ATDC5 cells as a model of cartilage extracellular matrix neosynthesis, maturation and assembly

Dafné Wilhelm, Hervé Kempf, Arnaud Bianchi, Jean-Baptiste Vincourt

Version of Record: https://www.sciencedirect.com/science/article/pii/S1874391920300865
Manuscript_b32450b127a4e5b8dd8ecf25f1f8b620

ATDC5 cells as a model of cartilage extracellular matrix neosynthesis, maturation and assembly

Wilhelm Dafné[1], Kempf Hervé[1], Bianchi Arnaud[1], Vincourt Jean-Baptiste[1,2]

[1] UMR 7365 CNRS-UL IMoPA, Vandoeuvre-lès-Nancy, France

[2] Proteomics core facility of UMS 2008 UL-CNRS-INSERM IBSLor, Vandoeuvre-lès-Nancy, France

Corresponding author: jean-baptiste.vincourt@univ-lorraine.fr; address: UMR 7365 CNRS-UL IMoPA, Biopole du Campus Brabois-Santé, 9, Avenue Forêt de Haye, CS 50184, 54505 Vandoeuvre-lès-Nancy Cedex France.

Key words: post-translational modifications; collagen; cartilage; proteoglycan; mass spectrometry; crosslink; extracellular matrix.

ABSTRACT:

Fibrillar collagens and proteoglycans (PGs) are quantitatively the major constituents of extracellular matrices (ECM). They carry numerous crucial post-translational modifications (PTMs) that tune the resulting biomechanical properties of the corresponding tissues. The mechanisms determining these PTMs remain largely unknown, notably because available established cell lines do not recapitulate much of the complexity of the machineries involved. ATDC5 cells are a model of chondrogenesis widely used for decades, but it remains described mostly at histological and transcriptional levels. Here, we asked to what extent this model recapitulates the events of ECM synthesis and processing occurring in cartilage. Insulin-stimulated ATDC5 cells exhibit up- or down-regulation of more than one-hundred proteins, including a number of known participants in chondrogenesis and major markers thereof. However, they also lack several ECM components considered of significant, yet more subtle, function in cartilage. Still, they assemble the large PG aggrecan and type II collagen, both carrying most of their *in vivo* PTMs, into an ECM. Remarkably, collagen crosslinking is fully lysyl oxidase (LOX)-dependent. The ATDC5 model recapitulates critical aspects of the cartilage ECM-processing machinery and should be useful to decipher the mechanisms involved.

Proteomics data are available *via* ProteomeXchange with identifier PXD014121.

Reviewer account details to get access to uploaded data *via* ProteomeXchange
Username: reviewer21888@ebi.ac.uk; **Password:** R4KJ8Ka8

ABBREVIATIONS: AB (Ammonium Bicarbonate); βAPN (beta-aminopropionitrile); CNBr (cyanogen bromide); dMMB (dimethyl methylene blue); ECM (extracellular matrix); FA (formic acid); GO (Gene Ontology); IAA (iodoacetamide); LOX (lysyl oxidase); PG (proteoglycan); PTM (post-translational modification); iATDC5 (insulin-stimulated ATDC5 cells); cATDC5 (control ATDC5 cells).

1. INTRODUCTION

The extracellular matrix (ECM) is primarily composed of proteoglycans (PGs) and fibrillar collagens, the most abundant proteins of the human body. It provides a critical scaffold for cells, is required for essential cell functions (motility, proliferation, survival and differentiation) [1] and supports most of the biomechanical properties of tissues and organs [2]. Although very distinct in their composition and elementary functions, both collagens and PGs consist in particularly large macromolecular complexes. Their building blocks carry multiple and tightly regulated post-translational modifications (PTMs) that perfectly cope with intrinsic needs and solicitations of each organ [3]. The crucial role of the ECM is highlighted by the wide spectrum of syndromes and diseases, from minor to lethal, that arise from genetic abnormalities in ECM proteins [4-6] and enzymes which modify them [7-9].

Over decades, remarkable improvements were achieved in describing ECM molecules [2, 10], notably through mass spectrometry [11-13] and characterizing the molecular mechanisms of their processing [14-17]. In vertebrates, cartilage and bone cells produce among the most abundant and dense ECM. *In vivo* high-throughput analysis of the chondrocyte transcriptome [18, 19] and proteome [20, 21] has brought a global picture of the major genes and proteins involved in cartilage ECM build-up. However, many of the processing mechanisms remain unsolved. In order to promote a more comprehensive characterization of the molecular mechanisms of cartilage ECM synthesis, the field would benefit from models of a lower complexity, such as established cell lines that recapitulate a substantial part of the sequential events occurring *in vivo*. Many established cell lines secrete procollagens and at least some of their processing intermediates into the medium. Nevertheless, very few are capable of assembling collagen into an insoluble network, and when they do, it is abnormally matured [22, 23], demonstrating how unsatisfactorily they recapitulate *in vivo* situations.

The mouse embryonal carcinoma-derived cell line ATDC5 is a widely used cellular model of chondrogenesis [24-31], referenced in more than 600 publications. It could potentially represent an instrumental model to investigate the mechanisms of ECM assembly *in vitro*. However, to date, the ATDC5 model has only been described at histological and transcriptomic levels [32-34]. Herein, for the first time, we inspected the secreted proteome of insulin-stimulated ATDC5 (iATDC5), compared to unstimulated ATDC5 (cATDC5) as controls, in order to determine to what extent these cells behave as a model of cartilage ECM synthesis.

2. EXPERIMENTAL PROCEDURES

2.1 Cell culture.

The ATDC5 cell line was obtained from the Riken Cell Bank (Tsukubai, Japan) at passage 12 and cultured at all steps at 37 °C in a humidified atmosphere of 5% CO_2 – 20% O_2. ATDC5 cells were cultured in expansion medium consisting in DMEM/F-12 containing 5 % FBS (11320-033 Gibco), 5.5 µg/ml human transferrin (T3309 Sigma), and 3.10^{-8} M sodium selenite (S5261 Sigma) as previously described [24]. Here, we used the most consensual differentiation method as follows: ATDC5 cells were plated at 3.10^5 cells/cm² in 12-well culture plates and grown in monolayer to confluence (day 0 of the experiment). From this point and for the next 14 days, cells were either maintained further in expansion medium (cATDC5) or switched to the same medium supplemented with 10 µg/mL insulin (I3146 Sigma) to induce chondrogenesis (iATDC5). From day 14 until the end of the experiment, DMEM/F12 was replaced with α-MEM medium (12561-056 Gibco) supplemented with 10 mM β-glycerophosphate (G9422 Sigma-Aldrich). Media were refreshed every other day. At days 3, 7, 14, 17, 21, 28 and 35, dedicated cultures were rinsed once and washed once for 5 min in PBS and then FBS–starved in medium containing similar transferrin and insulin concentrations for two more days, prior to analysis (allowing suitable analysis of secreted proteins). This procedure was not found to alter mitochondrial activity, as found earlier with primary cells [35].

2.2 Study design.

As the study required multiple analytical methods, which could not be performed from the same samples, 15 replicate culture wells were prepared for each experimental point to be investigated. Of them, 3 were used for mineral deposition analysis, 3 for proteoglycan content, 3 for DNA content, 3 for mRNA studies and 3 for proteome studies, allowing technical triplicate analysis in each analytical procedure. Most data reported here were obtained from a single experiment, all technical replicates were grown and processed at the same time, because our preliminary investigations suggested that independent experiments exhibited differences in the folds and exact time-points of regulations, although in other experiments the responses of iATDC5 followed a similar trend.

2.3 Differentiation and mineralization.

ATDC5 cells were stained with Alcian blue 8GX pH 0.1 (S/6-2 RA Lamb) and Alizarin red (A5533 Sigma-Aldrich) at various time points to detect respectively the proteoglycan

accumulation and mineralization. Cells were washed with PBS twice, then fixed with ethanol 90% at -20 °C until the end of the experiment, rinsed in PBS and stained with 0.1% Alcian blue in HCl 0.1M overnight, or 1% Alizarin red in distilled water (pH 4.1-4.3) for 45 sec, rinsed with distilled water, and photographed. Thereafter, Alcian blue was dissolved in 6 M HCl overnight and absorbance at λ = 600 nm was measured in 96-well plates using a Varioscan spectrophotometer (Thermofisher). Measurements of Alizarin red staining was performed as described [55].

2.4 Hoechst assay for DNA content.

Monolayer cell cultures were resuspended in 1 mL of 10 mM Tris, 1 mM EDTA, 0.1 M NaCl, pH 7.4 and lysed by two pulse-mode sonication burst cycles of 30 seconds each. Samples were centrifuged for 30 minutes at 4°C, 16 000g. 200 µL of 0.1 µg/mL Hoechst 33258 (H3569 Invitrogen) were added to 10 µL supernatants in a 96-well plates. Fluorometric measurements (348 nm excitation wavelength, 456 nm absorption wavelength) were performed with the Varioskan Flash (Thermo Scientific) running Skan it 2.4.3 software. Absolute quantification was determined using standard curves prepared from calf thymus DNA (D3664-2MG Sigma).

2.5 RNA extraction and reverse transcription-polymerase chain reaction analysis.

Total RNA was isolated from ATDC5 cells using RNeasy mini kit® (74104 Qiagen). 100ng of total RNA were reverse-transcribed using M-MLV reverse transcriptase (28025021 Invitrogen) with random hexamer primers, according to the manufacturer's instructions.

2.6 Real-time quantitative polymerase chain reaction.

Real time quantitative PCR was performed with standard curve calibration, using the following primers : RPS29_For, GGAGTCACCCACGGAAGTT; RPS29_Rev, GCCTATGTCCTTCGCGTACT; COL1A1_For, GTCCCGCTGGTCAAGATG; COL1A1_Rev, CTCCAGCCTTTCCAGGTTCT; COL2A1_For, TGGTATTCCTGGAGCCAAAG; COL2A1 Rev, ACCAGTTGCACCTTGAGGAC; COL10A1_For, TTCATCCCATACGCCATAAAG; COL10A1_Rev, AGGGACCTGGGTGTCCTC; ACAN_For, GCTGCAGTGATCTCAGAAGAAG; ACAN_Rev, GATGGTGAGGGAAGACCCTA; MMP13_For, ACTCAAATGGTCCCAAACGA; MMP13_Rev, GGTGTCACATCAGACCAGACC. The results were normalized to *RPS29* mRNA levels as described [36] and expressed as fold regulation over the control.

2.7 Protein fractionation.

All steps were performed at 4 °C unless stated otherwise. Centrifugations were all carried out at 16,000 g for 5 min. All fractions were stored at -80°C until further use. Soluble secretomes were collected, clarified and stored. Cell monolayers were washed twice in PBS, scrapped in 1 ml PBS-Triton X-100 1%, transferred to Eppendorf tubes and centrifuged. The soluble fraction was collected and its protein content was measured using Pierce® BCA Protein Assay Kit (23225 Thermo Scientific) as recommended by manufacturer's instructions. The insoluble fraction was washed once in 1 ml PBS-Triton X-100 1% and further extracted in 500 µl Guanidine Buffer [guanidine hydrochloride (G3272 Sigma-Aldrich), 6 M; Tris, 50 mM, DTT 50 mM, pH 8.0] overnight under rotation. The Guanidine-soluble fraction was collected. The remaining insoluble material consisted mostly in crosslinked collagen, which can be extracted only by proteolytic digestion. This was achieved by reaction with CNBr (cyanogen bromide, C91492 Sigma-Aldrich), which specifically cleaves peptide bounds C-terminal to methionine residues (and modifies methionine to homoserine lactone). To this aim, the guanidine-resistant material was washed once in guanidine buffer and twice in distilled water, then rotated overnight in 500 µl pure formic acid (FA; 405792 Carlos Erba). 1 volume water was added and the FA-soluble fraction was removed. Of note, this fraction contained insignificant protein amounts as judged when running evaporated samples on Coomassie gels. The remaining insoluble material was resuspended in 200 µl FA, 70% containing 0.125 M CNBr and incubated overnight on a rotator at room temperature in the dark under a chemical fume hood. At this point, no insoluble material was visible in any sample. 5 volumes of distilled water were added and samples were dried overnight prior to storage, effluents being collected in a cold trap containing excess NaOH.

2.8 Sample preparation for proteome studies.

Samples were all normalized based on the DNA content of the corresponding time points. Secretomes were normalized by dilution according to cell numbers into fresh expansion medium and precipitated using DOC 0.02% (D6750 Sigma)/TCA 15% (T6399 Sigma), rinsed twice in acetone, resuspended in 50 µl Urea Buffer (6M Urea, 50 mM Tris, pH 8.0) and stored until further use. Guanidine extracts were precipitated with 2 vol. ice-cold Ethanol overnight at -20 °C, rinsed twice in Ethanol, dried for 5 min and resuspended in an adequate volume of Laemmli loading buffer for normalization, the minimal volume being set to 100 µl. Samples were resuspended through 3 freeze/thaw cycles and 20 µl aliquots were Ethanol-precipitated, dried, resuspended in 20 µl Urea Buffer and stored until further use. Dried CNBr extracts

were directly resuspended in Laemmli Buffer, normalized and prepared for further use similarly. Prior to tryptic digestion of soluble secretome and Guanidine extracts, 20 µl samples prepared in Urea Buffer were treated at room temperature with 0.8 µl DTT 100 mM for 45 min, then with 0.8 µl IAA (iodoacetamide) 1 M for 45 min and finally with 0.8 µl DTT 1 M for 10 min. These steps were avoided for CNBr extracts. Samples were then diluted 10-fold in Tris, 50 mM, CaCl2 1 mM, pH 8.0, containing 200 ng sequencing grade trypsin (V5280 Promega) and incubated at 37 °C overnight. Digests were purified using C18 spin columns (89870 Thermofisher) according to manufacturer's recommendations, resuspended in 20 µl ACN 2 %, TFA 0.1 % and transferred to nanoHPLC sampler tubes.

2.9 Nano-HPLC setup for proteome investigations.

Fractionation was managed mostly as described in [37] with slight adjustments depending on sample types. For soluble secretomes and Guanidine extracts, samples were separated into 340 fractions using longer LC runs, the ACN gradients being lengthened to 60 min, in order to better cope with sample complexity. For CNBr extracts, considering the relatively low hydrophobicity of most collagen peptides, LC gradients were kept to 30 min and the 170-fraction collection was set between 17.5 and 34.5 min.

2.10 Relative peptide quantification based on label-free LC-MALDI.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium [38] *via* the PRIDE [39] partner repository with the dataset identifier PXD014121 and 10.6019/PXD014121. Measurements were performed mostly as described in [35]. Ratios and t-test p-values for all (m/z ; retention time) coordinates were calculated using ProfileAnalysis 2.3 (Bruker) by comparing MS ion chromatograms with the following filters: 50 ppm mass deviation; 30 sec retention time deviation; minimum 3 counts within one biological group. Most calculations were made between a given iATDC5 time-point and its control counterpart. However, the soluble secretomes of control cells were incrementally contaminated by FBS-derived proteins from day 21 to 35 (evidenced in figure 5. D), which were likely to lower the apparent relative amounts of individual proteins of interest in the corresponding samples. In order to avoid such bias, proteins found increased in iATDC5 cells at these time points were cross-validated by calculations of their changes in iATDC5 cells between successive time points, from day 14 to 35. All allowed calculations were exported to proteinscape 2.4 server.

2.11 Peptide, PTM and protein identifications.

All (m/z: retention time) coordinates found regulated with relative ratios below -2 or above +2 with significance (t–test p-value < 0.05) at any time point were selected by creating corresponding scheduled precursor lists in ProfileAnalysis, which were then used for automated MS/MS fragmentation in dedicated automatic methods in WARP-LC (Bruker). In addition, all masses found with s/n > 20 in one random sample of each analysis were also selected for MS/MS. All obtained MS/MS spectra together with the corresponding (m/z; retention time) coordinates were exported to Proteinscape server 2.4 and used for Swissprot mouse database interrogation through Mascot server 2.4. There were distinctions in database interrogation filters and data handling, depending on the considered extracts, which are detailed in Supplemental data legends. For global extract analysis, proteins were considered as identified with sufficient certainty if obtaining an FDR below 0.5 % using the random decoy strategy.

2.12 Protein ratio calculations and statistical interpretation.

Ratios and corresponding p-values were assigned to identified peptides in Proteinscape 2.4 using the following alinement criteria: m/z deviation < 0.03 Da, retention time deviation < 30 sec. For each protein identified from 4 or more peptides passing quality control, the two peptides exhibiting lowest and highest ratio values were excluded. In addition, when analyzing soluble secretomes, due to massive regulation of proline hydroxylation in fibrillar collagens (explained in results), all peptides corresponding to their triple helices were manually unselected, leaving for calculation only the peptides originating from C-propeptide regions. The median value of ratios observed for all remaining peptides of a given protein was used as the protein fold change and the median value of all remaining peptide p-values was used as the protein p-value for the observed change (if any). In the case of qualitative changes (peptides that were observed in all three replicates of a given group but lacking in the other), ratios and t-tests could not be calculated and therefore, these peptides were not integrated into the statistics of the corresponding protein. However, peptide qualitative changes were retained in the protein tables, as there were proteins described only by such peptides, reflecting potent, yet unquantifiable regulations. Such cases were considered whenever the corresponding protein was described by a minimum of two such peptides (all of them exhibiting coherent regulations).

2.13 SDS-PAGE and staining.

The same samples used in MS analysis were separated by SDS-PAGE on gradient 4-20% PROTEAN® TGX Stain-Free™ Gels (456-8095 BIO-RAD), fixed, washed in water and

stained either with colloidal Coomassie or dimethylmethylene Blue (dMMB, 03610 Polysciences) [40]. dMMB staining was chosen to estimate both the length of glycan chains (based on electrophoretic mobility) and their sulfation (based on staining intensity) and was performed as follows: 122 mg dMMB was first dissolved in 10 ml ethanol overnight under rotation and then settled for 2 hours. The staining solution consisted in 5 ml of the obtained soluble fraction diluted into 100 ml formic acid, 1mM, pH 3.5, 20% ethanol. Gels were stained overnight and washed in pure water several times and scanned on a Chemidoc XRS+ imager (BIO-RAD).

2.14 Western blotting analysis.

SDS-PAGE was performed as described in the previous section and proteins were transferred to polyvinylidene difluoride membranes within Trans-Blot® Turbo™ Transfer Packs (170-4157 BIO-RAD) with Trans-Blot® Turbo™ transfert system. Antibodies against the C-propeptide of type II procollagen as well as the corresponding HRP-conjugated secondary antibodies were described elsewhere [41]. Blots were developed with the enhanced chemiluminescence system (1705061 BIO-RAD) and acquired on a Chemidoc XRS+ imager (BIO-RAD). Loading controls were performed based on stain-free signal.

2.15 Polyacrylamide gel band analysis.

For gel band protein content analysis, gel pieces were washed as follows : three times in 200 µL of amonium bicarbonate (AB) 100 mM for 15 min, once in 200 µL of ACN 50%/AB 50% for 15 min, then once in 200 µL of AB for 15 min, and finally twice in 200 µL ACN for 3 min. Gel fragments were then sonicated for 2 min, and dried 30 min. 100 ng of trypsin was added in a total volume of 20 µl AB, 50 mM and digestion was performed overnight at room temperature. Peptides were extracted twice in 25 µL of TFA 1% /ACN 80% under sonication for 7 minutes, supernatants were pooled, dried and resuspended in 10 µL ACN 2% TFA 0,1% and sonicated for 30 sec. Supernatants were transferred to nano-HPLC sampler tubes and processed for MALDI analysis as described above without notion of relative quantification, ions being selected for fragmentation based on their S/N ratio, the minimal S/N threshold being set to 20.

2.16 Bioinformatic Gene Ontology Analysis.

Proteins identified by LC-MALDI-MS/MS were analyzed using DAVID Bioinformatics Resources v6.7 [42] to obtain a comprehensive description of the overrepresented biological processes and functional related groups of proteins within our data set. For DAVID analysis,

only FDR significant (p-value < 0.005) overrepresented terms were considered. As background, the default *Mus musculus* genome was used.

2.17 Abbreviations of protein names.

Most protein names reported here are abbreviated according to the Uniprot nomenclature. However, in the field of collagen biology, a specific nomenclature is preferred. For clarity, all abbreviated protein names were reported in Supplemental data S.1 together with full protein names and links to the corresponding Uniprot web pages.

3. RESULTS

3.1 The chondrogenic-like phenotype of iATDC5 confirmed at the proteome level by soluble secretome analysis.

After confirming that the phenotype of our samples were in accordance with the literature [26, 28] (Supplemental data S.2), we examined their soluble secretomes [43] as previously described [35, 37]. 72 proteins being significantly up- or down-regulated were revealed. To obtain a comprehensive description of the related groups of regulated proteins, we performed Gene Ontology (GO) annotation enrichment analyses, focusing on biological processes. **[figure 1]** Figure 1. A shows the significant top four enriched GO term categorization that classifies 33 proteins. Two of the most enriched terms are consistent with chondrogenic differentiation. Collagen fibril organization [GO: 0030199; including α1(II), α1(V), α2(V), α1(XI), α2(XI) collagens, but also PGCA, ANXA2, FMOD, LYOX], the most enriched of all, and ECM organization (GO: 0030198; SMOC2, CCD80, FBLN1, FBLN5, LEG3, NID2, POSTN, PGBM), with both early and late protein regulations, depicting one major valuable feature of iATDC5 cells: the ability to produce large amounts of ECM components. Noticeably, glycolytic process (GO: 0006096; ALDOA, ENOA, PGK1, PGAM1, KPYM, TPIS), the second most enriched term, was upregulated in iATDC5 immediately upon insulin stimulation and throughout the whole experiment.

Among the 72 regulated proteins in the soluble secretome, 42 proteins (figure 2. B) have previously been reported as regulated at the mRNA level in chondrogenic differentiation of ATDC5 cells [33] or at the protein level upon chondrogenesis *in vivo* [20] and *ex vivo* [21]. Among them, α1(II) and aggrecan (PGCA) were dramatically upregulated from day 7 (figure 1. B), demonstrating at the proteome level the chondrocyte-like expression profile of

iATDC5. Type XI collagen [α1(XI) and α2(XI)], although poorly used as a marker in the model, is a component of functional heterofibrils found *in vivo* in cartilage [44] and was observed concomitantly with α1(II). Noteworthy, among ECM-related components, there were not only up- but also dramatically down-regulated proteins, including Periostin (POSTN), Fibulin-1 (FBLN1), type V collagen [α1(V)] and type XII collagen [α1(XII)], suggesting that iATDC5 cells not only started synthesizing cartilage-related proteins but also ceased synthesizing a number of proteins related to their basal phenotype.

3.2 The composition of the guanidine-soluble ECM of ATDC5 cells.

Next, we examined guanidine extracts, which, when prepared from cartilage, are mostly composed of PGs, non-fibrillar collagens and accessory ECM proteins [20]. Guanidine extracts obtained from cATDC5 were negligible at time points earlier than day 7 of the experiment and therefore we started investigating iATDC5 vs cATDC5 regulation ratios from this time point. **[figure 2]** By day 17, however, the content of guanidine extracts from cATDC5 increased and reached total levels comparable to those of iATDC5 (figure 2. A), suggesting that switching the medium to αMEM favored their ECM synthesis. In this fraction, 52 proteins were found regulated (figure. 2. B), demonstrating both early and late ECM remodeling events. Regarding the common markers of chondrogenesis, guanidine extracts demonstrated again the upregulation of α1(II) and PGCA, suggesting that iATDC5 not only synthesized major cartilage components, but also integrated them to their ECM. In addition, this fraction revealed the upregulation of α1(X) (type X collagen), as early as day 7, while it was absent from soluble secretomes. However, although *COL10* mRNA levels were found 400-fold upregulated (supplemental data S.2, in agreement with the literature), it remained a very minor component of the ECM, as suggested by its relatively low identification score until the end of the experiment (not included in figure 2 but available from Supplemental data S.4). Among the 52 proteins found regulated in this fraction, 24 proteins were previously reported through proteome analysis of guanidine extracts prepared from chondrogenesis models, either *in vivo* [20] or *ex vivo* [21]. These included, in addition to the common markers of chondrogenesis, other fibrillar [α1(I), α2(I), α1(XII)] and non-fibrillar [α1(VI) and α2(VI)] collagens, PGs (PGS1, PGBM), accessory ECM proteins (NID2, SRPX2, EMIL1, TENA, TSP1), extracellular (LOXL4) and intracellular (G3P, PDIA3) enzymes, transporters (VDAC1, ADT1, ADT2), a chaperone (GRP78), but also cytoskeletal (TBA1B, VIM) and nuclear (LMNA) components. We also found a number of regulations not described earlier in

chondrogenesis (notably, RPN1, PDIA1, GRP75, ACTB, TBB5, TIMP3, ATPB), including several nuclear proteins (H11, H13, H13, H15, H2A1H and H2B1H, HNRH1, ROA1, ROA2, DHX9, RMXL1).

3.3 iATDC5 accumulate largely mature aggrecan in their ECM.

In the guanidine extract data set, peptide identification of aggrecan core protein was restricted to the G1, G2, and G3 globular domains, suggesting that interglobular glycosaminoglycan carrier domains must be heavily modified, preventing identification by LC-MS/MS [45]. To better characterize the length of glycan chains and sulfation of the PGs, we performed SDS-PAGE coupled to dMMB staining (figure 2. C). It did not reveal any staining against samples obtained from cATDC5 (figure 2. C) even at the latest time-points, while both iATDC5, as early as day 14, and bovine articular cartilage samples exhibited very high molecular weight PGs. The major corresponding protein was found by mass spectrometry to be aggrecan in both cases (data not shown). These findings were in accordance with the notion that aggrecan was largely glycosylated and sulfated (based on dMMB staining [40]). When the same samples were investigated after protein backbone removal, the observed patterns smeared over almost the entire lanes, reflecting the heterogeneity of glycosaminoglycan chain lengths. There were apparently less chains of the highest molecular weights from iATDC5 than from bovine cartilage however. Still, aggrecan of iATDC5 appears highly glycosylated and sulfated in a way that resembles the *ex vivo* sample, indicating that iATDC5 possess most of the naturally occurring PG processing machinery in cartilage.

3.4 Composition and regulation of the collagen fraction.

Up to day 14 of the experiment, cATDC5 did not contain any guanidine-resistant material. Therefore, the CNBr contents at earliest time-points could not be compared and we investigated only the composition of the corresponding iATDC5 CNBr extracts. **[figure 3]** These were almost uniquely composed of fibrillar collagens, with the only significant exception of Fibronectin (FINC, figure 3. A). Therefore, iATDC5 were capable of integrating fibrillar collagens into a solid network, as found *in vivo*. Based on the number of identified peptides for each of them, α1(II) came second, preceded only by α1(I), and representing about $1/4^{th}$ of the total collagen. From day 17 onward, cATDC5 started accumulating material in the CNBr fraction, enabling comparisons with iATDC5 (figure 3. B). In addition to α1(II), which was substantially upregulated at all time-points, type III collagen [α1(III)] was very robustly down-regulated. Others regulations, although less prominent, were concomitantly

evidenced, notably upregulation of α1(XI) and downregulation of FINC. Therefore, collagens are not only produced by iATDC5, but also crosslinked into the ECM.

3.5 Full inhibition of collagen crosslinking by βAPN.

In vivo, collagen crosslinking requires LOX-mediated oxidative deamination of key lysine residues [13]. iATDC5 cells were treated during the whole experiment with 0.2 mM beta-aminopropionitrile (βAPN), which irreversibly inhibits enzymes of the LOX family (figure. 3. C). In contrast to non-treated iATDC5, βAPN-treated iATDC5 exhibited a total shift of the collagen content from the CNBr digest to the guanidine extract, reflecting potent inhibition of collagen crosslinking. Therefore, in the ATDC5 model, collagen crosslinking appears to be fully dependent on LOX activity.

3.6 Type II procollagen proteolytic cleavage in iATDC5.

As the proteolytic processing of procollagen is a critical step of fibrillogenesis [46], we next questioned whether N- and C-propeptides were cleaved from the type II procollagen secreted by iATDC5. **[figure 4]** We found peptides that unequivocally demonstrated cleavage of both the N- and C-propeptides at their conventional locations (figure 4), but none corresponding to the uncleaved precursor. Indeed, our western blotting experiments using an antibody directed against the C-propeptide of type II procollagen revealed, as early as day 7 and until day 35, a single reactive species corresponding to the cleaved C-propeptide (Supplemental data 7). Altogether, these data indicate that synthesized α1(II) by iATDC5 is fully processed in terms of proteolytic cleavage.

3.7 The PTMs of type II collagen in iATDC5 are very similar to those found *in vivo.*

In order to further characterize type II collagen synthesized by iATDC5, we investigated its chemical modifications (figures 4&5). **[figure 5]** Remarkably, many of the peptides identified in this approach were found under several distinct chemical states. This finding highlights the heterogeneity of PTM combinations carried by the population of α1(II) molecules (proline hydroxylation, lysine hydroxylation, galactosylation and glucosyl-galactosylation). Overall, combining findings from the CNBr digests and soluble secretomes, 59 % of the triple helix sequence was covered by this analysis. This rate is insufficient to report on the exact PTM status of α1(II) in the model, especially considering that several regions of particular interest, such as the N-telopeptide and both GXKGHR motifs involved in cross-linking, were missing. Still, this level of sequence coverage allows to compare the observed PTMs to those found *in*

vivo (figure 5), taking as a reference the most exhaustive α1(II) PTM mapping reported to date, obtained from bovine cartilage pepsin extracts [47]. The vast majority of the PTMs found in iATDC5 was consistent with those reported *in vivo*.

4. DISCUSSION

Since its first description in 1990 [24], the ATDC5 model has become truly popular in the field of chondrogenesis and spread through worldwide laboratories, allowing pertinent functional characterization of multiple actors of chondrogenesis without the complexity of *in vivo* experiments [25, 27, 28, 51-53]. Noticeably, iATDC5 cells accumulate fibrillar collagen-resembling structures into their ECM, as shown by electron microscopy [31].

However, the ATDC5 proteome remained largely unknown until now. Here we investigated the time-course of proteome regulations occurring in iATDC5, taking cATDC5 as controls, in order to better describe the actual composition of their ECM, and found 106 proteins significantly regulated proteins, taking into consideration all time points and extracts (72 from soluble secretomes and 34 additional ones from Guanidine/CNBr extracts). Among them, 62 proteins had been described earlier in proteomics studies of chondrogenesis, either *in vivo* or *ex vivo* [20, 21], consolidating the biological significance of the model. Several of the presently identified proteins have not been characterized in chondrogenesis yet and ATDC5 cells may thus be considered as a good model to assess their functions in the future.

The present findings also indicate that a number of proteins considered important in chondrogenesis are missing or at unexpectedly low levels in this model. In particular, Cartilage Oligomeric Matrix Protein (COMP) is abundant (and functionally important) in cartilage. Indeed, it is among the most abundantly secreted proteins in primary chondrocytes [37, 54] and in mesenchymal stem cells undergoing chondrogenesis [35, 55]. Yet, it is absolutely missing in the present dataset. Likewise, other functionally important cartilage ECM components, notably, type IX collagen [56] and Matrilins [57], though less obviously evidenced by mass spectrometry in cell culture systems, are not found either. Type X collagen is found specifically in guanidine extracts prepared from iATDC5 (figure 2), but it remains close to the minimal detection threshold. These findings are worth being considered prior to using the model, especially when interested in the processes in which these proteins are involved. In particular, COMP, type IX collagen and Matrilins are believed to work

collectively at interconnecting ECM molecules within cartilage and their absence in the model is likely to alter the ECM ultrastructure [58].

Nevertheless, the model synthesizes and assembles into its ECM the major cartilage components, aggrecan and type II collagen. In the ATDC5 model, aggrecan resembles that found *in vivo* (figure 2). Type II collagen is fully cleaved proteolytically (figure 4). Its crosslinking is fully dependent on LOX activity (figure 3), like it is in cartilage [59]. These are appreciable behaviors of the model, in comparison to previously characterized cell lines [22, 60]. In addition, many of the post-translational modifications of type II collagen are similar to those observed *in vivo* (figure 5). Maturation processes of PGs and collagens involve multiple and coordinated steps, that remain largely uncharacterized, precisely because established cellular models that recapitulate them are lacking. In this respect, in the light of our findings, ATDC5 cells should be instrumental to better dissect the mechanisms of cartilage ECM assembly.

ACKNOWLEDGEMENTS: we are grateful to Frédéric Cailotto and David Moulin, from our laboratory, for their critical reading of the manuscript. The proteomics core facility of UMS 2008 UL-CNRS-INSERM IBSLor acknowledges the funding from CPER IT2MP_BIOS.

DECLARATION OF INTEREST: none.

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

5. BIBLIOGRAPHY

[1] Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. J Cell Sci. 2010;123:4195-200.

[2] Kadler KE, Baldock C, Bella J, Boot-Handford RP. Collagens at a glance. J Cell Sci. 2007;120:1955-8.

[3] Hudson DM, Garibov M, Dixon DR, Popowics T, Eyre DR. Distinct post-translational features of type I collagen are conserved in mouse and human periodontal ligament. J Periodontal Res. 2017;52:1042-9.

[4] Jarvelainen H, Sainio A, Koulu M, Wight TN, Penttinen R. Extracellular matrix molecules: potential targets in pharmacotherapy. Pharmacol Rev. 2009;61:198-223.
[5] Forlino A, Marini JC. Osteogenesis imperfecta. Lancet. 2016;387:1657-71.
[6] Wenstrup RJ, Florer JB, Davidson JM, Phillips CL, Pfeiffer BJ, Menezes DW, et al. Murine model of the Ehlers-Danlos syndrome. col5a1 haploinsufficiency disrupts collagen fibril assembly at multiple stages. J Biol Chem. 2006;281:12888-95.
[7] Schwarze U, Cundy T, Liu YJ, Hofman PL, Byers PH. Compound heterozygosity for a frameshift mutation and an upstream deletion that reduces expression of SERPINH1 in siblings with a moderate form of osteogenesis imperfecta. Am J Med Genet A. 2019;179:1466-75.
[8] Scietti L, Campioni M, Forneris F. SiMPLOD, a Structure-Integrated Database of Collagen Lysyl Hydroxylase (LH/PLOD) Enzyme Variants. J Bone Miner Res. 2019;34:1376-82.
[9] Hudson DM, Joeng KS, Werther R, Rajagopal A, Weis M, Lee BH, et al. Post-translationally abnormal collagens of prolyl 3-hydroxylase-2 null mice offer a pathobiological mechanism for the high myopia linked to human LEPREL1 mutations. J Biol Chem. 2015;290:8613-22.
[10] Theocharis AD, Skandalis SS, Gialeli C, Karamanos NK. Extracellular matrix structure. Adv Drug Deliv Rev. 2016;97:4-27.
[11] Wu JJ, Weis MA, Kim LS, Eyre DR. Type III collagen, a fibril network modifier in articular cartilage. J Biol Chem. 2010;285:18537-44.
[12] Hsueh MF, Khabut A, Kjellstrom S, Onnerfjord P, Kraus VB. Elucidating the Molecular Composition of Cartilage by Proteomics. J Proteome Res. 2016;15:374-88.
[13] Eyre DR, Weis MA, Wu JJ. Advances in collagen cross-link analysis. Methods. 2008;45:65-74.
[14] Heard ME, Besio R, Weis M, Rai J, Hudson DM, Dimori M, et al. Sc65-Null Mice Provide Evidence for a Novel Endoplasmic Reticulum Complex Regulating Collagen Lysyl Hydroxylation. PLoS Genet. 2016;12:e1006002.
[15] Pyott SM, Schwarze U, Christiansen HE, Pepin MG, Leistritz DF, Dineen R, et al. Mutations in PPIB (cyclophilin B) delay type I procollagen chain association and result in perinatal lethal to moderate osteogenesis imperfecta phenotypes. Hum Mol Genet. 2011;20:1595-609.

[16] Kalamajski S, Liu C, Tillgren V, Rubin K, Oldberg A, Rai J, et al. Increased C-telopeptide cross-linking of tendon type I collagen in fibromodulin-deficient mice. J Biol Chem. 2014;289:18873-9.
[17] Eyre DR, Weis MA. Bone collagen: new clues to its mineralization mechanism from recessive osteogenesis imperfecta. Calcif Tissue Int. 2013;93:338-47.
[18] Karsenty G, Wagner EF. Reaching a genetic and molecular understanding of skeletal development. Dev Cell. 2002;2:389-406.
[19] Clancy BM, Johnson JD, Lambert AJ, Rezvankhah S, Wong A, Resmini C, et al. A gene expression profile for endochondral bone formation: oligonucleotide microarrays establish novel connections between known genes and BMP-2-induced bone formation in mouse quadriceps. Bone. 2003;33:46-63.
[20] Kudelko M, Chan CW, Sharma R, Yao Q, Lau E, Chu IK, et al. Label-Free Quantitative Proteomics Reveals Survival Mechanisms Developed by Hypertrophic Chondrocytes under ER Stress. J Proteome Res. 2016;15:86-99.
[21] Wilson R, Norris EL, Brachvogel B, Angelucci C, Zivkovic S, Gordon L, et al. Changes in the chondrocyte and extracellular matrix proteome during post-natal mouse cartilage development. Mol Cell Proteomics. 2012;11:M111 014159.
[22] Fernandes RJ, Schmid TM, Eyre DR. Assembly of collagen types II, IX and XI into nascent hetero-fibrils by a rat chondrocyte cell line. Eur J Biochem. 2003;270:3243-50.
[23] Kerwar SS, Cardinale GJ, Kohn LD, Spears CL, Stassen FL. Cell-free synthesis of procollagen: L-929 fibroblasts as a cellular model for dermatosparaxis. Proc Natl Acad Sci U S A. 1973;70:1378-82.
[24] Atsumi T, Miwa Y, Kimata K, Ikawa Y. A chondrogenic cell line derived from a differentiating culture of AT805 teratocarcinoma cells. Cell Differ Dev. 1990;30:109-16.
[25] Shukunami C, Shigeno C, Atsumi T, Ishizeki K, Suzuki F, Hiraki Y. Chondrogenic differentiation of clonal mouse embryonic cell line ATDC5 in vitro: differentiation-dependent gene expression of parathyroid hormone (PTH)/PTH-related peptide receptor. J Cell Biol. 1996;133:457-68.
[26] Shukunami C, Ishizeki K, Atsumi T, Ohta Y, Suzuki F, Hiraki Y. Cellular hypertrophy and calcification of embryonal carcinoma-derived chondrogenic cell line ATDC5 in vitro. J Bone Miner Res. 1997;12:1174-88.
[27] Akiyama H, Shigeno C, Hiraki Y, Shukunami C, Kohno H, Akagi M, et al. Cloning of a mouse smoothened cDNA and expression patterns of hedgehog signalling molecules during

chondrogenesis and cartilage differentiation in clonal mouse EC cells, ATDC5. Biochem Biophys Res Commun. 1997;235:142-7.
[28] Shukunami C, Akiyama H, Nakamura T, Hiraki Y. Requirement of autocrine signaling by bone morphogenetic protein-4 for chondrogenic differentiation of ATDC5 cells. FEBS Lett. 2000;469:83-7.
[29] Yao Y, Wang Y. ATDC5: an excellent in vitro model cell line for skeletal development. J Cell Biochem. 2013;114:1223-9.
[30] Altaf FM, Hering TM, Kazmi NH, Yoo JU, Johnstone B. Ascorbate-enhanced chondrogenesis of ATDC5 cells. Eur Cell Mater. 2006;12:64-9; discussion 9-70.
[31] Newton PT, Staines KA, Spevak L, Boskey AL, Teixeira CC, Macrae VE, et al. Chondrogenic ATDC5 cells: an optimised model for rapid and physiological matrix mineralisation. Int J Mol Med. 2012;30:1187-93.
[32] Chen L, Fink T, Zhang XY, Ebbesen P, Zachar V. Quantitative transcriptional profiling of ATDC5 mouse progenitor cells during chondrogenesis. Differentiation. 2005;73:350-63.
[33] Chen L, Fink T, Ebbesen P, Zachar V. Temporal transcriptome of mouse ATDC5 chondroprogenitors differentiating under hypoxic conditions. Exp Cell Res. 2006;312:1727-44.
[34] Zhai Z, Yao Y, Wang Y. Importance of suitable reference gene selection for quantitative RT-PCR during ATDC5 cells chondrocyte differentiation. PLoS One. 2013;8:e64786.
[35] Henrionnet C, Gillet P, Mainard D, Vincourt JB, Pinzano A. Label-free relative quantification of secreted proteins as a non-invasive method for the quality control of chondrogenesis in bioengineered substitutes for cartilage repair. J Tissue Eng Regen Med. 2018;12:e1757-e66.
[36] Guibert M, Gasser A, Kempf H, Bianchi A. Fibroblast-growth factor 23 promotes terminal differentiation of ATDC5 cells. PLoS One. 2017;12:e0174969.
[37] Riffault M, Moulin D, Grossin L, Mainard D, Magdalou J, Vincourt JB. Label-free relative quantification applied to LC-MALDI acquisition for rapid analysis of chondrocyte secretion modulation. J Proteomics. 2015;114:263-73.
[38] Deutsch EW, Csordas A, Sun Z, Jarnuczak A, Perez-Riverol Y, Ternent T, et al. The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. Nucleic Acids Res. 2017;45:D1100-D6.
[39] Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019;47:D442-D50.

[40] Farndale RW, Buttle DJ, Barrett AJ. Improved quantitation and discrimination of sulphated glycosaminoglycans by use of dimethylmethylene blue. Biochim Biophys Acta. 1986;883:173-7.
[41] Vincourt JB, Etienne S, Cottet J, Delaunay C, Malanda CB, Lionneton F, et al. C-propeptides of procollagens I alpha 1 and II that differentially accumulate in enchondromas versus chondrosarcomas regulate tumor cell survival and migration. Cancer Res. 2010;70:4739-48.
[42] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44-57.
[43] Maeda N, Hagihara H, Nakata Y, Hiller S, Wilder J, Reddick R. Aortic wall damage in mice unable to synthesize ascorbic acid. Proc Natl Acad Sci U S A. 2000;97:841-6.
[44] McAlinden A, Traeger G, Hansen U, Weis MA, Ravindran S, Wirthlin L, et al. Molecular properties and fibril ultrastructure of types II and XI collagens in cartilage of mice expressing exclusively the alpha1(IIA) collagen isoform. Matrix Biol. 2014;34:105-13.
[45] Knudson CB, Knudson W. Cartilage proteoglycans. Semin Cell Dev Biol. 2001;12:69-78.
[46] Miyahara M, Njieha FK, Prockop DJ. Formation of collagen fibrils in vitro by cleavage of procollagen with procollagen proteinases. J Biol Chem. 1982;257:8442-8.
[47] Song E, Mechref Y. LC-MS/MS identification of the O-glycosylation and hydroxylation of amino acid residues of collagen alpha-1 (II) chain from bovine cartilage. J Proteome Res. 2013;12:3599-609.
[48] Hojima Y, van der Rest M, Prockop DJ. Type I procollagen carboxyl-terminal proteinase from chick embryo tendons. Purification and characterization. J Biol Chem. 1985;260:15996-6003.
[49] Van der Rest M, Rosenberg LC, Olsen BR, Poole AR. Chondrocalcin is identical with the C-propeptide of type II procollagen. Biochem J. 1986;237:923-5.
[50] Dombrowski KE, Prockop DJ. Cleavage of type I and type II procollagens by type I/II procollagen N-proteinase. Correlation of kinetic constants with the predicted conformations of procollagen substrates. J Biol Chem. 1988;263:16545-52.
[51] Zhang Y, Yang TL, Li X, Guo Y. Functional analyses reveal the essential role of SOX6 and RUNX2 in the communication of chondrocyte and osteoblast. Osteoporos Int. 2015;26:553-61.

[52] Lodewyckx L, Cailotto F, Thysen S, Luyten FP, Lories RJ. Tight regulation of wingless-type signaling in the articular cartilage - subchondral bone biomechanical unit: transcriptomics in Frzb-knockout mice. Arthritis Res Ther. 2012;14:R16.
[53] Castano Betancourt MC, Cailotto F, Kerkhof HJ, Cornelis FM, Doherty SA, Hart DJ, et al. Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. Proc Natl Acad Sci U S A. 2012;109:8218-23.
[54] Taylor DW, Ahmed N, Parreno J, Lunstrum GP, Gross AE, Diamandis EP, et al. Collagen type XII and versican are present in the early stages of cartilage tissue formation by both redifferentating passaged and primary chondrocytes. Tissue Eng Part A. 2015;21:683-93.
[55] Rocha B, Calamia V, Casas V, Carrascal M, Blanco FJ, Ruiz-Romero C. Secretome analysis of human mesenchymal stem cells undergoing chondrogenic differentiation. J Proteome Res. 2014;13:1045-54.
[56] Eyre DR, Apon S, Wu JJ, Ericsson LH, Walsh KA. Collagen type IX: evidence for covalent linkages to type II collagen in cartilage. FEBS Lett. 1987;220:337-41.
[57] Klatt AR, Becker AK, Neacsu CD, Paulsson M, Wagener R. The matrilins: modulators of extracellular matrix assembly. Int J Biochem Cell Biol. 2011;43:320-30.
[58] Brachvogel B, Zaucke F, Dave K, Norris EL, Stermann J, Dayakli M, et al. Comparative proteomic analysis of normal and collagen IX null mouse cartilage reveals altered extracellular matrix composition and novel components of the collagen IX interactome. J Biol Chem. 2013;288:13481-92.
[59] Ahsan T, Harwood F, McGowan KB, Amiel D, Sah RL. Kinetics of collagen crosslinking in adult bovine articular cartilage. Osteoarthritis Cartilage. 2005;13:709-15.
[60] Fernandes RJ, Schmid TM, Harkey MA, Eyre DR. Incomplete processing of type II procollagen by a rat chondrosarcoma cell line. Eur J Biochem. 1997;247:620-4.

FIGURE LEGENDS

Figure 1: Proteins found modulated in the soluble secretomes of the ATDC5 model. A. Gene Ontology focused on biological process of the significantly overrepresented terms in ATDC5 undergoing chondrogenic differentiation (FDR <0.005). Only the four highest overrepresented terms for biological process category are shown. Numbers embedded in the bars indicate the percentage of involved proteins within a term among all proteins within our data set. B. Time-dependent regulation of secreted proteins of iATDC5 versus cATDC5 are hierarchically clustered in the form of a two-dimensional heat map. Protein abbreviations are

according to the Uniprot database, excepted for collagen α-chains, which follow the collagen nomenclature (Supplemental data S.1). The scale bar indicates color coding in the range from one hundred-fold down-regulation to one hundred-fold up-regulation. Arrows pointing down or up indicate proteins found only in cATDC5 or iATDC5, respectively. Note that proteins found unregulated at any time point were not listed here. Proteins/genes previously reported *in vivo* [20], *ex vivo* [21] or *in vitro* [33] are highlighted in blue, grey and purple respectively. Lack of data is indicated by white boxes. Proteins implicated in at least one of four described biological processes are in bold. The whole output data are listed in Supplemental data S.3.

Figure 2: The overall ATDC5 phenotype revealed by guanidine extract analysis. A. Examination of the protein content of guanidine extracts obtained at major time points by SDS-PAGE coupled to Coomassie staining. The pattern shown here are representative of all replicates. This panel highlights the low content of extracts obtained at early time-points in cATDC5 and also serves as loading controls for panel C. B. The regulations of iATDC5 versus cATDC5 guanidine extracts are hierarchically clustered in the form of a two-dimensional heat map. Protein abbreviations are according to the Uniprot database, excepted for collagen α-chains which follow the collagen nomenclature (Supplemental data S.1). The scale bar indicates color coding in the range from one hundred-fold down-regulation to one hundred-fold up-regulation. Arrows pointing down or up indicate proteins found only in cATDC5 or iATDC5, respectively. Lack of data is indicated by white boxes. Note that proteins found unregulated at any time point were not listed here. The whole output data are listed in Supplemental data S.4. Proteins/genes previously reported *in vivo* [20], *ex vivo* [21] or *in vitro* [33] are highlighted in blue, grey and purple respectively. C. Examination by SDS-PAGE coupled to dMMB staining of the PG content of guanidine extracts obtained from iATDC5 at day 14 and 35, compared to chondroitin sulfate standards, guanidine extracts prepared from bovine articular cartilage as positive control and cATDC5 at day 35 as negative control, with or without proteinase K treatment.

Figure 3: Profound and biologically relevant regulation of collagen assembly by iATDC5 cells as revealed by CNBr extract analysis. A. Number of identified peptides (mascot score > 35) in CNBr extracts from iATDC5 at day 14. The corresponding output data are listed in Supplemental data S.5. B. Time-dependent regulations of CNBr extract contents during chondrogenic differentiation of ATDC5 cells. The regulation profiles of proteins found in CNBr extracts of iATDC5 versus cATDC5 are hierarchically clustered in the form of a two-dimensional heat map. The scale bar indicates color-coding in the range from one hundred-

fold down-regulation to one hundred-fold up-regulation. Arrows pointing down or up indicate proteins found only in cATDC5 or iATDC5, respectively. Lack of data is indicated by white boxes. Note that proteins found unregulated at any time point were not listed here. The corresponding output data are listed in Supplemental data S.6. C. βAPN fully inhibits collagen crosslinking. ATDC5 chondrogenesis was induced by insulin with or without βAPN, and cultures were analyzed for the production of covalently cross-linked ECM (bottom panel) and of non-crosslinked ECM (top panel) by SDS-PAGE coupled to Coomassie staining. Results shown here are from one replicate obtained from iATDC5 at day 21, but are representative of the whole experiment. The major content, as determined by mass spectrometry of tryptic peptides, of most prominent bands were annotated. In the bottom panel, annotations are according to the collagen cyanogen bromide peptide nomenclature.

Figure 4. Mass spectral evidence of type II procollagen maturation in iATDC5. (A) MALDI-TOF analysis of the LC fraction exhibiting maximal signal for a compound observed at m/z = 3016.37 at day 14 in the soluble secretome from iATDC5 (top panel) and its absence in the corresponding LC fraction obtained from cATDC5 (bottom panel). The m/z window was shorten to show the isotopic pattern. (B) TOF/TOF analysis of the major compound evidenced in A and mass spectral interpretation of the fragments, evidencing N-propeptide cleavage site. The fragmentation does not allow identifying the exact locations of hydroxylations, which could only be deduced from the fragmentations of the other peptide version, the precursor of which is seen in (A) at m/z = 3000.36, actually corresponding to the same peptide with 5 hydroxylations instead of 6. (C) TOF/TOF analysis identifying the C-propeptide cleavage site, also from the soluble secretome. (D) TOF/TOF analysis identifying a version containing hydroxylations and galactosyl-glucosylation at an internal triple-helical peptide (residues 786-818 according to the Uniprot precursor sequence), from the CNBr digest. In this case (and several others), the residue localization of hydroxylations could not be ascertained based on fragmentation only but was considered more likely at indicated residues based on their Y position in the GXY triplet. Bold residues indicate the positions of post-translational modifications.

Figure 5. PTMs of type II collagen synthesized by iATDC5. Residues are numbered according to their position in α1(II) precursor of mouse and bovine origin as annotated in the UniprotKB database. The PTMs annotated here for the bovine protein are according to [47] and serve as comparison to those found here in iATDC5. Residues annotated in grey correspond to peptides which were not identified. Residues annotated in italic correspond to

peptides which were identified only from soluble secretome analysis, the others being from CNBr extracts. Residues in red were found hydroxylated and those underlined were optionally modified. The prolines in green were found hydroxylated but not conserved between species. Yellow circles indicate lysine galactosylation and blue circles indicate additional glucosylation. Arrows indicate the theoretical proteolytic cleavage sites by N- and C-proteinases, respectively [48-50], respectively, the two of which were confirmed from soluble secretome analysis in the present study. The PTM residue positions shown here are those matching best peptide Mascot scores according to MSMS fragmentation. Many of the peptides were identified under several versions differing by the number of PTMs and there were cases of peptide versions for which the exact residue position of a given PTM could not be certified. In such cases, the PTM localization was considered more likely at a position found in other versions of the same peptide. The double arrows show a situation where the exact PTM localization could not be discriminated. The corresponding output data are listed in Supplemental data S.8 and Supplemental data S.9.

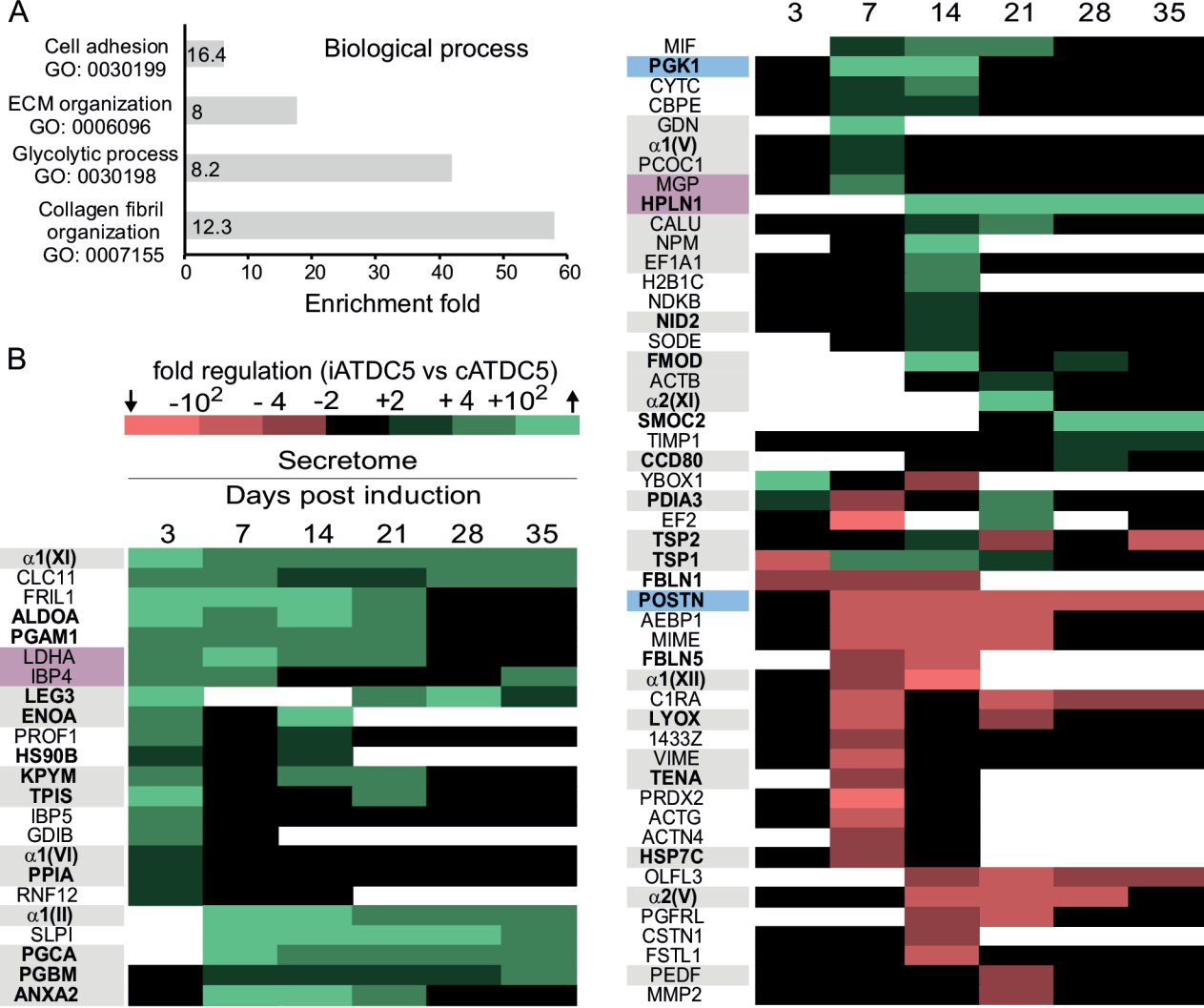

A
Biological process
Cell adhesion
GO: 0030199
16.4
ECM organization
GO: 0006096
8
Glycolytic process
GO: 0030198
8.2
Collagen fibril
organization
GO: 0007155
12.3
0
10
20
30
40
50
60
Enrichment fold
B
fold regulation (iATDC5 vs cATDC5)
-10^2
- 4
-2
+2
+ 4
+10^2
Secretome
Days post induction
3
7
14
21
28
35
α1(XI)
CLC11
FRIL1
ALDOA
PGAM1
LDHA
IBP4
LEG3
ENOA
PROF1
HS90B
KPYM
TPIS
IBP5
GDIB
α1(VI)
PPIA
RNF12
α1(II)
SLPI
PGCA
PGBM
ANXA2
3
7
14
21
28
35
MIF
PGK1
CYTC
CBPE
GDN
α1(V)
PCOC1
MGP
HPLN1
CALU
NPM
EF1A1
H2B1C
NDKB
NID2
SODE
FMOD
ACTB
α2(XI)
SMOC2
TIMP1
CCD80
YBOX1
PDIA3
EF2
TSP2
TSP1
FBLN1
POSTN
AEBP1
MIME
FBLN5
α1(XII)
C1RA
LYOX
1433Z
VIME
TENA
PRDX2
ACTG
ACTN4
HSP7C
OLFL3
α2(V)
PGFRL
CSTN1
FSTL1
PEDF
MMP2

Figure 1.

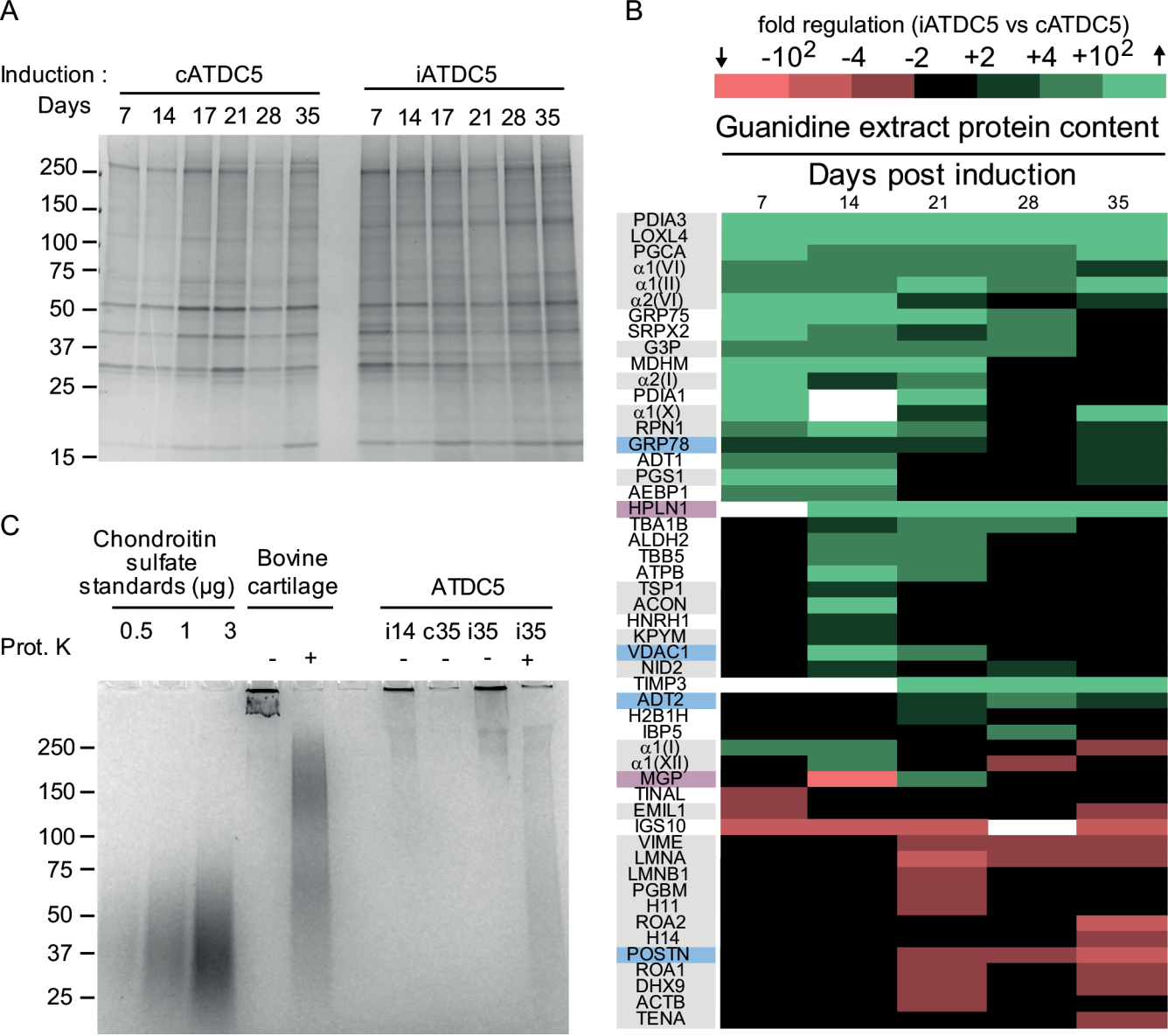

A
Induction :
cATDC5
iATDC5
Days
7 14 17 21 28 35
7 14 17 21 28 35
250
150
100
75
50
37
25
15
B
fold regulation (iATDC5 vs cATDC5)
-10^2 -4 -2 +2 +4 +10^2
Guanidine extract protein content
Days post induction
7 14 21 28 35
PDIA3
LOXL4
PGCA
α1(VI)
α1(II)
α2(VI)
GRP75
SRPX2
G3P
MDHM
α2(I)
PDIA1
α1(X)
RPN1
GRP78
ADT1
PGS1
AEBP1
HPLN1
TBA1B
ALDH2
TBB5
ATPB
TSP1
ACON
HNRH1
KPYM
VDAC1
NID2
TIMP3
ADT2
H2B1H
IBP5
α1(I)
α1(XII)
MGP
TINAL
EMIL1
IGS10
VIME
LMNA
LMNB1
PGBM
H11
ROA2
H14
POSTN
ROA1
DHX9
ACTB
TENA
C
Chondroitin sulfate standards (µg)
Bovine cartilage
ATDC5
0.5 1 3
i14 c35 i35 i35
Prot. K
- + - - - +
250
150
100
75
50
37
25

Figure 2.

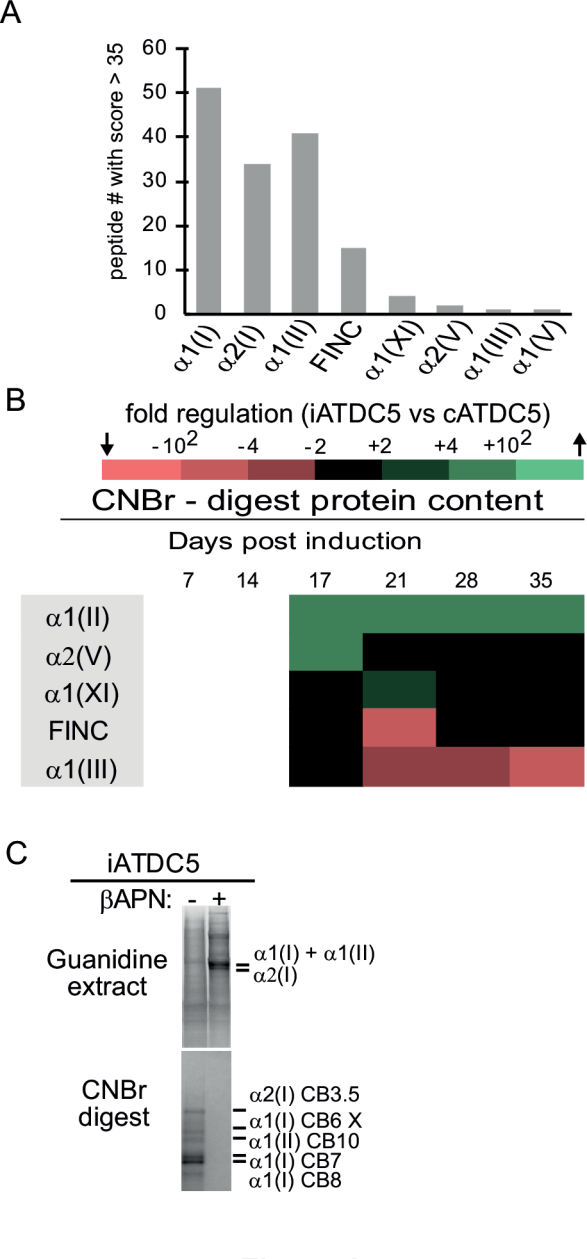

A
peptide # with score > 35
60
50
40
30
20
10
0
α1(I)
α2(I)
α1(II)
FINC
α1(XI)
α2(V)
α1(III)
α1(V)
B
fold regulation (iATDC5 vs cATDC5)
-10^2
-4
-2
+2
+4
+10^2
CNBr - digest protein content
Days post induction
7
14
17
21
28
35
α1(II)
α2(V)
α1(XI)
FINC
α1(III)
C
iATDC5
βAPN: - +
Guanidine extract
α1(I) + α1(II)
α2(I)
CNBr digest
α2(I) CB3.5
α1(I) CB6 X
α1(II) CB10
α1(I) CB7
α1(I) CB8

Figure 3.

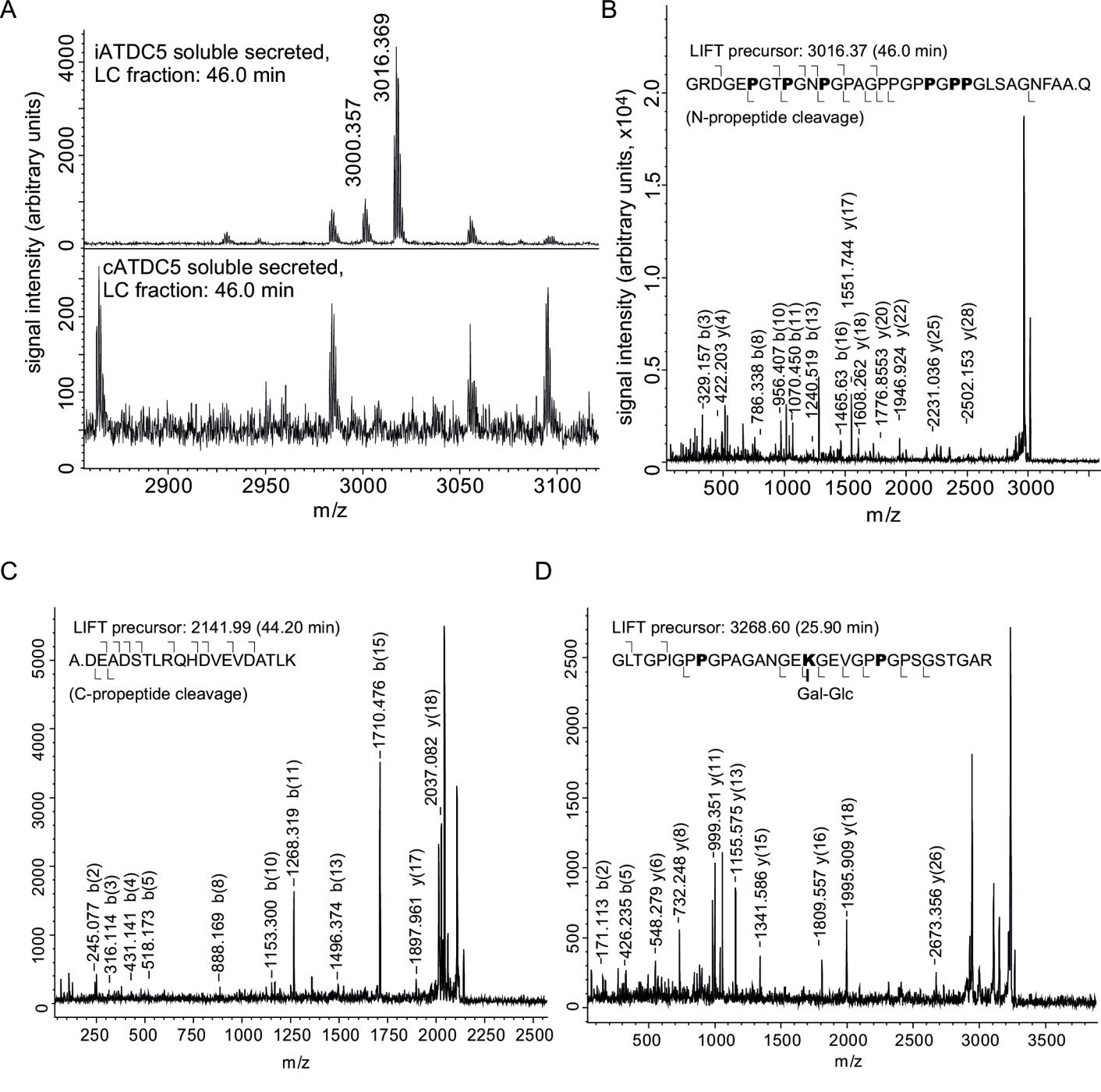
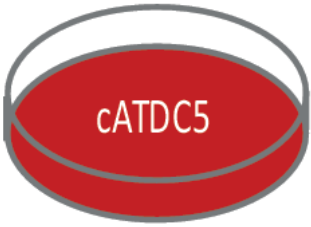

A
iATDC5 soluble secreted,
LC fraction: 46.0 min
3000.357
3016.369
cATDC5 soluble secreted,
LC fraction: 46.0 min
signal intensity (arbitrary units)
m/z
B
LIFT precursor: 3016.37 (46.0 min)
GRDGEPGTPGNPGPAGPPGPPGPPGLSAGNFAA.Q
(N-propeptide cleavage)
signal intensity (arbitrary units, x10⁴)
m/z
C
LIFT precursor: 2141.99 (44.20 min)
A.DEADSTLRQHDVEVDATLK
(C-propeptide cleavage)
m/z
D
LIFT precursor: 3268.60 (25.90 min)
GLTGPIGPPGPAGANGEKGEVGPPGPSGSTGAR
Gal-Glc
m/z

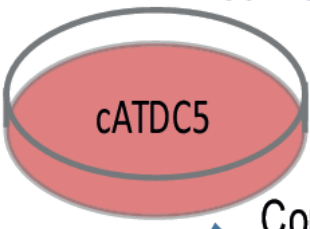
Figure 4.

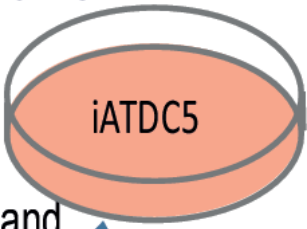
Figure 5.

Triplicate control and insulin-stimulated ATDC5 cells grown to day 3, 7, 14, 17, 21, 28, 35



switched to serum-free medium for 48 h





Controls and sequential extraction

Clarified secretomes

Triton - soluble

Guanidine - soluble

CNBr - digested

for each fraction at each time-point, precipitation, normalization, digestion and label-free quantification through LC-MALDI, iATDC5 vs cATDC5

Individual protein relative quant

COL2a1 PTM mapping

PG analysis