Okuyama, Y. Sawa, *Stem Cells Transl.Med.* **2015**, *4* 1258-1264.
- [40.] R. Kottler, M. Mank, R. Hennig, B. Muller-Werner, B. Stahl, U. Reichl, E. Rapp, *Electrophoresis* **2013**, *34* 2323-2336.
- [41.] N. Callewaert, H. Van Vlierberghe, A. Van Hecke, W. Laroy, J. Delanghe, R. Contreras, *Nat.Med.* **2004**, *10* 429-434.
- [42.] J. Schwarzer, E. Rapp, U. Reichl, *Electrophoresis* **2008**, *29* 4203-4214.
- [43.] C. B. Karsten, F. F. Buettner, S. Cajic, I. Nehlmeier, B. Neumann, A. Klippert, U. Sauermann, U. Reichl, R. Gerardy-Schahn, E. Rapp, C. Stahl-Hennig, S. Pohlmann, *J.Virol.* **2015**, *89* 11727-11733.

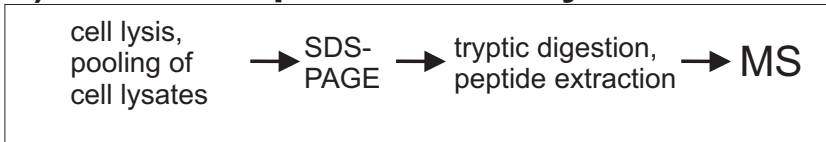
- [44.] J. E. Huffman, M. Pucic-Bakovic, L. Klaric, R. Hennig, M. H. Selman, F. Vuckovic, M. Novokmet, J. Kristic, M. Borowiak, T. Muth, O. Polasek, G. Razdorov, O. Gornik, R. Plomp, E. Theodoratou, A. F. Wright, I. Rudan, C. Hayward, H. Campbell, A. M. Deelder, U. Reichl, Y. S. Aulchenko, E. Rapp, M. Wuhrer, G. Lauc, *Mol.Cell Proteomics* **2014**, *13* 1598-1610.
- [45.] M. Acar, H. Jafar-Nejad, H. Takeuchi, A. Rajan, D. Ibrani, N. A. Rana, H. Pan, R. S. Haltiwanger, H. J. Bellen, *Cell* **2008**, *132* 247-258.
- [46.] T. V. Lee, M. K. Sethi, J. Leonardi, N. A. Rana, F. F. Buettner, R. S. Haltiwanger, H. Bakker, H. Jafar-Nejad, *PLoS.Genet.* **2013**, *9* e1003547.
- [47.] Y. C. Wang, J. W. Stein, C. L. Lynch, H. T. Tran, C. Y. Lee, R. Coleman, A. Hatch, V. G. Antontsev, H. S. Chy, C. M. O'Brien, S. K. Murthy, A. L. Laslett, S. E. Peterson, J. F. Loring, *Sci.Rep.* **2015**, *5* 13317.
- [48.] F. Alisson-Silva, R. D. de Carvalho, L. Vairo, K. D. Asensi, A. Vasconcelos-dos-Santos, N. R. Mantuano, W. B. Dias, E. Rondinelli, R. C. Goldenberg, T. P. Urmenyi, A. R. Todeschini, *Glycobiology* **2014**, *24* 458-468.
- [49.] S. A. Konze, S. Werneburg, A. Oberbeck, R. Olmer, H. Kempf, M. Jara-Avaca, A. Pich, R. Zweigerdt, F. F. Buettner, *J.Proteome.Res.* **2017**.
- [50.] F. Hirschhaeuser, H. Menne, C. Dittfeld, J. West, W. Mueller-Klieser, L. A. Kunz-Schughart, *J.Biotechnol.* **2010**, *148* 3-15.
- [51.] X. Lian, C. Hsiao, G. Wilson, K. Zhu, L. B. Hazeltine, S. M. Azarin, K. K. Raval, J. Zhang, T. J. Kamp, S. P. Palecek, *Proc.Natl.Acad.Sci.U.S.A* **2012**, *109* E1848-E1857.
- [52.] X. Lian, J. Zhang, S. M. Azarin, K. Zhu, L. B. Hazeltine, X. Bao, C. Hsiao, T. J. Kamp, S. P. Palecek, *Nat.Protoc.* **2013**, *8* 162-175.
- [53.] J. M. Sperger, X. Chen, J. S. Draper, J. E. Antosiewicz, C. H. Chon, S. B. Jones, J. D. Brooks, P. W. Andrews, P. O. Brown, J. A. Thomson, *Proc.Natl.Acad.Sci.U.S.A* **2003**, *100* 13350-13355.
- [54.] S. Kobayashi, T. Kohda, N. Miyoshi, Y. Kuroiwa, K. Aisaka, O. Tsutsumi, T. Kaneko-Ishino, F. Ishino, *Hum.Mol.Genet.* **1997**, *6* 781-786.
- [55.] D. Evseenko, Y. Zhu, K. Schenke-Layland, J. Kuo, B. Latour, S. Ge, J. Scholes, G. Dravid, X. Li, W. R. MacLellan, G. M. Crooks, *Proc.Natl.Acad.Sci.U.S.A* **2010**, *107* 13742-13747.
- [56.] R. Passier, C. Denning, C. Mummery, *Handb.Exp.Pharmacol.* **2006**, 101-122.
- [57.] M. L. Lepire, C. A. Ziomek, *Biol.Reprod.* **1989**, *41* 464-473.
- [58.] L. Wang, Y. Xue, Y. Shen, W. Li, Y. Cheng, X. Yan, W. Shi, J. Wang, Z. Gong, G. Yang, C. Guo, Y. Zhou, X. Wang, Q. Zhou, F. Zeng, *Cell Res.* **2012**, *22* 1082-1085.
- [59.] T. A. Prokhorova, K. T. Rigbolt, P. T. Johansen, J. Henningsen, I. Kratchmarova, M. Kassem, B. Blagoev, *Mol.Cell Proteomics.* **2009**, *8* 959-970.

- [60.] A. Bondue, S. Tannler, G. Chiapparo, S. Chabab, M. Ramialison, C. Paulissen, B. Beck, R. Harvey, C. Blanpain, *J.Cell Biol.* **2011**, *192* 751-765.
- [61.] H. Nakamura, R. N. Cook, M. J. Justice, *BMC.Dev.Biol.* **2013**, *13* 9.
- [62.] R. Toyoda, H. Nakamura, Y. Watanabe, *Gene Expr Patterns* **2005**, *5* 778-785.
- [63.] M. Baruscotti, A. Bucchi, C. Viscomi, G. Mandelli, G. Consalez, T. Gneccchi-Rusconi, N. Montano, K. R. Casali, S. Micheloni, A. Barbuti, D. DiFrancesco, *Proc.Natl.Acad.Sci.U.S.A* **2011**, *108* 1705-1710.
- [64.] T. S. Nair, K. E. Kozma, N. L. Hoefling, P. K. Kommareddi, Y. Ueda, T. W. Gong, M. I. Lomax, C. D. Lansford, S. A. Telian, B. Satar, H. A. Arts, H. K. El Kashlan, W. E. Berryhill, Y. Raphael, T. E. Carey, *J.Neurosci.* **2004**, *24* 1772-1779.
- [65.] M. S. Denzel, M. C. Scimia, P. M. Zumstein, K. Walsh, P. Ruiz-Lozano, B. Ranscht, *J.Clin.Invest* **2010**, *120* 4342-4352.
- [66.] S. Natunen, T. Satomaa, V. Pitkanen, H. Salo, M. Mikkola, J. Natunen, T. Otonkoski, L. Valmu, *Glycobiology* **2011**, *21* 1125-1130.
- [67.] Y. Zhao, Y. Sato, T. Isaji, T. Fukuda, A. Matsumoto, E. Miyoshi, J. Gu, N. Taniguchi, *FEBS J.* **2008**, *275* 1939-1948.
- [68.] H. Schachter, *Biochem.Cell Biol.* **1986**, *64* 163-181.
- [69.] C. Tagwerker, K. Flick, M. Cui, C. Guerrero, Y. Dou, B. Auer, P. Baldi, L. Huang, P. Kaiser, *Mol.Cell Proteomics.* **2006**, *5* 737-748.
- [70.] J. Cox, M. Mann, *Nat.Biotechnol.* **2008**, *26* 1367-1372.
- [71.] T. Hulsen, J. de Vlieg, W. Alkema, *BMC.Genomics* **2008**, *9* 488.
- [72.] A. Franceschini, D. Szklarczyk, S. Frankild, M. Kuhn, M. Simonovic, A. Roth, J. Lin, P. Minguéz, P. Bork, C. von Mering, L. J. Jensen, *Nucleic Acids Res.* **2013**, *41* D808-D815.
- [73.] R. Hennig, E. Rapp, R. Kottler, S. Cajic, M. Borowiak, U. Reichl, *Methods Mol.Biol.* **2015**, *1331* 123-143.
- [74.] A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, J. D. Marth, C. R. Bertozzi, G. W. Hart, M. E. Etzler, *Proteomics* **2009**, *9* 5398-5399.

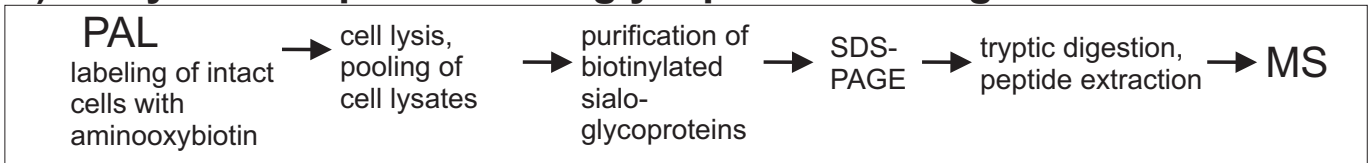
Figure 1



a) whole cell proteome analysis



b) analysis of captured sialo-glycoproteins using PAL



c) analysis of N-glycans via xCGE-LIF

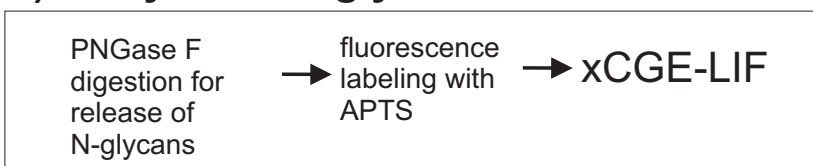


Figure 2

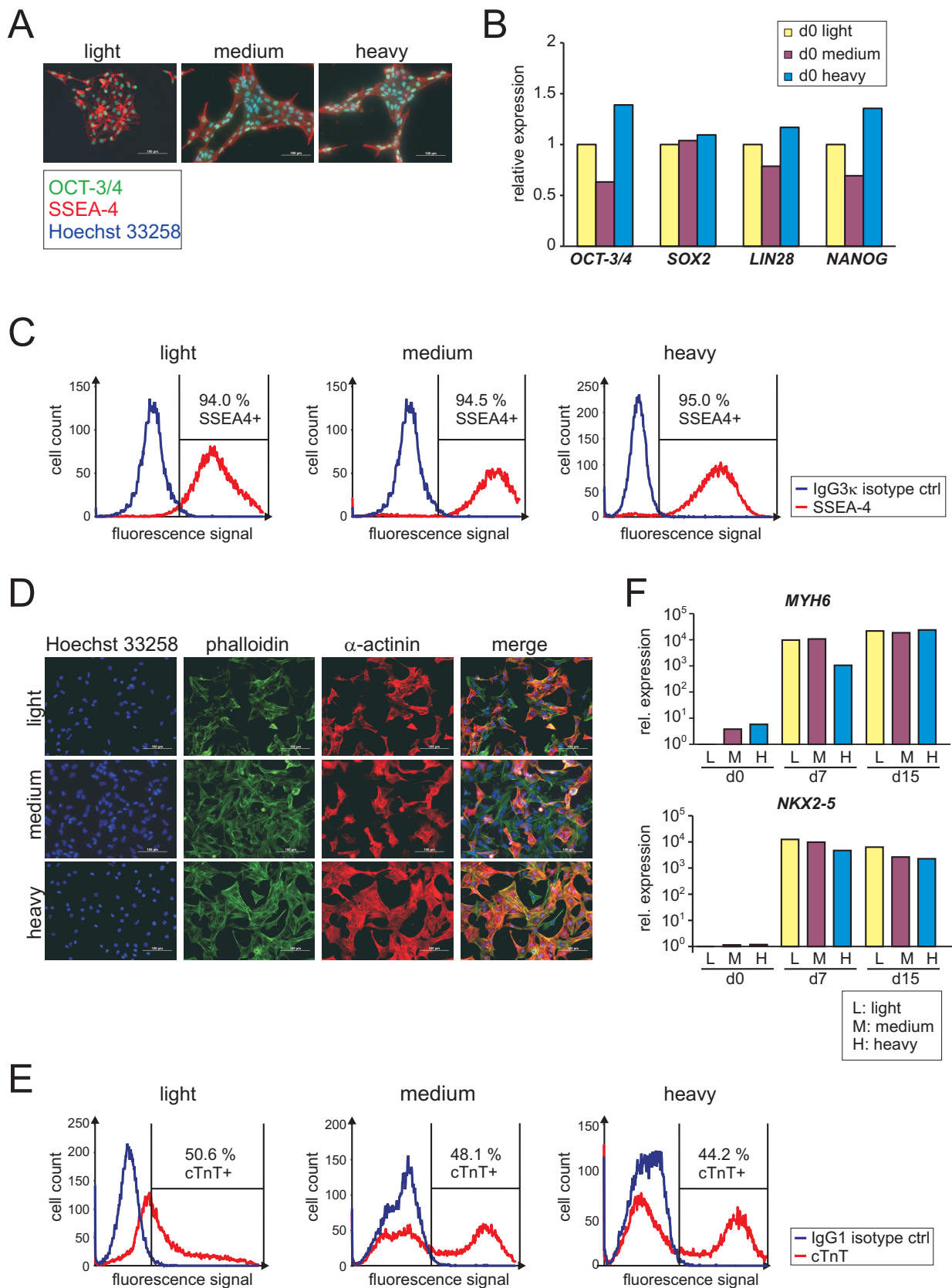


Figure 3

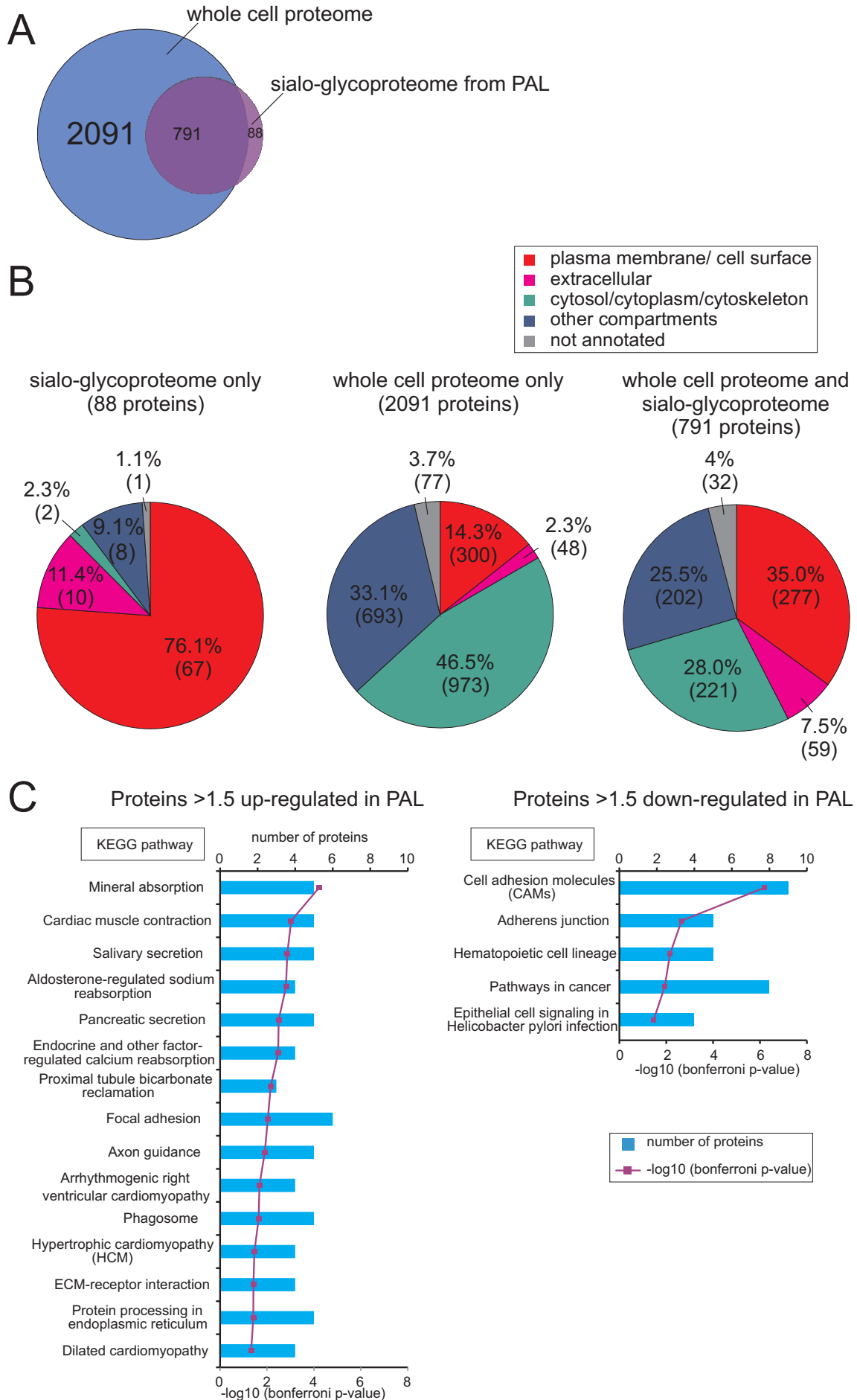


Figure 4

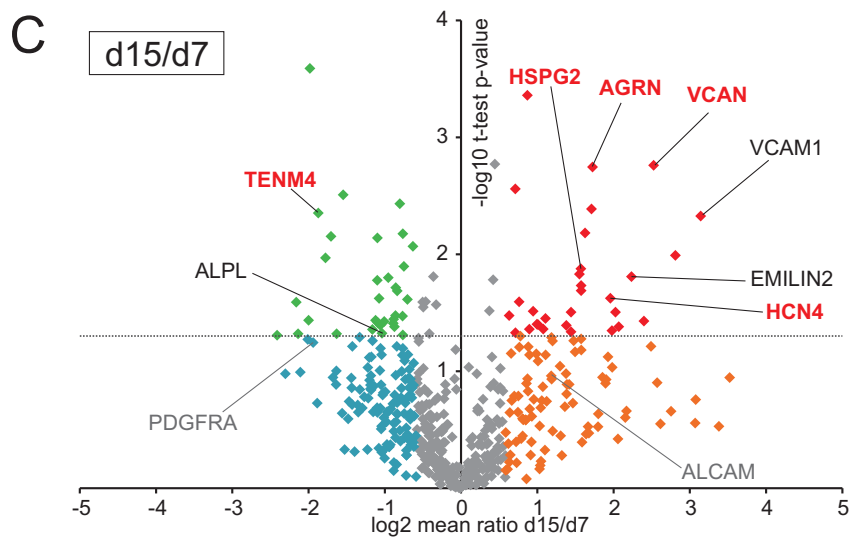
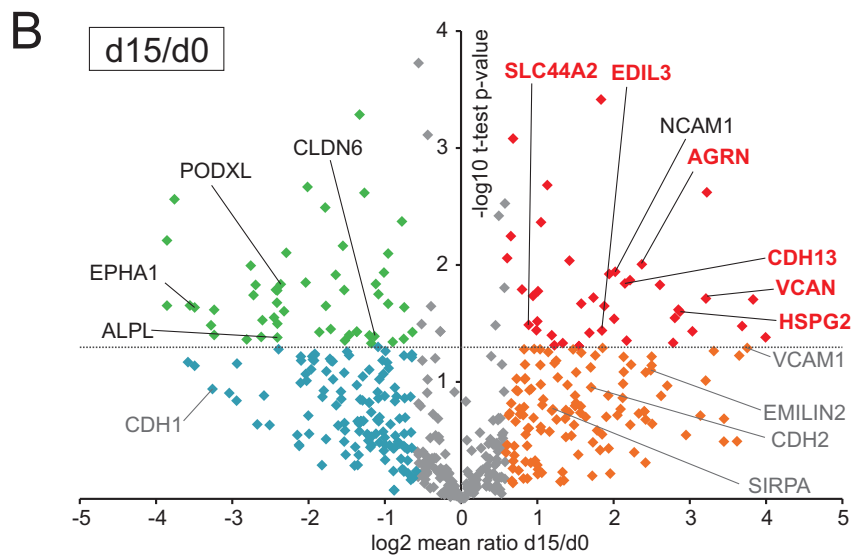
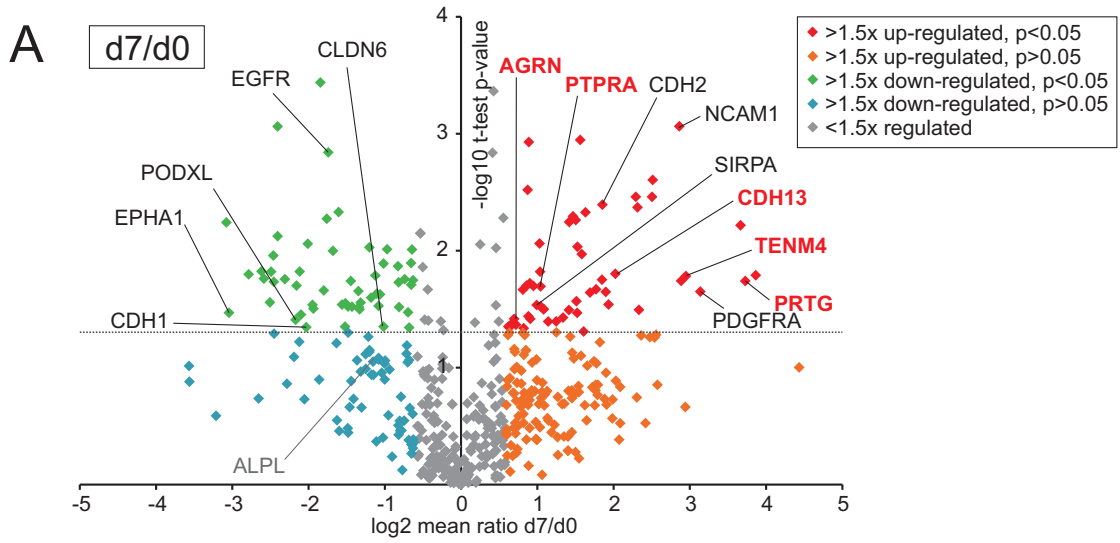


Figure 5

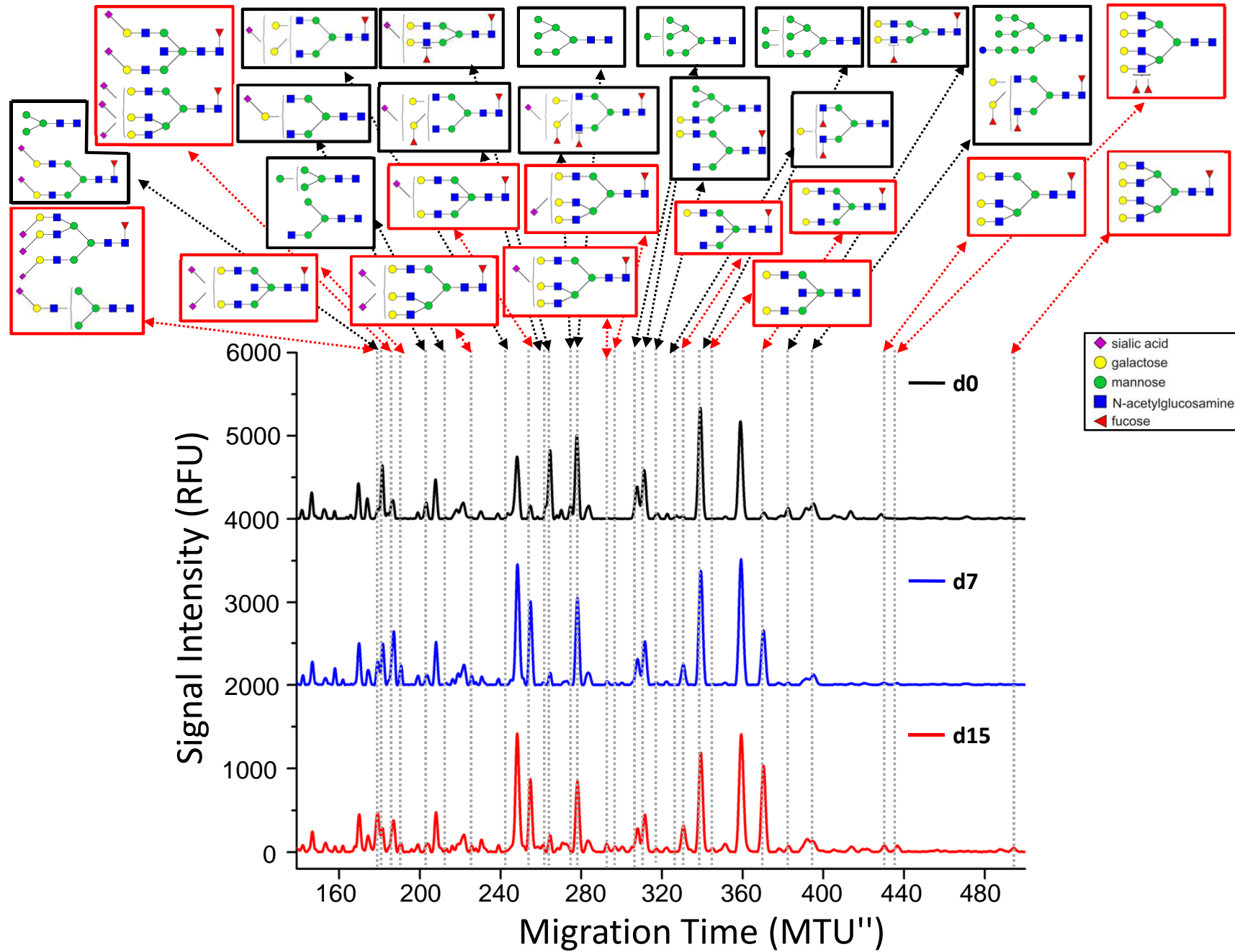


Figure 6

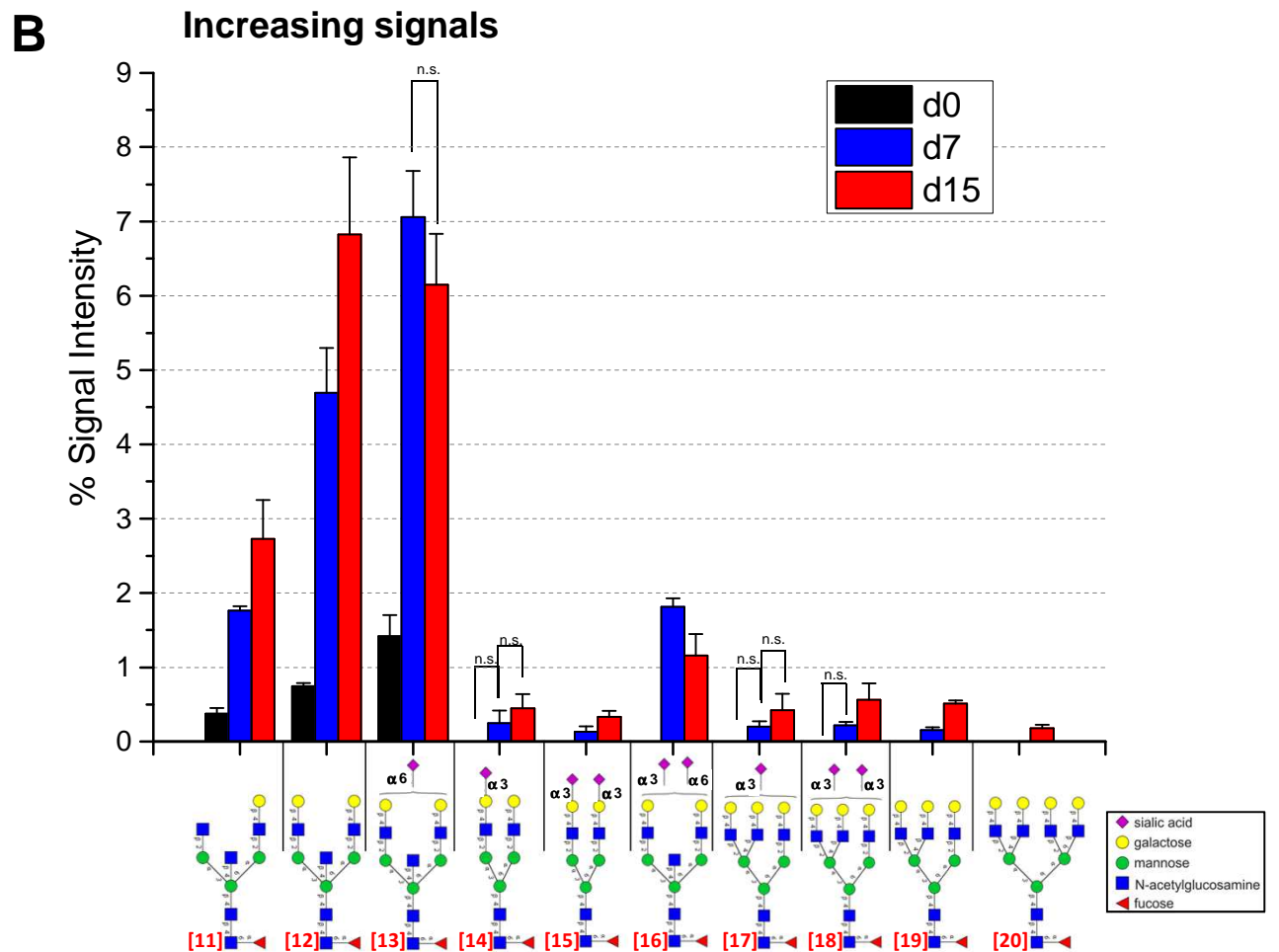
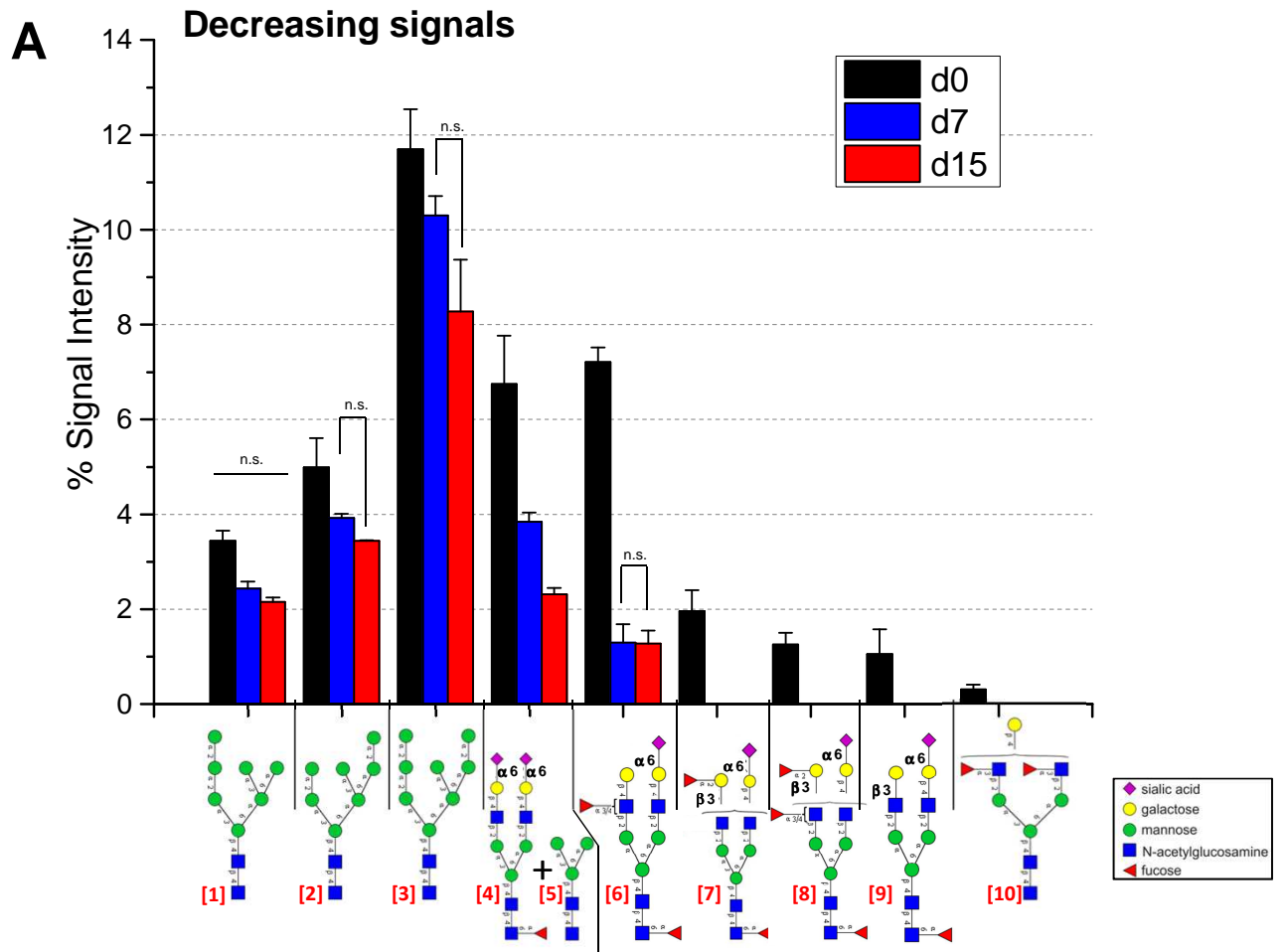
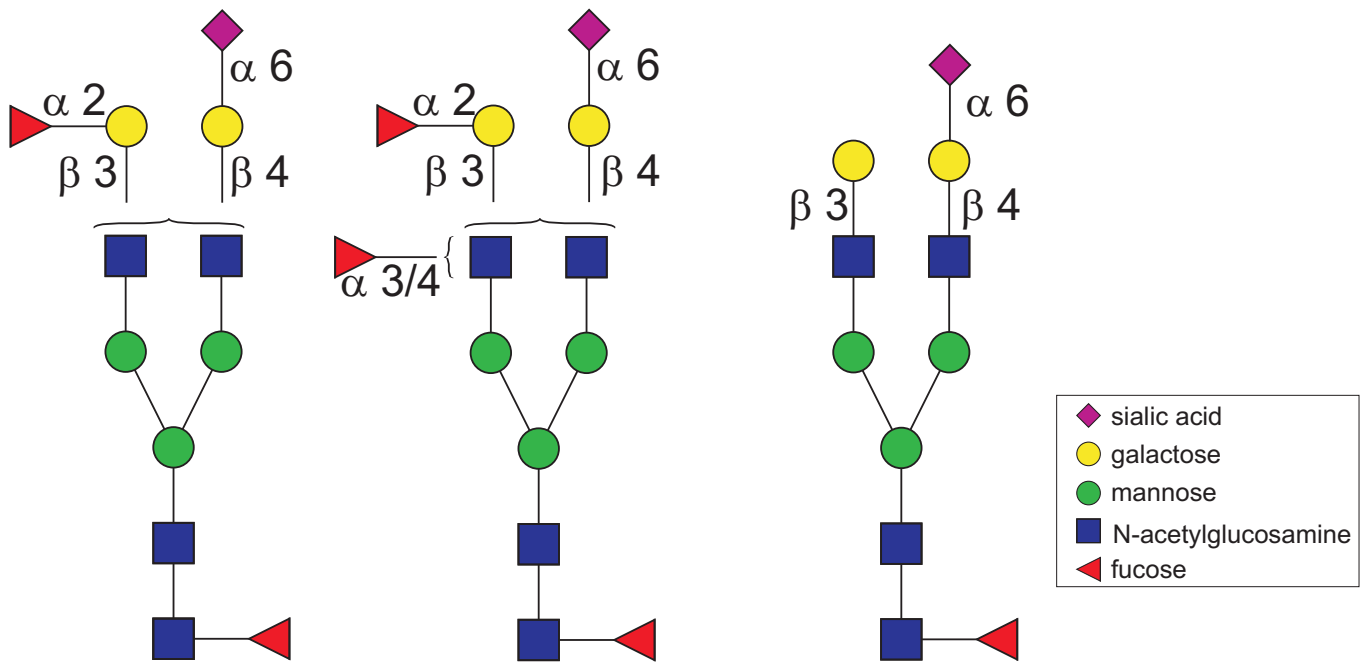


Figure 7



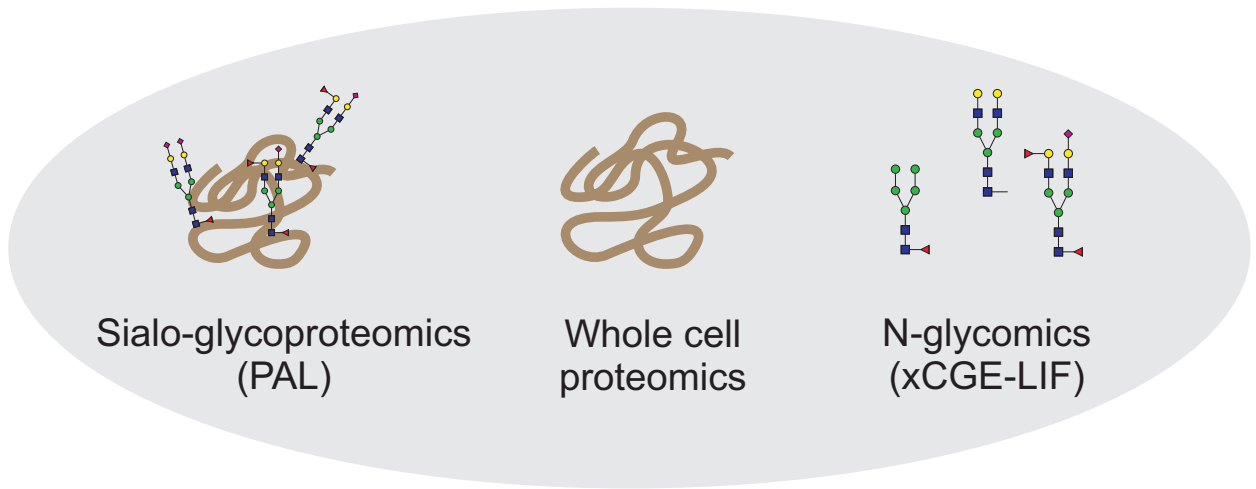
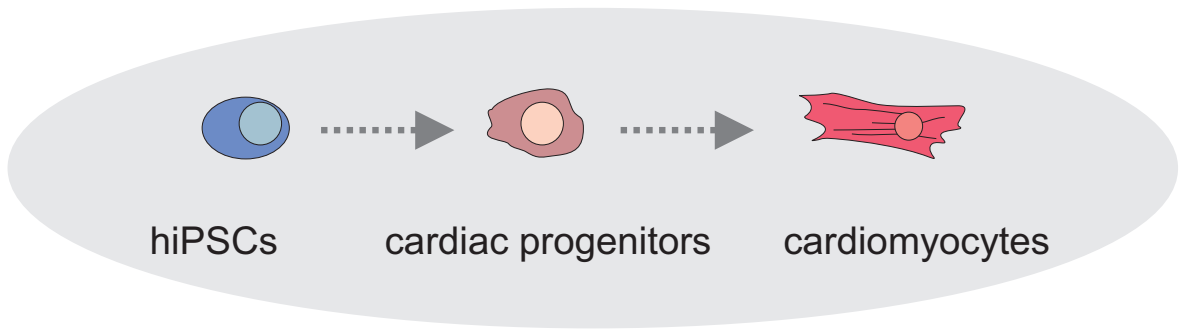


Figure S1

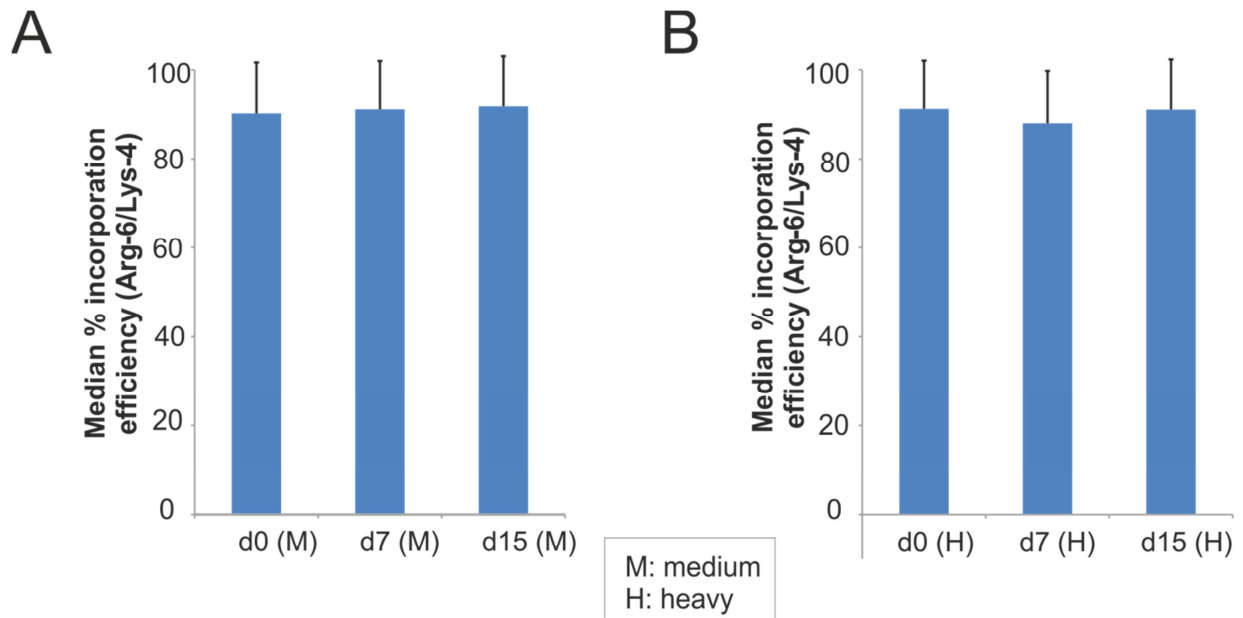


Figure S1. Determination of median SILAC amino acid incorporation efficiencies. (A) Median incorporation efficiencies of Arg-6 or Lys-4 for medium (M)-labeled samples from d0, d7 and d15. (B) Same presentation as in (A), but of Arg-10 or Lys-8 for heavy (H)-labeled samples.

Figure S2

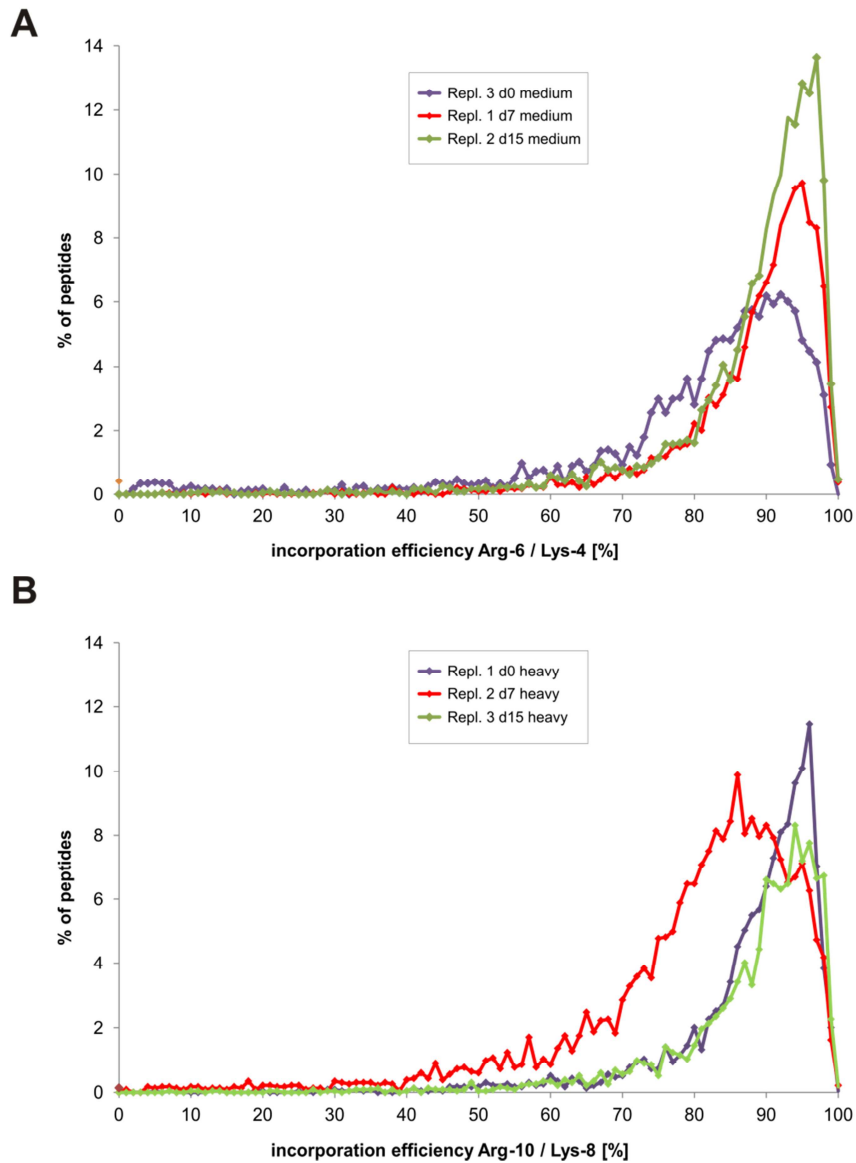
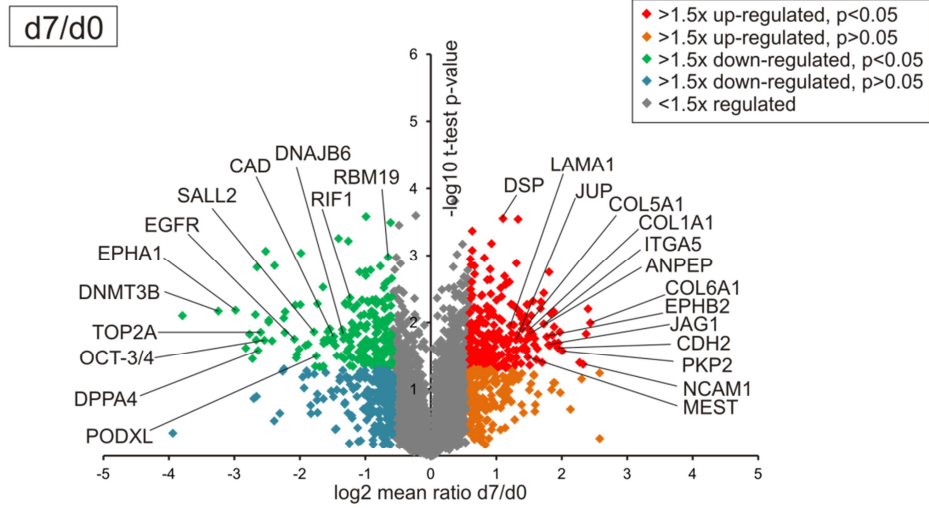


Figure S2. Presentation of SILAC amino acid incorporation efficiencies for peptides.

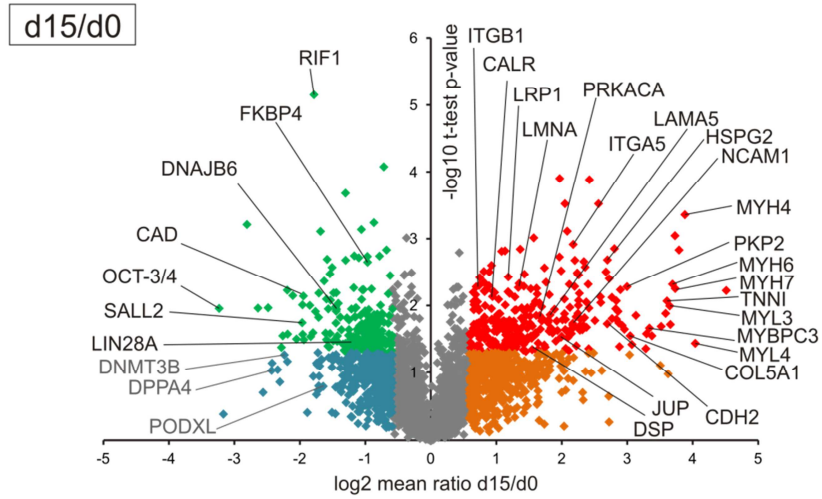
Incorporation efficiencies were calculated for the 50% most abundant peptides and were plotted in intervals increasing by 1% incorporation efficiency (x-axis) against the proportion of peptides (y-axis) belonging to the respective interval. (A) Incorporation efficiencies of Arg-6 or Lys-4 for medium (M)-labeled peptides from d0, d7 and d15. (B) Same presentation as in (A), but of Arg-10 or Lys-8 heavy (H)-labeled peptides.

Figure S3

A



B



C

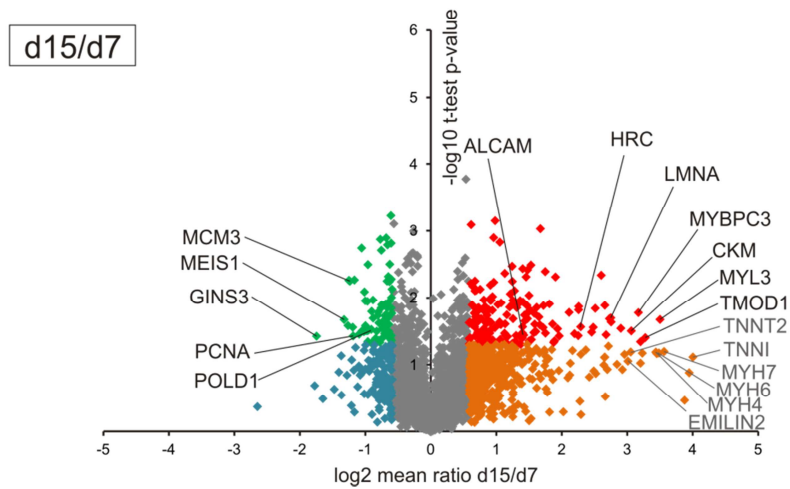


Figure S3. Regulation of proteins identified in the whole cell proteome analysis. (A) Volcano plot of all proteins of the d7/d0 comparison. The log₂ mean regulation ratio is plotted against the negative decadic logarithm of the p-value obtained from Student's t-test. (B, C) Similar presentation as in (A), but for proteins of the d15/d0 (B) and d15/d7 (C) comparisons. Proteins names mentioned in the manuscript text are indicated in black (significantly regulated) or grey (non-significantly regulated).

Figure S4

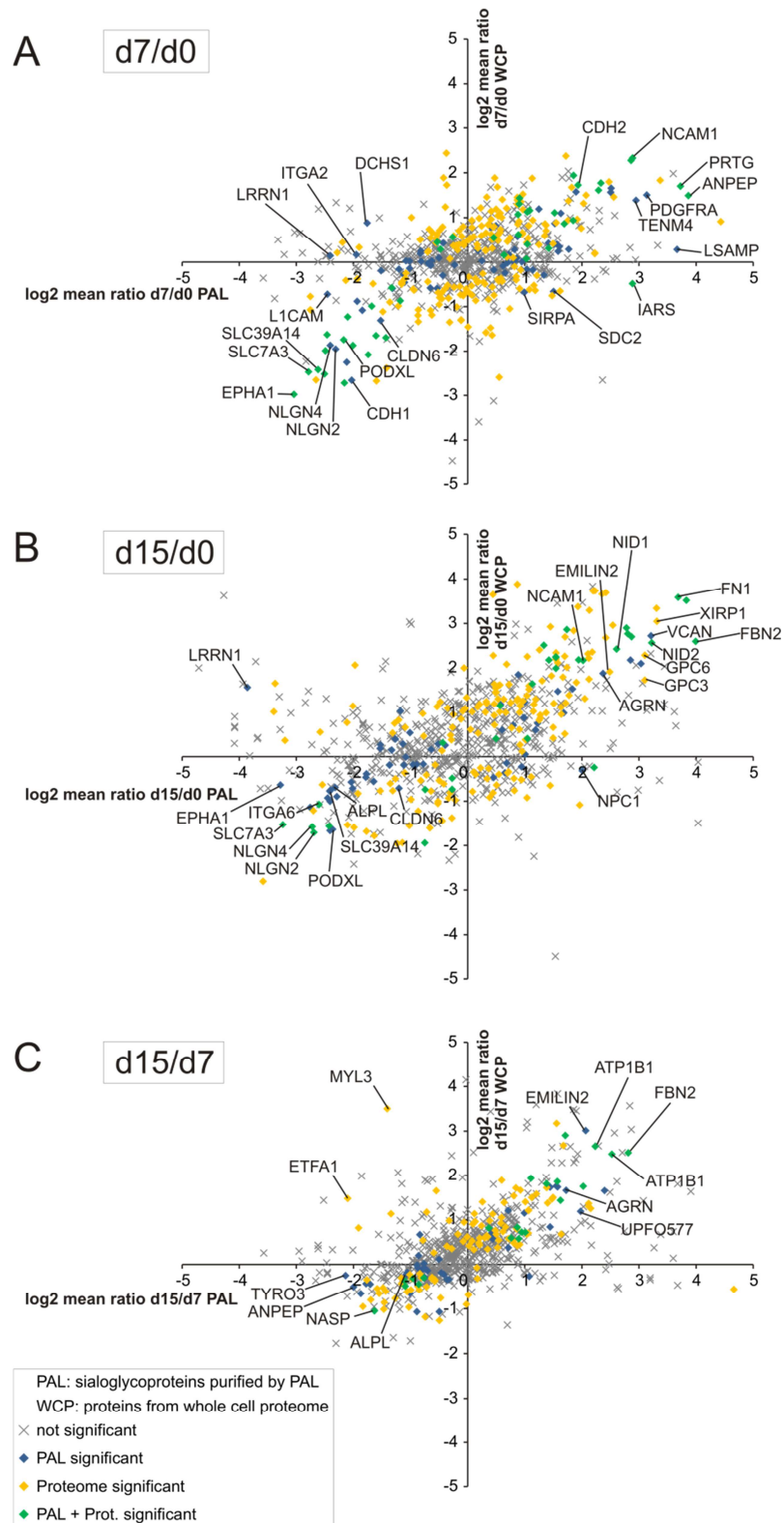
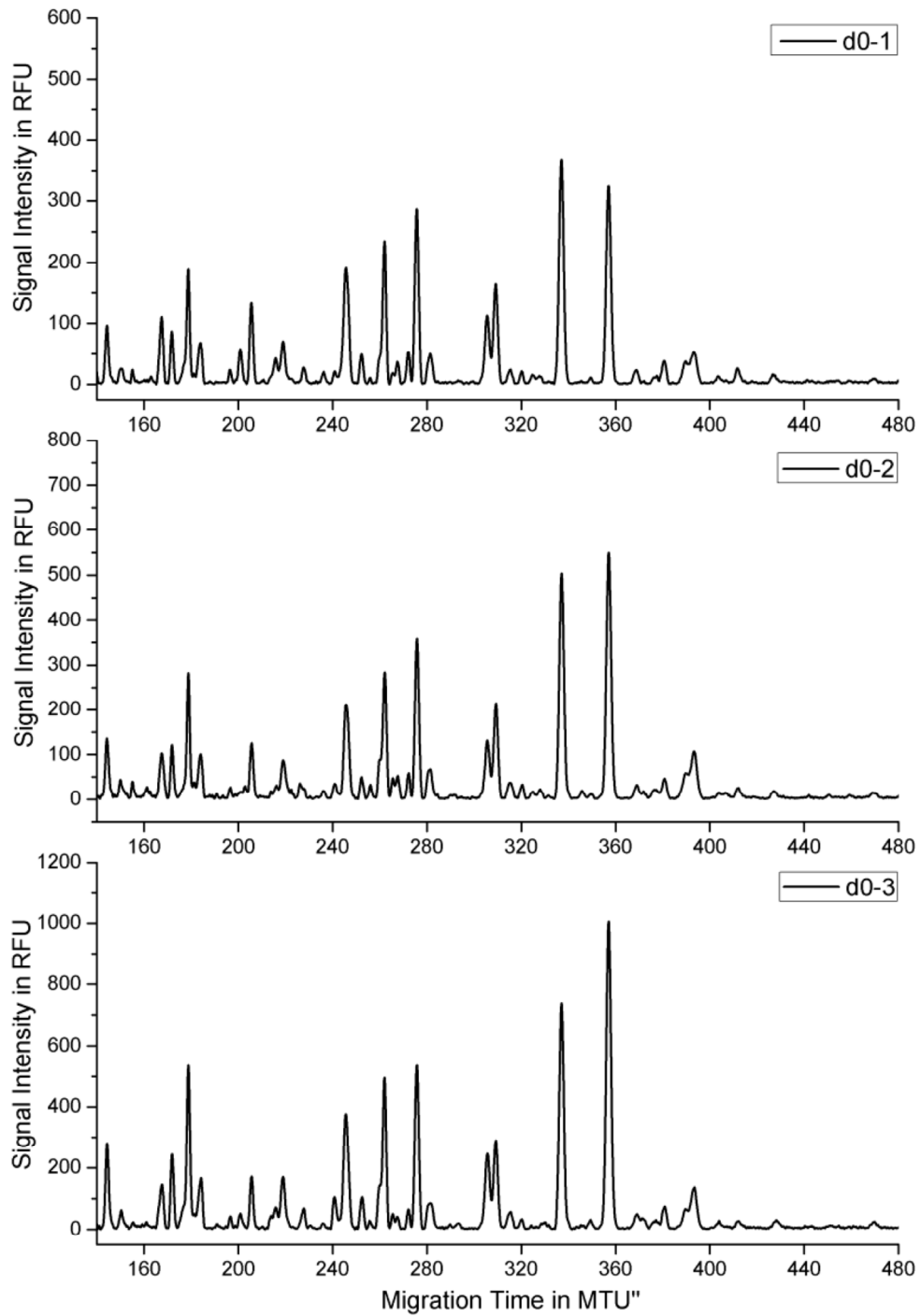


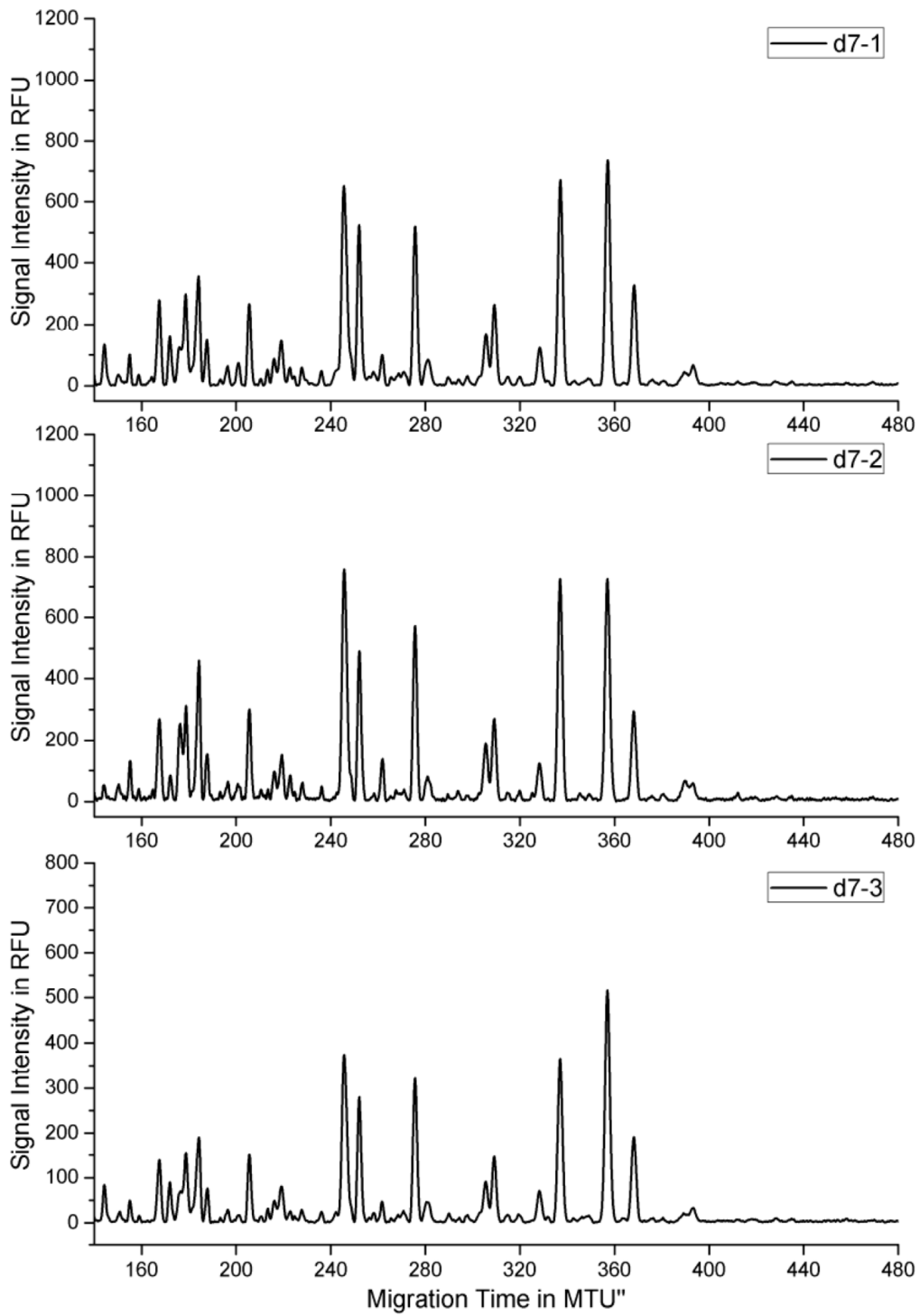
Figure S4. Comparison of regulation ratios of the whole cell proteome and the sialo-glycoproteome analysis. (A) Plot of the log₂ mean regulation ratios for the d7/d0 comparison of the whole cell proteome and the sialo-glycoproteome of all 791 proteins identified in both analyses. (B, C) Same presentation as in (A), but for the d15/d0 (B) and d15/d7 (C) log₂ mean regulation ratios.

Figure S5

A



B



C

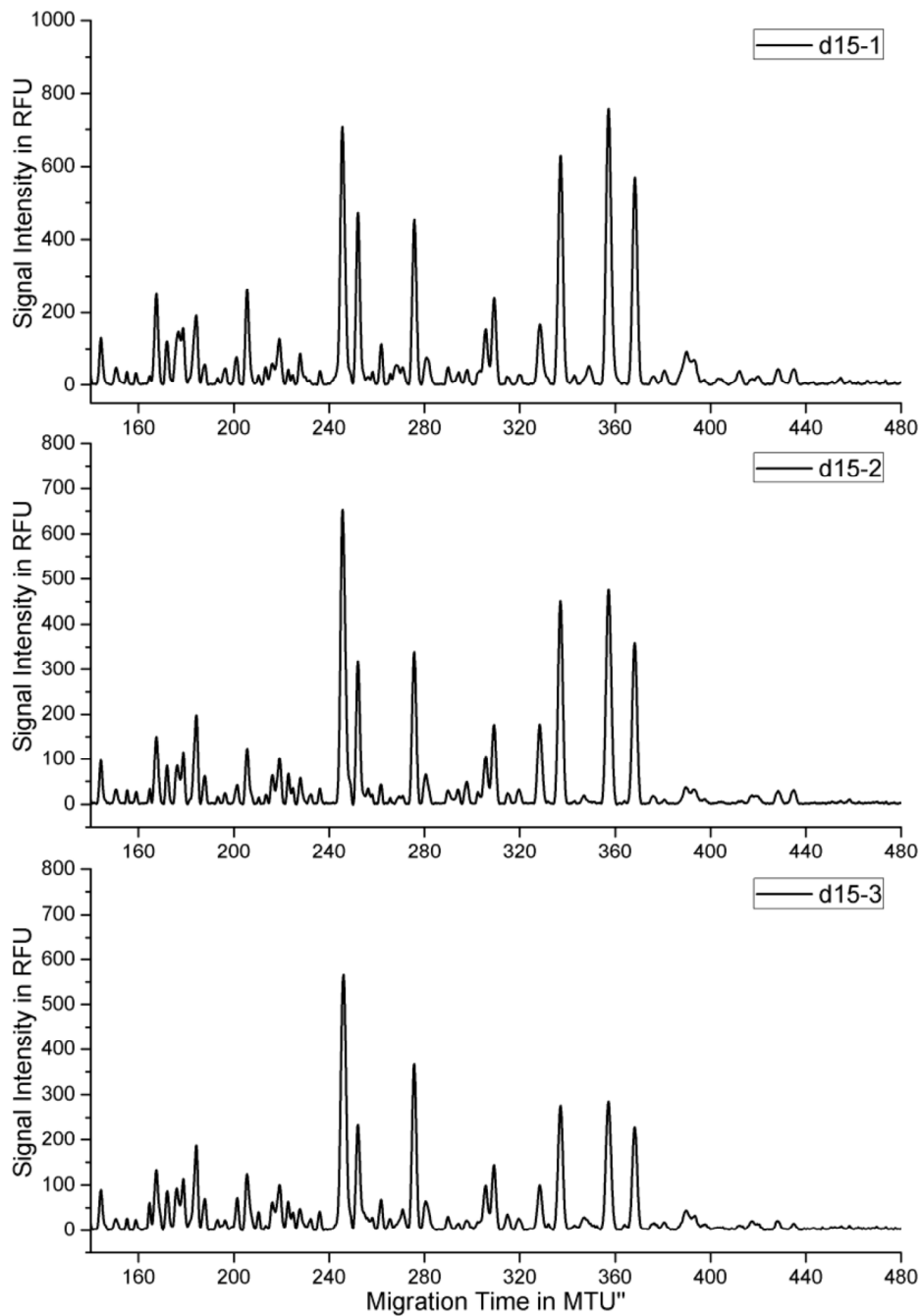
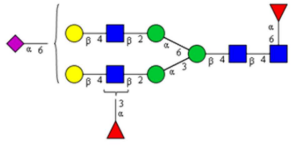
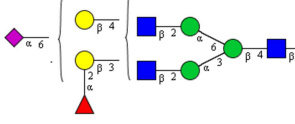
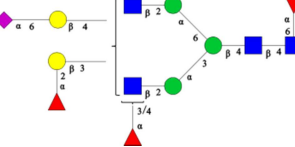
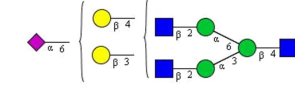
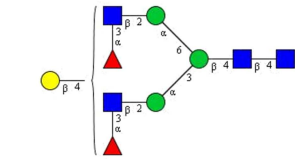
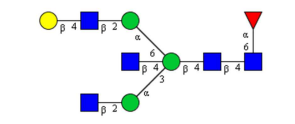
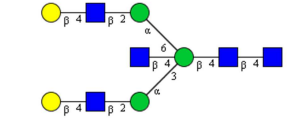
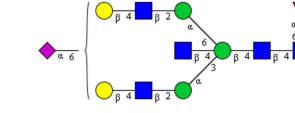
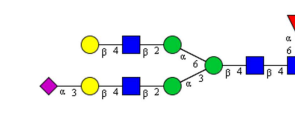
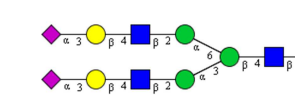


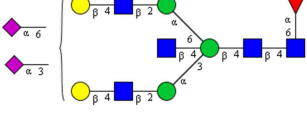
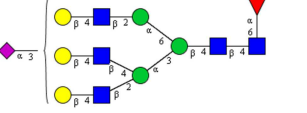
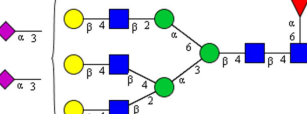
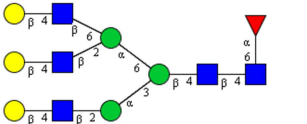
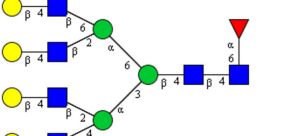
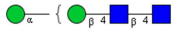
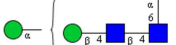
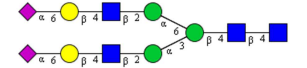
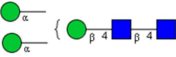
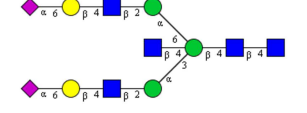
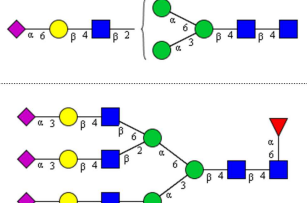
Figure S5. Individual xCGE-LIF electropherograms showing high congruency of the three biological replicates. (A) Electropherograms of the three d0 replicates. (B, C) Same presentation as in A, but for the d7 (B) and the d15 replicates (C).

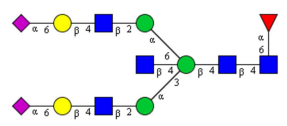
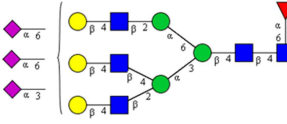
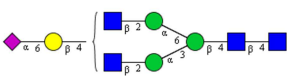
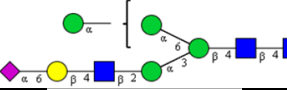
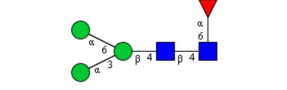
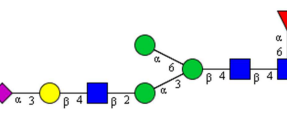
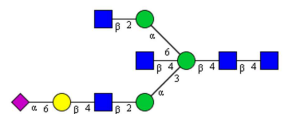
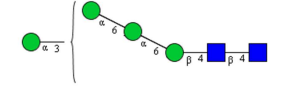
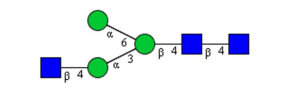
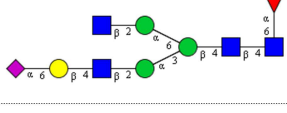
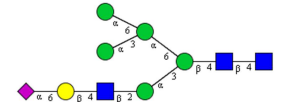
Table S3. Structures and relative intensities of N-glycans identified in d0, d7, d15 samples.

Intensities of all 62 quantifiable peaks were summed up (total peak intensity = 100%) and relative peak intensities were calculated for each peak as % of total peak intensity. Only peaks with a signal-to-noise ratio (S/N) ≥ 9 were selected for quantification. Peaks with $9 > S/N > 3$ were defined as not quantifiable. Peaks with $S/N \leq 3$ were not accounted for calculations. Values are average of three experiments \pm standard deviation. Glycans also depicted in Figure 6 are referred to with numbers in square brackets highlighted in red. Abbreviations: NQ: not quantifiable, ND: not detectable.

N-glycan Structure	% Intensity (\pm standard deviation)		
	d0	d7	d15
<p>[1]</p>	<p>3.44 (± 0.22)</p>	<p>2.40 (± 0.14)</p>	<p>2.13 (± 0.09)</p>
<p>[2]</p>	<p>4.99 (± 0.62)</p>	<p>3.85 (± 0.09)</p>	<p>3.40 (± 0.01)</p>
<p>[3]</p>	<p>11.70 (± 0.84)</p>	<p>10.12 (± 0.40)</p>	<p>8.18 (± 1.09)</p>
<p>[4]</p>	<p>6.74 (± 1.03)</p>	<p>3.77 (± 0.19)</p>	<p>2.29 (± 0.12)</p>
<p>[5]</p>			

<p>[6]</p> 	<p>7.21 (±0.31)</p>	<p>1.27 (±0.38)</p>	<p>1.26 (±0.27)</p>
<p>[7]</p> 	<p>1.97 (±0.44)</p>	<p>ND</p>	<p>ND</p>
<p>[8]</p> 	<p>1.25 (±0.25)</p>	<p>ND</p>	<p>ND</p>
<p>[9]</p> 	<p>1.06 (±0.52)</p>	<p>ND</p>	<p>ND</p>
<p>[10]</p> 	<p>0.31 (±0.10)</p>	<p>ND</p>	<p>ND</p>
<p>[11]</p> 	<p>0.37 (±0.08)</p>	<p>1.76 (±0.06)</p>	<p>2.72 (±0.53)</p>
<p>[12]</p> 	<p>ND</p>	<p>0.11 (±0.04)</p>	<p>0.20 (±0.15)</p>
<p>[13]</p> 	<p>1.42 (±0.28)</p>	<p>7.06 (±0.62)</p>	<p>6.15 (±0.69)</p>
<p>[14]</p> 	<p>ND</p>	<p>0.25 (±0.17)</p>	<p>0.45 (±0.19)</p>
<p>[15]</p> 	<p>ND</p>	<p>0.13 (±0.07)</p>	<p>0.33 (±0.08)</p>

<p>[16]</p> 	ND	1.81 (±0.11)	1.16 (±0.29)
<p>[17]</p> 	ND	0.20 (±0.07)	0.42 (±0.22)
<p>[18]</p> 	ND	0.22 (±0.04)	0.56 (±0.22)
<p>[19]</p> 	ND	ND	0.29 (±0.03)
<p>[20]</p> 	ND	ND	0.18 (±0.05)
	0.70 (±0.21)	0.81 (±0.09)	0.54 (±0.05)
	3.10 (±0.76)	3.82 (±0.18)	3.61 (±0.40)
	2.71 (±0.61)	1.20 (±0.47)	1.50 (±0.16)
	2.71 (±0.61)	1.20 (±0.47)	1.50 (±0.16)
<p>(detected in d7 and d15 only)</p> 	1.25 (±0.30)	2.64 (±0.92)	2.61 (±0.87)
	1.25 (±0.30)	2.64 (±0.92)	2.61 (±0.87)

(detected in d7 and d15 only)			
	2.31 (±0.22)	5.23 (±0.70)	3.78 (±0.81)
(detected in d7 and d15 only)			
	0.64 (±0.08)	0.78 (±0.10)	0.57 (±0.13)
	0.98 (±0.74)	0.55 (±0.33)	1.03 (±0.44)
	NQ	ND	ND
	3.34 (±0.95)	3.99 (±0.16)	3.18 (±0.61)
	ND	0.20 (±0.07)	0.42 (±0.33)
	0.45 (±0.18)	0.42 (±0.18)	0.45 (±0.09)
	0.94 (±0.16)	1.05 (±0.02)	0.75 (±0.13)
			
	1.91 (±0.28)	1.77 (±0.05)	1.82 (±0.21)
			

	ND	0.84 (±0.24)	1.10 (±0.48)
	0.99 (±0.14)	0.73 (±0.02)	1.04 (±0.15)
	ND	ND	0.39
	0.47 (±0.12)	0.58 (±0.04)	0.45 (±0.39)
	6.16 (±0.58)	11.08 (±0.30)	13.30 (±2.44)
	0.96 (±0.29)	ND	ND
	ND	0.41 (±0.09)	0.66 (±0.15)
	8.69 (±0.69)	8.17 (±0.25)	7.35 (±1.53)
	1.48 (±0.13)	1.11 (±0.07)	1.27 (±0.22)

	ND	0.32 (±0.07)	0.68 (±0.03)
	ND	0.18 (±0.04)	0.62 (±0.31)
	ND	0.33 (±0.06)	0.41 (±0.16)
	0.79 (±0.16)	0.29 (±0.06)	0.48 (±0.18)
	NQ	ND	ND
 (detected in d0 only)	0.29 (±0.08)	0.21 (±0.03)	0.37 (±0.29)
	12.82 (±1.97)	11.76 (±1.82)	8.94 (±1.82)
	0.74 (±0.04)	4.69 (±0.61)	6.83 (±1.04)

	<p>0.29 (±0.22)</p>	<p>ND</p>	<p>ND</p>
	<p>0.44 (±0.12)</p>	<p>ND</p>	<p>ND</p>
	<p>1.09 (±0.11)</p>	<p>0.20 (±0.06)</p>	<p>0.34 (±0.16)</p>
	<p>1.18 (±0.24)</p>	<p>0.64 (±0.19)</p>	<p>0.96 (±0.25)</p>
	<p>2.07 (±0.42)</p>	<p>0.86 (±0.11)</p>	<p>0.77 (±0.18)</p>
<p>(detected in d0 only)</p>			
	<p>0.32 (±0.06)</p>	<p>NQ</p>	<p>0.12 (±0.06)</p>
	<p>0.59 (±0.26)</p>	<p>0.17 (±0.10)</p>	<p>0.26 (±0.20)</p>
	<p>ND</p>	<p>0.15 (±0.04)</p>	<p>0.51 (±0.04)</p>

	ND	0.15 (±0.01)	0.48 (±0.15)
--	----	-----------------	-----------------

Table S4. Labeling pattern of SILAC samples

	Replicate 1	Replicate 2	Replicate 3
d0 label	Heavy (#1)	Light (#2)	Medium (#3)
d7 label	Medium (#4)	Heavy (#5)	Light (#6)
d15 label	Light (#7)	Medium (#8)	Heavy (#9)

denotes the independent sample number

Table S5. Antibodies used for immunofluorescence microscopy and flow cytometry.

Primary antibodies					
Target	Host species (isotype)	Supplier	Order #	Clone	Dilution
OCT-3/4 / Octamer-binding protein 3/4	mouse (IgG _{2b})	Santa Cruz Biotechnology	sc-5279	C-10	IF: 1:50
SSEA-4 /Stage-specific embryonic antigen-4	mouse (IgG _{3κ})	Bio-Legend	330401	MC-813-70	IF: 1:100 FC: 1:100
α-actinin (sarcomeric)	mouse (IgG ₁)	Sigma-Aldrich	A7811	EA-53	IF: 1:500
cardiac Troponin T (cTnT)	mouse (IgG ₁)	Thermo Fisher Scientific	MA5-12960	13-11	FC: 1:100
IgG1 isotype control	mouse (IgG ₁)	DAKO	X0931	-	FC: 1:100
IgG _{3κ} isotype control	mouse (IgG _{3κ})	BioLegend	401301	MG-335	FC: 1:100
Secondary antibodies					
Target species (isotype)	Host species	Supplier	Order #	Fluorochrome	Application: Dilution
mouse IgG _{2b}	goat	Molecular Probes	A 21141	Alexa Fluor [®] 488	IF: 1:500, FC: 1:1200 (against anti-OCT-3/4 (IF), anti-SSEA-4 (FC) and anti-cTnT (FC))
mouse IgG (H+L)	goat	Jackson Immuno Res. Labs	115-165-003	Cy [™] 3	IF: 1:500 (against anti-SSEA-4, anti-α-actinin)

IF: immunofluorescence microscopy; FC: flow cytometry

Table S6. Oligonucleotides used as primers for qPCR.

Target gene name / full name	Primer name	Primer sequence (5' to 3')
Housekeeping gene		
<i>ACTB</i> / actin, beta	ACTB_fw_qPCR	ATGCAGAAGGAGATCACTGC
	ACTB_rev_qPCR	AAGCATTTCGCGGTGGACGAT
Pluripotency markers		
<i>OCT3/4 / POU5F1</i> / POU class 5 homeobox 1	OCT4_fw	AGAAGGAGAAGCTGGAGCAA
	OCT4_rev	CTTCCCAAATAGAACCCCA
<i>SOX2 / SRY</i> (sex determining region Y)-box 2	SOX2_fw	GGACACTGGCTGAATCCTTCC
	SOX2_rev	CTCGCTGATTAGGCTCCAACC
<i>LIN28</i> / lin-28 homolog A	LIN28_fw	ATGGAGAAAACCCGGTACGC
	LIN28_rev	TTTTGCGTGAGTGTGGATGG
<i>NANOG</i> / Nanog homeobox	NANOG_fw	TTGAGGAGCAGGCAGAGTGG
	NANOG_rev	TGCATTTGGACAGAGCATGG
Cardiomyocyte markers		
<i>NKX2-5</i> / NK2 homeobox 5	NKX2_5_fw	AGAAGACAGAGGCGGACAAC
	NKX2_5_rev	CGCCGCTCCAGTTCATAG
<i>αMHC / MYH6</i> / myosin, heavy chain 6, cardiac muscle, alpha	aMHC_fw	ATTGCTGAAACCGAGAATGG
	aMHC_rev	CGCTCCTTGAGGTTGAAAAG
<i>MYL4</i> / myosin light chain, embryonic	MYL4_fw	GAGAGAAGATGACTGAGGCT
	MYL4_rev	CCTGACATGATGTGCTTGAC