**Proteomic analysis of regulatory T cells reveals the importance of Themis1 in the control of their suppressive function**

Fanny Duguet1,2,3,4, Marie Locard-Paulet1,3, Marlène Marcellin1,3, Karima Chaoui1,3, Odile Burlet-Schiltz1,3, Olivier Andreoletti5, Renaud Lesourne2,3,4, Anne Gonzalez de Peredo1,3,§,&, Abdelhadi Saoudi2,3,4,§,&

1Institut de Pharmacologie et de Biologie Structurale, Universite de Toulouse, CNRS UMR 5089, 31077, Toulouse Cedex, France

2UMR Inserm, U1043, Toulouse, 31300, France

3UMR CNRS, U5282, Toulouse, 31300, France

4Université de Toulouse, UPS, Centre de Physiopathologie de Toulouse Purpan, Toulouse, 31300, France

5UMR INRA ENVT 1225, Interactions Hôtes Agents Pathogènes, Ecole Nationale Vétérinaire de Toulouse, 31000 Toulouse, France.

**Running title**: Label-free proteomics of regulatory T cells

&Contributed equally to this work

§**Corresponding authors:**

Abdelhadi Saoudi, e-mail: [abdelhadi.saoudi@ inserm.fr](mailto:abdelhadi.saoudi@%20inserm.fr) ; Phone: +33 562 744 512

Anne Gonzalez de Peredo, gonzalez@ipbs.fr; Phone +33 561 175 541

**Abbreviations**: Treg: regulatory T cells; Tconv: conventional CD4+ T cells; Foxp3: forkhead box P3; GFP: Green Fluorescent Protein; XIC: Extracted Ion Chromatogram; SDS PAGE: Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis; TFA: Trifluoroacetic Acid; CID: Collision Induced Dissociation; FDR: false discovery rate; GO: Gene Ontology; LN: lymph node; MS: mass spectrometry

**Keywords**: regulatory T cells, mass spectrometry, proteomics, label-free quantification, Themis1, TCR pathway

# **Abstract**

Regulatory T cells (Treg) represent a minor sub-population of T lymphocytes of paramount importance for establishing self-tolerance, controlling inflammatory responses and maintaining immune homeostasis. Here, we present a large-scale quantitative proteomic study that allowed defining a specific “signature” of Treg and identifying some candidate proteins that may contribute to their suppressive function. Treg and conventional T lymphocytes (Tconv) sub-populations were sorted by flow cytometry with high purity, and global proteomic analysis was performed by single-run nanoLC-MS/MS on a fast-sequencing Q-Exactive mass spectrometer. The chromatographic setup was optimized to improve peptide fractionation on 50cm columns, and increase the depth of proteome analysis to more than 3000 proteins, starting from low-numbers of primary cells. Label-free quantitative methods based on MS signal analysis were implemented for the comparison of protein expression profiles in Treg versus Tconv. Besides “historical” proteins that characterize Treg, our study identified numerous new proteins that are significantly up- or down-regulated in Treg versus Tconv. We focused on Themis1, a protein particularly under-expressed in Treg, and previously proposed as a new target gene that could participate in the pathogenesis of immune diseases. By over-expressing Themis1, we provided evidence for its functional importance in Treg through *in vitro* co-culture tests or animal models of inflammatory bowel disease. Thus, label free quantitative methods allowed better characterizing the Treg cell lineage, and deciphering the molecular mechanisms underlying the potential role of Themis1 in these cells.

**Introduction**

Regulatory T cells (Treg) are a subset of CD4+ T cells that are characterized by the expression of the transcription factor Foxp3. They play a central role in maintaining peripheral immune tolerance and preventing autoimmune diseases (Sakaguchi et al., 2008). This is best exemplified by the severe systemic autoimmunity and lymphoproliferative disease observed in Treg deficient Scurfy mice and in human IPEX patients carrying non-functional or hypomorphic alleles of the FOXP3 gene (Brunkow et al., 2001; Fontenot et al., 2003; Hori et al., 2003; Wildin et al., 2001). Furthermore, the quantitative or qualitative defect in Treg cells has also been implicated in the development of several common autoimmune and inflammatory diseases. In addition to the maintenance of immunological tolerance, Treg population can also be exploited to establish immunologic tolerance to transplanted tissues [Nadig, 2010 #134]. This has led to an increasing interest in the possibility of using Treg as a target for therapy to preserve and restore tolerance to self-antigens (in autoimmunity), to allergen (in allergy) and to alloantigens (in transplantation). However, this ambitious aim needs to take in account the fact that reaching an excessive Treg activity could coincidently impair immunity toward pathogens and tumors. Therefore, a better understanding of how Treg function, how they are activated/inactivated, and how their effector mechanisms work will be necessary to avoid potential negative side effects.

In this study, we compared the whole proteomes of CD4+Foxp3+ Treg (including both CD25+ and CD25- Treg) and CD4+Foxp3- T cells and performed functional studies on Themis1 as an interesting identified candidate. An important challenge in this context was to optimize proteomic analytical depth, starting from the low protein amounts obtained from highly purified primary murine Treg cells. In order to increase the analytical coverage of a whole cell proteome, a commonly used approach is to pre-fractionate the protein mixture and perform nanoLC-MS/MS analysis of each fraction separately. Although label-free workflows are sensitive to technical variability of upstream processing steps, 1D gel pre-fractionation by SDS-PAGE was previously shown to be compatible with label-free quantitation (provided proper integration and normalization of signal intensity across fractions and samples was performed) [Cox, 2014 #4;Gautier, 2012 #1], and it was successfully used in several large-scale label-free quantitative proteomics studies [Gautier, 2012 #1;Luber, 2010 #5]. Alternatively, improved LC conditions based on long LC columns operated on high-pressure chromatographic systems, with extended gradient times, were also described to improve peptide separation and were shown to successfully achieve deep proteome coverage in single-run analysis, without pre-fractionation [Nagaraj, 2012 #7;Thakur, 2011 #6]. We herein used such an approach, which allowed us to mine in-depth the Treg proteome and extensively identify the protein markers that discriminate this T cell population from conventional T cells (Tconv). We present here this specific signature of Treg cells, characterized at protein level. Comparison with transcriptomic data, as well as with the other recently published proteomic dataset on CD4+CD25+ T cells, allowed to define a very robust set of protein markers that characterize the Treg subset of T lymphocytes. Among all protein markers identified, Themis1 appeared as a protein particularly down-regulated in Treg cells, and was selected for further validation studies. We showed both *in vi*tro and *in vivo* that overexpression of Themis1 in Treg leads to an increase of their suppressive functions suggesting that this protein represents a checkpoint control of the suppressive function of Treg.

**Experimental procedures**

**Mice.** Treg cells were purified from DEREG mice expressing a diphtheria toxin receptor-enhanced green fluorescent protein (DTR-eGFP) fusion protein under control of the endogenous Foxp3 promoter [Lahl, 2007 #151] (kindly provided by Dr. Tim Sparwasser - Hannover Medical School, Germany). Transgenic mice expressing Themis1 under the control of the human CD2 gene promoter (Themis1-Tg) were used for functional studies (kindly provided by Dr Renaud Lesourne), and *Rag2-/-* mice. All the experiments were conducted with sex and age-matched mice between 6 and 10 weeks old housed under specific pathogen-free conditions at the INSERM animal facility (Zootechnie UMS-006; accreditation number A-31 55508), which is accredited by the French Ministry of Agriculture to perform experiments on live mice. All experimental protocols were approved by the local ethics committee and are in compliance with the French and European regulations on care and protection of the Laboratory Animals (EC Directive 2010/63).

**Antibodies and flow cytometry analysis.** The fluorescent conjugated antibodies used for flow cytometry were purchased from Biosciences and BD Biosciences: anti-TCR (H57-597), anti-CD4 (RM4-5), anti-CD25 (PC61), anti-CD62L (MEL-14), anti-CD45RB (16A), anti-CD39 (Duha59), anti-CD73 (…), anti-PD-1 (J43), anti-CTLA4 (UC10-4F10-11), anti-GITR (DTA-1) and anti-CD127 (SB/119). Intracellular staining was performed using the Foxp3 staining buffer kit (e-Biosciences) according to the manufacturer’s instructions before data acquisition on a BD LSRII flow cytometer and analysis with the FlowJo software.

**Purification and sorting** **primary mouse T cells**. Spleen and lymph nodes were harvested from DEREG mice and lysed in ACK buffer and washed twice. CD4+ T cells were enriched with the Dynal mouse CD4 cell negative isolation kit (Invitrogen) according to the manufacturer’s instructions. CD4+ T cells were then labeled with anti-TCR, anti-CD4 and anti-CD62L for FACS sorting of CD4+CD62LhighGFP- T cells (Tconv) or CD4+CD62LhighGFP+ T cells (Treg).

**Sample preparation for proteomic analysis**. CD4+CD62LhighGFP- and CD4+CD62LhighGFP+ cells were lysed in 4% SDS, 150 mM Tris pH8 and sonicated. Protein concentrations in each samples were determined using a detergent-compatible assay (DC assay, Bio-Rad) and total protein amounts were adjusted across samples. Cysteine residues were reduced by addition of 25 mM final of dithiothreitol for 5 min at 95°C, and alkylated by addition of iodoacetamide at a final concentration of 90 mM for 30 min at room temperature in the dark. For each protein sample, 10-15 ug were loaded onto 1D SDS-PAGE gel (stacking 4% and separating 12% acrylamide). For one-shot analysis of the entire mixture, no fractionation was performed, and the electrophoretic migration was stopped as soon as the protein sample entered the separating gel. The gel was stained with Coomassie Blue, and a single band, containing the whole sample, was cut. Alternatively, electrophoretic migration was performed to fractionate the protein sample into 9 gel bands. Gel slices were washed with two repeated cycles of 15 min incubations at 37 °C: 100 mM ammonium bicarbonate and 100 mM ammonium bicarbonate/acetonitrile (1:1). The proteins were then digested overnight at 37°C with 0.6 µg of modified sequencing grade trypsin (Promega) in 50 mM ammonium bicarbonate. The resulting peptides were extracted from the gel by incubation in 50 mM ammonium bicarbonate for 15 min at 37°C followed by two incubations in 10% formic acid/acetonitrile (1:1) for 15 min at 37 °C. The resulting peptides were dried in a SpeedVac, and resuspended with 17 µl of 5% acetonitrile, 0.05% TFA before being subjected to nanoLC-MS/MS analysis.

**NanoLC-MS/MS Analysis.** Peptides were analyzed by nanoLC-MS/MS using an UltiMate 3000 RSLCnano system (Dionex, Amsterdam, The Netherlands) coupled to an LTQ-Velos Orbitrap or a QExactivePlus mass spectrometer (ThermoScientific, Bremen, Germany). 5 µl of each sample were loaded onto a C-18 precolumn (300-µm inner diameter x 5 mm, Dionex) at 20 µl/min in 5% acetonitrile, 0.05% Trifluoroacetic Acid (TFA). After 5 min of desalting, the precolumn was switched online with the analytical C-18 column (75 µm inner diameter x 15 cm, PepMap C18, Dionex) equilibrated in 95% solvent A (5% acetonitrile, 0.2% formic acid) and 5% solvent B (80% acetonitrile, 0.2% formic acid). The peptides were eluted using a 5 to 50% gradient of solvent B at 300 nl/min flow rate, with gradient length as specified in the text. The mass spectrometer was operated in data-dependent acquisition mode with the XCalibur software. For LTQ-Velos Orbitrap runs, MS survey scans were acquired in the Orbitrap on the 300-2000 m/z range with a resolution of 60000 and the 10 most intense ions per survey scan were selected for CID fragmentation and analysis in the linear trap. For Q-ExactivePlus runs, survey MS scan were acquired in the Orbitrap on the 350-2000 m/z range with a resolution of 70000, the 10 most intense ions per survey scan were selected for HCD fragmentation and resulting fragments were analyzed at a resolution of xxx in the Orbitrap.

**Proteins identification and quantification.** Raw MS files were analyzed by MaxQuant version 1.5.2.8. Data were searched with the Andromeda search engine against Mouse entries of the Swissprot protein database (UniProtKB/Swiss-Prot Knowledgebase release 2014\_09, Mouse taxonomy, 16699 entries). The search included methionine oxidation and protein N-terminal acetylation as variable modifications, and carbamidomethylation of cysteine as a fixed modification. Validation was performed through a false discovery rate set to 1% for proteins and peptides determined by target-decoy search in MaxQuant (with a minimum length of 7 amino acids and a minimum Andromeda score of 20 as additional filters for peptide validation). Specificity of trypsin digestion was set for cleavage after K or R, and two missed cleavages were allowed. The precursor mass tolerance was set to 20 ppm for the first search and 4.5 ppm for the main Andromeda database search. The mass tolerance in MS/MS mode was set to 0.8 Da for LTQ-Velos data and 20 ppm for Q-Exactive data. For label-free relative quantification of the samples, the “match between runs” option of MaxQuant was enabled with a time window of 2 min, to allow cross-assignment of MS features detected in the different runs. The minimal ratio count was set to 1 for calculation of LFQ intensities.

**Statistical analysis of the proteomic data set.** (Marie action)

**Proliferation analysis by Cell Trace Violet staining.** For proliferation assays, Tconv and Treg populations were sorted based on the expression of the endogenous CD25 and CD62L markers. CD4+CD62LhighCD25- naive T cells purified from wild type (WT) mice were cultured in 96-well round-bottomed plates in the presence of CD4+CD62LhighCD25bright Treg cells purified either from WT mice or Themis1-Tg mice. To assess the number of cell divisions, CD4+CD62LhighCD25- T cells were initially resuspended at 10x106 cells/mL in phosphate-buffered saline (PBS), incubated with 2 mM of Cell Trace Violet (life Technologies) at 37°C, and washed in 5% FCS-PBS. Cell Trace Violet-labeled cells were cultured with CD4+CD62LhighCD25bright T cells at different ratio, in presence of 8 µg/mL of anti-CD3 (biolegend), and 12x104 irradiated splenocytes. The incorporation of Cell Trace Violet was measured by flow cytometry after 3 days of co-culture.

**Quantification of mRNAs by real-time PCR.** Total RNA was extracted using the RNeasy kit (Qiagen, Germantown, MD, USA) according to the manufacturer’s instructions. cDNA generated by SuperScript III (Life Technologies Ltd) were analyzed using primers for the indicated genes. Real-time PCR was performed using SYBR green. Results were normalized to ß2-microglobulin expression levels. The primers used were as follows:

Themis1 F: 5’-TGAAATCCAAGGTGTGCTGA-3’;

Themis1 R: 5’-CGTCCGTAGACAGCAACTGA-3’;

ß2m F: 5’–ACATACGCCTGCAGAGTTAAGCAT-3’;

ß2m R: 5’–CGATCCCAGTAGACGGTCTTG- 3’

qPCR was performed using LightCycler 480 (Roche).

**Immunoblot analysis.** Total cellular proteins from purified Tconv and Treg were extracted by the addition of a twice-concentrated lysis buffer (2% *n*-dodecyl-β-maltoside, 10% glycerol, 100 mM Tris [pH 7.5], 270 mM NaCl, 1 mM EDTA, [pH 8]) supplemented with protease and phosphatase inhibitors. After 10 min of incubation on ice, cell lysates were centrifuged at 20000 *g* for 15 min at 4 ̊C. Eluates were completed with laemmli buffer and resolved by SDS-PAGE followed by Western Blotting on Immobilon-P membranes (Millipore) using Themis1 antibody (Millipore) as well as Gapdh (Sigma) for loading control. Immunoreactive bands were detected by chemiluminescence with the SuperSignal detection system (Pierce Chemical, Rock-ford, IL).

**Induction and evaluation of colitis.** CD4+ T cells were enriched from spleen and lymph node cell suspensions using a CD4 negative selection kit (Dynal) before staining with anti-TCR, anti-CD4, anti-CD25 and anti-CD45RB mAbs. CD4+CD25-CD45RBhigh T cells and CD4+CD25bright T lymphocytes were then separated by fluorescent cell sorting. Colitis was induced in 6-week-old *Rag2-/-* mice by intravenous injection of 4x105 naive CD4+CD25-CD45RBhigh WT T cells in 100 µL of PBS. In the cotransfer experiments 2x105 or 105 CD4+CD25bright from WT or Themis1-Tg mice were cotransferred with 4x105 naive CD4+CD25-CD45RBhigh WT T cells. Macroscopic colonic tissue damage was evaluated using a scale ranging from 0 to 11 as follows: erythema (0 to 2), edema (0 to 2), strictures (0 to 3), ulceration (0 or 1), mucus (0 or 1), and adhesion (0 to 2). Colon wall thickness was measured with an electronic caliper. Colonic tissue specimens were excised 2 cm proximal to the cecum and immediately transferred into 10% formaldehyde to be embedded in paraffin. Five-micrometer colonic sections were then stained with haematoxylin and eosin. Damage was evaluated on a scale ranging from 0 to 8 (Olivier action). Cellular infiltration and mucosal alteration were graded from 0 to 3 (absent, mild, moderate, and severe), and submucosal edema was scored from 0 to 2 (absent, moderate, and severe).

**Data analysis.** Data are expressed as mean ± s.e.m. The GraphPad Instat statistical package was used for statistical analyses (GraphPad Software, Inc., La Jolla, CA USA). Results were compared using Mann–Whitney test. Results were considered statistically significant when the p-value was <0.05. \*: p<0.05; \*\*: p<0.01; \*\*\*: p<0.001.

**Results**

**Purification and mass spectrometry analysis of regulatory and conventional T cell subpopulations**. Originally, CD4+ Treg cells were identified by the constitutive expression of CD25 (ref). However, both the existence of CD25- Treg cells (ref) as well as the expression of CD25 on activated Tconv (ref) limits the interpretation of data obtained by using CD25 as marker of Treg. The most unambiguous Treg cell marker currently known is Foxp3, which is expressed specifically in mouse CD4+ Treg cells. To define a specific “signature” of the Treg subpopulation at proteomic level, we used the DEREG mouse model expressing a diphtheria toxin receptor-enhanced green fluorescent protein (DTR-eGFP) fusion protein under control of the endogenous Foxp3 promoter [Lahl, 2007 #151]. In this model, Treg specifically express DTR-eGFP, allowing fluorescent detection and very efficient purification of Foxp3+ Treg cells by FACS sorting independently of CD25 expression. Post-sorting verification by intracellular staining of Foxp3 indicated a purity >99% in Foxp3+ cells among all sample preparations (**Fig. S1)**. Although selection through the GFP marker increases the recovery yield of Treg cells, this population represents only a very minor percentage of the total CD4+ T cells, and we typically isolated around 150-200 x 103 Treg cells per mouse. In order to optimize protein extraction, we lysed the cells with a strong detergent concentration (4% SDS) combined to sonication. This led to 5 to 10 µg of proteins per sample. Total protein amounts from Treg and Tconv cell populations were then adjusted, samples were trypsin-digested and analyzed in parallel by mass spectrometry using the label-free quantitative analytical workflow described in **Fig. 1A**. .**Fig. 1B** illustrates the different experimental workflows we tested in order to increase our coverage of the Treg and Tconv proteomes: (i) a direct, one-shot analysis of the total peptide mixtures obtained from whole cell lysates in routine LC-MS conditions using a 15 cm nanoLC column and 120 min-long gradient time on a LTQ-Velos instrument (see Experimental procedures); (ii) a workflow including an additional pre-fractionation step through 1D SDS-PAGE separation; (iii) a workflow based on the one-shot analysis of total peptides using an in-house packed 50 cm column and a longer gradient (480 min) on the LTQ-Velos. While the first workflow only allowed the identification of less than 2000 proteins, the two latter experimental designs clearly improved the proteomics depth to >3000 proteins **(Fig1B**). The identification gain associated with the fractionation workflow was smaller (maybe due to the low amounts of starting material), and better results were obtained with single-run analysis and a longer gradient. In addition, total analytical time was much reduced with the latter workflow, which was therefore selected for the comparative analysis of Tconv and Treg proteomes. **Fig. 1C** summarizes the analysis conditions of 7 independent sample preparations of Treg and Tconv cells populations, processed and trypsin-digested as described above, and analyzed by single-run LC-MS/MS on a 50 cm reversed-phase column. The use of a Q-Exactive Plus mass spectrometer for 5 of these biological replicates allowed us to shorten the gradient to 300 min while retaining an analytical depth comparable to the one obtained on a LTQ-Velos instrument with lower sequencing speed (data not shown). All analytical runs (biological and up to 3 technical replicates from the 2 cell types - a total of 28 long gradient LC-MS runs -) were simultaneously processed through MaxQuant, which performed realignment and LFQ normalization of MS intensity signals for all identified proteins. Database search with Andromeda (1% FDR validation at peptide and protein level) led to the identification of around 3500 protein groups in each biological experiment, and finally to more than 55000 peptides and 4346 proteins groups identified in the whole study. The dataset has been deposited to the ProteomeXchange Consortium [Vizcaino, 2014 #8] via the PRIDE partner repository with the dataset identifier PXDxxxxx.

**Identification of Treg markers***. Because we specifically purified primary Treg cells based on their Foxp3 expression, our data set is not biased by the presence of CD25+ Tconv cells. For the same reasons, our population of Tconv cells is free of CD25- Treg cells that can be present in previous similar analysis. This allowed us* tto identify proteins specifically up- or down-regulated in Treg compared to Tconv cells and generate an updated list of Treg cell markers.

A two-sided paired t-test was performed on the 7 biological replicates and proteins presenting a Benjamini-Hochberg adjusted p-value under or equal to 0.05 and a minimum 2-fold change between Treg and Tconv were considered significantly different between the two cell populations. According to these criteria, we identified a total of 168 differentially expressed proteins, including 117 proteins more abundant in the Treg and 51 proteins less abundant in Treg compared to Tconv (**Fig. 2A)**. . The most differentially regulated proteins between Treg and Tconv are presented in the **Figure 2B**. Most of the “historical” markers commonly utilised to characterize Treg cells are found amongst the 40 most up-regulated proteins in Treg cells compared to Tconv cells (log2-transformed ratio between Treg and Tconv of 2 to 6) (**Fig. 2C).** These included firstly FOXP3, the master transcription factor regulating Treg phenotype and function [Fontenot, 2003 #58]; The interleukin-2 receptor subunit alpha (IL2RA/CD25) is commonly used as a specific surface marker for detection and purification of naive Treg cells [Sakaguchi, 1995 #10;Dieckmann, 2001 #13]. We also identify the zinc finger proteins Helios (IKZF2) [Thornton, 2010 #52] and Eos (IKZF4) [Pan, 2009 #16]; the tumor necrosis factor superfamily receptors Tnfrsf4/OX40 [Vu, 2007 #17] and Tnfrsf18/GITR [McHugh, 2002 #10]; and cell surface proteins such as Neuropilin1 (NRP1) [Sarris, 2008 #18], P2X purinoceptor 7 (P2RX7) [Hubert, 2010 #19], the Folate receptor 4 (Folr4/JUNO), and the ecto-5'-nucleotidase (NT5E/CD73) [Deaglio, 2007 #15].

**Fig3A** illustrates schematically the proteins showing expression changes in CD4+Foxp3+ Treg compared to CD4+Foxp3- Tconv, as characterized in the present study, and their broad functional classification based on GO annotation. We could detect many cell surface membrane proteins up-regulated in Treg, the most induced ones being the classical markers quoted above (CD25; P2X7; OX40; JUNO; GITR; CD73). In addition to those, integrin beta-1 (CD29) was also found to be strongly up-regulated, as well as a panel of cell surface molecules with smaller, but significant expression ratios including SLAF1/CD150, FAS/CD95, CD2, CD44 and CD48. Regarding nuclear proteins, in addition to the transcription factors Foxp3, Helios and Eos, we also detected the DNA-binding factors IRF4, the NF-kappa-B subunits REL and NFKB2/p52 as over-expressed in Treg. Conversely, the transcription factors SATB1, TCF7, LEF1 and Ikaros were found down-regulated. Other intracellular molecules involved in various functions such as cytoskeleton organization, metabolic processes, cell cycle control or apoptotic processes were also found as differentially expressed in the two cellular subsets of CD4+ T cells. Markedly, only a few proteins involved in TCR signaling showed expression changes, including the actin-uncapping protein Rlptr (up-regulated in Treg), and also the TCR signaling molecules ITK and Themis (both down-regulated in Treg). Besides the molecules already described in the literature to be involved in Treg commitment or function, our study identified new proteins with a marked different expression level in Treg compared to Tconv, and which function was not characterized in these cells. Some examples of such uncharacterized markers illustrated respectively in **Fig3B** and **3C**.

**Overexpression of Themis1 by Treg enhanced their suppressive functions.** As shown above, Themis was one of the most strongly down-regulated proteins in Treg, and one of the few proteins involved in the TCR pathway showing a modified expression level in these cells. Because differences in TCR signaling have been characterized before between Tconv and Treg cells, and may be important for Treg development and function, we wondered whether the specific down-regulation of Themis in the CD4+Foxp3+ subset may have functional relevance. In agreement with LC-MS analyses (**Fig. 4A)**, we confirmed by immunoblot that Treg express lower levels of Themis1 protein than Tconv (**Fig. 5B)**. In addition, we showed that Treg express low amount of Themis1 mRNA as compared to Tconv demonstrating a transcriptional regulation of this gene in Treg (**Fig. 5C)**. To examine the functional impact of Themis1 expression on Treg development and suppressive functions, we used Themis1-Tg mice that overexpress Themis1. Of note, Treg from Themis1-Tg mice express high but similar levels of Themis1 as WT Tconv (**Fig. 5B)**. We compared by flow cytometry the frequency, the absolute numbers and the phenotype of Treg between WT and Themis1-Tg mice and showed that the overexpression of Themis1 has no major impact on the development of Treg in the thymus (**Fig. 5D)**, on the homeostasis of Treg in the spleen (**Fig. 5E)** and on the expression of known markers that characterize Treg such Foxp3, CD25, CD39, CD73, CD103, PD-1, CTLA-4, GITR and CD127 (**Fig. 5F)**.

To examine the effect of Themis1 overexpression on the immunosuppressive function of Treg cells *in vivo*, we used a well-defined adoptive transfer model, which involved cotransfer of CD45RBhigh naive CD4+ T cells with Treg cells purified from the WT or Themis1-Tgmice (**Fig. 6A)**. First, we controlled that Themis1 overexpression had no impact on the percentage of Foxp3 expressing cells among CD4+CD25bright sorted Treg, nor on the level of Foxp3 expression (**Fig. S2)**. As expected, adoptive transfer of CD4+CD25-CD45RBhigh naive T cells into *Rag2*-/- mice induced severe signs of colitis and weight loss, starting from week 3 after transfer (**Fig. S3A and B)**. This was associated with pronounced inflammatory cellular infiltration in the mucosa and submucosa, as well as significant epithelial destruction (**Fig. S3C)**. These pathological phenotypes were efficiently blocked when the naive T cells were co-transferred with WT Treg cells or Themis1-Tg Treg at optimal ratio of 1:2 between Treg and naive T cells (**Fig. S3)**. To reveal a possible difference in the suppressive capacities of WT Treg and Themis1-Tg Treg, we chose a suboptimal 1:4 ratio between Treg and naive T cells, that does not prevent colitis, but induces a certain degree of inflammation suitable to detect differences in the suppressive capacity of Treg subsets. In this setting, we showed that Themis1-Tg Treg cells exhibited an enhanced ability to suppress the pathological activity of the naive T cells compared with the WT Treg cells at macroscopical level (**Fig. 6C)**. This observation could be due to enhanced numbers of Themis1-Tg Treg cells in the recipient mice and/or enhanced suppressive activity of these cells. To directly assess the immune-suppressive activity of Themis1-Tg Treg cells, we used an *in vit*ro suppression assays using Treg cells purified from the spleen and LN of WT and Themis1-Tgmice (**Fig. 6D).** Themis1-Tg Treg cells showed an enhanced ability to suppress TCR-stimulated proliferation of coexisting CD4+ T cells at different ratio (**Fig. 6E and 6F)**. Collectively, these data demonstrate that overexpression of Themis1 in Treg enhanced their suppressive functions both *in vitro* and *in vivo* suggesting that this protein represents a checkpoint control of the suppressive function of Treg.

**Discussion**

Due to the technical improvements of mass spectrometer devices (higher sensitivity, dynamic range, sequencing speed), but also to the optimization of upstream separative methods (liquid nanochromatography), shotgun proteomic approaches applied to the large-scale analysis of complex proteomes are continuously progressing in terms of proteomic depth and accuracy of relative quantification. They have been successfully applied in recent years in the field of immunology to characterize cellular subsets, immune cell differentiation, or specific responses following stimulation, on different cell types such as dendritic cells (Luber et al, Immunity 2010, PMID: 20171123; Schlatzer et al, J Immunol Methods. 2012 PMID: 21945394), macrophages (Becker et al, PlosOne 2012PMID: 22428014), natural killer cells (Scheiter et al, MCP 2013, PMID: 23315794) or cytotoxic T cells (Hukelmann et al, Nat Immunol 2016, PMID: 26551880). Here, we used such methods to elucidate the proteomic composition of two distinct immunological subsets of CD4+ T cells, i.e. Foxp3 +/ CD4+ T cells and Foxp3 -/ CD4+ T cells, with very different functional properties. Using a chromatographic setup optimized for peptide fractionation on 50cm columns, we could identify more than 3000 proteins/sample, starting from unfractionated lysates obtained from low-numbers of cells. This experimental workflow involving minimal sample processing and fractionation provided good reproducibility and was compatible with the use of label-free quantitative methods based on MS signal analysis. Moreover, it could be implemented in relatively short time, which allowed the analysis of 7 biological replicates to strengthen the statistical validation of proteomic variations. We could globally quantify more than 4000 proteins in the two cellular subsets, a depth enabling the detection of most classical protein markers already used to characterize Treg cells. Through label-free quantification, we could readily measure the differential expression pattern of these markers, being either cell surface proteins (CD25; Nt5e/CD73; P2X7; Tnfrsf4/OX40; Folr4/JUNO; Tnfrsf18/GITR/CD357) or nuclear proteins (Foxp3, Helios/IKZF2, Eos/IKZF1), which confirmed the efficiency of the quantitative approach. In addition, we could characterize significant variation of the expression level of a large number of additional proteins, defining a very specific signature of the Treg subset of T cells. Some of them have been already described as playing a role in Treg function, consistent with their specific up or down regulation in these cells. This is the case for example of Rlptr, an actin-uncapping protein that was recently identified through a mutagenesis screen and was shown to be essential for costimulation of the TCR via CD28, possibly due to its ability to couple CD28 to PKC-θ and Carma1 (PMID: 23793062). Interestingly, the development of thymic Treg cells was severely impaired in mice bearing a mutation on Rlptr, a phenotype reminiscent to that of Cd28−/− mice. In addition, Rlptr was reported to be downregulated in biopsies from psoriatic patients (PMID: 15588584), and polymorphisms in Rlptr are associated with ankylosing spondylitis in humans (PMID: 19854717), suggesting that it may play a role in autoimmunity. Consistent with the fact that Rlptr is up-regulated in Treg (this study), these data indicate that this protein may participate in the development and function of regulatory T cells. Interestingly, other proteins known to be involved in the regulation of actin structure organization where found to be either up-regulated in Treg (Coro2a, Capg) or down-regulated (Fam101b/RefilinB,…), suggesting a particularly important role in these cells of the mechanisms driving cytoskeleton architecture and actin remodeling, which are known to be actively taking part in TCR signalling.

Other proteins with no known function in Treg were also characterized with strong differential expression ratios. In most cases, little functional information is available on these proteins, particularly in the context of T cell biology. Among them, we found for example Fam129a/Niban, Gsto1, Ladinin1 as highly up-regulated in Treg, while Pdlim4 or Tdrp were strongly down-regulated.

One noticeable marker is the protein Niban/Fam129a, which showed the most pronounced expression ratio in the Treg/Tconv comparison and was consistently detected in all Treg samples…….

This may open the way to further functional characterization of these proteins and of their role in the suppressive function of Treg.

In the course this study, two other proteomic studies on regulatory T cells were reported, and we thus compared our data to these recently published datasets. Starting from FACS sorting based on the CD25 marker, Barra et al performed a comparative study of CD4+/CD25+ and CD4+/CD25- murine T cells, using isotopic dimethyl labeling and separation of peptides with offgel isoelectrofocusing (PMID: 26324778). Using this pre-fractionation approach, they could identify about 5000 proteins from 4 biological experiments, and quantify more than 4000 of them in at least 2 replicates. On the other hand, Procaccini et al performed a study on human Treg cells isolated from human blood, and identified about 1900 proteins after MudPit analysis, which were compared with those found in Tconv cells using an in-house method for semi-quantitative analysis (PMID: 26885861). This later study reported differences mainly related to cellular metabolism, with an upregulation of glycolysis-related enzymes in freshly-isolated human Treg cells (e.g; GAPDH, PGK1, TALDO1 or ALDOA), while Tconv appeared to express higher amounts of mitochondrial proteins (e.g.: IDH2, ACO2, CS, ETFA, VDAC1, SLC25A3, ATP5F1, …). In our study, we did not detect such major metabolic changes in the two subsets of CD4+ T cells isolated from mouse, with most glycolytic and mitochondrial enzymes and transporters quantified with a ratio close to 1 in the comparison of Treg and Tconv performed here. This may possibly be due to the different nature and origin of the Treg samples analyzed, which may present discrepant metabolic status, or alternatively to the differences in the techniques of protein analysis and quantification used in the two studies. Conversely, we could recapitulate in CD4+ /Foxp3+ murine cells many protein variations described by Barra et al in CD4+/CD25+ murine cells. Particularly, among proteins showing the most pronounced down-regulation ratio in Treg, we indeed detected the transcription factor Tcf7 that was selected by these authors as a promising candidate and shown to play a role for restricting the differentiation into the Treg lineage during thymic development. In order to get an integrated view of of these two proteomic datasets, we performed a systematic comparison of the protein lists and associated quantitative ratios measured in the two studies. In addition, we also compared these data with previous transcriptomic results reported for murine Tconv/Treg comparison, in order to pinpoint possible differences between protein and mRNA data, but also to highlight the most robust and characteristic markers of the Treg phenotype. In order to avoid selection biases due to the statistical methods used in the different studies to call a protein differentially expressed between Treg and Tconv, we applied a simple selection criteria based on a 2-fold expression change between the two subsets, and retrieved the proteins quantified as variant in at least 2 of the 3 considered studies (proteomic dataset 1, this study, 7 biological replicates; proteomic dataset 2, Barra et al, 4 biological replicates; transcriptomic dataset, 3 biological replicates). Altogether, this comparison (Sup data x) allowed obtaining a synthetic representation of expression changes at protein and mRNA level in Treg. Through the compilation of the results obtained for several biological replicates and different methods of Treg isolation, it was possible to identify the markers that are very consistently identified in all studies and many replicates, and should be considered as strong phenotypic markers, and possibly, as promising candidates for functional studies when their role in Treg is not described yet. In addition, the integration of these results also allowed to rescue some interesting candidates that would not be identified or not considered as statistically significant a particular single study.

**Figures legends**

**Figure 1**: **Extensive label-free quantitative analysis of conventionnal and regulatory T cells proteome.** (A) General analytical workflow based on cell sorting by flow cytometry using the DEREG mouse model and parallel proteomic analysis of Tconv and Treg cell populations by nanoLC-MS/MS and label-free quantification. (B) Test of different sample pre-fractionation and nanoLC procedures in order to increase proteomic depth starting with low amounts of protein material. Samples were analyzed on a LTQ-Velos Orbitrap instrument, and proteins were identified through database search with Andromeda under MaxQuant. Validation of identifications was performed at 1% FDR at peptide and protein level. (C) Protein identification results obtained for 7 independant biological experiments using workflow 3 (single run analysis on either LTQ-Velos or Q-Exactive orbitrap instruments, long gradient times).

**Figure 2**: **Proteomic analysis of Foxp3+ Treg cells***.* (**A**) Volcano plot (-log10(p-value) versus log2(fold change)) showing proteins up-or down-regulated in Treg compared to Tconv (results from 7 biological experiments). Proteins with absolute fold change≥2 and BH adjusted p-value≤0,05 are shown in yellow. Proteins considered as well-characterized markers of Treg are shown in red (up-regulated proteins) and blue (down-regulated proteins). (**B**) Heat map of intensities for proteins respectively up-regulated (upper map) or down-regulated (lower map) in Treg compared to Tconv. LFQ intensities retrieved from MaxQuant were respectively plotted for the 7 replicates Tconv (left) and Treg (right) samples. Proteins are ranked by decreasing mean intensity in ??? (**C**) Histogram illustrating the fold change values calculated between Treg and Tconv for the 4166 proteins identified and quantified in the whole experiment (grey). Known Treg markers (up- or down-regulated proteins, indicated in red and blue respectively) are found among the proteins with most extreme fold changes, together with new proteins with uncharacterized role in Treg. The 50 proteins with most extreme fold changes and with BH adjusted p-value ≤ 0,05 are represented in the enlarged views.

**Figure 3**: **Schematic representation of the Treg proteomic signature**

**Anne and Marie action**

**Figure 4**: **Comparison of transcriptomic and proteomics Treg datasets.**

**Anne and Marie action**

**Figure 5**: **Overexpression of Themis1 does not affect the development and the phenotype of Treg at steady state***.* **(A)** Themis1 quantification analyzed by using LFQ intensity in Treg and Tconv. **(B)** Immunoblot analysis of Themis1 and Gapdh expression by Treg and Tconv isolated from WT mice and Themis1-Tg mice (left panel); densitometric quantification of those results (right panel). **(C)** Themis1 mRNA expression by Treg and Tconv presented as relative ratio to ß2-microglobulin. (**D**) Thymus and (**E**) spleen were collected from WT mice (n=5) and Themis1-Tg mice (n=7). Frequency of Foxp3+ among CD4+ T cells, absolute numbers of CD4+ Foxp3+ regulatory T cells and median fluorescence intensity (MFI) of Foxp3 in Treg from these tissues were calculated. (**F**) Expression of selective markers on CD4+ Foxp3+ T cells in the spleen from WT mice and Themis1-Tg mice (n=4 per group). ns, non significant. Data are representative of seven (**A**), two (**B**) or five independent experiments (**C**). \*\**p ≤0.01;* \*\*\**p ≤0.001.*WT mice (WT) and Themis1-Tg mice (Tg).

**Figure 6: Overexpression of Themis1 improves the suppressive function of Treg***.* (**A**)CD4+CD25-CD45RBhigh naive T cells were sorted from WT mice and injected into *Rag2*-/- mice to induce colitis. In addition, CD4+CD25bright T cells (Treg) were sorted either from WT mice or Themis1-Tg mice and were cotransferred in a ratio of 1:4. (**B**) Mice were monitored for their body weight for 8 weeks. (**C**) The intensity of disease was assessed 8 weeks after T-cell transfer by measuring macroscopic appearance and wall thickness of the colon and histological analysis of tissue damage. Data are representative of three independent experiments with 8 to 10 mice per group. \**p ≤0.05*; \*\**p ≤0.01;* \*\*\**p ≤0.001.* (**D**) Suppressive activity of CD4+CD25bright T cells (Treg) from WT and Themis1-Tg mice was assessed in coculture experiments using Cell Trace Violet (CTV)-labeled naive Tconv from WT mice as effector cells stimulated with plate-bound anti-CD3 and irradiated syngeneic APC. After 3 days of coculture, Tconv proliferation is assessed by CTV dilution. (**E**) Representative histograms of the proliferation of Tconv alone or in the presence of the indicated ratios of Treg from WT or Themis1-Tg mice. Percentages indicate the proportion of CTV low effector cells. Results are representative of one among seven independent experiments (**D**) Suppressive functions of Treg from WT or Themis1-Tg mice at different ratios. Data are expressed as percentage of inhibition and presented as mean values ± SEM obtained from seven experiments. \**p ≤0.05*; \*\**p ≤0.01;* \*\*\**p ≤0.001.*

**Supplementary Figure legends**

**Figure S1**: **Purification of Treg and Tconv from DEREG mice.** CD4+ T cells were isolated from lymph nodes and spleen of DEREG mice and cell-sorted on the basis of GFP expression: Treg are CD4+GFP+ cells (P2) and Tconv are CD4+GFP- cells (P1). Post-sort P1 and P2 cell populations were controlled for intracellular expression of Foxp3by flow cytometry.

**Figure S2: Analysis of Foxp3 expression by purified naïve CD4 T cells and Treg.** (**A**) CD4+CD25-CD45RBhigh naïve WT T cells and CD4+CD25bright WT and Themis1-Tg Treg cells were FACS-sorted and controlled for intracellular Foxp3 expression by flow cytometry. (**B**) Mean fluorescence intensity (MFI) of Foxp3 in Treg isolated from WT and Themis1-Tg mice

**Figure S3:** **Analysis of the suppressive function of WT and Themis1-Tg Treg *in vivo***. CD4+CD25-CD45RBhigh naïve T cells were sorted from WT mice and injected into *Rag2*-/- mice with or without CD4+CD25bright T cells (Treg) purified either from WT mice or Themis1-Tg mice at a classical ratio of 1:2. Inflammatory bowel disease development was monitored using body weight lost (**A**). The intensity of disease was assessed 6 weeks after T-cell transfer by measuring macroscopic appearance and wall thickness of the colon (**B**)and histological analysis of tissue damage (**C**). The results are from one experiment including 6 to 10 mice per group. \*\**p ≤0.01.*

**Figure S4:** **Overexpression of Themis1 has no impact on the expression of selective marker of Treg upon their activation.** CD4+CD25bright T cells (Treg) were stimulated with plate-bound anti-CD3 and irradiated syngeneic APC for 3 days and expression of selective markers were assessed on CD4+ Foxp3+ T cells from WT mice and Themis1-Tg mice (n=5 per group). ns, non significant. MFI: Mean fluorescence intensity

**Acknowledgements:**

We acknowledge Drs. Anne Dejean and Daniel Gonzalez-Dunia for their valuable inputs on the manuscript. The authors acknowledge the technical assistance provided by Fatima-Ezzahra L’Faqihi-Olive, Valérie Duplan-Eche and Anne-Laure Iscache from the CPTP Cytometry facility; Florence Capilla and Christine Salon from Inserm US006 hitstopathological platform, and the personnel of Inserm US006 Anexplo/Crefre animal facility.

This work was supported by INSERM, Association Française contre les Myopathies and Région Midi-Pyrénées. FD was supported by grants Région Midi-Pyrénées and ARSEP. AS is supported by Centre National de la Recherche Scientifique.

**References**